Journal of Indian School of Political Economy

Vol. XXXIV, Nos. 3-4
July-December 2022

A Journal devoted to the Study of Indian Economy, Polity and Society

INDIAN SCHOOL OF POLITICAL ECONOMY

JOURNAL OF INDIAN SCHOOL OF POLITICAL ECONOMY is devoted to the study of the Indian Economy, Polity and Society. Emphasis is primarily on reviewing developments since Independence with roots in the British administration where relevant. However, papers with a similar focus but not necessarily reviewing developments since Independence will also be considered. When a review is based on statistical data, full statistical base data are presented as far as possible.

ANNOUNCEMENT

Correspondence regarding subscription to the Journal, etc., should be addressed to the Director, Indian School of Political Economy, 'Arthabodh', 968/21-22, Senapati Bapat Road, Pune-411 016 (India). e-mail: ispepune@gmail.com. Details regarding subscription rates are given on the backside of the Contents page.

Editorial communications should continue to be sent to the Editor, *Journal of Indian School of Political Economy*, at the above address. Comments on articles and documentation appearing in the Journal are welcome.

JOURNAL OF INDIAN SCHOOL OF POLITICAL ECONOMY

A Journal Devoted to the Study of Indian Economy, Polity and Society

Vol. XXXIV **July-December 2022** **Nos. 3&4**

Annual Subscription Rates

Journal of Indian School of Political Economy, Pune

New annual subscription rates effective beginning with issues of Volume XXVII (2015) of the Journal:

	India (Rs)	**Other Countries (US $)**
For Universities and Institutes other than Colleges	1000.00	100.00
For Individuals and Colleges	500.00	50.00

Payments should be made by Demand Draft or cheques, payable to *Indian School of Political Economy, Pune,* payable at par in Pune.
Please mail your orders/enquiries to

Indian School of Political Economy,
'Arthabodh', 968/21-22, Senapati Bapat Road,
Pune 411 016 (India). Phones: (020) 25657132, 25657210. e-mail: ispepune@gmail.com

Editor's Note: University and college teachers and students of Economics, Political Science and Sociology/Social Anthropology are invited to send to us questions of wider interest on the subjects of their study. We shall endeavour to publish in subsequent issues of the journal answers to selected questions received by us. This will form a separate section of the journal. The authors of the selected questions shall receive complimentary copies of one year's issues of the journal, as a token of our appreciation.

ANNOUNCEMENT

INDIAN SCHOOL OF POLITICAL ECONOMY,

is pleased to announce the following scheme recently introduced for subscription of the ***Journal of Indian School of Political Economy (JISPE),*** for the Block Period of Three years at a substantial concession resulting in savings of 20% to the new subscribers as under, provided that the subscription amount is remitted in lumpsum. If you have already remitted the subscription for long term, it will be revised as below only after completion of your present term of subscription.

Subscription Category	*Yearly Subscription*	*Three Years' lumpsum Subscription*
Universities & Institutes	*Rs. 1000/-*	*Rs. 2400/-*
Individuals & Colleges	*Rs. 500/-*	*Rs. 1200/-*

For more details write or send email to:
The Administrative Officer,
Indian School of Political Economy,
'Arthabodh', 968/21-22, Senapati Bapat Road,
Pune 411 016 (India). Phones: (020) 25657132, 25657210.
e-mail: ispepune@gmail.com; Wesite: ispepune.org.in

EDITORIAL

In our efforts to reduce delays in our publication, we have improved by bringing to you the 3rd and 4th issues of 2022 as a combined one. These contain papers relating to a very pertinent issue regarding Industrial location in India and factors contributing to such decisions (in two parts). Further, a detailed study of this aspect and related issues in the context of the auto- component industry (in three parts) in India also appears. And we are confident that, in 2023, issues will come out regularly with no delays.

S. Sriraman
April 2023

TEMPORAL ANALYSIS OF INDUSTRIAL LOCATION IN INDIA PART I CONCENTRATIONS AND DISPERSAL

Sharadini Rath*

It is well known that industrial locations in India exhibit severe clustering, leading to significant regional imbalance in formal sector industrial employment. In this paper we examine the temporal aspects of how decisions regarding industrial locations are made by capital. We use location address of individual plants in the sample of completed industrial projects from Projects Today over the period 2002 to 2018. It is found that there are 3 distinct periods of project completion, slow growth for the initial 6 years, then accelerating growth for the next 5 years, and then decelerating growth with a sharp downturn by 2017-18. Taking into account the time lag between decision on location and project completion, this trajectory matches the shape of GDP growth rate curve over this period. We also show that during the accelerated growth period, there is dispersal of industrial location to low industrialised regions, expanding cluster peripheries, while growth within clusters also continues. During the contraction period, locations tend to withdraw back towards the cluster cores, and cluster peripheries contract. However, some of the dispersed locations continue to attract location. With insights from field work, we also show that individual locations can behave very differently from each other over the same time period, depending on many factors, such as local entrepreneurship, availability of land and infrastructure, local political forces, and competition between states leading to shifting locations across states. So while clustering remains a powerful centripetal force for industrial location, good economic growth and local factors can be equally strong centrifugal forces leading to dispersal.

---

* Sharadini is Rath affiliate fellow, Indian School of Political Economy, Pune

Acknowledgements: Mahratta Chamber of Commerce, Industry and Agriculture (MCCIA), Pune, provided access to the Projects Today data, without which this analysis would not have been possible. Indian School of Political Economy has provided the academic base for this work, along with logistical support during field work. Ashish Andhale has given able research assistance in all areas of this work. Comments on the work from time to time by Abhay Tilak, Nilakantha Rath and Vikaschandra Chitre have helped shape the analysis and pose new perspectives. Comments from Sanjoy Chakravorty on an earlier version were useful. Learning geographical analysis tools from Ujaval Gandhi made for a sharp learning curve and new methods of analysis. All are gratefully acknowledged.

The Editor, JISPE is gratefully acknowledged for extensive comments which have improved this paper considerably.

1 Introduction

The Economic Survey for India 2014-15 [GOI, 2015] was presented in the form of an analytical view of the Indian economy, rather than the usual 'factual' presentation. Chapter 7 was titled 'What to make in India? Manufacturing or Services?'. While it posed an apparently available binary choice, the contents of the chapter gave a concise outline of past contributions of both sectors, and their possible futures in the then recently announced 'Make in India' programme of the new government. It was clear that registered manufacture was not performing its main function of providing employment to the vast mass of low skilled labour force that was growing apace in numbers. Both organised manufacture and services were essentially employing higher skills, leading to high labour productivity, while at the same time not really improving their contribution to the GDP significantly. Most low and unskilled labour was employed in the unorganised sector, whose productivity was low and not likely to improve. Share of employment in organised manufacture had never crossed about 6 percent of total at its peak since 1985, and output and employment had remained stagnant over the period. The Survey came to the conclusion that in large parts of India, there was in reality a trend of de-industrialisation, without ever achieving the peaks that countries like South Korea saw in the late 1980's. Creating employment for the low skilled labour force was never going to be possible with the organised sector. So as a solution in the long term, the Survey proposed that skill development programmes would result in a labour force more useful to organised manufacture, but this would entail the 'loss of one or two generations' being locked into low skill and low wage work. It was not clear in the Survey what assurances could be given about the state of manufacture, and its skill and labour requirements two generations in the future.

While the Economic Survey [GOI, 2015] had state level insights, it did not address sub-state level issues within this picture. Regional imbalance has been studied in various ways for a long time in India [Ahluwalia, 2000; Shetty, 2003; Bhattacharya and Sakthivel, 2004; Das and Barua, 1996] and has been a policy focus at least until the reforms in 1991. The economic geography of India has been studied with data available from various government agencies, such as sample surveys of the National Sample Survey Organisation, Annual Surveys of Industry, Population Census [Chandrashekhar and Sharma, 2014; Saikia 2009] and always defined by administrative boundaries, States, districts, towns and in one case the postal code.

There is sufficient evidence from Annual Surveys of Industries, available at most at district level, and Population Census worker data, that concentrations of manufacturing and other non-farm employment have grown steadily over the past few decades. Independently, there is evidence that certain sectors within manufacture show significant clustering, exhibiting specialisation [Chakravorty and Lall, 2007].

There is a long history of theories of industrial location [Weber, 1929], [Predöhl, 1928, 371-390]; [Hotelling, 1929, Pp. 41-57]; [Lösch, 1954]; [Smith, 1966, Pp. 95-113]; [Hoover, 1948]; [Isard, 1949, Pp. 476-506]. Over the past 20 years some observed phenomena such as clustering and related formation of agglomerations have seen sustained attention [Chakravorty and Lall, 2007]. They largely give a reasonably clear analysis of the advantages that firms in clusters see, *after the cluster has been formed.* Survey methods usually do not give insights to the parameters these models use, and emphasise highly local and/or highly personal aspects of decision making at firm level, as we shall show in Section 6 of the paper.

However, secondary, post facto, data tells us that there are indeed advantages to being within a cluster or concentration, in close proximity to markets and other industries, sometimes of the same type, but at the very least among many others. As more and more industries locate in a given area, the peripheries of these concentrations push outward, increasing their geographical size.

The sites of these large concentrations usually have historical roots [Chakravorty and Lall, 2007]. They also tend to be in and/or around large agglomerations, whose formation is usually a case of cumulative causation which means the more industries locate in a given area, the more the urban centre grows with other economic activities in the services sector, and creates a positive feedback loop of localisation and urbanisation economies, making it attractive for further industrial location, thus growing into an agglomeration. The effect of this process is economic growth in conjunction with mediated urbanisation economies, regional inequality and unbalanced growth. Marshall [1890] gives an outline of localisation economies. More recently, Krugman [1991, Pp. 484-499; 1996, Pp. 137-150] and Porter [1996] focus on increasing returns to scale in clusters.

What is not clear is: when does location of industries get dispersed? What are the conditions required for industries to move into new, previously less used, or even unused areas, far from existing concentrations? In other words, what required for regional imbalances in industrial location to be corrected? There are some interesting models of the centripetal and centrifugal forces that in different combinations of the underlying variables give rise to greater concentration or some degree of dispersal. The most dominant are the initial conditions of the region, basically historical disadvantages. In addition, transport cost, returns to scale, proximity to markets, mobility of labour and capital tend to be centripetal, leading to greater concentration. Diseconomies of agglomeration and increase in export share tend to be centrifugal.

According to Krugman and Livas [1996], transport costs are minimised when firms produce mainly for a domestic market, by locating close to consumers. Strong forward and backward linkages are formed that lead to agglomeration economies in historically advantaged regions. However, when external trade is liberalised, producers shift from domestic to international markets, forward and backward linkages are weakened, reducing the centripetal forces of agglomerations, and firms can disperse, creating a core-periphery model. Paluzie [2001, Pp. 67-85] showed that this is not necessarily true. Modifying the Krugman and Livas [1996] model of New Economic Geography, this analysis showed that when labour is mobile within countries, protectionist policies do not cause regional imbalances. Also, liberalisation of trade actually accrues most to regions with some initial advantage, reinforcing existing agglomerations.

Empirical work by Sjöberg and Sjöholm [2004, Pp. 287-310] and Chakravorty et. al., [2005, Pp. 331-353] sheds light on the difficulties of reconciling these theoretical constructs with reality. In an interesting study on India, which comes closest to the analysis we will present here, Chakravorty et. al., [2005] looked at firm level data for locations within some metropolitan regions and concluded that land market rigidities trump localisation economies. In fact, state intervention in land markets is the major centripetal force that aids agglomeration formation and has the greatest influence on firm level location choice. The Sjöberg and Sjöholm [2004] study of Indonesia showed that industry did not disperse during a major period of trade liberalisation, and in fact, existing agglomerations grew rapidly. They point at the spatial configuration within the country, historical advantages and issues of access to provide possible explanations. Basically, the Paluzie [2001] conclusions seem to be borne out by these studies.

Most empirical work on industrial clustering relies on data defined by administrative boundaries and uses tools such as the Herfindahl index or Moran's I [Chakravorty and Lall, 2007; Zhang et. al., 2008]. These methods use some attribute of the plants, such as employment or financial parameters, to study if plants with similar attributes locate together, with the geographical extent being defined by administrative boundaries.

This paper is written in two parts. Part I studies the geography of industrial locations, while Part II studies the geography of real estate developments and manufacture that use large land areas. Data is taken from a single source, over the same time period. We present, in two parts, an analysis that does away with the confines of administrative boundaries, using locations of individual plants alone to construct a spatio-temporal density based (not attribute based) picture of location choice. We show how dominant concentrations cut across regions, and bind together transcending state policy initiatives. We also show that there are indeed conditions under which significant geographical dispersal can take place, even as the cores of the concentrations grow. On the other hand, we examine the development of new locations and show that they have strong local imperatives, and can be explained only to a limited extent by any existing economic model. An economic geography approach is necessary to understand how they develop. Many of these processes operate simultaneously in time over different geographical scales.

In Part I, Section 2 gives the details of the database used and its context within the overall Indian manufacture scenario. Section 3 gives the analysis of spatio-temporal evolution of plant location, linking the case of dispersal of industrial location to macroeconomic parameters, and the essential role played by large urban agglomerations in the economic geography of industrialisation within the core - periphery picture. Section 4 identifies geographical clusters from the plant level location data, their evolution over time, and processes by which new locations might be developing. Section 5 analyses the impact of plant clustering on employment in the formal sector manufacture by linking the plant location and product data to employment data from Annual Surveys of Industry to show that clustering in employment follows plant clustering, and dispersal usually takes place in products that require lower skill levels. Section 6 takes up the issue of the role played by local factors in the development of new locations, or the decline of existing ones within the same time period, in the same state, governed by the same policy and economic environment. Interview based studies in two locations in Maharashtra give insights that greatly complicate the theorising of industrial location choice, cluster formation and employment generation. The role of land availability, state level vigilance regarding the performance of industrial locations, the role played by inter-state competition, along with broader factors of technology changes in manufacture are brought out to show that location models can go only so far in creating a policy direction for governments. Ultimately, close monitoring, engagement with local entrepreneurs, and long range planning are required to successfully not only keep industrial development within the state, but also to keep it on the growth path. Section 7 summarise and concludes Part I.

Given the pivotal role of urban agglomerations in the industrial geography of India in Part I, we take up a geographical analysis of the development of large land area projects for housing, manufacture and services in India over the past 20 years. It attempts to establish an evidence base for their location choices and how they correlate with trends in employment and preferential location of State provision of infrastructure and basic amenities. This study uses data from Projects Today, along with the Population Census of India and Special Economic Zones. The study is based on for locations such as Hyderabad in Telangana State, and Jamshedpur and Bokaro in Jharkhand State. Interviews with State Government and industry officials in Jharkhand, along with those with representatives of large real estate developers are used to understand how location choices are linked to State action on infrastructure and basic amenities.

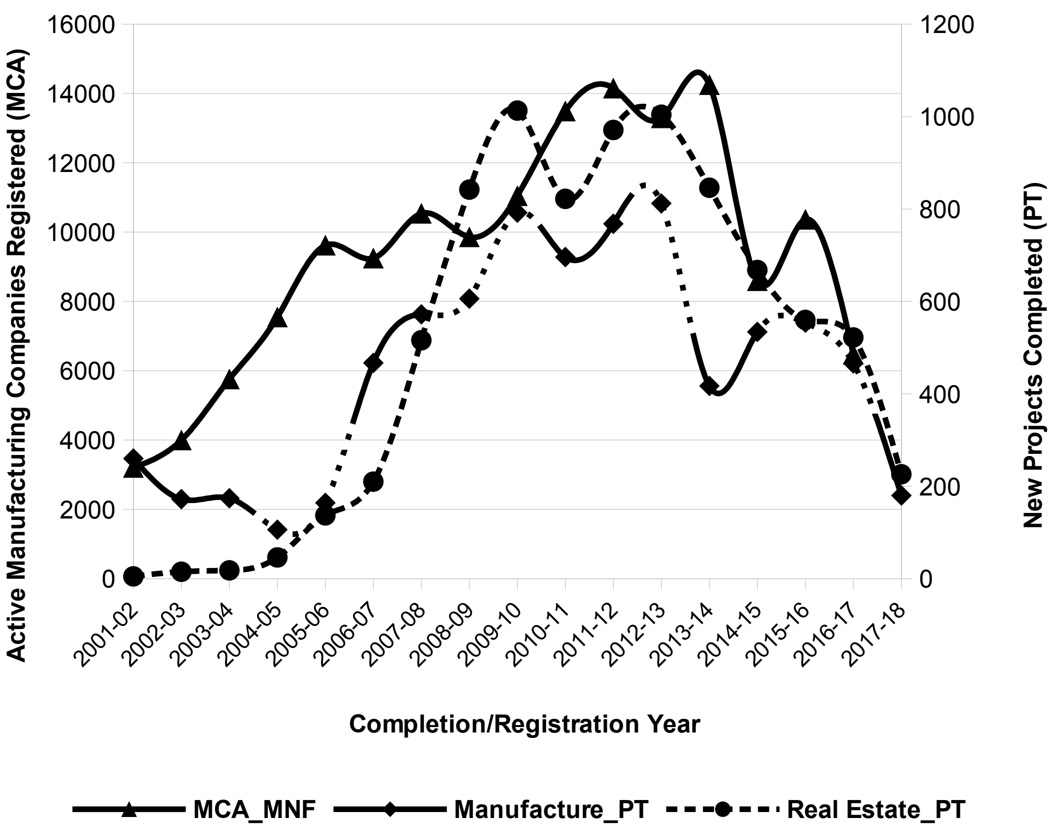
We now turn, to begin with, to plant level analysis of manufacture in India.

1 Plant Level Analysis of Indian Manufacture:

The basis of the analysis is plant level location data from 2002 to 2018 for the formal sector, registered manufacture from Projects Today (2002-2018). The data gives a geographically uniform sample of all new plants that were completed in each year. It gives the date of completion, address of the plant, and the product that will be manufactured so they can be classified at the 2-digit level as per the standard National Industrial Classification [2008] followed by Annual Surveys of Industry. However, we have neither employment nor data such as output and finances about these plants since these are newly completed plants. What we do have is a revealed time line of location choice and type of product when the location choice was made and what the plant would produce.

We have used geo-spatial tools such as the Quantum Geographic Information System software and Google Application Programming Interface tools to locate the completed manufacturing plants on the map of India by geo-locating their addresses into a latitude - longitude pair. Doing this opens up a wide array of geo-temporal analyses without any reference to administrative boundaries. We then look at what this data reveals in all its available details and then try to see into which theories they fit, to what extent and how.

Figure 1. Number of Completed Projects in Manufacture (Manufacture_PT) and Real Estate Construction (Real Estate_PT) from Projects Today (PT) and Active Manufacturing Companies Registered in each Year (Ministry of Corporate Affairs, GoI) (MCA_MNF) from 2001-02 to 2017-18



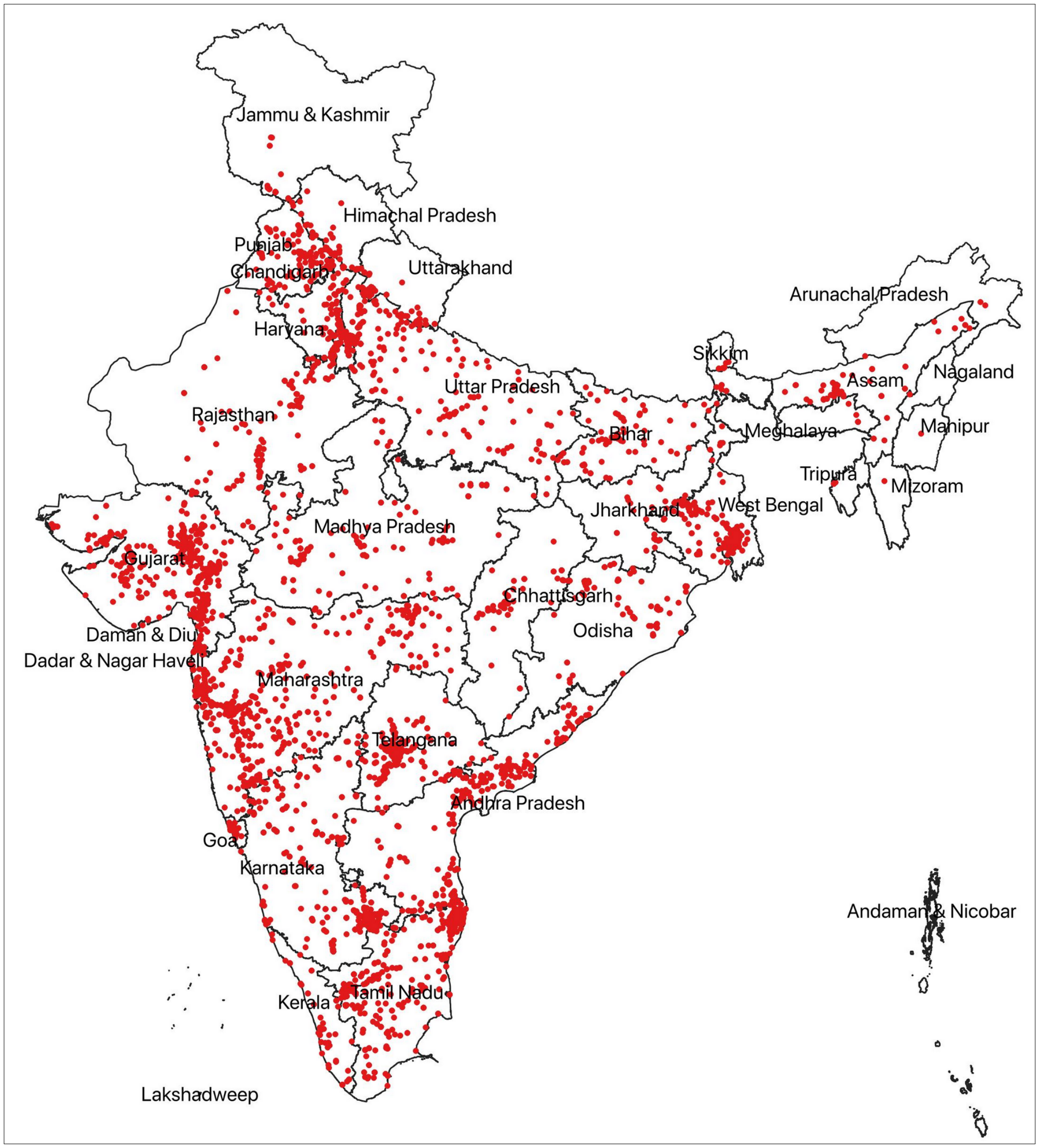
Source: Projects Today and Ministry of Corporate Affairs, GoI.

To make sure that the sample from Projects Today represents a reasonable picture of new formal sector manufacturing developments in India, we have used data from the GOI [2002-2018] that gives number of active manufacturing companies registered every year. Figure 1 gives 3 time series: number of completed manufacturing plants and real estate projects from Projects Today and number of active manufacturing companies registered. Though the peaks in the period 2009-10 to 2013-14 in data sets from Projects Today and company registration are slightly shifted, the Projects Today sample appears to be a reasonable representation of new registrations during this period.

Map 1 gives the actual plant locations of the Projects Today data for all completed manufacture plants from 2002 to 2018. It is clear that the distribution of manufacturing in India is heavily clustered. There is broad preference for a north-west - west - south alignment to these locations. The east, except for a patch in West Bengal is not the preferred area. Also, all southern states with significant concentrations of plants show severe regional disparities. Map A1 (Appendix) shows that while road infrastructure is necessary for plant location, it is by no means sufficient. For example, Uttar Pradesh, Bihar and eastern Maharashtra appear to have dense highway infrastructure, but low incidence of plant location.

The obvious concentrations in Map 1 are all centred around large metropolitan cities; Delhi, Chandigarh and Ludhiana in Punjab, Ahmedabad and Surat in Gujarat, Mumbai, Thane and Pune in Maharashtra, Hyderabad in Telangana, Chennai and Coimbatore in Tamil Nadu and Bangalore in Karnataka. In the east there is only Kolkata in West Bengal. There are two corridor locations, one in coastal Andhra Pradesh, and another between Ahmedabad through Surat in Gujarat to Mumbai in Maharashtra along the west coast. It is clear that there are two other corridor developments taking place: from Pune, Maharashtra towards Hyderabad, Telangana to the east, and towards Bangalore, Karnataka, to the south, and from Bangalore, Karnataka towards Chennai, Tamil Nadu to the east. The region between Chennai and Coimbatore in Tamil Nadu is also seen to be preferred significantly. There is some corridor development along the eastern edge of Rajasthan, along the Delhi - Mumbai highway. This pattern in Map 1 overlaps with the pattern giving district share of all India Other Workers (non-agricultural) in the Population Census 2011 [GOI, 2011] shown in Map A2 (Appendix). But the match is even better with the pattern in Map A3 (Appendix) of district share of all India workers in all hiring non-agricultural enterprises from the Economic Census of 2013 [GOI, 2013] which leaves out most of the unorganised sector, and is more pertinent to the sample in Projects Today of registered manufacture.

Map 1. Location of completed manufacturing plants from Projects Today from 2002 to 2018, with states of India (Total plants = 7232)



Source: Author’s calculation from Project Today data.

2 Spatio-Temporal Evolution of Plant Location

That Indian manufacture has continued to be spatially concentrated over the period 2002 to 2018 is clear from these patterns of locations of new plants. This data gives us the opportunity to see how this pattern has evolved over time. In order to do this analysis, we form 3 groups for different time periods based on the manufacturing curve in Figure 1 as follows:

Time period 1 - Group 1: 2002-03 to 2007- 08 (1343 factories, 6 years)

Time period 2 - Group 2: 2008-09 to 2012-13 (4246 factories, 5 years)

Time period 3 - Group 3: 2013-14 to 2017-18 (2151 factories, 5 years)

Total plants = 7740

Group 2 period saw a peak in terms of new factory completion, Group 1 saw a growth, and Group 3 a near precipitate decline by 2017-18.

A State and year group-wise percentage shares of all plant locations (Table 1) shows that Gujarat has the largest share of these locations. It has a similar share when compared with Maharashtra in the 2nd time period while in the 3rd period, it is share keeps growing while that of Maharashtra slows down considerably. Most of the growth in Gujarat during the 2013-18 has been driven by chemical and textiles industries. The 6 southern States of Maharashtra, Telangana, Andhra Pradesh, Karnataka, Tamil Nadu and Kerala made up 42.54% share of total plants. Bihar is the only State that keeps growing throughout the entire three time periods even though, in totality, it had a very small share in terms of industrial location. We will come back to this observation later in the paper.

Table 1. State and Year group-wise break-up of percentage share of new plant locations in descending order of total (2002-18) share

State	Group 1 2002-08	Group 2 2008-13	Group 3 20013-18	Total 2002-18
(1)	(2)	(3)	(4)	(5)
Gujarat	1.99	7.67	7.09	16.76
Maharashtra	2.53	7.84	4.41	14.79
Tamil Nadu	2.71	4.48	1.95	9.14
Karnataka	1.62	3.06	1.82	6.51
Andhra Pradesh	1.08	2.95	2.22	6.25
West Bengal	1.62	2.51	1.25	5.38
Telangana	1.30	2.58	1.10	4.97
Uttar Pradesh	1.18	2.19	1.27	4.64
Uttarakhand	1.42	2.75	0.30	4.47
Rajasthan	0.75	1.99	1.38	4.13
Haryana	0.91	1.95	1.20	4.06
Madhya Pradesh	0.43	1.95	1.11	3.49
Punjab	0.68	1.27	0.73	2.68
Himachal Pradesh	0.95	0.75	0.13	1.84
Odisha	0.21	0.88	0.48	1.58
Bihar	0.03	0.41	1.00	1.44
Chhattisgarh	0.46	0.64	0.33	1.42
Assam	0.19	0.57	0.53	1.28
Dadara & Nagar Havelli	0.31	0.65	0.19	1.15
Kerala	0.30	0.41	0.17	0.88
Jharkhand	0.20	0.34	0.17	0.71
Jammu & Kashmir	0.20	0.41	0.09	0.70
Goa	0.14	0.21	0.04	0.40
NCT of Delhi	0.07	0.21	0.11	0.40
Daman & Diu	0.11	0.19	0.03	0.33
Puducherry	0.07	0.10	0.04	0.21
Meghalaya	0.04	0.04	0.06	0.14
Sikkim		0.06	0.03	0.09
Tripura		0.06	0.03	0.09
Arunanchal Pradesh		0.04		0.04
Manipur			0.01	0.01
Mizoram			0.01	0.01
Nagaland			0.01	0.01
Total	21.53	49.19	29.29	100.00

Source: Author's calculations from data given in Figure 1

Gujarat was 3rd in new plant locations during Group 1 phase, but went on to attract a large proportion of plants in Groups 2 and 3, overtaking Tamil Nadu and Maharashtra. Gujarat and Bihar are the only states that show a sustained increase in new plants over the entire period. All other states saw increased locations in Group 2 and then decline during Group 3 phase.

Spatially, we are in a position to postulate from the pattern emerging in Map 1 that most major concentrations are anchored to metropolitan areas. Using the Euclidean distance between the geographical centroid of major cities and the plants as the basic distance metric, we choose a radius around each city to define its effective range of 'agglomeration influence'. Within this radius, we calculated quartiles of the distribution of distance from the city to the plants (Table 2). We then provided weights to the interquartile distance by the number of plants in each range, normalised this by the maximum distance of the distribution and plotted these to understand the temporal dynamics of regarding location choices, i.e.,

Plant_Share = % share of plants in the interquartile range of total plants in the radius of agglomeration influence

Weighted Occupancy = Interquartile Distance (km) * Plant_Share / Maximum Distance (km)

We undertook this analysis for three major metropolises; Mumbai, Chennai and Delhi. Their agglomeration impact radii are defined as follows:

Mumbai: 400 km radius of agglomeration influence, covering the relevant portion of state of Maharashtra and the Union Territory of Dadra and Nagar Haveli on its northern border

Chennai: 350 km radius of agglomeration influence, covering relevant portions of Tamil Nadu and Andhra Pradesh

Delhi: 400 km of radius of agglomeration influence, covering relevant portions of the states of Uttar Pradesh, Haryana, Rajasthan, Madhya Pradesh, Uttarakhand, Himachal Pradesh and Punjab.

These radii of agglomeration influence were chosen to cover the state in which the city is located, and portions of adjoining states that are seen to form a continuum of locations of new plants from Map 1.

Table 2. Quartiles of distribution of distance of plants from city within radius of agglomeration influence for each year group

City	Year	Minimum	Q1	Median	Q3	Maximum
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mumbai	2002-08	3.05	54.98	115.33	156.45	351.64
Mumbai	2008-13	3.16	82.26	130.87	258.85	398.54
Mumbai	2013-18	1.74	109.00	153.53	258.85	385.45
Chennai	2002-08	1.11	30.11	254.17	288.50	348.73
Chennai	2008-13	1.11	42.01	210.02	289.04	349.42
Chennai	2013-18	3.26	49.25	139.85	281.42	346.74
Delhi	2002-08	9.39	67.33	197.88	259.07	397.22
Delhi	2008-13	7.86	62.14	153.59	235.17	398.65
Delhi	2013-18	7.86	43.73	111.38	255.72	399.53

Source: Author’s calculation from Project Today data.

Weighting the interquartile distances with respective plant occupation share and normalising with the Maximum gives us the real picture of how location choices were made in these periods. Table 3 gives the interquartile distance, plant share and weighted occupancy and Figures 2, 3 and 4 give charts for weighted occupancy for the 3 metropolises.

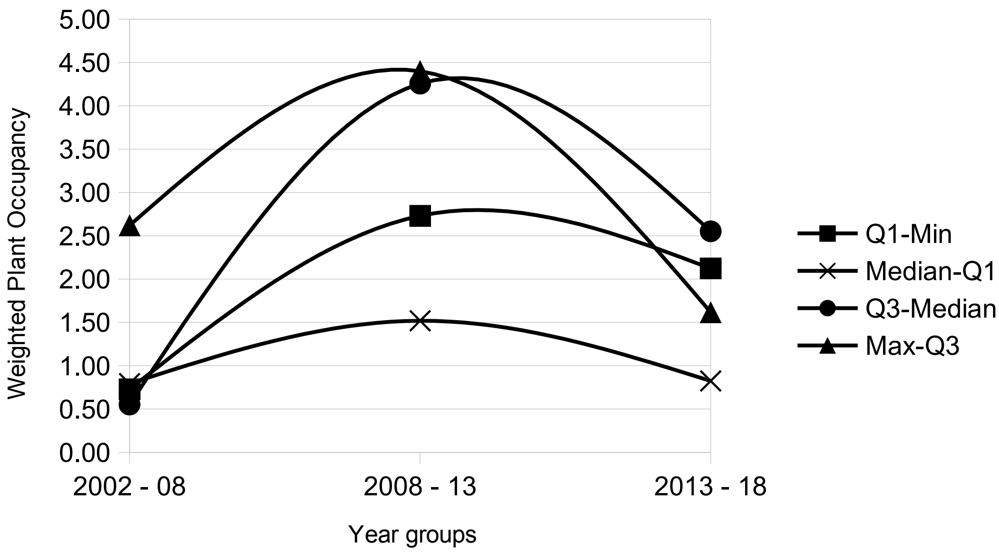
The period 2008-13 in which the number of new industrial plants completed peaked, also saw an effective dispersal of industry, relative to the nearest metropolis. It is remarkable that this happened in all quartile ranges. From being close by to being far away, all scales in a 350 to 400 km radius of these cities showed increased incidence of new location. In other words, the core grew, and simultaneously the periphery was pushed outward bringing in new areas into the agglomeration influence radius. What is even more remarkable is that in the third period, 2013-18, as number of new plants declined, their locations contracted geographically. The new locations retreated inward, toward the agglomeration. The outer periphery, which had seen more new locations during 2008-12, were no longer considered attractive. The whole process contracted geographically and with the number of new plants declining, their locations came shifted inward towards the metropolis.

Table 3. Interquartile distance, plant share and weighted occupancy of distribution of distance of plants from city within radius of agglomeration influence for each year group

City	Year Group	Q1-Min	Median-Q1	Q3-Median	Max-Q3
(1)	(2)	(3)	(4)	(5)	(6)
			Interquartile Distance (km)		
Mumbai	2002-08	51.93	60.35	41.11	195.20
Mumbai	2008-13	79.10	48.62	127.97	139.69
Mumbai	2013-18	107.25	44.53	105.32	126.60
Chennai	2002-08	29.01	224.05	34.33	60.23
Chennai	2008-13	40.90	168.01	79.02	60.38
Chennai	2013-18	45.99	90.60	141.56	65.32
Delhi	2002-08	57.94	130.55	61.19	138.15
Delhi	2008-13	54.28	91.45	81.58	163.48
Delhi	2013-18	35.88	67.65	144.34	143.81
			Plant Share (%)		
Mumbai	2002-08	4.92	4.62	4.72	4.72
Mumbai	2008-13	13.76	12.45	13.25	12.55
Mumbai	2013-18	7.63	7.13	9.34	4.92
Chennai	2002-08	6.91	6.02	6.47	6.47
Chennai	2008-13	12.37	11.15	11.71	11.71
Chennai	2013-18	6.91	6.69	6.80	6.80
Delhi	2002-08	6.64	6.64	9.47	3.75
Delhi	2008-13	13.43	12.44	12.72	12.86
Delhi	2013-18	6.01	5.23	5.30	10.81
			Weighted Occupancy		
Mumbai	2002-08	0.73	0.79	0.55	2.62
Mumbai	2008-13	2.73	1.52	4.26	4.40
Mumbai	2013-18	2.12	0.82	2.55	1.62
Chennai	2002-08	0.57	3.87	0.64	1.12
Chennai	2008-13	1.45	5.36	2.65	2.02
Chennai	2013-18	0.92	1.75	2.78	1.28
Delhi	2002-08	0.97	2.18	1.46	1.30
Delhi	2008-13	1.83	2.85	2.60	5.27
Delhi	2013-18	0.54	0.89	1.91	3.89

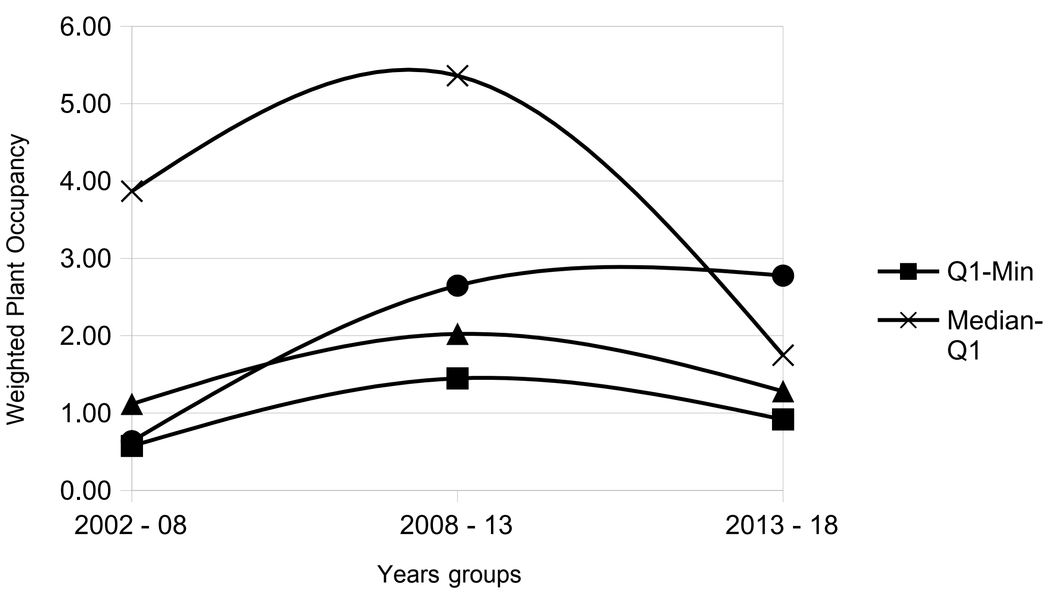
Source: Author’s calculation from Project Today data

Figure 2. Weighted Plant Occupancy for Mumbai (radius 400 km)



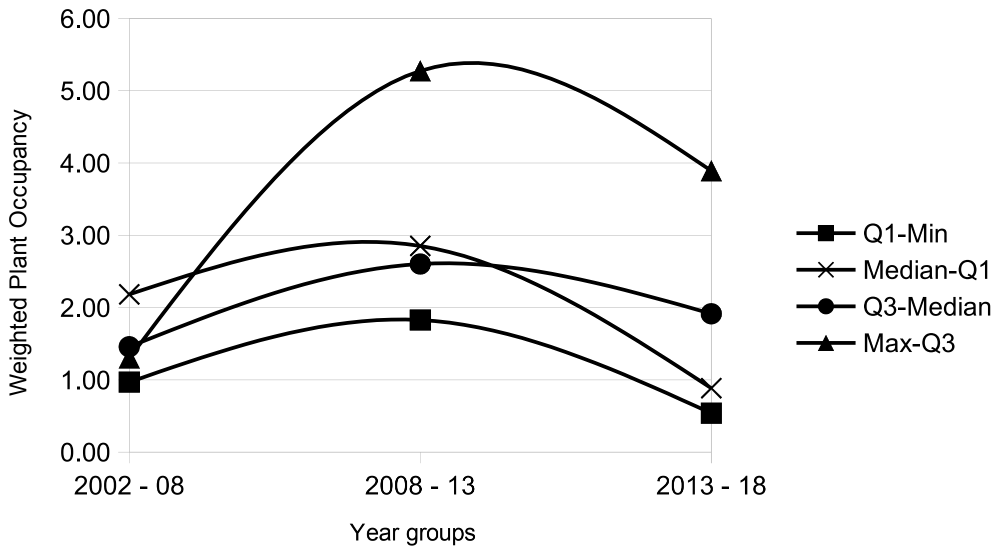
Source: Author's calculation from Project Today data.

Figure 3. Weighted Plant Occupancy for Chennai (radius 350 km)



Source: Author's calculation from Project Today data.

Figure 4. Weighted Plant Occupancy for Delhi (radius 400 km)

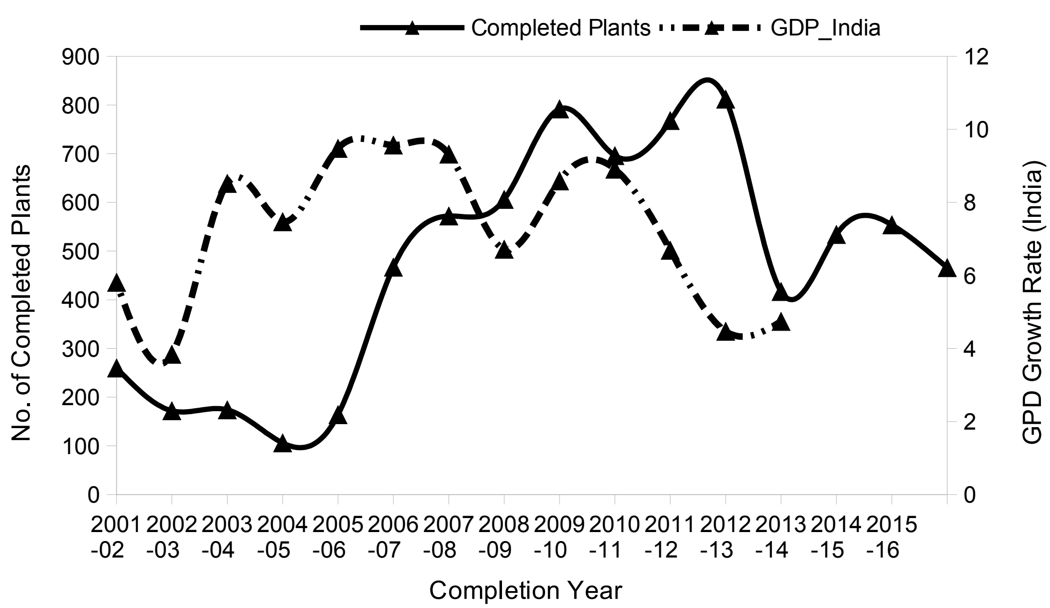


Source: Author's calculation from Project Today data.

Manufacturing industries that dispersed the most produced, in descending order, food and beverages, chemical products, basic metals, non-metallic minerals and, to a smaller extent, textiles. Industries requiring more skilled labour such as electronics and electricals, pharmaceuticals, vehicle and transport, etc., stayed within the core of the agglomeration. The expanded periphery, largely in rural areas or smaller urban centres, saw relatively lower skill requirement for the dispersed plants. It is most likely that these industries would have good access to rural markets. Some sense of this was given in Ghani et. al., [2012]. Though done for a different period, 1989-2005, they concluded that the formal sector manufacturing was becoming 'less urbanised', while the unorganised sector was becoming 'more urbanised'.

We also looked for some macroeconomic indicator of the broader economy that can explain the context of the basic pattern of new plants in Figure 1. We observed that it is quite simply the GDP growth rate Figure 5 gives the results for GDP for India at 2004-05 base year. According to industry owners, a medium to large size industrial plant takes anywhere from 2 to 4.5 years from announcement to completion of construction, depending on the complexity of civil and machinery requirements. This fits in well with the forward shift of the completed plants curve from the GDP growth curve. Decisions are made about investing in new plants in good years and many of them are located in dispersed areas, where the risk of being further away from the main market is perhaps slightly higher. But since economic prospects look bright, they are considered worth taking. Given the fact that their skill requirement is low and rural markets pick up in good times, these are not difficult decisions to make. It is also possible that this investment has a strong local flavour, with local entrepreneurs sensing a good opportunity to set up businesses as we will show in Section 6.

Figure 5. GDP growth rate (2004 - 05 base) for India and number of completed plants in each year



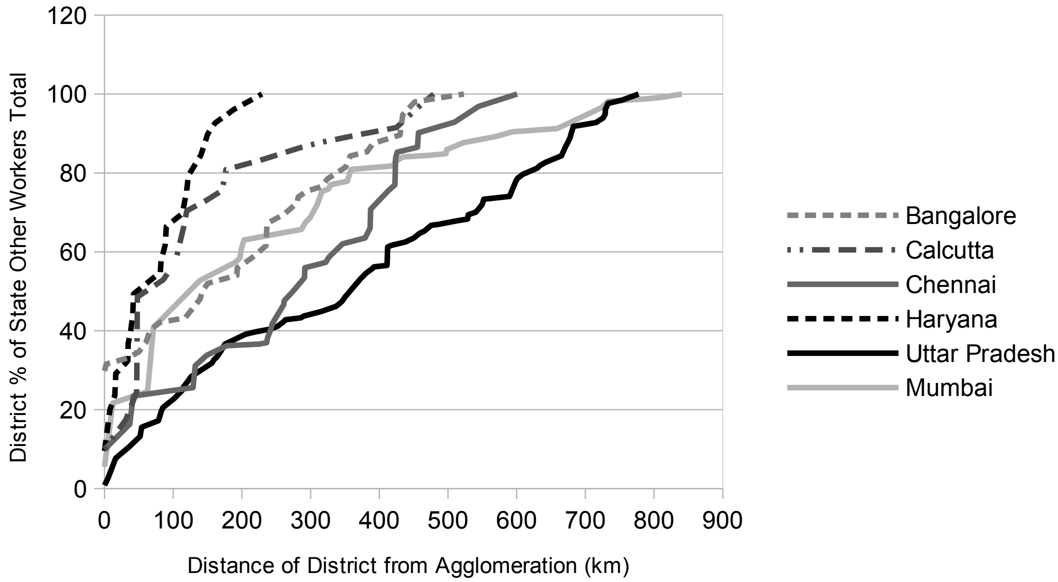
Source: Author's calculation from Project Today data, GDP data from Reserve Bank of India.

The issue of land regulation brought up by Chakravorty et. al., [2005] is more complex. In good years, land price or availability may not be important factors. But for cores to grow and locations to contract towards metropolises in poor years, when higher land prices prevail, might seem counter-intuitive. But it is clear that land price and transport cost are secondary issues compared to markets. In low times, it is possible that rural markets are also low, so capital relies on the fairly steady all-weather urban markets, and continues locating near them. With states creating industrial estates by acquiring land and providing water and power, it is possible that the dispersed locations have easy access to these facilities locally. It is also noted that industrial location outside industrial estates is most prevalent close to and within agglomerations, rather than in distant areas. Municipal corporations routinely take zoning decisions to modify existing guidelines to accommodate demand from industry. So not only is land available, it is being used at relatively high prices in and close to agglomerations, both in good and poor economic periods. In some sense, urbanisation economies and urban markets trump all other considerations. Urbanisation economies include not only access to the urban market, but also a mixed industrial environment, varied labour pool, services, but also most importantly, urban amenities for the workforce. Transport costs for intermediate inputs and raw materials become secondary considerations, when the location externalities maximise returns.

The overall performance of the economy as a whole and its perceived prospects in the medium term seem to play a crucial role in location decisions which give rise to dispersal. When the outlook is positive both processes take place simultaneously, cores of concentrations become more dense, while the periphery also pushes outward and gets more populated, in effect, increasing the influence area of the urban agglomeration.

In order to quantify the place of the metropolis in the location preferences of industry, we did a small geographical calculation for Other Workers GOI [2011]. We measured the Euclidean distance of the centroid of the metropolis to the centroids of districts around it. We attributed the percentage share of district Population Census Other Workers in the respective State total to these centroids. Then plotted the cumulative share of Other Workers as one travels away from the metropolis in widening circles. Figure 6 gives the results.

Figure 6. Cumulative district % share of State total of Other Workers (Population Census 2011) with distance from main agglomeration



Source: Author's calculations from Population Census 2011 data.

This exercise has been done for Bangalore (within state of Karnataka), Mumbai (within state of Maharashtra), Chennai (within state of Tamil Nadu), Calcutta (within state of West Bengal), and using Delhi as the metropolis, within the states of Haryana and Uttar Pradesh have been plotted (distance from Delhi is normalised to the nearest district in these states). It becomes clear that the cumulative share of Other Workers reaches at least 40% within 200 kms of the metropolis. In Maharashtra, this share reaches 50% by 200 km from Mumbai and 80% by 350 km, showing the severity of regional imbalance within the state, and its western heft. The same is seen in the case of Calcutta and West Bengal, with 80% share reached within 200 km. In Haryana, Other Workers take 60% share within 100 km from its border with Delhi. One of the more linear profiles is for Tamil Nadu (Chennai). And given the wider dispersion that the plant data shows in different parts of this State (refer to Map 1), this fits in the picture of a State as widely urbanised with perhaps the least regional inequality among the more industrialised states.

This demonstration of the primacy of agglomerations agrees broadly with the conclusions of Chakravorty et. al., [2005] that more than localisation, it is urbanisation economies that matter for industrial location. It quantifies their influence radii, and also supports the choice of these radii for different metropolises in the geographical analysis. It also explains the broad economic impetus that creates the simultaneous processes of core growth and dispersal.

3. Cluster formation in manufacturing industry:

Another type of analysis that can done with the plant level data from Projects Today is to examine if there is significant cluster formation, and the size, shape and locations of these clusters. Industrial clustering is broadly understood to mean a dense and geographically closely situated collection of industrial plants. Marshall [1890] analysed 'industrial districts', Jacob [1969] linked such clusters to existence urban centres, Porter [2000, Pp. 77-90] attempted to define a cluster in terms of its component firms/activities. However, Martin and Sunley [2003, Pp. 5-35] pointed out that none of these studies actually prescribed how to define the geographical extent of a cluster (in kms). They also pointed out that while clusters did confer certain advantages on firms within them, they also had disadvantages, such as labour cost inflation, inflation of land and housing cost, and widening of income disparities.

In this section we take up the task of quantifying the geographical size of cluster of the Projects Today plants, their locations and properties. This analysis can also help to study the areas of eastern India that are far away from major agglomerations, but might still have developed industrial clusters in preferred regions.

We use the algorithm of Density Based Spatial Clustering of Applications with Noise (DBSCAN) [Ester et. al., 1996, Pp. 226-231] to identify clusters using locations of individual plants. This method does not restrict the shape of clusters, does not require number of clusters to be specified, and deals very well with outliers or noise (non-cluster) points. Since it is a density-based algorithm, it is particularly suitable for the analysis of plant locations.

DBSCAN requires two input parameters:

Epsilon (km): Defines the radius of interest around each point

Minimum Density Threshold (MinPts): Number of points that must fall within Epsilon of each point for it to be considered sufficiently dense.

We have used the knee point method[1] to calculate Epsilon for the three year groups and the full data set, and the optimum value for MinPts. These values for the three year groups and for the full period are given in Table 4 as follows:

Table 4. Cluster parameters for all years groups and full period

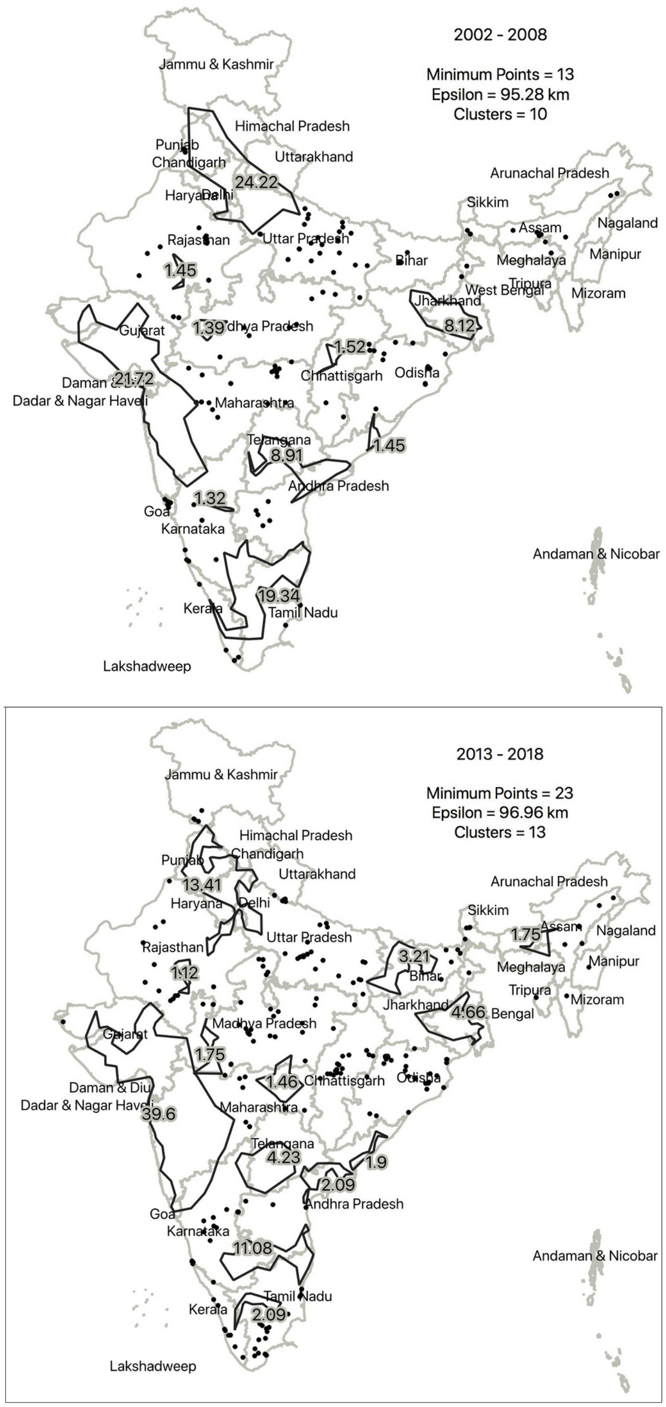
Year Groups	Epsilon (km)	MinPts	No. of Clusters	% Non-Cluster Plants
(1)	(2)	(3)	(4)	(5)
2002-08	95.28	13	10	10.56
2008-13	94.11	30	12	9.57
2013-18	96.96	23	13	11.66
2002-2018	64.89	30	17	9.61

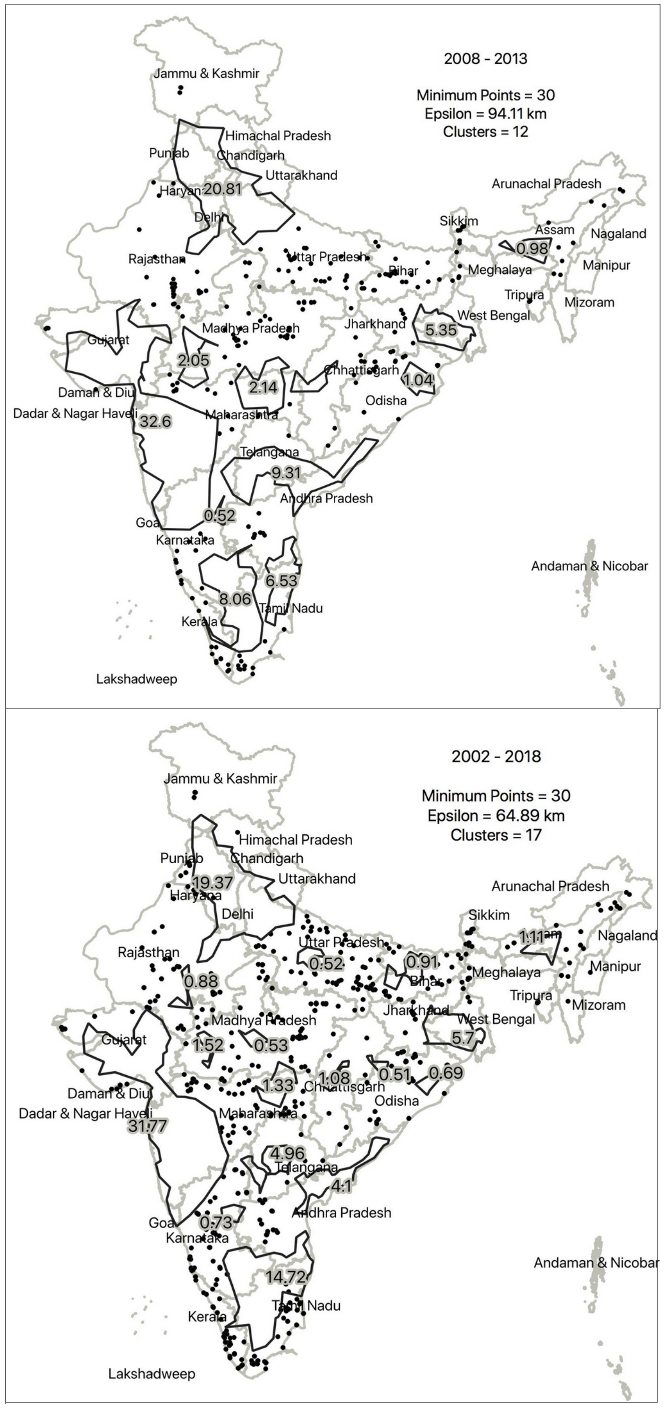
Source: Author's calculation Project Today data.

The clusters identified with this method are given Map 2. The cumulative result for cluster formation for the entire time period 2002 - 2018 shows that as plants got added over the years, locations became denser which meant a decreasing Epsilon and increasing MinPts, thereby giving rise to more clusters, many of which were formed during the peak of 2008-13.

Map 2 given below shows that some of the clusters are truly large and they persist from one time period to another. They are multi-state and multi-district with a large cluster around Delhi with a 19.37% of new plants, the west coast cluster stretching from Mumbai to Ahmedabad with a 32% share of new plants, the Chennai cluster with encompassing Bangalore and Coimbatore with a share of 14.72% share, and other smaller clusters such as, the cluster in coastal Andhra Pradesh and in around Hyderabad in Telangana and the small cluster around Calcutta. Some of these are fixed features of the industrial geography of India, and others have evolved over many decades. Karan and Jenkins [1959, Pp. 269-273] have plotted the geography of industrial workers in India. If we compare their Figure 1 with Map A2 (Appendix) of district share of Other Workers from Population Census 2011 in the Appendix in this paper, there is remarkable overlap: the Mumbai - Pune -Surat - Ahmedabad region in the west, the Chennai - Coimbatore - Bangalore region in the south, coastal Andhra Pradesh, and the region around Kolkata extending into Jharkhand. The picture in Map 1 from the Projects Today data deviates from 1959 on two main fronts: the decline of Kolkata and its surrounds

Map 2. Clusters (DBSCAN) for each period and for the entire 15 years





Source: Author's calculation Project Today data.
Note: Clusters are labeled with their share of the new industries completed in that period. The minimum points used and the calculated Epsilon, along with number of clusters are also given. The dots are industry plants that were outside of any cluster (non-cluster).

(including Jharkhand) and the area around Kanpur as major industrial hubs, and the significant rise of the border regions of Delhi in Uttar Pradesh and Haryana.

From this analysis, it was possible to see clearly the dynamics of industrial location in the eastern part of the country where there are fewer instances of major agglomerations, and industrial location is sparse. But by 2018, in the aggregate, Bihar has a small cluster around Patna. It is made up largely of the food and beverage industry. This cluster started growing in the peak period of 2008-13 and unlike other locations, it continued to grow even in the downturn of 2013-18 and has become a small but significant location (also seen in Table 2). This is also true of Lucknow in Uttar Pradesh, Guahati in Assam, Bhuvaneshwar in Odisha and Bhopal in Madhya Pradesh. They all had small clusters around large regional cities by 2018. If the logic of dispersal taking place in good economic times holds true, it is unlikely any more new clusters will form in the current economic conditions given the low GDP numbers overall.

Having looked at the broad clustering phenomena and their causes, we will complete the analysis of the Projects Today data by looking at the product profile of the clusters and links to employment generation.

4 Clustering in Industry: Type and Employment

In this Section, we will deal with the issue of employment generation in the formal sector industry. As mentioned in Section 1, the Economic Survey [2015] spoke of the possibility of de-industrialisation. The Projects Today data does not come with employment figures for newly completed plants, but the product information is given in some detail. It will be interesting to see if the product type can be used to in conjunction with worker data from Annual Surveys of Industry data over this period to see which of these plants are likely to create the most employment, and which clusters.

There is detailed information in Projects Today data regarding the product that each plant will produce. We have used the Nation Industrial Classification (NIC) 2008 scheme to assign a 2-digit classification to each factory. Accordingly, we use the following classification scheme:

15 - Manufacture of Food Products and Beverages

16 - Manufacture of Tobacco Products

17 - Manufacture of Textiles

18 - Manufacture of Wearing Apparel Dressing and Dyeing of Fur

19 - Tanning and Dressing of Leather, Manufacture of Luggage, Handbags, Saddlery, Harness and Footwear

20 - Manufacture of Wood and Products of Wood and Cork, Except Furniture, Manufacture of Articles of Straw and Plating Materials

21 - Manufacture of Paper and Paper Products

22 - Publishing, Printing and Reproduction of Recorded Media

23 - Manufacture of Coke, Refined Petroleum Products and Nuclear Fuel

24 - Manufacture of Chemicals and Products

25 - Manufacture of Rubber and Plastic Products

26 - Manufacture of Other Non-Metallic Mineral Products

27 - Manufacture of Basic Metals

28 - Manufacture of Fabricated Metal Products, Except Machinery and Equipments

29 - Manufacture of Machinery and Equipments (National Electric Code (NEC)

31 - Manufacture of Electrical Machinery and Apparatus N.E.C.

34 - Manufacture of Motor Vehicles, Trailers and Semi-Trailers

35 - Manufacture of Other Transport Equipment

36 - Manufacture of Furniture; Manufacturing N.E.C.

Table 5 gives the breakup of the industry types in the Projects Today factories. With small variations, this order of the industry type is persistent in all 3 period groups. The only source of data that is compatible and can be used to link this profile with employment is the Annual Surveys of Industry data (ASI) [GOI, ASI, various years]. We use ranking to highlight the consistency between these two data sets. We rank the following in descending order for the 2-digit classification:

Number of factories in each year group: Group 1 (2002-08), Group 2 (2008-13), Group 3 (2013-18)

Number of factories in ASI for years: 2002-03, 2010-11, 2015-16
Number of workers in ASI for years: 2002-03, 2010-11, 2015-16

The ASI years have been chosen to be representative of the Projects Today period groups.

Table 5. Product type for Projects Today factories 2002-2018 sorted by descending order of no. of new factories

NIC 2008 2-Digit	Industry Type	Projects Today Factories
(1)	(2)	(3)
24	Chemicals and Pharmaceuticals	1479
15	Food and Beverages	1096
27	Basic Metals	896
17	Textiles	716
26	Nonmetallic Minerals	512
29	Machinery and Equipment	463
31	Electricals and Electronics	440
25	Rubber and Plastic	433
34	Motor Vehicles	341
28	Fabricated Metals	209
21	Paper and products	191
18	Wearing Apparel	151
36	Furniture and others	75
35	Other Transport	66
23	Petroleum and products	51
20	Wood and products	41
19	Leather and others	36
22	Publishing and Printing	32
16	Tobacco products	4
	Total	7232

Source: Author's calculation Project Today data.

In Table 6, we order the ranks of all these variables by the rank of number of factories in each 2-digit type in Group 1. The ranking of number of factories by 2-digit classification varies slightly between Projects Today and ASI. However, the variation is mostly within the top 10 ranks by Group 1.

Rank correlation coefficients between number of Projects Today new plants and ASI factories and workers are as follows:

Group 1 (2002-08) PT new plants & ASI 2002-03:

ASI Factories = 0.78, ASI Workers = 0.72

Group 2 (2008-13) PT new plants & ASI 2010-11:

ASI Factories = 0.85, ASI Workers = 0.79

Group 3 (2013-18) PT new plants & ASI 2015-16:

ASI Factories = 0.84, ASI Workers = 0.81

Total period (2002-18) PT new plants & ASI 2015-16:

ASI Factories = 0.79, ASI Workers = 0.83

There is strong rank correlation between number of new Projects Today plants in each year group, as well as the entire period of Projects Today, and rank of factories and workers in ASI years representative of the groups. This shows that along with factories, employment is also highly clustered, maintaining the ranking of industry type by employment.

Table 6. Rank comparison between Projects Today (PT) industry type with ASI, in number of factories and workers

NIC 2-digit	Industry Type	PT Year Group	ASI Fac tories	ASI Workers						
		Group 1	Group 2	Group 3	2002-03	2010-11	2015-16	2002-03	2010-11	2015-16
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
24	Chemicals and Pharma	1	1	1	4	4	5	3	4	3
27	Basic Metals	2	3	3	8	8	8	6	3	5
15	Food and Beverages	3	2	2	1	1	1	1	2	2
17	Textiles	4	4	4	2	3	3	2	1	1
31	Electricals and Electronics	5	8	8	9	9	10	12	12	12
26	Nonmetallic Minerals	6	5	5	3	2	2	4	5	4
29	Machinery and Equipment	7	6	7	5	6	7	8	9	8
34	Motor Vehicles	8	9	9	14	12	12	10	6	6
25	Rubber and Plastic	9	7	6	7	7	6	11	10	10
18	Wearing Apparel	10	12	12	11	10	11	7	7	7
28	Fabricated Metals	11	11	10	6	5	4	9	8	9
21	Paper and products	12	10	11	10	11	9	13	16	16
36	Furniture and others	13	13	13	17	16	16	16	15	14
35	Other Transport	14	14	14	18	18	18	14	14	15
19	Leather and others	15	18	18	16	15	14	15	13	13
20	Wood and products	16	17	16	13	14	15	19	19	19
23	Petroleum	17	15	15	19	19	19	18	18	18
22	Publishing and Printing	18	16	17	12	13	13	17	17	17
16	Tobacco	19	19	19	15	17	17	5	11	11

Source: Author's calculation Project Today data, and GOI, ASI, (for 2002-03, 2011-12, 2015-16.

Table 7. Projects Today industry types by cluster and non-cluster regions 2002-18

NIC 2-digit	Industry Type	Cluster	Non-cluster	Total
(1)	(2)	(3)	(4)	(5)
24	Chemicals and Pharma	1337	114	1451
15	Food and Beverages	866	193	1059
27	Basic Metals	771	82	853
17	Textiles	648	52	700
31	Electricals and Electronics	436	19	455
29	Machinery and Equipment	413	18	431
26	Nonmetallic Minerals	395	99	494
25	Rubber and Plastic	378	33	411
34	Motor Vehicles	336	3	339
28	Fabricated Metals	192	10	202
21	Paper and products	166	17	183
18	Wearing Apparel	136	11	147
36	Furniture and others	70	4	74
35	Other Transport	63	3	66
23	Petroleum	42	9	51
20	Wood and products	37	4	41
19	Leather and others	31	3	34
22	Publishing and Printing	27	2	29
16	Tobacco	4		4
Total		**6348**	**676**	**7024**

Source: Author's calculation Project Today data

If we now look at which industry types form the clusters and which ones dominate in the non-cluster regions, Table 7 shows the spread. While about the same order of industry types is maintained in the cluster and non-cluster regions, Food and Beverages and Non-metallic Minerals tend to have a larger presence in non-cluster regions. This is quite logical since food, beverages and bricks have large rural markets, and brick kilns are also located largely in rural areas due in input requirements.

This analysis shows that along with factories, employment in manufacture is also clustering quite severely, creating dense areas of formal sector workers. If Ghani et. al., [2012] are right, and even the informal sector is becoming more urbanised, then the regional imbalance in industrial employment in general looks to be in a dire state. We will now examine how local factors play a strong role in determining success or failure in regard to an industrial location. The case study undertaken relates to the State of Maharashtra.

5 Local Factors in Location Decisions in Maharashtra

We have seen in the preceding Sections that industrial location in influenced by overall economic conditions as seen by their promoters. In good economic times, both the cores and peripheries of agglomeration grow, while in poor times, there is a tendency to retreat to the core, to take advantage of urban markets. However, these are broad conclusions. This Section reports findings of a field study done in Maharashtra to understand how different locations are influenced by the same economic conditions and the role played by local factors in creating completely different outcomes over the same period.

Maharashtra has been one of the top industrial states in India, as well as one of the most urbanised. In the early part of the decade of 2001-10, there was convergence between the growth of existing urban centres and their associated industrial activity, measured by employment. Mumbai remained the preeminent centre of gravity of the state but was faced with a declining industrial profile and urban population growth. Over time, industry had moved inland from this coastal metropolis to adjoining Thane, Pune on the eastern side of the western ghats and smaller centres further east such as Nashik, Aurangabad, Ahmednagar, Kolhapur and Solapur. These areas saw growing urban populations along with industrial growth.

We used data from two Economic Census in 1998 and 2005 [GOI, 1998; 2005] to understand the details of how industry, its location and employment behaved over this period. This particular period was not a good time for industry in Maharashtra from the employment perspective.[2] Employment size of large factories declined and most growth in numbers was in the owner operated category, followed by those with up to 10 workers. This phenomenon, however, did not lead to an aggregate growth in employment, which either declined or stagnated. The real growth in employment was in the service sector with trading being a large contributor. The number of manufacturing enterprises (registered and unregistered) during 1998 and 2005 (Economic Census) increased by 14.77%, largely those employing 0-10 workers, with all other categories showing negative growth. However, number of total workers dropped by 9% and number of hired workers by about 12%. This was also the period in which the use of contract workers became standard industry practice and their numbers increased by leaps and bounds. From 1998 to 2011, ASI data (on registered enterprises) showed that contract workers in manufacture increased by about 200%, going from 1.32 lakhs to over 4 lakhs, while directly employed workers declined by about 4%. ASI data also shows that over these years, large industry had continued to employ on an average about 57% of all industry workers, regular and contract, and pay about 70% of all wages, even though their per unit employment had fallen.

Map 3 shows the overlap of district share of all India Other Workers [Population Census, 2011], Projects Today plant locations and locations of industrial infrastructure created by the Maharashtra Industrial Development Corporation (MIDC), a parastatal of the State government. Creation of industrial estates by the states is a common policy across India. It is meant to ease the entry of capital into the state by keeping land and related industrial infrastructure ready to be leased. It is also used to direct investment into areas of low industrialisation by offering incentives in addition to land and other amenities. The State spends significant resources in creating this infrastructure.

However, as seen from Map 3, there has been little impact of this policy when it comes to directing enterprises to low industrialised regions of the state. The central and eastern parts of the state are very sparsely populated by new industrial plants by 2018.

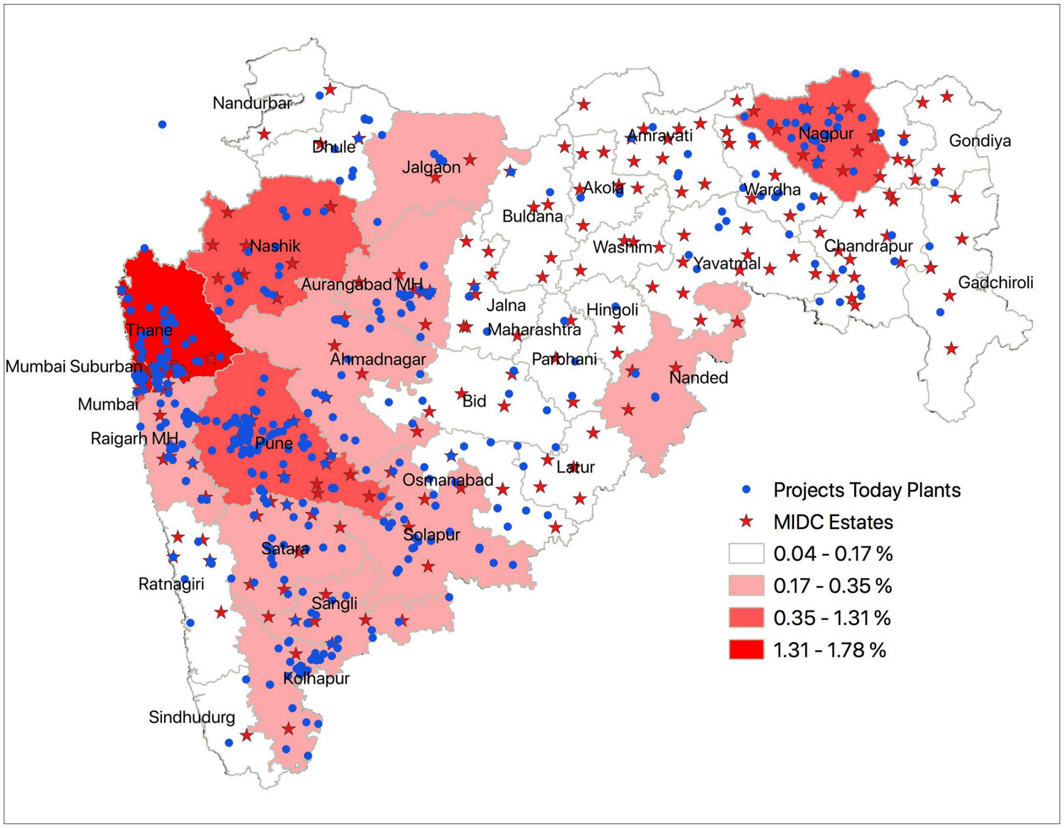
Such persistent regional imbalance in creating non-farm employment led us to pose questions that could possibly lead to some understanding of the processes that influence such persistence.

The following questions were asked:

1 Do all regions, districts and locations within the state behave roughly in a way similar to the aggregate picture?

2 If they do not, why not?

Map 3. Projects Today plants, MIDC Estates, and district % share of all India Other Workers (Population census 2011) for Maharashtra state



Source: Author's calculation Project Today, GOI [2011] and MCCIA [2002-18]

The function of the State, both State and Central governments, in India changed fundamentally with the reforms of 1991 [Saikia, 2009] The State withdrew from its function of the geographical control over where capital should locate. Among the many characterisations of the reforms, its geographical aspect remains perhaps the least appreciated, that industrial policy was no longer in the domain of state/national development policy for lagging regions. The two had decoupled and capital was given freedom to locate at will, wherever it found the most increasing returns. This had the effect of putting individual States in charge of managing their own industrial development. However, there was now a fundamental re-working of the relationship between State governments and capital. Where the State could dictate certain terms before, now capital could make some demands of the state, if it wanted to encourage industrial investment. States were more focused on attracting industrial investment, even if it meant allowing capital to locate where they could maximise returns. As shown by Saikia [2011, Pp. 197-214], this has led to increasing existing concentrations, at the cost of addressing regional imbalances. The previous section, giving analysis of locations, shows the results of this changed dynamic.

Economic literature [World Development Report, 2009] giving primacy to urban metropolises in GDP growth, the inevitability of industrial concentrations around such locations, the role of such growth in increasing overall income and reducing poverty, and therefore by implication, the lack of a necessity to even characterise regional imbalance as something that policy should address, helped State policy makers to swing their attention to attracting capital the best they could.

What was the geography of the response of entrepreneurs and industrial investors to this changed system? The aggregate picture outlined above says little changed and if anything, became more of the same with existing concentrations growing further. The western districts of Maharashtra continued to hold primacy. The decline of Mumbai could be explained by increasing costs and lack of space to expand, which in its turn explained the rise of adjoining Thane and Pune (and their rural peripheries). However, upon closer inspection, some interesting stories of individual locations behaving quite differently from the average emerged. Of these, we examined two districts Jalna and Nashik (see Map 3).

The Economic Census [GOI, 1998; 2005] data for the years 1998 and 2005 was used to initially identify these two districts. Significant disparities in their trajectory of industrial development were noticed and accordingly field work was undertaken to understand the root causes. This work was done during the period 2012-14, and is being reported here to show that location choices have many facets, some of them extremely local. Some of the insights were updated in 2020 with earlier respondents.

Jalna is a backward district in the middle of the State of Maharashtra in the eastern part of Marathwada. Jalna is largely a trading town. It is one of the 10 least developed districts in Maharashtra as gauged by its human development index [GOM, 2012] This perennially dry district is the home of one of the most successful seed companies in India, Mahyco which has its research centre located there. Milling of pulses and oil, and tobacco processing were the main businesses here until 1998 when one steel rolling mill began operations employing about 60 workers. By 2005, there were 36 scrap to steel processing units and rolling mills, each employing more than 50 workers. Jalna's agricultural produce processing industry has grown as well. However, the focus will be on the steel mills, all of which are located in the MIDC estates adjoining the town. This is our first case study.

Nashik city situated about 220 kms north-north east of Mumbai, just past the western ghats, is a much larger city than Jalna and known for its fertile hinterland engaged in horticulture and other produce. The district is among the top 10 on the human development index of Maharashtra. A famous pilgrimage centre, it also had large public sector industries like Hindustan Aeronautics and National Security and Currency Presses, all of which made Nashik an industrial town by the end of the 1960's. With the development of two large MIDC estates adjoining the city limits, Nashik soon had a mix of diverse types of industry, and was well known for luggage, electrical goods, auto and pharmaceuticals. Many multinationals made Nashik their base over the years. In other parts of the district, Malegaon and Sinnar are smaller industrial locations. The district also boasts of wine producers. This apparently thriving district showed a deep decline in industrial employment during the period 1998 to 2005. From having 523 units with more than 50 workers in 1998, the numbers went to just 165 of such industries in 2005, with a 40% drop in total employment. This can be seen to have happened to some of the biggest employers in 1998: textiles, leather goods, chemicals, fabricated metals, rubber and plastics, motor vehicles, other transport equipment. The only growth that took place was in basic metals (steel and rolling mills) and electrical machinery. Nashik went from having 7.4% of the total manufacture employment in the state in 1998, to just 4.8% in 2005.

To answer the first question, not all locations in the State followed the same pattern showing overall decline largely defined by the three big ones, Mumbai, Thane and Pune. Jalna bucked the trend, and Nashik slid far more than the average. An old established hub on the decline, and new growth in an unlikely place, both mediated by the same set of policies and economic conditions.

To answer the second question as to what factors could result in such industrial behaviour in these two districts, we decided to go and talk to industry owners in these districts to understand their experience and observations. We mainly interviewed industry owners, along with labour contractors and consultants, and trade unionists. We shall start with Jalna.

a Jalna

Jalna is, in some sense, a classic case of the processes of localisation of industry. According to Marshall [1890], localisation externalities include intra-industry collaboration, reduced transport costs for inputs as well as products, and access to a specialised and therefore, more efficient labour pool. These lead to better productivity of all firms within the location. In the late 1970's, a trading family in Jalna educated their son to be an engineer. They decided that to take advantage of this new skill, it would be good to help him put up an industrial unit. Thus, they decided that it would be worth their while helping him to work by building up an industry. A family friend informed them that he had obtained a licence for a rolling mill (making construction steel products from prefabricated ingots) to be set up in Jalna. He had applied for a license (Central Government was the licensing authority then) for some other product, but had been given one for a rolling mill, an industry about which he knew nothing at all. However, since an engineer was to be involved, he was willing to form a partnership, if they wished to come on board. Land was available in the first phase of the State industrial estate set up near Jalna town, and that is when the partnership became viable and the process of setting up a rolling mill started. The promoters travelled all the way to Punjab, the then centre of steel processing in India, to learn the business of building a rolling mill, and to buy the necessary equipment. They never really looked back from this starting point. The incubation period of steel in Jalna was rather long, about 20 years, till the end of the 1990's. Following this example, yet another group of trader families[3] with an engineer son decided they would also start an industry. When they presented a proposal for a pulse mill to a local accountant for help with finances and government clearances, he told them that a steel or rolling mill would be better. With the existence of a couple of local rolling mills reputed to be doing well, yet another set of local entrepreneurs jumped into the fray, travelling to Pune to learn the processes of manufacturing. At this time, there seems to have been a sense that this was an industry that could make real profits while being located in Jalna, and more and more trading based investors came forward to become manufacturers. In a short span of 7 years, 1998 to 2005, 36 mills came into existence. The cluster had a turnover of about Rs. 6000 crore by 2012 There are 6 steel brands in this cluster now. While they compete in the market, for very large orders they can come together to fulfil the demand.

This cluster formation process has all the classic ingredients for success: labour market pooling, provision of non-traded intermediate inputs, and information or technological spill-overs. In the course of a set of riveting interviews with 8 firm owners, this whole story unfolded to us like a classroom lecture in the theory of localisation [Marshall, 1890].

We now turn to point out some unique features of the development of steel industry in Jalna. The role of trade surplus being invested in industry is crucial. Many families have come together to share the risk and put up the initial capital. Bank loans have been routinely employed even with high interest rates. There has also been a conscious effort to acquire the latest technology to cut energy and labour costs. The later entrants especially, post 2003, reinvested very large sums in this effort. There have also been significant cases of local technological innovations, almost all of them related to reduction in energy costs, which were then shared with others.

Other than having created the MIDC estate near Jalna town, this development of the steel industry can be traced to no other specific policy initiative of the State government. The availability of large land parcels in the largely empty MIDC estate has been useful since the rolling mills require land to sort scrap and store ingots and finished steel products.

Electricity costs constitutes the lion's share of input costs at 60%. Labour is a distant 5-7%. Other inputs including coal consume the rest. The industry needs 24 hour power supply since shutting down a blast furnace is completely non-viable. Once shut down, it requires months to bring it back to optimum production conditions. For several years from about 2005 onwards, furnace plants ran every-day for 24 hours. The stability of dedicated power supply within the MIDC estate has been an absolutely crucial factor for this. There appears to have been no shortage of power. Coal is used to reheat ingots for rolling and comes from Chandrapur, situated all the way in the east on the Andhra Pradesh border. We came to understand that there is some thought of bringing in technology that would help bypass the use of coal altogether which might prove successful too.

Most of the labour for the steel industry in Jalna is available locally, from within the district, and most of it hired on a contractual basis. Only workers who deal with the high heat conditions of the furnaces come from Bihar or UP. Labour union activity appears to be very low and there seems to have been no real conflict. It was interesting to note that there were no hired managers working in the plants. Owners run their plants on a daily basis. One reason given for this was that controlling cost was extremely important and had to be managed on a daily basis, a job best not left to hired managers. Skilled labour requirement is low, again mostly provided by families of the owners themselves.

A network of transporters brings scrap from all over the country, from Mumbai, Kanpur, Hyderabad and Chennai, to give just a few names. They also take away finished steel. The market for Jalna steel is mostly in the metropolises of western Maharashtra, but product is also sent to Chennai and Bangalore in the south. Transport costs are factored into the price of scrap on the one hand, and the price of finished steel on the other. Transport contractors, who are mostly local truck owners, handle all needs. Road infrastructure could be better, but its present condition has certainly not hampered growth. In fact, transport cost is not really a factor that the industry owners are concerned with. This traffic of both inputs and output using the same transporters helps to create a no-empty movement of vehicles that cuts costs for all parties involved.

A combination of strong local entrepreneurship, local factors of land and infrastructure, and location advantages are the key features emerging from a case study of the Jalna steel industry.

b Nashik

Unlike Jalna, Nashik is not a single industry case. Broadly speaking, the industrial map of Nashik can be said to be one consisting of large industry and its vendors ranging in size from medium to micro, and associated services. There is a combination of reasons, in such a scenario, for the stagnation or decline of this location. Each reason that was corroborated by more than one respondent will be presented separately. The interviews in Nashik were conducted across industry types with both small and medium units participating. The following respondents were interviewed:

a. 11 medium to large size industry owners, of varying products, 5 vendors of intermediate inputs, and 6 making stand alone products

b. 15 members of an association of small and medium industry owners, all vendors of intermediate inputs

c. 2 labour contractors who supplied workers at all skill levels to all size industry, both handling upward of 500 persons on their rolls; 1 labour consultant who worked with industry to understand requirements and create plans for fulfilling them; 1 labour union organiser with a long association with public sector unions and later working with unionising private sector workers.

d. Head of the local MIDC office in Nashik, and other local government officials in the Department of Industry.

We start by looking at labour issues in Nashik.

1 Labour

Traditionally, Nashik had a large component of migrant workers, mainly from UP, Bihar and Rajasthan, and then from surrounding districts and other regions of across Maharashtra. There were two big shocks to the established system of industrial workers of Nashik. The first was 'militant unionism' in the mid to late 1990's that led to a number of closures, and the industry turned to contract labour on a large scale to avoid further conflict. The second was the sons-of-the-soil movement that was started by a political party which by around 2002 drove away nearly the entire lot of workers hailing from northern States out of Nashik, many of whom never returned. Since then, Nashik suffered from a chronic shortage of workers in all skill categories well into the 2010's. There appears to be a consensus among the interviewees that this resulted in at least a 30% worker deficit in Nashik. In the case of the skilled category, there was a unanimous view that fresh recruits who were coming out of engineering colleges were of low quality, many of whom are virtually unemployable. As a result, there was even a move by the local industry to collaborate with local colleges to improve training at the degree level while, at the same time, a collective effort was also made to start a training centre with a view to improve skill levels.

Moreover, since the early 2000s, there has been restructuring of labour needs largely mediated by infusion of new technology. Large industry is increasingly outsourcing production jobs that require precision machinery to produce intermediate inputs of a specified quality, while largely keeping assembly lines in-house. This has meant that skilled labour (and jobs) have spread into small and medium industries, who were unable to afford hiring such people at expected high salaries that large industry would have normally paid. As a consequence, small industry has become the training ground for larger units, while suffering high rates of attrition, as the best trained workers move to better pay in larger industry. Since the need for regularly employed skilled workers had reduced, the proportion of low to unskilled contract workers in large industry has increased. This analysis was corroborated by small and medium industry owners and labour contractors. The result was an overall reduction in aggregate worker numbers.

To summarise, on the labour front there were two main issues that have contributed to the picture of Nashik between 1998 and 2005: labour shortages due to local political factors, restructuring of labour use by large industry and unfulfilled need for high quality of skilled workers. There was a persistent sense expressed by all those interviewed that among the younger generation there was 'an unwillingness to work'. This was variously blamed on better working conditions in the service sector for similar wages, recent good performance of the surrounding agricultural belt, increasing aspirations due to urbanisation and so on.

2 Land

Issues relating to land were considered the most pressing of the problems that probably resulted in the slow-down of industry activity in Nashik. Since the early 2000's, the State had not been able to make available significant amounts of land for expansion of the existing industrial units or for new entrants. Price of industrial land had also increased substantially to make the threshold cost for new entrants prohibitive. As a result of high prices, there was even speculation in industrial land, with owners of closed units holding on to their land in the hope of getting better prices. Small industry owners who wished to expand were frustrated enough to demand that MIDC should create a per square meter productivity criterion (not really defined) used to put a ceiling on how much land large industry could 'hoard'. Medium size industries had their expansion plans undertaken in units dotted all over the industrial estates in a disjointed manner, due to lack availability of contiguous large plots elsewhere.

Lack of land for expansion and high prices of available land was the reason for some crucial large industry downsizing or moving out of Nashik altogether, dealing a severe blow to the prospects of its existing large vendor base. Several large industries, mainly in electrical machinery and goods, moved away from Nashik from about 2008. Many moved to Baroda in Gujarat. We visited Baroda for some insights in this process. Out of the large array of issues that were mentioned by various types of industry owners in Baroda, the main reason for this relocation of large industry emerged to be the availability of land on a scale that would satisfy their expansion needs for the next 20 years. Gujarat had spent significant amount of resources to persuade industry to locate in newly created industrial estates, specifically tailored for the needs of large industry. In fact, giving land for the Tata's Nano car factory to move from West Bengal after it ran into land related conflicts there, was considered one example of 'poaching' from other states (This was a clear example of the stiff competition between states to attract industrial investment and location. We found many large industries in the new industrial estates of Baroda that had relocated there from other states, including Maharashtra, all giving the same reason of land availability.)

While in Jalna availability of MIDC land led to creation of a new cluster, in Nashik the lack of planning for future land requirements of existing industry and for new entrants, led to a downturn in the fortunes of this location. According to the MIDC representative in Nashik, it was not difficult to acquire new land, but the process was highly time consuming. But the State seems to have underestimated the time scales over which it would be required.

3 Post reform emergence of competition

This was one aspect of the industrial performance in Nashik that spoke directly of the new business environment after the reforms. It is mentioned here briefly since we have very little concrete data to support it but was corroborated by some very thoughtful industry owners.

The one effect of reforms was that large industry could source their inputs from a variety of vendors. Those who had put up industry in prime locations within the Nashik industrial estates on the basis of licenses in an earlier era, suddenly found themselves having to compete with a new brand of entrepreneur on the one hand, and new, highly specific demand from large clients, with increasing emphasis on quality control, along with capital intensive technology, on the other. Many of them simply collapsed since they had no capacity to compete in this new environment. The creation of a competitive market after reforms appears, at least partly, to have been a corrective mechanism weeding out inefficient production. The shift towards increasing automation of the industrial processes seems to have coincided with the issues of labour and land outlined above and seems to have peaked around the early part of the decade, 2001-10.

Given the large industry - vendor nature of intra-industry relationships in Nashik, the decisions of large capital regarding location were always going to prove pivotal. Their decision to locate here had led to its growth, and it seems, their decision to move away had led to its decline. A few enterprising vendors followed them to these new locations. But there was a realisation that transport cost was not a consideration any longer, and many vendor bases were becoming pan-India. Thus, it was necessary for vendors to look beyond their own backyards to get new business, not an easy proposition for the small and medium industries.

We also came across one case of a large luggage producing industry moving its entire production line to China in this period. This led to a distressing effect on its large and well-established vendor base and was yet another contributory factor to unemployment, migrants leaving, and overall decline.

From the above analysis, it appears that Maharashtra failed to please anyone to properly develop Nashik as a good location, especially from a longer term perspective; not the large industry, not the medium and small vendors, and not the workers, who lost jobs in the process. Being focused on metropolises does not mean that other existing locations should be ignored. Nashik fell through the large cracks in the State's industrial policy, perhaps pointing to a certain complacency on its part to view itself as an 'industrialised state'.

The case studies reveal that local factors are paramount in assessing the potential of a particular location for industry and many of these are unquantifiable especially in a policy process that is highly State centric. In addition, to lump them under the rather vague heading of 'transport costs' reported in past modelling studies is to override policy corrections that States need to make and also to be vigilant about continuing to nurture existing locations in the long term.

Conclusions

In this paper, we present a spatial analysis of industrial location in the Indian context that does not confine itself to State administrative boundaries. This has been done using information about date of completion of industrial projects, addresses and products for individual plants to construct a spatio-temporal picture of location choice over the period 2002 - 2018.

Location decisions appear to align closely with the overall economic outlook, both current and near-term future. It is shown that in good economic times, many processes of industrial location operate simultaneously. So the Krugman and Livas [1996] process that sees dispersion due to liberalised trade policies can be applied to the peripheries of major concentrations in India pushing outward during good economic growth, taking advantage of the large rural market in which demand has improved. However, this process of dispersal actually reverses, and locations contract inward towards agglomerations in times of low economic growth. The Paluzie [2003] process relating to the cores of historical concentrations growing due to mobility of labour also operates and they become more dense. The primacy of agglomerations in generating positive externalities of both localisation and urbanisation economies [Chakravorty et. al., 2005], along with increasing returns to scale remains the strongest centripetal force. But once smaller locations get established in good times, they have the potential to become important regional foci for future growth since they now have the advantage of 'history'. Our cluster analysis gives a picture of this cumulative process that operates over long time periods in apparently sparse locations, as is the case of Bihar. The large clusters in India transcend State boundaries showing that State policies might have limited impact on persuading capital to locate in 'backward' areas. Their persistence also shows that capital has very clear ideas about where to locate. We also show that along with plant location, industrial employment is clustering quite severely.

This kind of analysis without the confines of administrative boundaries, and crucially, a clean temporal data set that allows for compelling interpretations is unusual. But with private firms gathering such data on large scale, it might be possible to study these patterns more closely, with more parameters such as employment and capital investment. It is bound to pay rich dividends in proving new insights.

The case studies of Nashik and Jalna in Maharashtra showed that this narrative can be complicated by local factors, a powerful force sometimes creating new locations, and at others leading to decline of existing ones. Different locations may behave very differently in response to the same overall economic conditions and policies, depending on local factors. Economic location theories do not take into account these local factors. But they still give a structure to work with, so that patterns that have been developing over a long period of time can be analysed with a view to learning from their complex history.

The question that now needs to be addressed, in the context of what the Economic Survey [2015] held and that was quoted at the start of this paper, is what kind of policy at the State or the Central government level can support these various processes. The goal of reducing regional inequality and creating better skilled employment in low industrialised areas can be served only if such analysis is available at a highly disaggregate level. Simply training workers for an as yet unseen skill requirement is too simplistic an idea in such a complex geography. Given rapid automation and declining requirement of a skilled workforce in manufacture over the past two decades, it is not clear what role only skill development can play in creating large scale employment in the formal sector.

NOTES

1. The knee point method finds the optimum values of Epsilon and Minimum Points by calculating the distance of each plant to its nearest n plants, then sorting and plotting the result on a line chart. The point at which this line shows a knee, gives the best value for Epsilon. Minimum Points is chosen to suit the calculated value of Epsilon.

2. Unit level data for Maharashtra from Economic Census of 1998 and 2005 and Annual Surveys of Industry (ASI) from 1998-99 to 2010-11, Population Census data for 1991, 2001, 2011.

3. The first family came from the Aggarwals, a trading community from Haryana, settled for several generations in Jalna. The second case involves the Maheshwari Jain community from Rajasthan, again traders in the region for several generations. These are the only two communities in the steel industry in Jalna.

PART II
EVIDENCE BASE FOR AGGLOMERATION FORMATION: PUBLIC GOODS, PRIVATE INVESTMENT

Sharadini Rath*

This paper relates to a geographical analysis of the development of large land area projects for housing, manufacture and services in India over the past 20 years. It establishes an evidence base for their location choices and how they correlate with trends in employment and State provision of basic amenities. For most part, large housing projects are highly focused in and around major urban agglomerations, catering to their significant service economies. Most large-scale service sector developments such as Special Economic Zones (SEZs) also prefer these locations. Only large-scale manufacture is seen to go further afield, and is not really served by major housing developments. The infrastructure and civic amenities that the State creates and maintains in large cities is also one of the major considerations for locations of many economic activities. While the locations of present agglomerations in India have deep historical roots, their continued growth and success are due to the exercise of this feedback loop: the more capital locates there and employment increases, the more the State has to spend on their amenities, to remain attractive for capital to move in, and so on. The location choices of large capital investments are important for the State to keep in mind in developing infrastructure to suit their needs, if it wants to see employment, and tax and non-tax revenue generation. The Jamshedpur model of creating a township for steel in the early 20th century and later for motor vehicles and other related products that goes beyond the development of a housing colony for its employees is not favoured currently by either major housing or industrial investors. From extensive fieldwork in Jharkhand and interviews with developers across India, it is clear that the governance of such entities is considered to be unsustainable for private capital. There is also no evidence that using this model can sow the seeds of new agglomerations or even towns with a wide economic base.

---

* Sharadini Rath is Affiliate Fellow of Indian School of Political Economy, Pune.
Acknowledgements: This work was supported by a research grant from Azim Premji University, Bangalore. Indian School of Political Economy, Pune, provided academic base. The Mahratta Chamber of Commerce, Industry and Agriculture, Pune, made available the Projects Today data. Officials in Jharkhand government in Ranchi, Jamshedpur and Bokaro shared their knowledge of many issues involved in complex governance situations. Ujaval Gandhi helped with geographical tools, data sources and many solutions to using them effectively. Ashish Andhale gave able research assistance. Abhay Tilak gave introductions to large developers in the Pune area and Ramanath Jha gave insights into policy and management issues of SEZs. Inventicon Business Intelligence created the opportunity to attend their conference on Integrated Townships. Contacts made in this conference within the Indian real estate community were invaluable in getting the industry view. All are gratefully acknowledged.

The Editor, JISPE is gratefully acknowledged for extensive comments which have improved this paper considerably.

1. Introduction:

In Part I, we showed that while the centripetal force of urbanisation economies retains its primacy, periods of good economic growth can act as a significant centrifugal force resulting in dispersal of industrial location, pushing out the boundaries of existing clusters, and even seeding new ones in previously low industrial regions. In Part II, we look at the geography of capital investment in large land area projects in housing, manufacture and services. We also look at how these patterns correlate with State provision of basic education and health services.

There have been some parallel developments in India since the turn of the millennium:

1. The rapid growth of the non-farm economy.
2. The enormous growth of the existing metropolitan agglomerations in the country in terms of populations, with a diversity of enterprises and employment.
3. The concurrent expansion of the ambit of private sector activity in all types and forms of construction projects, including very large ones.
4. The creation of policy instruments at the State and Central government levels to, variously, a) regulate existing forms of such private sector activity, b) respond to the demand for new forms, and c) innovate ab initio in the urban space to create economic growth impetus. Policy on governance of these new developments has been a logical corollary of these initiatives.

To create a context as to how the State views agglomerations, it is good to keep in mind that their revenue generating ability is hefty. We have estimated that by 2013-14, commercial tax from Bangalore Urban district alone made up 63% of total commercial tax collection in the State of Karnataka, 41% of its total tax revenue and 26% of total State revenue receipts. If other taxes and tariffs from Bangalore Urban are taken into account, this percentage share increases further. This is likely to be the case with all major agglomerations in India. They have become one of the mainstays for the source of redistributive expenditure that States undertake on basic amenities in the rest of the State. Under such circumstances, there is every reason for the State to create conditions to keep private capital satisfied in their location choices. The objective of this study was to examine and understand the links among the parameters stated above. The analysis lays down evidence for patterns of urbanisation that are driven by two different models of economic development: the industrial township built around a single large enterprise (public or private sector) such as the steel townships of Jamshedpur and Bokaro in Jharkhand, and the more recent organic agglomeration processes that have created large metropolitan areas like Hyderabad.

This analysis pertains to the period 2001 to about 2018. The data used here looks at the following parameters of urban growth:

1 Population
2 Employment
3 Real Estate Developments
4 Health services 5 Educational services

Data sources are as follows:

1 Population Census of 2001 [GOI, 2001] and [GOI, 2011
2 Economic Census of 2005 and 2013 [GOI, 2005; GOI, 2015]
3 Projects Today data for real estate projects over the period 2002 to 2018
Special Economic Zones data from Government of India
4 Field visit insights from Ranchi, Jamshedpur and Bokaro[1]
5 Interviews with real estate companies operating in many states of India[2]

The sites we examine are the following:

At the district level:

a An all-India perspective

For geographies being analysed at taluk/tehsil/mandal levels:

b Districts of Medak, Hyderabad, Rangareddy, Mahabubnagar and Nalgonda in Telangana state
c Districts of Bokaro, Ranchi, Saraikela Kharsawan and Purbi Singhbhum in Jharkhand state

The study methodology makes extensive use of geographical analysis, using a distance metric to examine patterns of locations of investment, economic activity, employment, and basic amenities Insights from these patterns are used in conjunction with interviews with stakeholders to create a picture of choices being made by private capital for location, the response of the State in facilitating these with a view to deal with their consequences in terms of growth of populations. We begin with the all-India perspective

2. The All India Context:

To start developing the chain of evidence to understand which factors have gone hand in hand to create the geography of private greenfield developments, we started by looking at where they are predominantly located in India at the district level. We use Projects Today [2002 to 2018] data for this purpose. This data base follows a sample of capital investment projects from their inception to completion in real estate, manufacture, and service sectors. We understand from the analysis that the sample is geographically uniform. The data gives the following information about the projects:

1 Address of the project location
2 Completion year

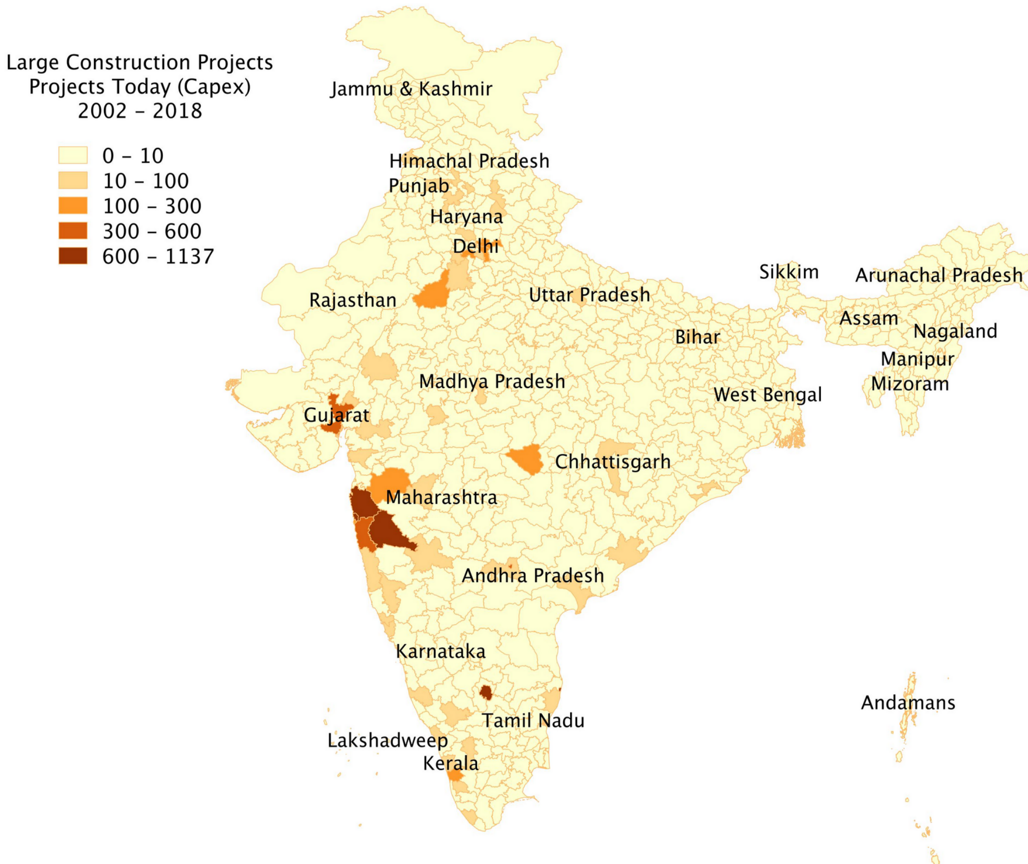
Other information such as size of investment and land size was patchy and not usable. However, for manufacture, the product information was complete.

This data gave us a valuable tool to do a spatio-temporal analysis by using geographical methods. We have used the project address to geolocate the real estate projects, and used the completion year in conjunction, to reveal how the location choices of real estate development companies have been shaped by broader economic considerations.

It should be made clear that the data on land use is very patchy, and many times impossible to corroborate, as is the data on the cost of these projects. Given these limitations, the only way to interpret this data was to understand their intensity of geographical location. Since private investment in developing residential properties is predominantly predicated on demand for such housing, it then indicated to us where such demand has been estimated to grow sufficiently to give good returns on investments.

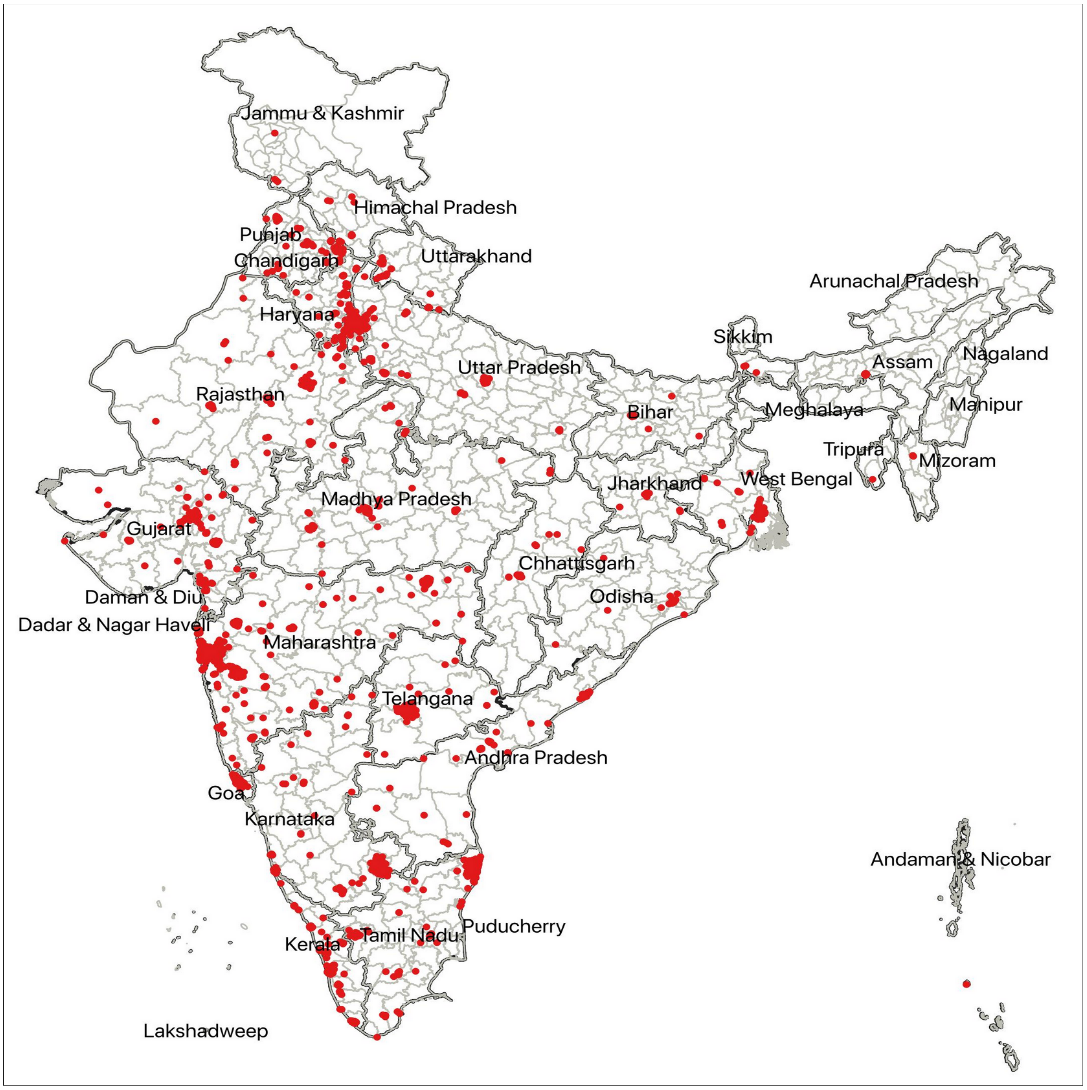
The predominantly western and southern locations of these projects becomes clearer by looking at Maps 1 and 2 and the top States and districts for their locations in Table 1. Of the total 8464 real estate projects obtained from Projects Today, just 4 districts, Thane and Mumbai Suburban in Maharashtra, Bangalore in Karnataka and Chennai in Tamilnadu, make up a 41.5% share.

Map 1. Number of completed construction projects from 2002 to 2018 (Projects Today)



Source: Author’s calculations from Projects Today data.

Map 2. Locations of Real Estate Projects (2002-2018) (Projects Today)



Source: Author's calculations from Projects Today data.

Table 1. Top states and districts for real estate projects (Total Projects = 8464)

State	No. of Real Estate Projects	% share in Total	District	State	No. of Real Estate Projects	% share in Total
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Maharashtra	3734	43.99	Thane	Maharashtra	1137	13.39
Karnataka	962	11.33	Mumbai Suburban	Maharashtra	909	10.71
Tamil Nadu	765	9.01	Bangalore	Karnataka	852	10.04
Gujarat	505	5.95	Chennai	Tamil Nadu	625	7.36
Telangana	413	4.87	Pune	Maharashtra	619	7.29
Kerala	369	4.35	Hyderabad	Telangana	360	4.24
Uttar Pradesh	350	4.12	Ahmedabad	Gujarat	345	4.06
West Bengal	307	3.62	Raigarh MH	Maharashtra	325	3.83
Haryana	272	3.20	Kolkata	West Bengal	269	3.17
Total	**7677**	**90.43**	Nashik	Maharashtra	252	2.97
			Ernakulam	Kerala	211	2.49
			Nagpur	Maharashtra	184	2.17
			Mumbai	Maharashtra	173	2.04
			Jaipur	Rajasthan	163	1.92
			Gurgaon	Haryana	153	1.80
			Ghaziabad	Uttar Pradesh	144	1.70
			Gautam Buddh Nagar	Uttar Pradesh	140	1.65
			Total		6861	80.82

Source: Projects Today [2002-2018].

Maharashtra by itself, with the addition of Pune, Raigarh, Nashik and Nagpur districts accounts for an overwhelming 44% share of all real estate projects completed in India over the period 2002-2018 and followed by Projects Today. The top districts that make up 80% of the total projects either have a metro city, or are on the borders of one.

Of the total 8464 projects, 8401 were residential developments of various kinds; apartment complexes, bungalows, affordable and lower income housing, etc. Of these, 120 were classified as ’township’, most probably meaning that they had other amenities included, and were large enough in area to require environmental clearance at either Central or State level. However, they were all predominantly residential in purpose.

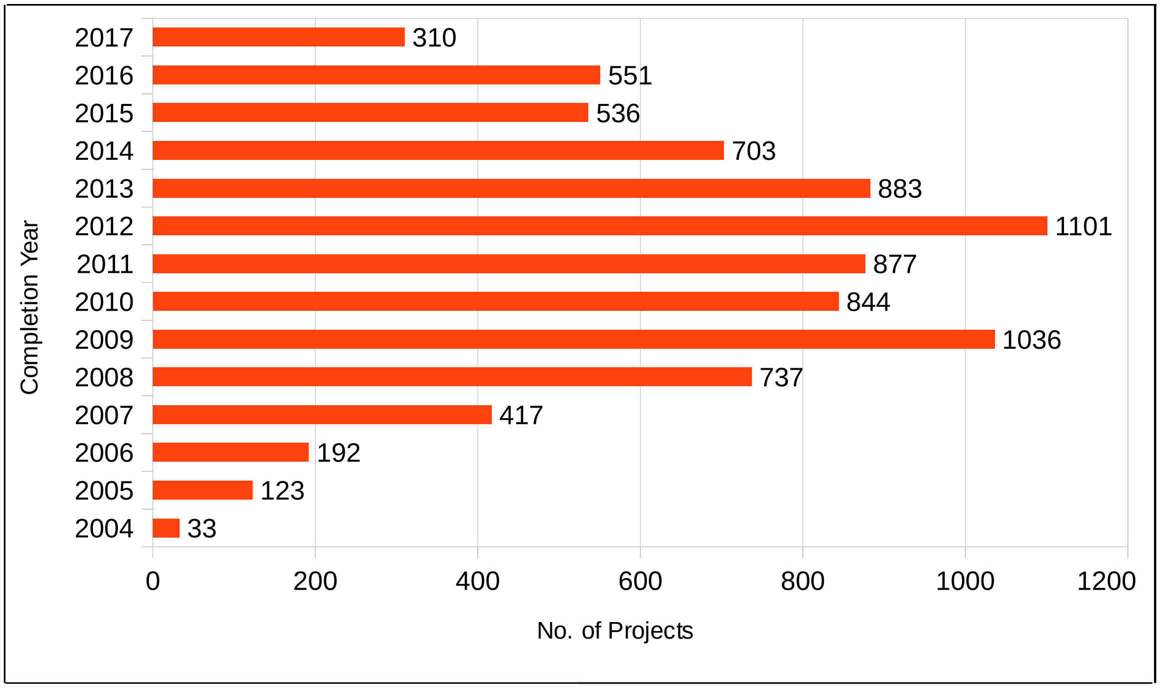
Using the completion year of the projects, we looked at the temporal aspects of real estate development. Figure 1 shows that there has been a surge of real estate development activity starting in 2007 with 417 projects and peaking in 2012 with 1101 projects completed. The fall in construction projects over the next four years has been precipitous, and by 2017, levels have fallen below those a decade earlier to just 310 projects completed.

This pattern matches almost exactly that of completed plants in manufacture in Part I. It also matches with the GDP growth curve over this period, if the time lag between location decision and project completion is taken into account. Capital investment decisions are based on medium term prospects of getting good returns in all sectors.

To have been completed in 2012-13, the investment decisions and processes of land acquisition for real estate projects would have started at least in 2007-08, if not earlier for larger projects. So just before or at the time of the global financial collapse of 2008, real estate investors appear to have been quite optimistic about the Indian economy, as far as demand for housing was concerned However, from about 2011-12, this has clearly changed, with projects being reduced quite drastically. Maharashtra had 175 projects completed in 2007, peaking at 477 in 2012, and falling to 161 in 2017. Karnataka, Tamilnadu, Gujarat, and other significant States showed the same trend. While CAPEX (capital expenditure) data of this type shows the behaviour of large investors, it is, of course, possible that local activity in building small residential properties with relatively smaller requirement of capital might have continued. However, the incremental housing capacity added would also have been much smaller.

These patterns also match Map 2A (in Appendix in terms of non-agricultural employment in 2011 especially seen in parts of Punjab, Haryana and regions of western UP bordering Delhi, the Jaipur region in Rajasthan, parts of Gujarat, western Maharashtra, Hyderabad and environs, Bangalore, Chennai and environs, southern areas of Tamilnadu around Coimbatore, parts of coastal Kerala around Ernakulam, and a cluster in southern West Bengal. These have been the dominant areas of non-farm employment since 2001, and remain so in 2011 also. It should be noted that a map of the share of non-farm workers in total workers at district level also gives the same regional spread (Map 2A in Appendix). The dominance of peninsular India in non-farm employment persists. The absolute number of these workers has gone up across the board as population has increased, but this increase has been most significant in districts which have large metro cities.

Figure 1. Spread of real estate projects by year of completion from 2004 to 2017 (Projects Today)



Source: Author's calculations from Projects Today data.

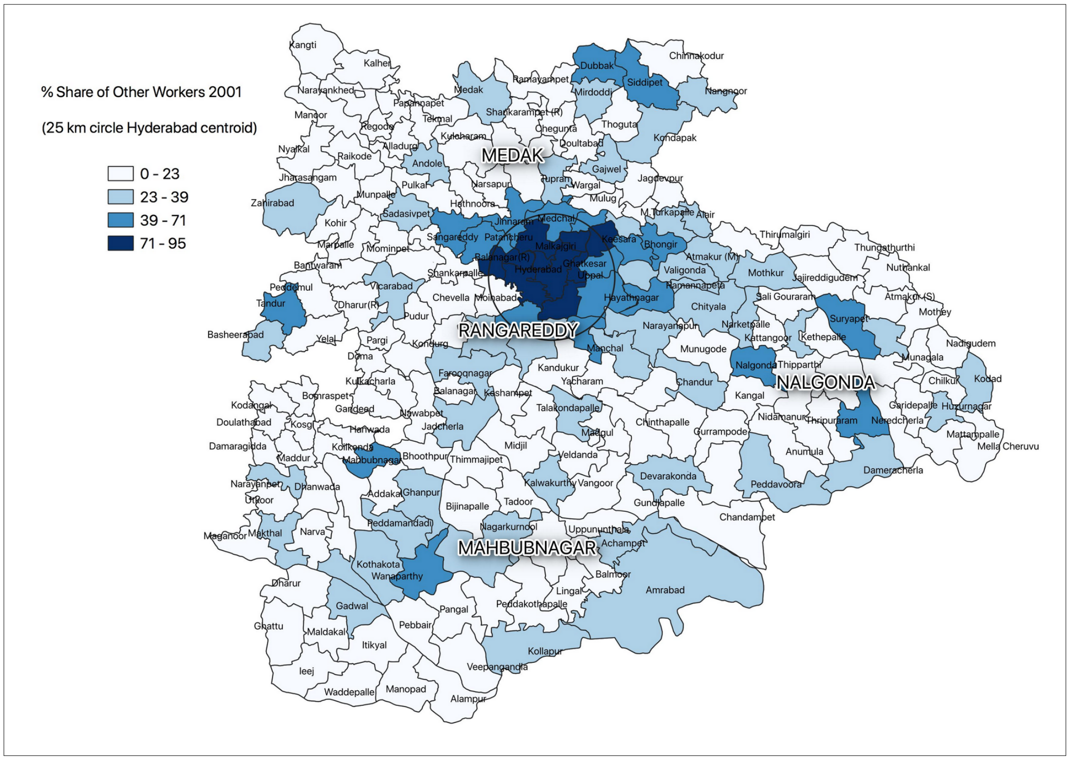
Map 3A (Appendix) gives the number of workers in non-primary sector enterprises that hire at least one person from the 2013 Economic Census [GOI, 2013]. This brings this basic regional spread into further focus. Real estate, largely housing, projects are concentrated where non-farm employment has peaked. It is also clear that the growth in such projects from 2007 to 2013 was well chosen by location to coincide with the growth of not just urban, but predominantly metro agglomeration economies as measured in employment.

After geolocating the real estate projects, their distance from the respective district headquarters was measured (using the centroid of the city municipal area). Geolocation was possible for 7423 projects out of the total 8401 housing projects. Table 2 gives the breakup.

Table 2. Distance of real estate housing projects from respective district headquarters (Projects Today)

Distance from District HQ (km)	No. of Projects	% share in Total
(1)	(2)	(3)
0 to 10	4869	65.6
10 to 25	1395	18.8
25 to 50	905	12.2
50 to 100	252	3.4
Total	**7423**	**100**

Map 3. Taluk level percentage share in total of non-farm workers in Polulation Census 2001 (25 km circle around centroid of Hyderabad district



Source: Author's calculations from Population Census 2001.

It becomes clear that an overwhelming majority of these projects were within a 25 km radius of the district headquarters. All metros are also district headquarters, and sometimes even State capitals. Their metropolitan regions can extend far beyond municipal boundaries, as much as 25 to 30 km in radius. In fact, it is very likely that boundaries of metropolitan regions are notified by taking into account planned and completed real estate housing developments.

At the all India level, we finally dealt with one last category of large land development: the Special Economic Zones. These entities are areas of economic activity along the guidelines outlined in the SEZ policy. The most recent list of notified SEZs was taken from the Press Information Bureau [GOI, 2022].

Along with the address and type of economic activity, this list also gives the area of the SEZ, the only data set in which this has been made available.

None of the SEZs in this list claimed to have any housing facilities within their boundaries. The list was used to do the same exercise of finding out how far they are from their respective district headquarters. Of the total 374 listed notified SEZs, it was possible to geolocate only 171. It is possible that the remaining ones, while notified, have not actually come into existence as working entities (one such example of the notified but non-working SEZ was seen during a field visit in the vicinity of Bangalore).

For the 171 SEZs that were identifiable, Table 3 gives their distance spread from their respective district headquarters, by land area. The Table also gives the breakup by two broad types of economic activity in the SEZ: manufacture and IT/ITES. Of the complete list of 374, 194 are IT/ITES and 153 manufacture. Given that originally, the SEZ policy wanted to recreate the manufacturing export zone models in China, it is interesting to note that more than half of them are now IT/ITES. The manufacturing SEZs tend to be more distributed up to about 100 km from the district HQ and have larger areas. The IT/ITES ones are more closely clustered within a 25 km radius and are smaller in area. This is again quite logical given that manufacturing activities tend to require far more space than service activities of the IT/ITES variety Even though both are likely to be greenfield projects, their space requirements are very different, as are the needs of their specific operations, such as plant activity, warehousing and goods transport for manufacture, and office space for IT/ITES. For large land parcels to be available, and at a reasonable price, it is necessary to go further out from city limits.

Of the 83 IT/ITES SEZs, 17 are in Bangalore, of which 12 are less than 20 hectares in area and located within 25km of the city. This is the largest such cluster in the country. In manufacture, Andhra Pradesh has the largest number, with 20 spread mostly over the coastal districts of Nellore, Vizianagaram, Vishakhapatnam and Kakinada.

In fact, the preference of coastal districts for manufacturing SEZs has also led to 16 in Gujarat and 14 in Tamilnadu, mostly in coastal areas. This makes sense, since SEZs were originally meant to house large export manufacturing enterprises, with concessions in tax and labour laws. Vicinity to ports can be a significant location factor. However, the large share of IT/ITES among SEZs is indicative of perhaps the basic failure of this policy aim.

Table 3. Distance of SEZs from respective district headquarters by area of SEZ

Distance to District HQ (km)	Area (HA) - IT/ITES						
	10 & less	10 to 20	20 to 50	50 to 100	100 to 500	500 to 1000	1000 & more
0 to 10	9	24	8	3	2		
10 to 25	2	13	10	1			
100 & more		1					
25 to 50	1	6	2		1		
Total	**12**	**44**	**20**	**4**	**3**		
Distance to District HQ (km)	Area (HA) - Manufacture						
		10 to 20	20 to 50	50 to 100	100 to 500	500 to 1000	1000 & more
0 to 10	10 & less	8	4	3	10	1	1
10 to 25	1	2	2	1	5		2
100 & more		2			4		
25 to 50		5	1	3	11	1	5
50 to 100	1	1	1	1	8	1	4
Total		**18**	**8**	**8**	**38**	**3**	**12**

It also becomes clear from this analysis that real estate housing developments are not really catering to the large land developments in manufacture but are more geared towards the location of service sector companies and their employees, Both are located with about 25 kms from major urban agglomerations.

In the Indian context, agglomeration formation has increasingly relied on IT/ITES and related services to generate demand. The tweaking of the SEZ policy in 2006 [GOI, 2006] accommodated some realities of this aspect, that services was the bulwark of growth, especially urban growth. And this sector wanted to be located within an agglomeration, much more than the large area based manufacture sector. A diversified labour pool, access to urban amenities, and increasing returns to scale were all being realised. Service enterprises with smaller land needs closer to metros were accommodated within the SEZ policy over a period of time. Policy was tweaked to needs of demand, not the other way round.

The analysis so far gives an idea of the links between large scale real estate activities, employment concentrations, periods of economic growth and types of economic activity over the past two decades.

Employment, Public Goods and Private Capital:

In this section will demonstrate the link between the geographical pattern of location of capital and employment with State provision of basic amenities like health and education. This will be done on a much smaller scale in terms of the following districts in Telangana and Jharkhand:

Telangana: Hyderabad and surrounding 4 districts of Mahabubnagar, Ranga-reddy, Nalgonda and Medak.

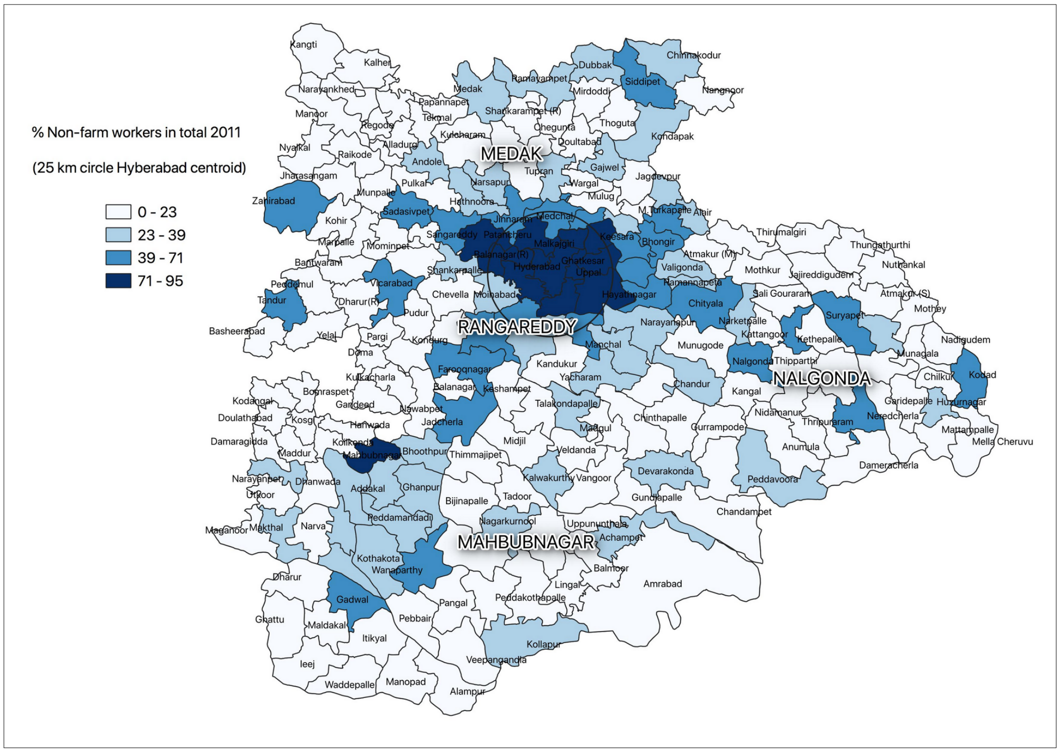
Jharkhand: Ranchi, Purbi Singhbhum, Saraikela Kharsawan, Bokaro.

Maps 4, 5 and 6 (given below) show the close relationship between the growth of non-farm workers in Hyderabad and surrounds and real estate developments in the region. More specifically, Maps 4 and 5 show that for Hyderabad, the growth of non-farm workers seems to have spread preferentially in the west-by-north-west direction of Hyderabad, into Serilingampally and further and to a lesser degree to the east. Map 6 shows the geocoded locations of the real estate developments in Projects Today. These show a clear preference for the western region, where much of the IT related service sector has located, all well within a 25 km radius centred on Hyderabad.

A geographical analysis of this type also gives a clear picture of where the State is creating basic amenities. Map 7 shows beds per 10,000 population in government hospitals at the taluk/mandal level. The central place of Hyderabad as State capital and other district head-quarters in healthcare provision becomes very clear.

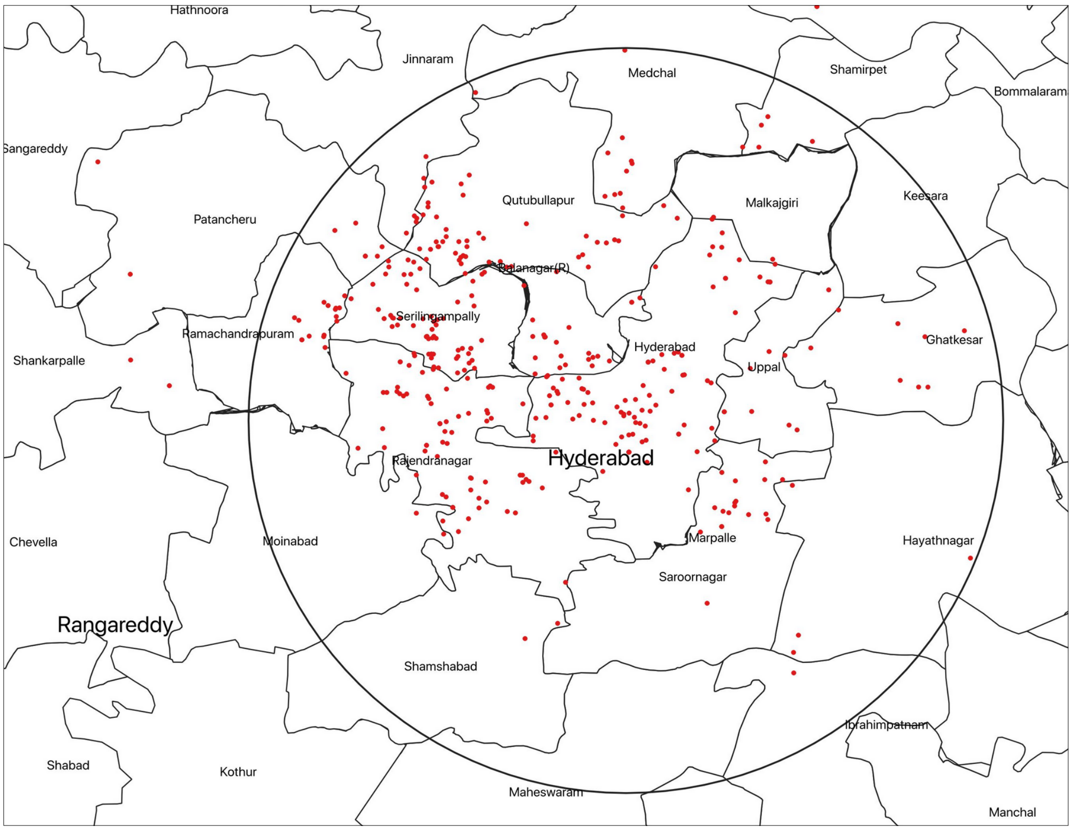
The State does indeed treat agglom-erations, as also capital cities and district headquarters, preferentially even in basic amenities. Similar analysis based on Economic Census data shows that the gap left by the State in basic health and education services is filled by the private sector, in areas poorly served by the state. However, location choices by large pri-vate capital are made leveraging the vital presence of the State in agglomerations to maximise its returns to scale.

Map 4. Taluk level percentage share in total of non-farm workers in Population Census 2011 (25 km circle around centroid) Hyderabad district



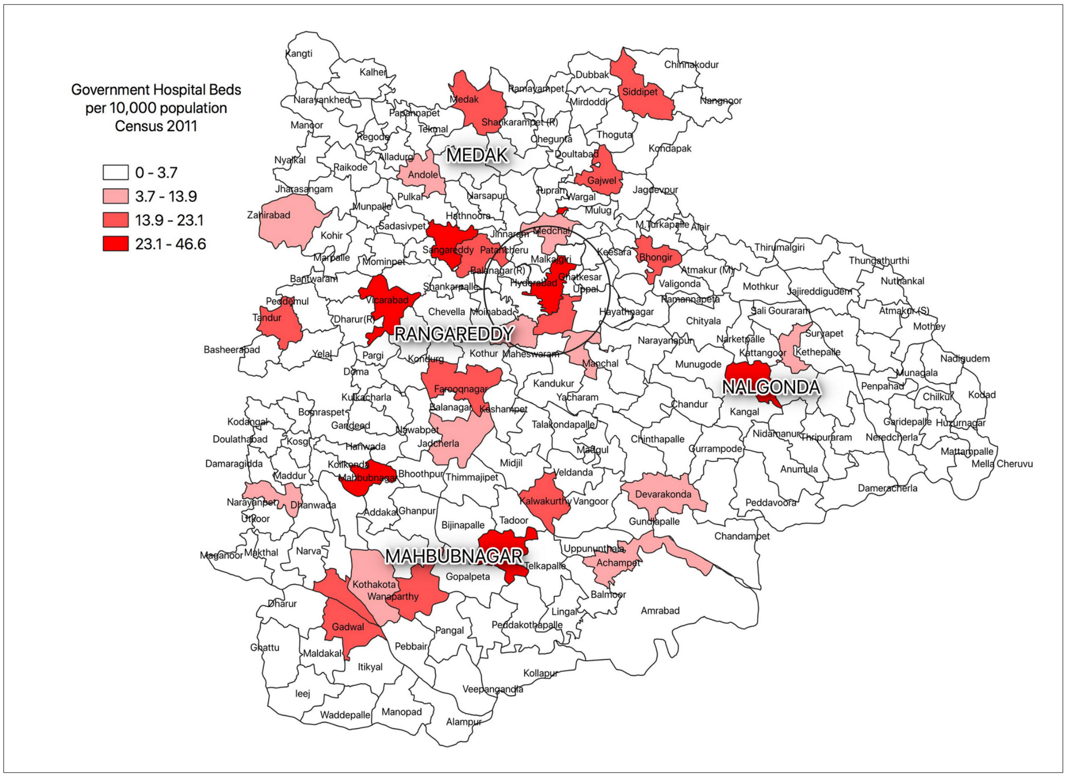
Source: Author’s calculations from Population Census 2011.

Map 5. Geocoded locations of Projects Today real estate developments in 2002-2018 (25 km circle)



Source: Author's calculations from Projects Today data.

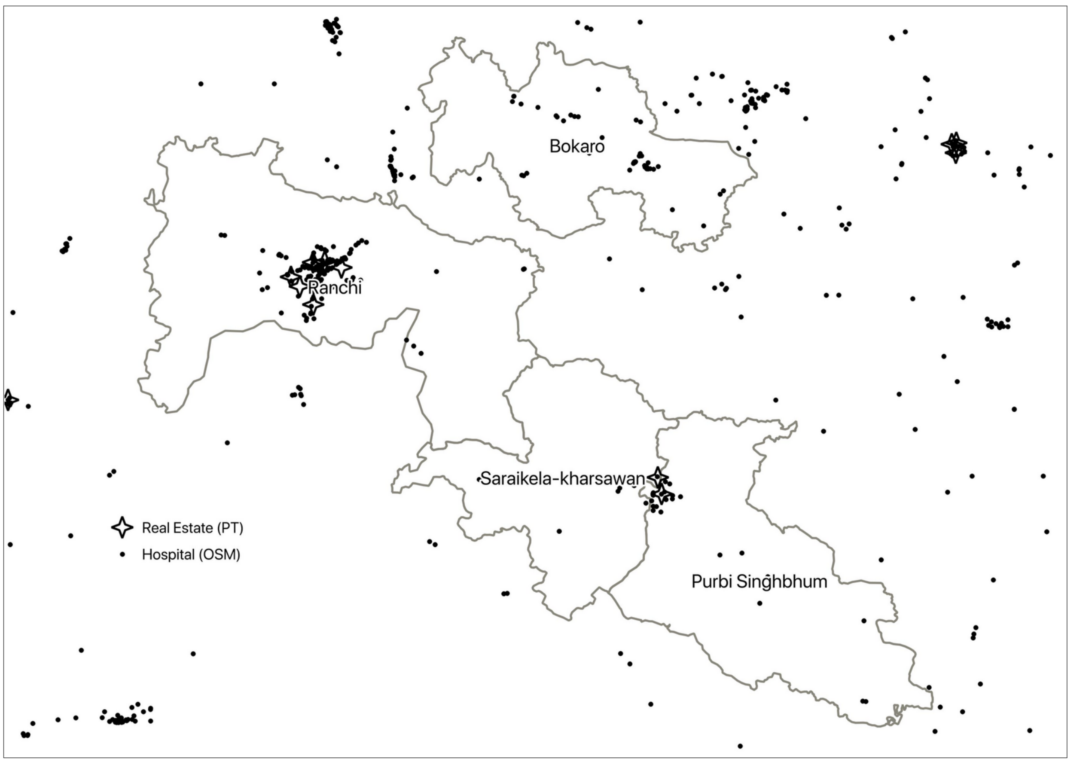
Map 6. Government hospital beds per 10,000 population for taluk/mandal from Population Census 2011



Source: Author's calculations from Population Census 2011.

Map 7. Projects Today real estate projects (star) and Open Street Maps

Hospital locations (dot) in 4 districts of Jharkhand



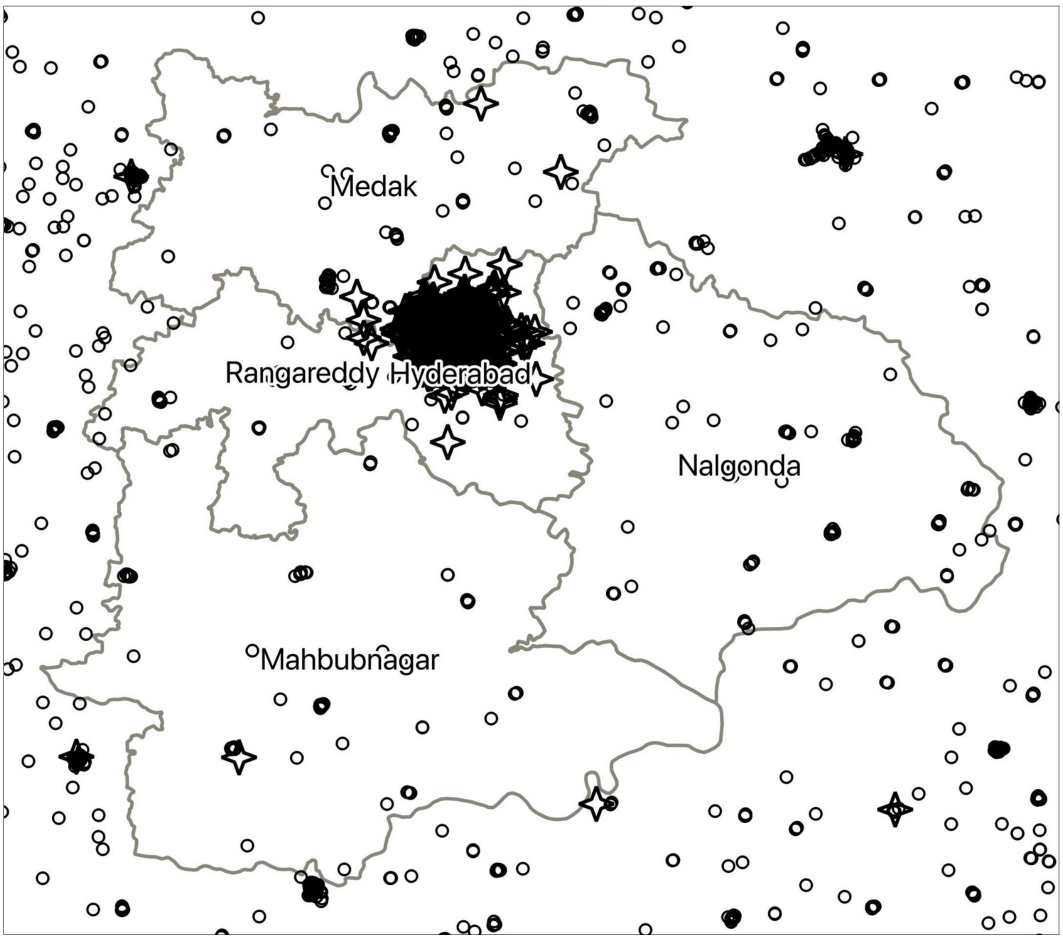
Source: Author’s calculations from Projects Today data and Open Street Maps.

Table 4. Some representative amenities in all towns (of all sizes) in 5 districts of Telangana and 4 districts of Jharkhand, from Town Amenities for Census 2011, categorised by the distance of the towns from their respective district headquarters

State	Distance to respective District HQ	No. of Towns	No. of Towns (%)	Total Town Population (%)	No. of Hospital Beds (%)	No. of Doctors (%)	No. Govt. Polytechnics (%)	No. Private Polytechnics (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Telangana (5 districts)	0 to 10 km	14	15.91	44.63	75.08	80.54	32.14	45.45
	10 to 30 km	30	34.09	41.99	10.66	8.31	25.00	36.36
	30 to 50 km	12	13.64	2.87	2.80	1.94	10.71	9.09
	Above 50 km	32	36.36	10.52	11.46	9.21	32.14	9.09
Telangana	**Total**	**88**	**100**	**100**	**100**	**100**	**100**	**100**
Jharkhand (4 districts)	0 to 10 km	16	25.00	74.37	81.32	74.84	60.00	100.00
	10 to 30 km	24	37.50	13.08	7.29	14.61	40.00	0.00
	30 to 50 km	12	18.75	7.34	5.66	4.42	0.00	0.00
	Above 50 km	12	18.75	5.21	5.72	6.13	0.00	0.00
Jharkhand	**Total**	**64**	**100**	**100**	**100**	**100**	**100**	**100**

From a slightly different geographical analysis, Table 4 gives the patterns for some of the basic amenities provided by the State at town level using the metric of distance from the respective district headquarters for 5 districts of Telangana and 4 districts in Jharkhand. This Town Amenities data from Population Census 2011 presents evidence of the degree to which private capital relies on State provision of basic amenities to make location choices. The provision of polytechnics shows that the State has wider geographical presence to provide such services to populations away from cities. There is no evidence to suggest that private capital can be relied upon to create wide ranging amenities and services away from cities and agglomerations. To illustrate this point, we looked at the overlap of large real estate projects and hospital (public and private, from Open Street Map amenities) locations in the 5 districts around Hyderabad and the 4 districts in Jharkhand in Maps 8 and 9. The presence of large real estate in Jharkhand is most significant around the State capital Ranchi, and to a much smaller degree near Jamshedpur (the border of Saraikela Kharsawan and Purbi Singhbhum). There are no such projects in Bokaro. The overall pattern of district headquarters being the locations for amenities, public and private, also becomes clear with these maps.

Map 8. Projects Today real estate projects (star) and Open Street Map Hospital locations (circle) in districts of around Hyderabad



Source: Author's calculations from Projects Today data and Open Street Maps

In the next Section, field insights from Jamshedpur, Bokaro, and other inputs from across India are given with a view to understand how private capital, old and new, is thinking about their role in structuring large developments.

Jamshedpur, Bokaro and Others:

Fieldwork was done in Jharkhand in Ranchi, Jamshedpur and Bokaro. In Ranchi, a range of state government officials spoke freely and clarified many aspects of their views on the place of Jamshedpur and Bokaro, the two industrial cities centred around steel production, within the governance structure of Jharkhand. Bokaro and Jamshedpur were also visited and officials of Bokaro Steel City and senior officials in Tata Iron and Steel Company (TISCO), who were in charge of running Jamshedpur, were interviewed.

There is a fundamental difference between Bokaro and Jamshedpur, in that the land given to the Steel Authority of India Limited (SAIL) for the Bokaro steel plant is transferred to them in perpetuity, as long as it is used for industrial purposes. The land on which the Tatas built Jamshedpur is leased to them, and the lease has to be renewed every 30 years. This means that their relationship with the State is completely different. Bokaro Steel City is basically what one can call an industrial township; the steel plant, its peripherals and large housing colonies for employees only, with amenities built for all its residents. It is run entirely by SAIL Jamshedpur has this core with Tata Steel, Tata Motors and other Tata companies, their housing and amenities, which they administer themselves. But, it has in in addition the outgrowth around this core, which is administered by a Notified Area Committee (NAC). The NAC consists of nominated members from Jharkhand government and TISCO. There is no NAC in Bokaro and the Bokaro Steel township is administered entirely by the Steel Authority of India.[3]

The population of Bokaro Steel City was about 4 lakhs in 2011, while that of Jamshedpur (and its outgrowth) was 6.7 lakhs. Both steel companies have one major factor in common: steel production has automated to a significant extent and their employment per million ton of steel has dropped precipitately over the previous 20 years, while production has steadily increased. Both were in the process of reducing their housing colonies by various means, including demolition. The major difference is that unlike Jamshedpur, Bokaro had not taken an interest in town development outside their housing colony, and there is very little outgrowth. The nearest municipal area is Chas, which is across a river, and has no administrative relationship with Bokaro.

It was clear from interviews with officials in the Urban Development Department (UDD) (in Ranchi that TISCO wanted a change in the arrangement of the lease. They were in favour of retaining the lease area on which the industrial plants and their associated housing amenities stood, and withdrawing from the NAC, with promises to continue providing utilities such as water and power as far as possible to the outgrowth outside the industrial township. The State agreed with this broad approach. They preferred a notified industrial township for TISCO and a municipal corporation combining the Jamshedpur outgrowth and some neighbouring smaller towns. They also wanted TISCO to hold to their promise of supplying water and power to the municipal corporation area in the long term with from the State budget. The long overdue lease renewal negotiation of TISCO with the Jharkhand government that had just concluded then had still not quite clarified this reorganisation of activities. One of the reasons for this delay was the fact that there is no working template for such a governance mechanism in India. There is not a single notified industrial township as per the 74th Constitution Amendment Act (CAA) in India and the UDD was in the process of trying to work out how to set it up without any precedent. At the time when field work was conducted, the reorganisation of the governance of Jamshedpur was still a work in progress.

The Jharkhand Industrial Infrastructure Development Corporation (JIIDCO) had another concern in this complex situation. They had spent money to set up the Adityapur Industrial Area close to Jamshedpur so that companies could leverage their proximity to this industrial area and grow into a true industrial cluster. However, TISCO officials made it clear that while they had supply contracts for both steel, motor vehicle and other parts with companies in Adityapur, the bulk of their intermediate inputs came from all over India, and even from other countries. Further, if as planned, TISCO made their new steel plant near Paradeep on the Odisha coast, the main base to take advantage of the thermal coal they are importing from Australia, it might adversely impact the prospects of the Adityapur industries. In interviews with these industry owners, it was clear that they were aware that supply chains for intermediate inputs for large industry were now global and physical proximity presented a smaller advantage. There was a sense that unless something changed dramatically, it was unlikely that

Adityapur would become a significant industrial location. And it was complicated by the lease negotiation.

There is an JIIDCO industrial estate near Bokaro Steel City (BIADA) also, which has an agreement with Bokaro Steel to set up ancillaries. There is an understanding that Bokaro Steel with give 30% of vendor tenders to units in BIADA. However, there are only about 120 medium sized such industries in BIADA. Such a formal understanding does not exist in the Adityapur industrial estate with Tata Steel or Tata Motors.

It is well known that the steel industry as a whole does not create a vendor base that has the capacity to become a broad industrial location. The vendors in Adityapur are vendors of Tata Motors, not Tata Steel. The motor vehicle industry, unlike steel, is heavily dependent on a highly specialised and broad vendor base. But Adityapur, in spite of being literally in the backyard of Tata Motors, has not grown into such a vendor cluster.

The opinion of State stakeholders across all departments was that neither Bokaro nor Jamshedpur have contributed to regional development. There is no overall industrial development in the State. In addition, they are left with a 'problem' in Jamshedpur, of an unruly urban outgrowth for which they have to find post-facto governance solutions. The Jamshedpur model is anything but a straightforward template for governance scale creation for other large manufacture or services developments. Status of land lease agreements, the evolution of manufacturing production technology and the viability of in terms of far flung supply chains facilitated by low transport costs were the factors that radically changed a long-held status quo and made a complex governance realignment necessary in Jamshedpur.

To get more insights and also to enable a closer, inside view of the real estate sector, the real estate industry conference in Mumbai on the topic of integrated townships in 2017 was extremely useful. It was attended by representatives of companies that are known for large urban housing developments across India. There were also participants from SEZs, and old public sector industrial townships similar to Bokaro, such as CIDCO in Mumbai Urban integrated townships have usually involved residential housing (apartments and/or houses) with amenities such as retail space, healthcare, schools and entertainment, for which there are many different internal arrangements for maintenance and lease/rental, with no mandate for 'governance'. Participants from SEZs were clear that while they maintained cordial relationships with towns/villages in their vicinity, they had no interest in taking on their governance which was best left to the gram panchayats or municipalities. They were also opposed to allowing housing for employees within their boundaries since this necessitated and brought in a completely different set of needs, which would not be sustainable. In fact, a senior official of a large southern SEZ went so far as to say that the Jamshedpur model should not be followed by large scale manufacture or service sector in the future. This was also the sentiment echoed by representatives of decades-old public sector industrial townships similar to Bokaro around Mumbai. It was becoming difficult to maintain the housing colonies in these townships, and there was no question of taking on formal governance of mixed towns.

Conclusion:

Large real estate housing projects have followed a temporal pattern that is similar to the manufacture developments in Part I. They show growth when they perceive good economic years in the short term, and decline when the outlook is poor.

There is strong geographical evidence from multiple data sources that large real estate development is deeply connected to hotspots of non-farm employment generation, for the large part in and around urban agglomerations in India. All of these agglomerations have long historical roots and have been formed by a steady feedback mechanism of location preferences of private capital, and State provision of basic amenities in response. Both capital and State are seen to reap rich revenue rewards from this decades long process; the former with good returns to scale and the latter from increased tax and non-tax revenues. In recent years, large scale service sector has played an increasing role in this process of agglomeration formation, locating close to large metros. It is also seen that large manufacturing and new models such as SEZs are not really served by the location choices of real estate, since they tend to need large land areas that makes them difficult to locate near metros due to high land prices.

State provision of infrastructure, such as industrial estates, roads and other amenities in non-agglomeration areas does not attract location of either manufacture or services, and as a result housing development is also at a low intensity in these areas.

There is no evidence that the Jamshedpur model of urban development leads to seeding of new agglomerations. In fact, industrial townships are seeing the effects of increasing automation in production, leading to lower workforce requirements. The rationale behind the creation of these townships has been overtaken by the pace of technological change, and they are no longer considered viable models for urbanisation in low industrialised areas and their economic development, even if private capital could be persuaded to locate there.

Geographical analysis, in conjunction with secondary data from government and other sources, can show revealed patterns of location choices by different stakeholders in the urbanisation processes. Planning for the consequences of these choices by the State can perhaps be made easier with such analysis.

NOTES

1. Interviews with: Jharkhand State Government officials in the Departments of Revenue, Urban Development and Industry in Ranchi; state officials in charge of industrial estates in Bokaro and Adityapur; senior executives in Tata Steel; some industry owners in Adityapur and Bokaro industrial estates; officials of the Bokaro steel plant and industrial township; commissioners of Jamshedpur and Chas municipalities

2. Attendance at the Inventicon Business Intelligence conference on Integrated Townships, Mumbai, 2018. Participants were representatives of major real estate development companies, some special economic zones and public sector industrial townships. Presentations and interviews.

3. A Notified Area Committee is created to administer an urban area that is transitioning from an industrial township to a municipality. It is created by a gazette notification by the State government. The NAC is not elected, but entirely nominated. It carries out all functions of administering the area under its jurisdiction with the help of state government appointed staff, similar in structure to a municipality.

REFERENCES

Isard, W., 1949; The general theory of location and space-economy, *The Quarterly Journal of Economics,* Vol. 63, No. 4.

Ahluwalia, M. S., 2000; ‘Economic Performance of States in Post-reforms Period’, *Economic and Political Weekly,* May.

Bhattacharya, B. B. and S. Sakthivel, 2004; ‘Regional Growth and Disparity in India: Comparison of Pre- and Post-Reform Decades’, *Economic and Political Weekly,* Vol. 39, No. 10, March.

Chakravorty, S., J. Koo, D. V. Lall, 2005; Do localization economies matter in cluster formation? Questioning the conventional wisdom with data from Indian metropolises, *Environment and Planning,* volume 37.

Chakravorty, S. and D.V. Lall, 2007; Made in India, The Economic Geography and Political Economy of Industrialization, Oxford University Press, New Delhi.

Chandrashekhar, S., A. Sharma, 2014; On the Spatial Concentration of Employment in India, *Economic and Political Weekly,* Vol. 49, No. 21.

Das, S K and Alokesh Barua, 1996; ‘Regional Inequalities, Economic Growth and Liberalisation, A Study of the Indian Economy’, *The Journal of Development Studies,* Vol. 32.

Ester, M., H. Kriegel J., Xu, X. Sander, 1996; Simoudis, Evangelos; Han, Jiawei; Fayyad, Usama M. (eds.). *A density-based algorithm for discovering clusters in large spatial databases with noise,* Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96). AAAI Press.

Ghani, E., Goswami A. Grover and W. R. Kerr, 2012; Is India’s Manufacturing Sector Moving Away from Cities? Policy Research Working Paper 6271, The World Bank.

GOI, (ASI, various years) *Annual Surveys of Industry* (various years). Central Statistical Office, Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

GOI, (various years) *National Industrial Classification,* Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

GOI, 2001; *Population Census,* Registrar General, Census Operations, Ministry of Home Affairs, Government of India, New Delhi.

GOI, 2002-2018; Number of Active Manufacturing Companies registered every year, Ministry of Corporate Affairs, Government of India, New Delhi.

GOI, 2005; *Economic Census,* Central Statistical Office, Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

GOI, 2006; *Special Economic Zones,* Press Information Bureau (Research Division), Ministry of Information and Broadcasting, Government of India, New Delhi.

GOI, 2011; *Population Census,* Registrar General, Census Operations, Ministry of Home Affairs, Government of India, New Delhi.

GOI, 2013; *Economic Census,* Central Statistical Office, Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

GOI, 2015; *Economic Survey 2015,* Department of Economic Affairs, Ministry of Finance, Government of India, New Delhi.

GOI, 2022; *Special Economic Zones,* Press Information Bureau (Research Division), Ministry of Information and Broadcasting, Government of India, New Delhi.

GOK, 2014; *State Budget Documents,* Commissioner for Commercial Taxes, Bangalore, Karnataka and Karnataka

GOM, 2012; *Maharashtra Human Development Report: Towards Inclusive Human Development,* Yashwantrao Chavan Academy of Development Administration, Pune, Maharashtra, India.

Hoover, E. M., 1948; *The location of economic activity,* New York: McGraw-Hill.

Hotelling, 1929; Stability in Competition, *The Economic Journal,* Vol. 39, No. 153.

Jacob, 1969; The Economy of Cities, New York: Random House.

Karan, P. P. and W. M. Jenkins, 1959; Geography of Manufacturing in India, *Economic Geography,* Vol. 35, No. 3.

Krugman, P., 1991; Increasing returns and economic geography, *Journal of Political Economy,* Vol. 99.

Krugman P, R Livas, 1996; Trade policy and the third world metropolis, *Journal of Development Economics*, 49.

Lösch, A., 1954; *The economics of location,* New Haven: Yale University Press.

Marshall, A., 1890; Principles of Economics. Macmillan, London.

Martin, R. & P. Sunley, 2003; Deconstructing clusters: chaotic concept or policy panacea? *Journal of Economic Geography,* 3.

MCCIA, 2002-18; *Year Books, Annual Reports,* etc. Mahratta Chamber of Commerce and Industry, Pune.

Paluzie, E., 2001; Trade policy and regional inequalities, Papers in Regional Science 80.

Porter, M., 2000; Clusters and the New Economics of Competition, *Harvard Business Review*.

Porter, M. E., 1990; "The competitive advantage of nations", First Edition, New York, The Free Press, New York.

Predöhl, A., 1928; The theory of location in its relation to general economics, *Journal of Political Economy,* Vol. 36, No. 3.

Projects Today, 2002-2018; Division of Economic Research India, Pvt. Ltd., Mumbai.

Saikia, D., 2011; The pattern of industrial location in India under liberalization: an analysis of the organised manufacturing industries, *International Research Journal of Management and Business Studies* Vol. 1, No. 7.

Saikia, Dilip, 2009; Industrial Location in India Under Liberalisation, Munich Personal RePEc Archive (MPRA).

Shetty, S. L., 2003; Growth of SDP and Structural Change in State Economies: Interstate Comparisons, *Economic and Political Weekly*, Vol. 38, No. 49, December.

Sjöberg O., F. Sjöholm, 2003; Trade Liberalization and the Geography of Production: Agglomeration, Concentration, and Dispersal in Indonesia's Manufacturing Industry, *Economic Geography*, Vol. 80, No. 3, July.

Smith, D. M., 1966; A theoretical framework for geographical studies of industrial location, *Economic Geography*, 42.

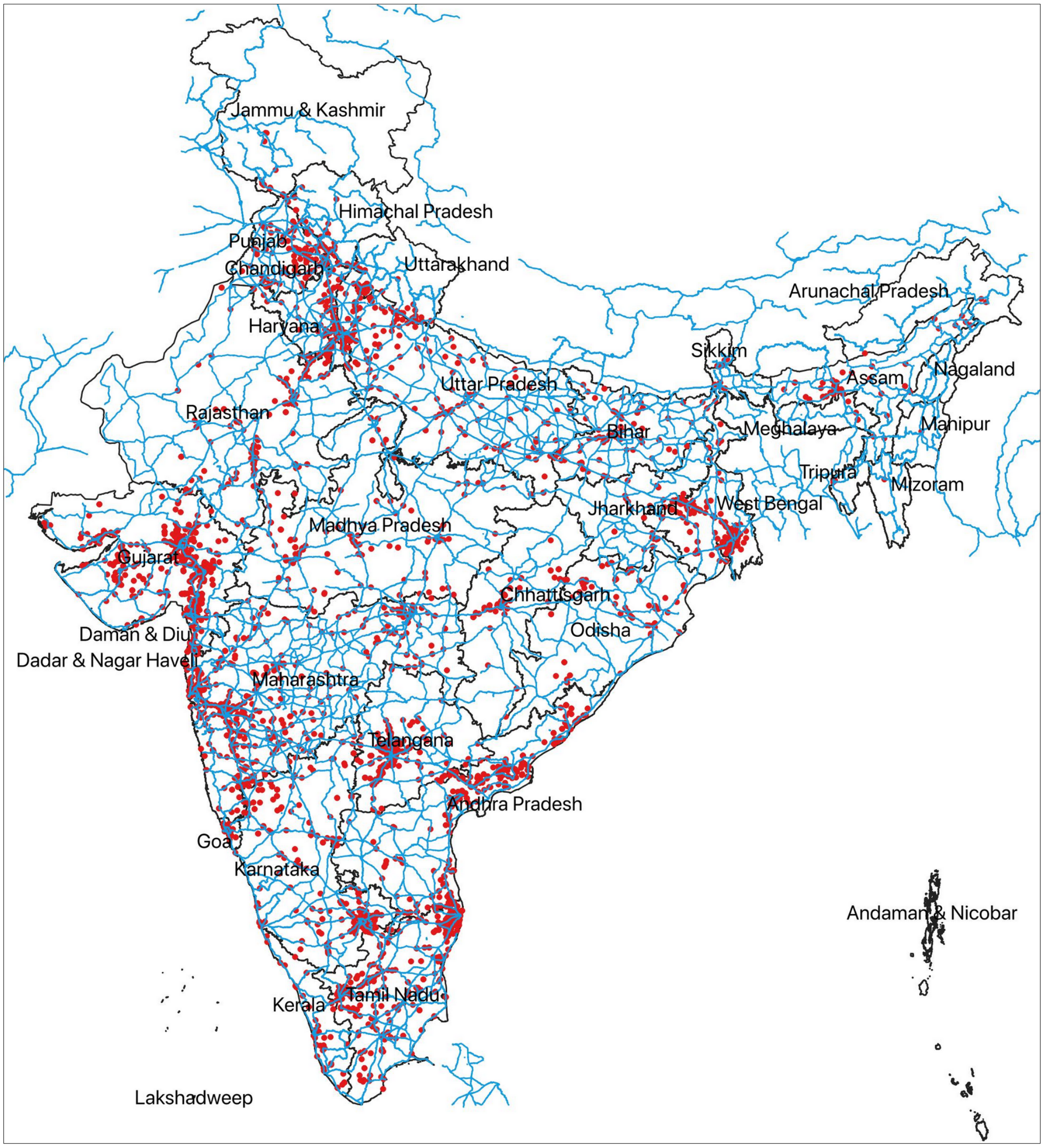
Special Economic Zones in India, Ministry of Commerce & Industry, Department of Commerce, Government of India, New Delhi.

WDR, 2009; *World Development Report:* Reshaping the Economic Geography, World Bank, Washington.

Weber, A., 1929; Theory of the location of industries, University of Chicago Press.

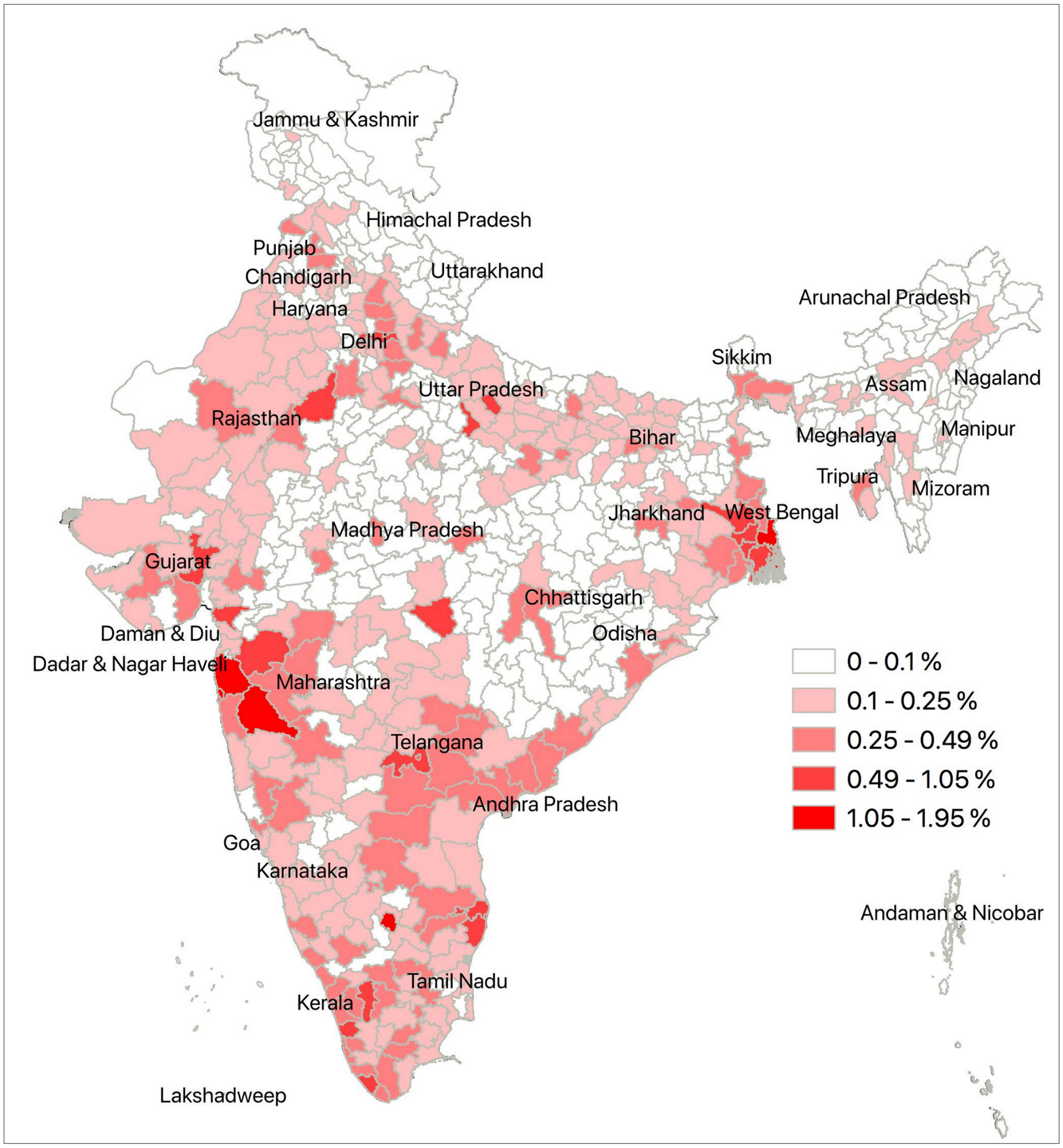
Zhang, C.S., Lin Luo, Weilin Xu, Valerie Ledwith, 2008; Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway Ireland, Sci. Total Environ.

Map A1. Manufacture plants (Projects Today) and highway infrastructure (Open Street Maps)



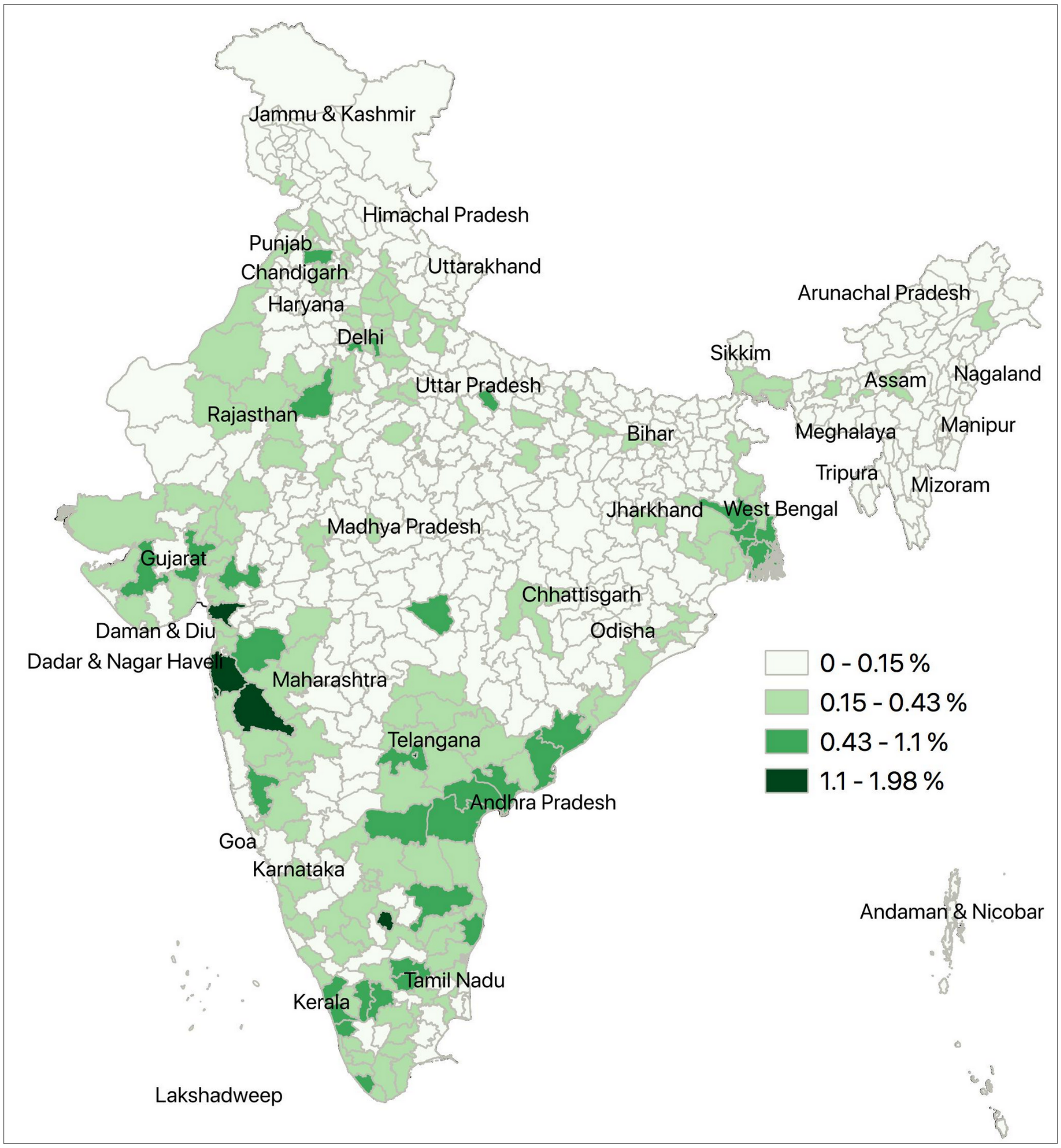
Source: Author's calculations from Project today data, and Open Street Maps.

Map A2. District percentage share of All India Other Workers in Population Census 2011



Source: Author’s calculations from Population Census 2011 data.

Map A3. District percentage share of All India Workers in all hiring enterprises Economic Census 2013



Source: Author's calculations from Economic Census 2013 data.

CLUSTER FORMATION IN THE AUTO COMPONENT INDUSTRY IN INDIA

Ashish Andhale and Sharadini Rath

The rise of the automobile industry has been one of the success stories of Indian manufacturing since the late 1990s and at the beginning of the new millennium. The GDP contribution of the sector increased from 2.77% in 1992-93 to 7.1% in 2015-16 (at current prices) and accounted for about 49% of manufacturing GDP in 2015-16. Auto component manufacturers led the bulk of the growth in this sector, making it the third largest employer in the formal manufacturing sector. However, after reaching a peak in 2017-18, the sector has seen a sustained slowdown due to weak global markets, increased oil prices, and impact of GST. From 2020, Covid-19 has compounded the stress on this sector. In this paper, we attempt to analyze the spatial patterns of plant location in regard to the auto-component industry in India. This analysis reveals that this sector exhibits significant clustering, while cutting across administrative boundaries and that these clusters geographically overlap with urbanization patterns in India. It is also seen that the greater urbanization in the southern States of India explains many aspects of this location pattern including provision of infrastructure. A larger presence of vehicle assembly plants, greater depth in ancillarisation, and far greater geographical dispersion of the industry mean that the south effectively constitutes a large agglomeration with the north of India viewed as an export market where the level of absolute demand is much higher. Further, this paper also analyses trends in the in data of automobile sector in India in regard to domestic sales, export, production and employment between 2001 to 2020. Our analysis reveals that production and sale of motor vehicles are highly correlated with employment in their manufacture. Using this, we estimate unemployment resulting from the slowdown seen after 2017-18 and also during later restrictions due to Covid-19. The figures for 2020-21 present a very gloomy employment picture for urban areas where the clusters in this sector are located.

Keywords: Automobile Industry, Industrial Clusters, Urbanization, Employment, Covid-19

Ashish Andhale is PhD Scholar at Gokhale Institute of Politics and Economics, Pune.
Sharadini Rath is Affiliate Fellow at Indian School of Political Economy, Pune, India.
Acknowledgements: Authors would like to thank the Automobile Component Manufacturers Association (ACMA), India for making available their membership directory, without which this analysis would not have been possible.
Authors are gratefully acknowledge the useful comments and suggestions given by Shri Abhay Tilak, Dr. Naresh Bodkhe, and Prof. Nilakantha Rath on the draft of the manuscript. They are thankful to Mr Ujaval Gandhi for his help in the use of geographical tools.

1. Introduction:

The evolution of the automobile industry in India happened over many time phases [Ranawat and Tiwari, 2009]. The first phase (1947 to 1965) was characterized by policies such as 'protection from foreign competition, emergence of licensing regulations and push for indigenization'. The second phase (1966-1979) witnessed 'further increase in regulations and disparate growth among different segments of the industry'. These first two phases could be termed as regulatory phases for the industry. In these phases, the implementation of indigenisation policy resulted in a vertically-integrated industry structure. A protected market combined with the lack of auto component supplier power led to the concentration of bargaining power with the automobile manufacturers. Also, much foreign exchange was spent on importing critical components. In order to resolve these concerns, the government in 1960 adopted policies emphasising encouraging the development of auto-component ancillaries. This resulted in the beginning of the development of a separate auto-component sector in India. In 1965, the government reserved 60 to 80 auto components for exclusive manufacture by small-scale units under the government's policy of supporting the small-scale sector. Also, earlier in 1956, the government imposed protection rates of tariff on small-scale ancillary units for manufacturing aftermarket vehicle components [Ranawat and Tiwari, 2009].

The third phase (1980-1990) related to relaxation in regulations and saw the entry of many Japanese collaborators in automobile manufacturing. This phase could be called 'a limited liberalization phase'. In the early 1980s, the government made policy decisions to infuse various fuel-efficient technologies and promote competition in the auto industry. This policy promoted relaxations in regard to new entrants, foreign equity collaborations and imports of technology and machinery. As a result, several joint ventures were established between Japanese and Indian companies for technology transfer and equity participation. In this period, there was an increase in vehicle models available to the Indian consumer. Also, there was an improvement in product technology and quality in the case of both, assemblers and component suppliers. The product designs changed to include components made of aluminium, plastics and fibre plastics which helped reduce the weight of vehicles resulting in increased fuel efficiency [Narayanan, 1998, Pp. 215-228]. Further, Japanese collaborators brought different and world class management practices into the industry. Further, a significant promotion of export of auto-components also took place in this phase [Ranawat and Tiwari, 2009]. The fourth phase (1991 onwards) started with the historic economic reforms of the liberalisation process in India in the early 1990s. The liberalisation of the automobile industry significantly altered its development trajectory via policy initiatives like delicensing, 51% foreign direct investments through the automatic route, relaxations for critical imports and suspension of local content requirements [Ranawat and Tiwari, 2009]. The results of these policy decisions were seen in further restructuring of the industry which resulted in the emergence of a fiercely competitive domestic market in terms of price and quality. Technology acquisition improved the performance of established firms while the number of new firms also increased [Narayanan, 1998].

In this paper, we examine some significant aspects related to the automobile industry in India classified as 'The Manufacture of Motor Vehicles, Trailers and Semi-Trailers'. in MOSPI [2008]. The industry has been showing consistent growth since 2000. It witnessed the highest percentage rise of workers (193.54%) from 2004-05 to 2016-17 as compared to other industries, and has added the greatest number of formal sector manufacturing workers, next only to the chemical and wearing apparel industry. Hence, the automobile industry has played a vital role in employment generation in India. It is observed that 76.26 per cent of workers in the automobile industry were in the auto component industry (in terms of Vendors who supply automobile parts) in 2004-05, which rose to 79.02 per cent in 2016-17, with motor vehicle manufacturers (Assemblers) contributing to the rest (around 21 per cent) of the employment in 2016-17 in the industry. Thus, workers in the component industry were 3.76 times more than in the case of assembler manufacturers.

The auto component industry of India witnessed rapid growth between 2004-05 and 2016-17 when the total output of this sector increased from Rs. 13.44 crores in 2004-05 to Rs. 28.47 crores in 2016-17, (i.e., by nearly 110 per cent), registering an average annual growth rate of 6.45 per cent. Gross value added doubled from Rs. 3.08 crores to Rs. 6.16 crores in during this period, an average annual growth rate of 5.95 per cent. Also, the number of registered factories nearly doubled, from 2,969 in 2004-05 to 5,775 in 2016-17 and the number of employees tripled during this period. This buoyant growth in the auto component sector happened due to the expansion of the domestic market and policy of promotion of exports to the external replacement market [Uchikawa and Roy, 2013].

One feature of this sector has been a high degree of geographical clustering. Clustering is the tendency of vertically and/ or horizontally integrated firms in related lines of business to concentrate geographically [OECD, 1999]. In India, in terms of location, almost all manufacturing activities have demonstrated a tendency to reflect very uneven patterns of locations over space, with a few regions being specifically preferred [Rath, 2022], the auto component industry stands out in this aspect as being particularly prone to develop in clusters. It would be useful at this stage to present a brief review of literature on the formation of clusters in in regard to the manufacturing sector. We now turn to this review.

The formation of clusters and their benefits has been discussed in the literature for a very long time. Marshall [1890] pointed out that industry districts closely linked to proximity favour intra-industry collaborations, reduce transport cost of inputs/output and allow firms to benefit from more specialized and efficient labour markets. Such externalities are known as Marshall's localization economies. Jacobs [1969] presented a view on urbanization economies also known as Jacob's externalities in which he highlighted the role of cities in the development of clusters. Furthermore, Becattini [2006, pp. 664-672] pointed out the idea of 'New Industrial Districts' where an industrial district forms a socio-economic entity with the presence of both community of people and the population of firms in one area. Sources of externalities are the governance structure and external economies in terms of information flow. Porter's definition of a cluster is perhaps most often used in cluster studies. According to Porter [Porter, 2000, Pp. 77-90], "clusters are geographic concentrations of interconnected companies, specialized suppliers, service providers, firms in related industries, and associated institutions, (e.g., universities, standards agencies, trade associations) in a particular field that compete, but also cooperate" (p. 16). A cluster, narrowly defined, is only a small part of the regional economy but it is often central to the transmission of knowledge and innovation to individual enterprises or entrepreneurs. Therefore, clusters and cluster policy are considered an efficient tool in transferring policy into economic growth and competitiveness and play an increasingly prominent role in business development and industrial policy [Brown, et. al., 2013, Pp. 3-10].

On the other hand, some insights from the literature show that the cluster has not been a well-defined concept and clusters differ in forms [Markusen, 1996, Pp. 293-313] and exist at different scale levels (a locality, a district, a state, or a nation can each be a cluster). According to Meine Van Dijk and Sverrisson [2003, Pp. 183-206] the spatial connotations of the cluster concept are rather vague in the literature. Martin and Sunley [2003, Pp. 5-35] criticise the Porter formulation of clusters. According to them, the geographical scale of 'clusters' was not defined by Porter. They pointed out that the key weakness being that there was nothing in the concept itself to indicate its spatial range or limits, or whether and in what ways different clustering processes operate at different geographical scales'. Further, Martin and Sunley [2003] also discussed the disadvantages of clusters which include labour cost inflation, inflation of land and housing costs, widening of income disparities, along with over specialization and foreign takeovers. They questioned the promotion of clusters as a policy instrument for industrial development.

The objective of this study is to examine and make an attempt to understand in some detail the spatial patterns of the auto component industry in India. In this context, the concept of clustering is used to connote the spatial and sectoral convergence of plants in the auto component industry. We identify clusters of the auto-component industry, followed by an understanding of key geographical and structural characteristics of the clusters, along with their regional differences which can ultimately provide insights into the differential nature of clustering in various regions. This study examined plant level data available in the Auto Component Manufacturers Association (ACMA) membership directory [ACMA, 2016]. The ACMA publishes an annual directory of member companies called "Buyers Guide". Information from the 2016 edition of this publication comprises the basic data used for the development of plant level database for this study. Plant level location, employment and product data are used for the purpose of the analysis. This analysis is restricted to the auto component industry of passenger vehicles and commercial vehicles segments only. In other words, the data set includes those companies and their plants which supply products to passenger and commercial vehicle manufacturers only. We use data relating to locations of both the main plant and other plants of 725 firms giving us location information for a total of 1,685 auto component manufacturing plants in India. The street addresses of all these plants were geo-coded using Google Application Programme Interface (API) to generate their longitude and latitude coordinates. Accordingly, this set of 1,685 coordinates forms the main basis of the spatial analysis of the auto component industry sector in this paper.

The paper is organized in the following manner. Section 2 presents the results of identification of clustering using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. After identifying auto component clusters, Section 3 provides the regional context for analysis of clusters. Section 4 discusses their characteristics and their relationship with urbanisation. Section 5 presents the links of the clusters to highway infrastructure and Section 6 gives an analysis of the impact on employment in this sector with the emergence of Covid-19 pandemic. Finally, Section 7 summarises the contents of the paper and provides our concluding remarks.

2. Density Based Spatial Clustering of Applications with Noise (DBSCAN):

From the literature perspective, it is seen that clusters have been examined using case studies and primary surveys. These provide detailed information on the relationships within a specific geographic area. However, the findings may not hold across the whole of a country. Other methodologies used in the literature use data on concentration of economic activity within administrative boundaries (using Gini coefficient, Location Quotient, Ellison-Glaeser index, etc.). However, it is a generally accepted feeling that clusters mostly form across multiple areas of administrative boundaries. Analysis restricted to local boundaries, therefore, may not provide proper evidence of these clusters as their effects get diluted across different areas. Variation within boundaries is also lost under this approach. Certain sectors may concentrate around particular infrastructure. Such precision is not captured in traditional analysis [DBEISUK, 2017].

In order to overcome the limitations of traditional cluster analysis methodologies DBEISUK [2017] used Density-Based Spatial Clustering of Applications with Noise DBSCAN [Ester, et. al., 1996, Pp. 226-231] as the primary technique for the analysis of identification of clusters. We follow this technique for our analysis. This technique is briefly described below.

DBSCAN identifies high concentrations of data points without relying on existing administrative boundaries. It uses the locations of individual industry units to form clusters. Some advantages of DBSCAN clustering are that it does not restrict the shape of the resulting clusters and it does not need the number of clusters to be specified in advance (as in K-means clustering). Also, a robust approach to outliers is another advantage of DBSCAN clustering. It does not force every point into clusters but allows points to be defined as 'noise', if points do not meet cluster requirements, whereas K-means forces every point into defined clusters. Thus, DBSCAN is a density-based clustering algorithm. Given a set of points, it gathers them together such that closely distributed points are packed together (points with many nearby neighbours) and assigned a label. Points marked as outliers (Noise) lie away from the clusters with nearest neighbours too far away [Q Rong, J. Yan, and G. Guo, 2004, Pp. 45-46]. Therefore, we get cluster and non-cluster points as two different categories in terms of distribution of points.

We need to set values of two parameters to determine how the geographic points' distribution can be separated into clusters. Epsilon (Eps in km) is considered the radius of each cluster which also defines its shape. Minimum Points (MinPts) is the density parameter, giving the smallest number of points in a cluster, which must be larger or equal to one [Chang, et. al., 2016, Pp. 37-40].

To begin with, all plant locations are labelled core points or noise points. The algorithm then iterates to make sure that all core points satisfy both conditions, i.e., they fall within distance Eps of more than MinPts. All core points which fall within each other's area of connectivity as defined by Eps are grouped together to form a cluster. Those points that do not form sufficiently dense clusters with other core points are labelled noise, or non-cluster, points. Some studies show how DBSCAN clustering works in their respective analysis [Ester, et. al., 1996; Wan & Wang 2020, P. 465; and Chang et. al., 2016].

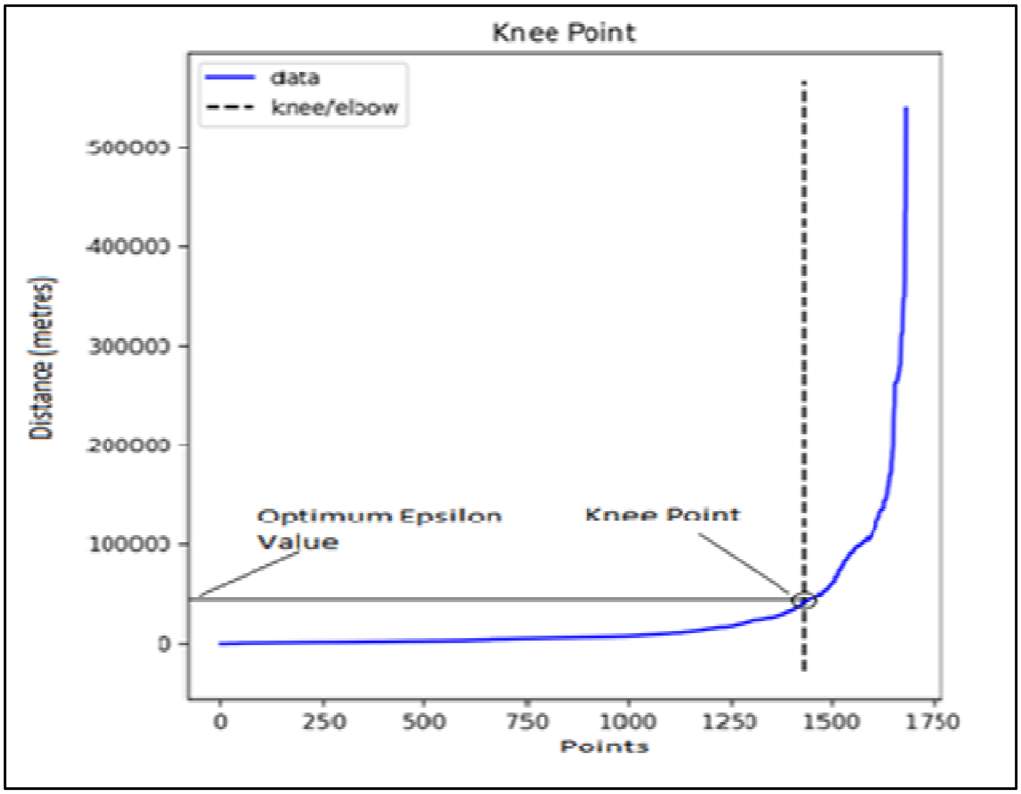
It is important to determine the optimum value of parameters Eps and MinPts to maintain the precision of the clustering analysis. Some studies use the trial-and-error method with multiple values, whereas we adopt the elbow or knee curve method to find optimum Eps (ε) value and adjust Minpts based on the optimum Eps (ε) [Rahmah and Sitanggang, 2016 p. 12012]. We find the optimal value for Eps (ε) by calculating the distance to the nearest n points for each point, sorting and plotting the results on the line chart. Then we find where the change is most pronounced on the line chart, and the knee is the optimal value of Eps (ε).

In Figure 1 given below, we see the knee point chart with plant points plotted on X axis and distance between them on Y axis for ACMA dataset of auto component plants.

Epsilon (Eps) = 41.93 km, corresponding to the knee point

We choose MinPts = 15 plants in a cluster, for this value of Eps

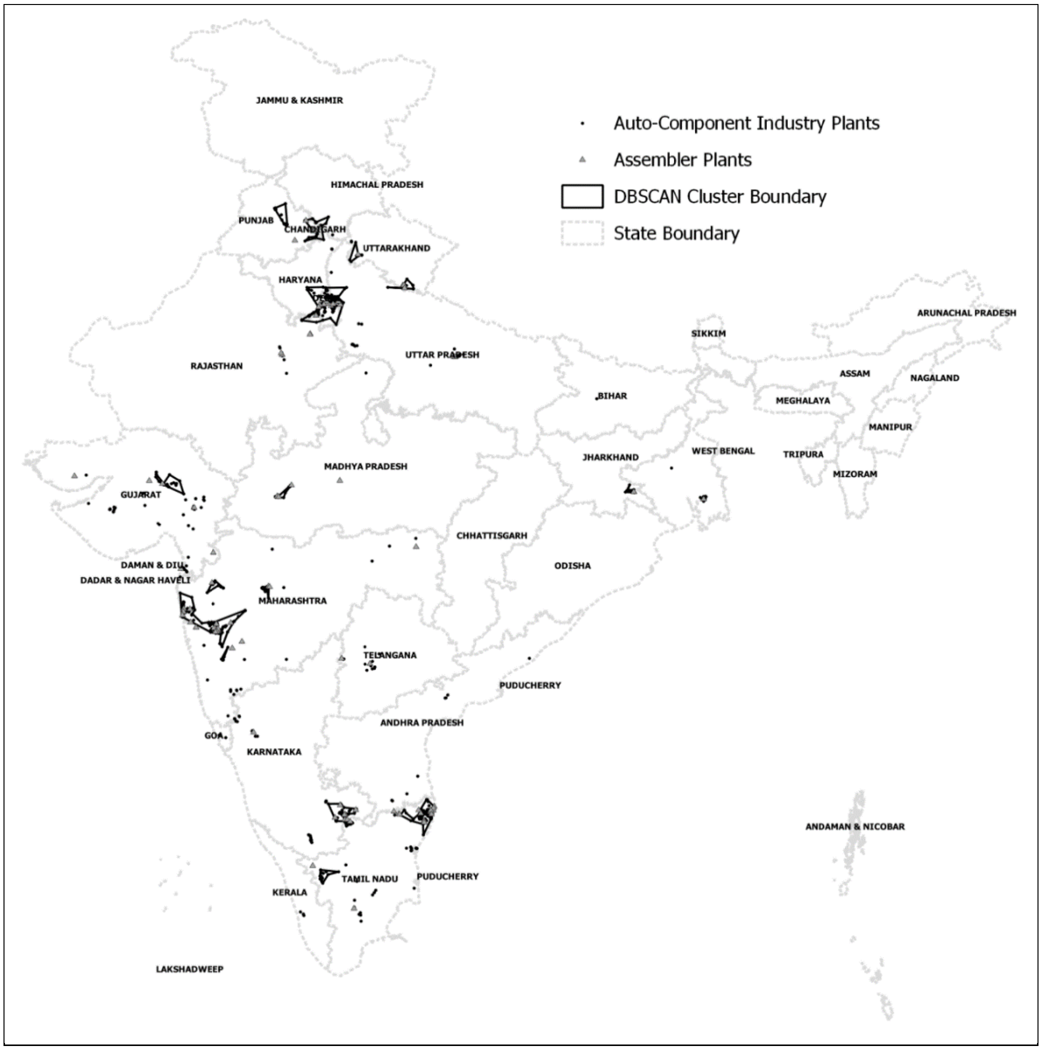
Figure 1. Knee point for ACMA data



Source: Author's Calculations.

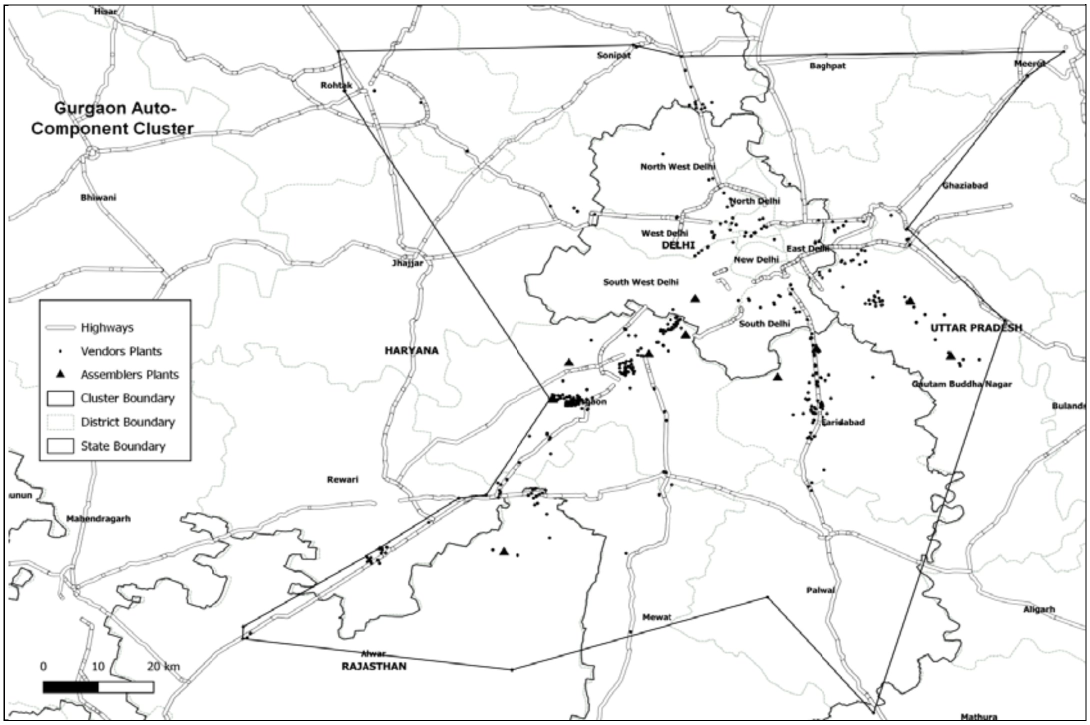
Using these values, DBSCAN generated sixteen auto-component clusters in India for the (Tier-I) plants in the ACMA directory. Nine clusters spread across the northern region and remaining seven clusters in the southern region. Map 1 give below shows all sixteen clusters in India. It is clear that there are no major clusters in the whole of eastern India. The presence of auto component industry is largely concentrated along a western alignment from north to south.

Map 1. DBSCAN Clustering: Sixteen Auto-component Clusters in India

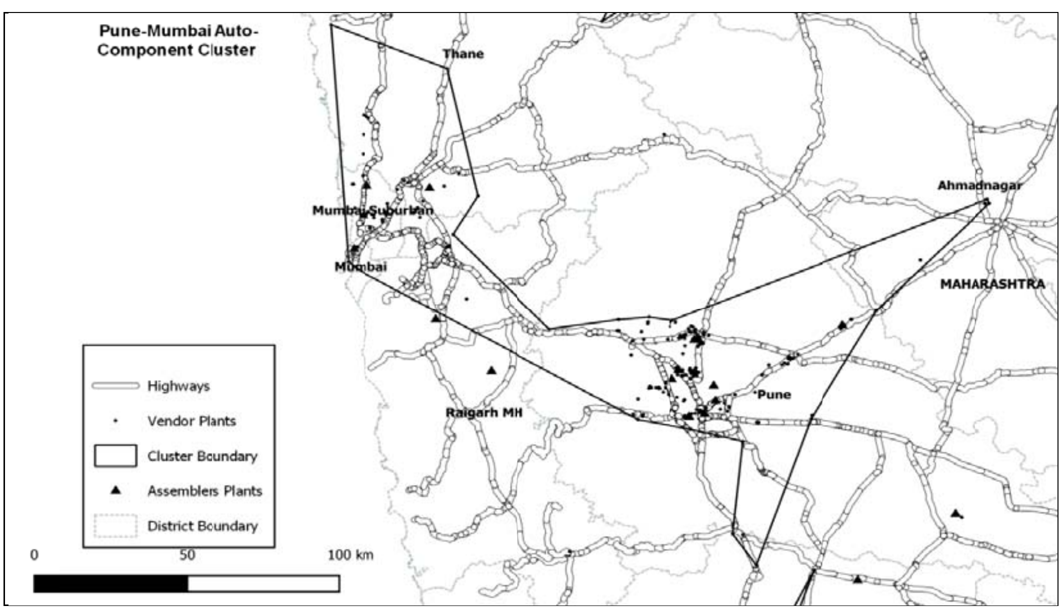


There are four large major clusters: 1) Gurgaon Cluster in North 2) Pune-Mumbai Cluster in South 3) Chennai Cluster in South and 4) Bangalore cluster in South. These four clusters hold majority share with 64.75% of total plants. These are seen in Maps 2, 3, 4 and 5 show extent of spread of these four major clusters in the vicinity.

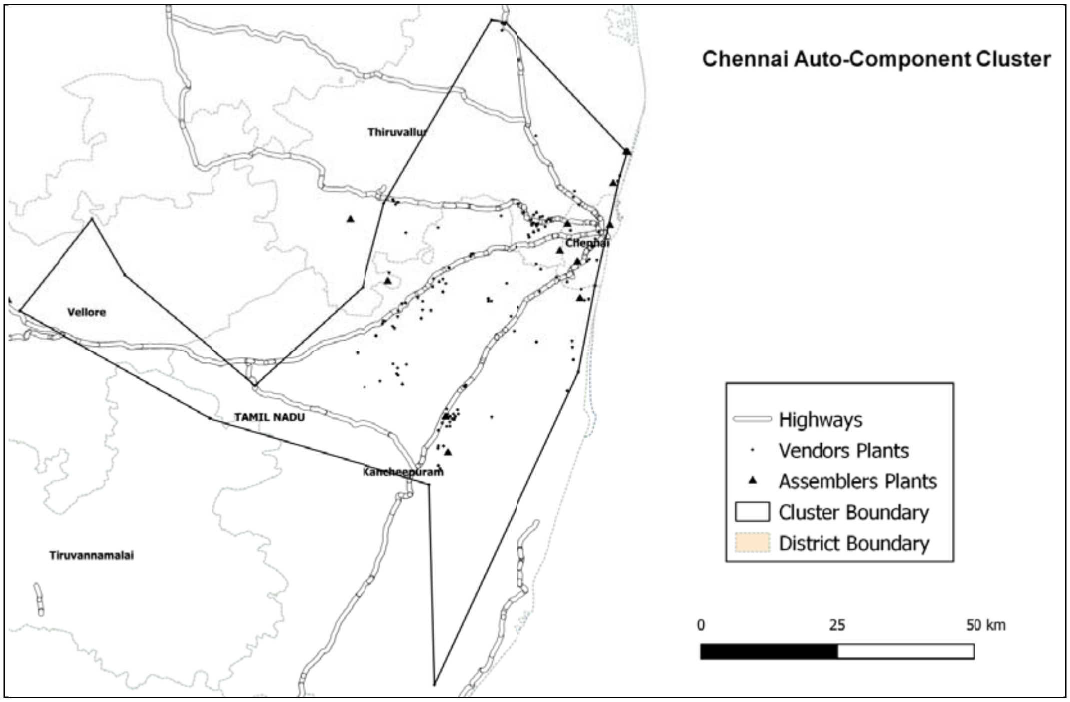
Maps 2. Gurgaon Auto-Component Cluster



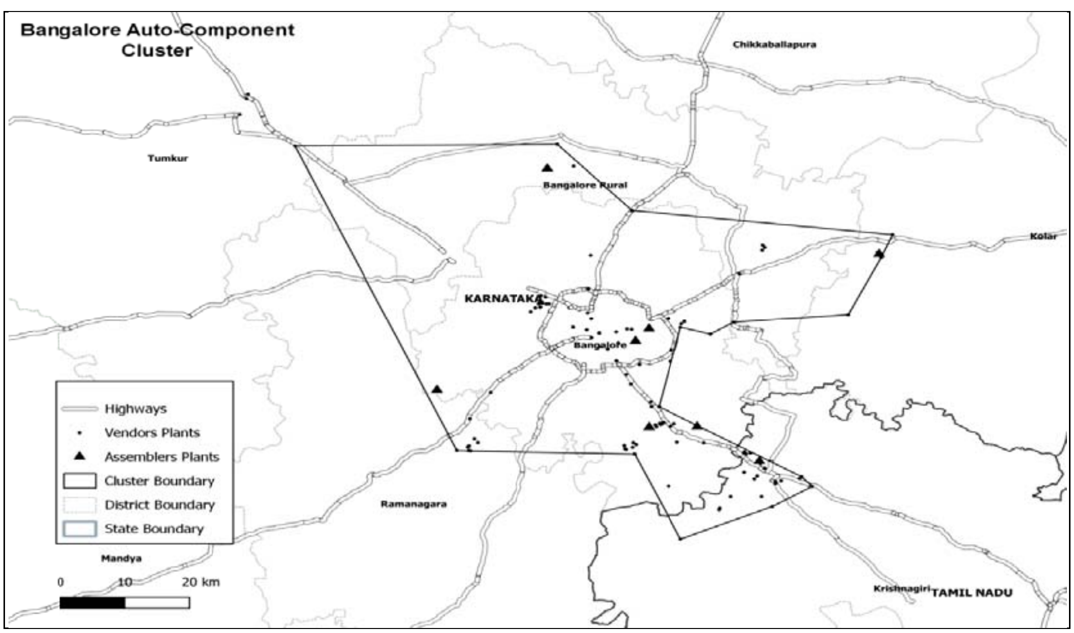
Map 3. Pune-Mumbai Auto-Component Cluster



Map 4. Chennai Auto-Component Cluster



Map 5. Bangalore Auto-Component Cluster



There are also other medium and small size clusters in India. In this, Ludhiana, Ahmedabad, Jamshedpur, Udham Singh Nagar, Haridwar, Solan, Kolkata and Dhar clusters are spread across the northern region, whereas, Aurangabad, Nashik, Satara and Coimbatore clusters are spread across the southern region of the country.

Table 1 gives the results of DBSCAN clustering analysis for auto component industry from the ACMA directory across India. There are 16 clusters, 9 across the northern region with 803 plants, and 7 across the southern region with 697 plants. While the northern region clusters are spread across 12 States and 38 districts, with a share of 47.66% of total plants, the southern region clusters are spread across 3 States and 24 districts with a share of 41.36% of total plants.

From a district perspective, we noted that there are 4 districts in India which have 10-11 plants that are non-clusters. These are Rajkot and Vadodara in Gujarat, Mysore in Karnataka, and Kolhapur in Maharashtra. This means that the choice of 15 minimum plants in a cluster was reasonable. Choosing less than 15 plants would have split the larger clusters into smaller ones, and the geographical extent and connectivity of large clusters would have been lost. The all India share of non-cluster plants is 10.92%. The north has a share of 43% of all non-cluster plants, spread across 9 States and 28 districts. These are situated in States that also have clusters, such as Gujarat, Haryana, Himachal Pradesh, Rajasthan, Uttar Pradesh and Uttarakhand. The south has a non-cluster plant share of 57%, spread across 10 States/UTs and 41 districts. In fact, all the States/UTs in the south have some presence of the auto component industry spread widely across many districts even though it has clusters in only 3 states. This shows that there is far greater dispersal of the auto component industry across the south than in the north, as would be expected with the larger degree of urbanisation in the south.

While the north has more auto component plants than the south, it has fewer assemblers than the south. It is worth noting that 41.5% of assembler plants are not within the boundaries of any auto component cluster. The history of auto component clusters shows that their growth has been supported by large automakers like Maruti Udyog in Gurgaon and Tata Motors in Pune [Okada and Siddharthan, 2008]. They have historically formed around the assembly plants of these companies. As demand for cars started growing around 2005, new auto makers entered the Indian market. It is possible that these new assembler plants found it easiest to locate on the

Table 1. State wise Composition of Total Plants, Plants in Clusters and Non-cluster Plants

Cluster No	Clusters Name	No. States	States	No. Districts	Region	Total Vendor Plants	% Vendor Plants (India Total)	% Vendor Plants (N/S Total)	No. Assembler Plants	% Assembler Plants (N/S Total)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1	Gurgaon Cluster	4	HR, DL, RA, UP	13	North	526	31.22	65.50	10	45.45
2	Udham Singh Nagar Cluster	2	UTT, UP	4	North	68	4.04	8.47	2	9.09
3	Ludhiana Cluster	1	PJ	5	North	44	2.61	5.48		
4	Jamshedpur Cluster	1	JH	3	North	37	2.20	4.61	2	9.09
5	Haridwar Cluster	1	UTT	3	North	32	1.90	3.99	1	4.55
6	Ahmedabad Cluster	1	GU	3	North	29	1.72	3.61	1	4.55
7	Solan Cluster	3	HP, PJ, CHD	5	North	27	1.60	3.36	3	13.64
8	Dhar Cluster	1	MP	3	North	22	1.31	2.74	2	9.09
9	Kolkata Cluster	1	WB	3	North	18	1.07	2.24	1	4.55
10	Pune- Mumbai Cluster	1	MH	7	South	256	15.19	36.73	14	42.42
11	Chennai Cluster Bangalore	1	TN	7	South	190	11.28	27.26	10	30.30
12	Cluster Aurangabad Cluster	2	KA, TN	5	South	120	7.12	17.22	9	27.27
13	Nashik Cluster Coimbatore	1	MH	1	South	55	3.26	7.89		
14	Cluster	1	MH	1	South	39	2.31	5.60		
15	Satara Cluster	1	TN	2	South	22	1.31	3.16		
16		1	MH	1	South	16	0.95	2.30		
Cluster North Total	9	12		38		803	47.66		22	23.40
Cluster South Total	7	3		24		697	41.36		33	35.11
Cluster Total Non-	16	15		62		1501	89.08		55	58.51
Cluster North Non-		9		28		79	4.69	42.93	18	19.15
Cluster South Non-		10		41		105	6.23	57.07	21	22.34
Cluster Total		19		69		184	10.92	100.00	39	41.49
Total Plants						1685	100.00		94	100.00

Source: Author's calculations based on ACMA data.

peripheries of existing large auto component clusters. So, these clusters appear to be acting as centrifugal forces for the location of auto assembler plants.

The northern region clusters also have a very small range. The adjoining States of Haryana, Delhi, Punjab, Uttarakhand, western Uttar Pradesh, north Rajasthan and Chandigarh make up the States that have 5 of the 9 north region clusters, with a 46% share of all cluster plants. Uttarakhand has a share in two of the north region clusters. This is a clear case of the impact of policy on the location choices of industry. The Hilly Region Policy of the Government of India [GOI, 2003], put in place in 2003, giving special incentives for companies located to the hill states has resulted in these clusters forming here. Some companies moved their plants from other parts of India to this area to take advantage of these incentives.

In the south, Tamil Nadu has a part in 3 clusters, Chennai, Coimbatore and Bangalore. All have long histories of heavy industrialisation. Chennai and Bangalore are metro cities. In fact, the Chennai cluster has spread north to the border of Andhra Pradesh and west towards Bangalore across the Karnataka border. We feel that that the industrial corridor between Chennai and Bangalore will soon become a geographic continuum thereby creating a single large cluster encompassing these two major metro areas.

The Dhar (Madhya Pradesh), Jamshedpur (Jharkhand), Kolkata (West Bengal) and Ahmedabad (Gujarat) clusters are very small compared to Gurgaon. Two of the smallest clusters are - Kolkata in the east and Satara in the west. It is a commentary on the demise of the auto industry in West Bengal, once the home of Hindustan Motors and the Ambassador car, but is now the smallest auto cluster in India. Satara can be considered to be in the backyard of the Pune cluster, with the large city of Satara to support it. Jamshedpur is the home of Tata Motors for the past nearly seven decades, but it can be seen that it supports only a small cluster in its vicinity. With Kolkata declining in its importance in the auto sector, the impact has been felt by this location in its backyard.

3. The Regional Context of the Auto Component Industry in India

From observations of the pattern of the auto-component industry that was given in Section 2, we can see some interesting features emerging in a regional context. In order to analyse these, we set up the regional context in this Section.

We have divided the country into two regions for this purpose: North and South.

The South region consists of the States of Maharashtra, Telangana, Andhra Pradesh, Tamil Nadu, Karnataka, Kerala and Goa, and the Union Territories of Daman and Diu, and Dadra and Nagar Haveli.

The North region consists of all other States and UTs of India.

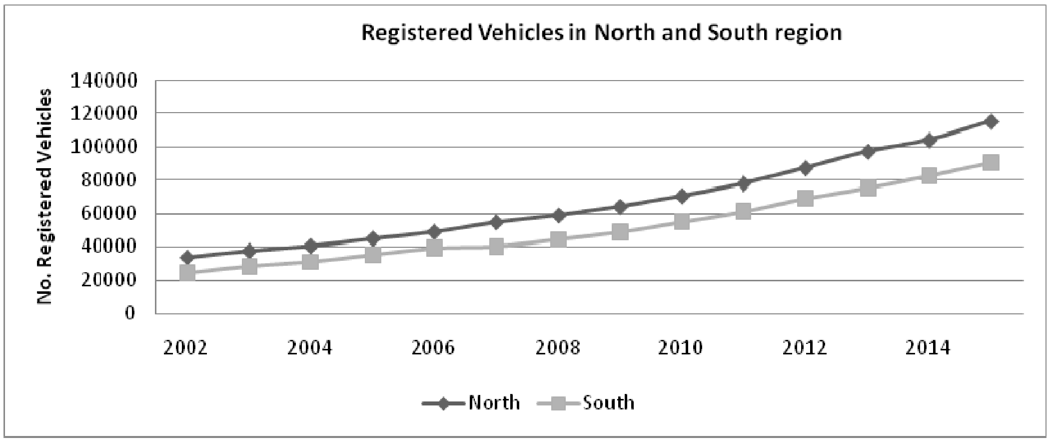
There is a significant demographic difference between the northern and southern regions of India as defined above. The share of urban population in the north is about 34%, while that in the south is about 54% as per Population Census 2011 [Population Census, 2011, RBI].

A motor vehicle sector specific analysis can be done to highlight the context of these regional differences. We use Ministry of Road Transport and Highways data [GOI, 2004-05; 2014-15] to get a time series of total vehicle registrations in all states for each year (cumulative), and aggregate it into North and South regions. We then use population data from the Population Census 2011 [GOI, 2001 and 2011] to calculate past demand for motor vehicles in these regions for each year in the period 2002 - 2015.

Figure 2 shows the absolute demand in terms of total motor vehicle registration in the northern and southern regions from 2002 to 2015. Absolute demand for vehicles in the north region is consistently higher than in the south region. North's demand was 1.27 times more than the South's in 2015. In addition to the higher demand, the population of the north region is 2.16 times greater, and area covered (Sq km) is 1.62 times more than south region. This shows that in absolute terms north is a bigger market for motor vehicles than the south. So, for motor vehicle industry in the south, north presents an effective large export market.

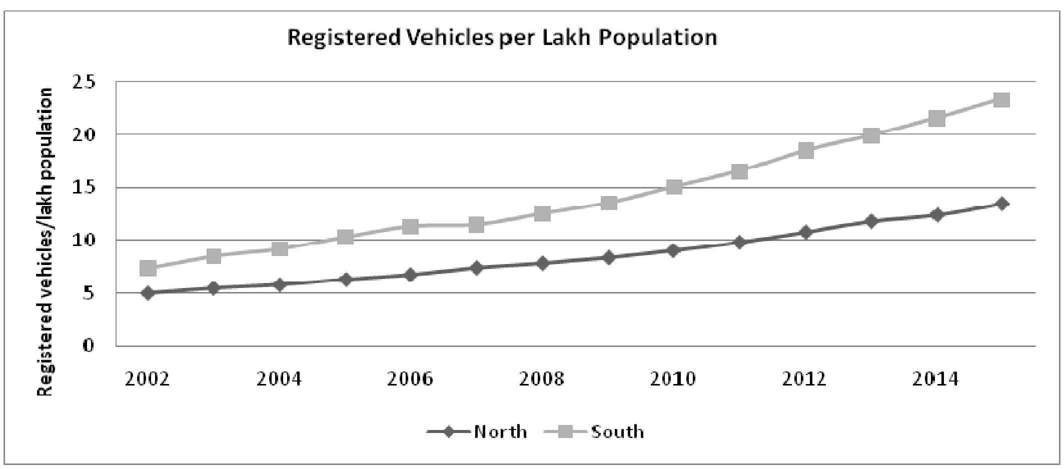
Figure 3 shows vehicles registered per lakh population in north and south for the period 2002 to 2015. This shows relative demand and purchasing capacity for vehicles in the two regions. The south has always had higher purchasing capacity for vehicles than the north, and the difference has been growing over the years. By 2015, 23.36 vehicles per lakh population were registered in the south, whereas 13.49 vehicles per lakh population were purchased in the north. Considering registrations to be coterminous with purchases, then, the south can be seen as an effective agglomeration with greater urbanisation and higher purchasing capacity.

Figure 2. Total Motor Vehicle Registration in north and south region



Source: Derived from MORTH (2004-05 to 2014-15)

Figure 3. Motor Vehicle Registration per lakh population in north and south region



Source: Derived from MORTH, [2004-05 to 2014-15], Population Census, [2001 & 2011]

Some fundamental trends of the northern and southern regions can be defined here are shown by looking at their respective shares in terms of 3 parameters: Gross State Domestic Product (GSDP), Per capita income State wise (GSDP /Total Population). Table (2) shows that over the period 2004-05 to 2015-16, share of south in gross total GSDP has increased by 2.63 per cent, in total per capita GSDP it has increased by 0.96 per cent, while its share of total population has decreased by 1.25 per cent. In the south there is a declining in the share of the population, but an increase in its share of the Indian economy. This picture is consistent with the vehicle registration data for north and south, with the south showing greater purchasing capacity normalised for population. This is shown in Figures 2 and 3.

Table 2. Share of North and South Regions in Gross Total GSDP, Total Per Capita GSDP and Total Population

	2004-05		2015-16	
	North	South	North	South
(1)	(2)	(3)	(4)	(5)
GSDP/Capita	65.34	34.66	64.38	35.62
Gross GSDP	58.49	41.51	55.86	44.14
Population (%)	69.08	30.92	70.31	29.69

Source: Reserve Bank of India.

Further, another aspect of this north-south comparison can be brought out by bringing in factory data from the Annual Surveys of Industry [EPWRF, 2015-16]. There is a hierarchy of plants in the automobile vendor sector. Factories that supply parts directly to automobile assemblers are classified as Tier-I. Tier-II and Tier III are plants that supply intermediate inputs to Tier-I and Tier-II plants respectively. ACMA [2016] directory states that all listed firms in their directory supply their output to at least one assembler. Therefore, we classified and assumed these firms and their plants as Tier-I suppliers. We use data for number of motor vehicle vendor factories (National Industrial Classification (NIC), 3-digit 291, 293) [GOI, 2016]. The ASI data frame includes only those registered establishments, which employ 10 or more workers with electricity or 20 or more without electricity, in accordance with the Section 2m (i) & 2m (ii) of Factories Act, 1948 [GOI, 1948]. Therefore, we assume that ASI data of auto-component industry counts all Tiers of the auto component industry, I, II and III, which are registered. With this, we can use the following formula to get an idea of the plant multiplier effect. We define this multiplier as the number of Tier-II and III plants for each Tier-I plant.

Plant Multiplier = (ASI Factories - ACMA Plants) / ACMA Plants (for year 2015-16).

Based on this definition of the multiplier, we find that:

Plant Multiplier (at the all-India level) = 2.38.

Plant multiplier (for south region) = 3.20

Plant multiplier (for north region) = 1.63

This shows that the southern region has generated almost twice the number of Tier-II and III plants for every Tier-I plant as compared to the northern region. Since these Tiers are defined by employment size of factories, this reflects more depth in ancillarisation in terms of factory size in the south than in the north.

Another factor that has significance for location of industry is the degree of urbanisation in the region. Map 6 shows the district level proportion of all-India total urban population. The overlap between all the auto component clusters and the map of urban India is almost complete.

Map 6. District % share of all India Urban Population

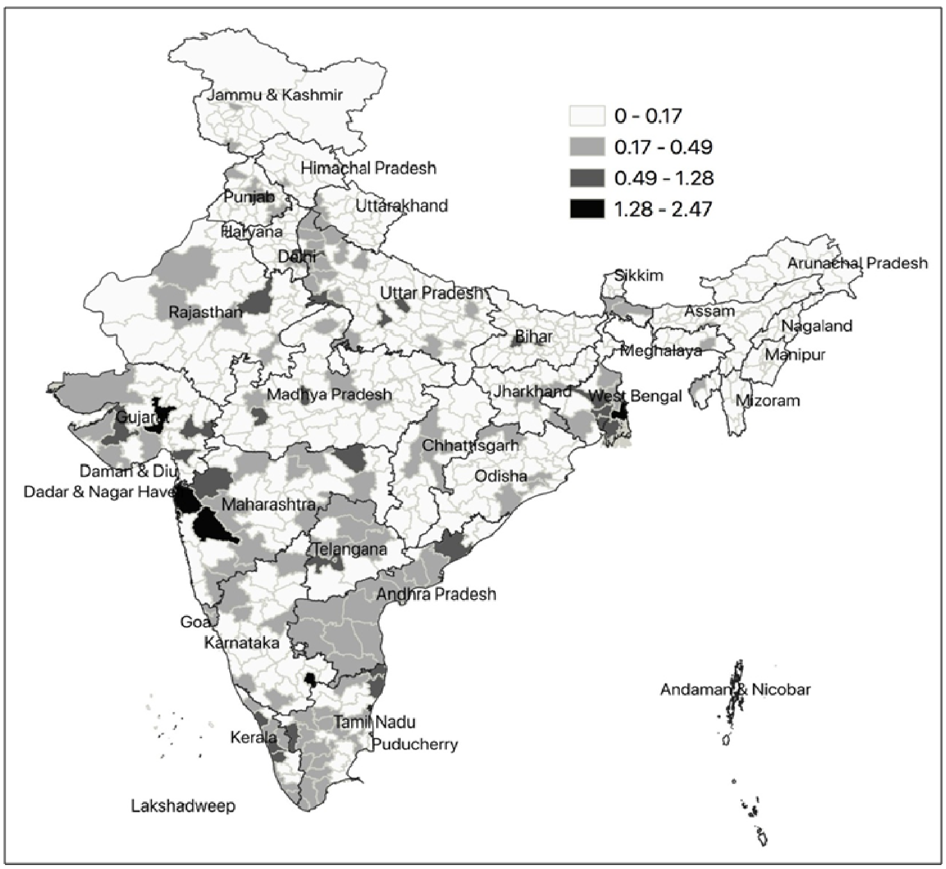


Table 3 links the degree of urbanisation to the densest parts of the clusters. The main districts, containing the largest proportion of plants in the cluster, are identified and their proportion of urban population is calculated from the Population Census 2011. It becomes clear that the densest parts of the four major clusters are located in districts that also have high proportions of urban populations, upward of 60%, with most of them containing a metro city, except for Gurgaon. Some of the smaller clusters are also located in districts with large cities. There are 4 small clusters located in districts with small urban populations; Solan and Dhar in the north, and Aurangabad and Satara in the south in Maharashtra. All these clusters are located inside state industrial estates.

Table 3. Degree of Urbanisation in Vicinity of Clusters

Cluster No.	Clusters Name	Main District in cluster	% Cluster plants in main district	% Urban population of the district	% Vendor Plants (India Total)
(1)	(2)	(3)	(4)	(5)	(6)
1	Gurgaon Cluster	Gurgaon	41.63	68.82	31.22
2	Udham Singh Nagar Cluster	Udham Singh Nagar	76.47	35.58	4.04
3	Ludhiana Cluster	Ludhiana	77.27	59.14	2.61
4	Jamshedpur Cluster	Purbi Singhbhum	86.49	55.55	2.20
5	Haridwar Cluster	Haridwar	90.63	37.77	1.90
6	Ahmedabad Cluster	Ahmedabad	86.21	84.05	1.72
7	Solan Cluster	Solan	55.56	17.70	1.60
8	Dhar Cluster	Dhar	68.18	18.91	1.31
9	Kolkata Cluster	Kolkata	72.22	100.00	1.07
10	Pune-Mumbai Cluster	Pune	79.69	60.891	5.19
11	Chennai Cluster	Chennai	48.42	100.00	11.28
12	Bangalore Cluster	Bangalore	71.67	90.94	7.12
13	Aurangabad Cluster	Aurangabad	100.00	6.37	3.26
14	Nashik Cluster	Nashik	100.00	42.53	2.31
15	Coimbatore Cluster	Coimbatore	95.45	75.83	1.31
16	Satara Cluster	Satara	100.00	18.98	0.95

Source: Author's calculations and Population Census [GOI, 2011]

Solan has been the beneficiary of the Hilly Regions policy. Dhar, Aurangabad and Satara all have some histories of industrial development, part of it driven by local factors, and partly by state encouragement of cluster formation in growth sectors, away from the major industrial centres. However, they still fall within the western regions of Madhya Pradesh and Maharashtra. These locations have taken advantage of the broader western alignment of industrial development.

Of the districts with about 10-11 non-cluster plants, Vadodara, Rajkot, Kolhapur and Mysore, we know that the plants in Kolhapur are located inside a state industrial estate. It is likely that this is the case with most of the others. All are located near major cities. There are 12 non-cluster plants in the vicinity of Hyderabad city in Telangana state.

Our analysis suggests that it is quite appropriate to view the southern region as a single agglomeration with the north region as a large export market for four-wheeler passenger vehicles. Widespread urbanisation economies in the south have created room for the auto component industry to thrive, not just in large clusters, but also in smaller patches. These, in turn, have acted as centrifugal forces for the location of auto assemblers on their peripheries possibly resulting in the reach of this cluster formation in the future.

The next two Sections discuss two other aspects relating to formation of auto-component clusters. We first study the role of physical infrastructure (highways) in cluster formation, and its differential provision for areas in the country that link high performing areas in Section 4. In Section 5 we look at the impact of Covid disruptions on this high performing sector and, in turn, their effects on employment.

4. Role of Highway Infrastructure in Location of the Auto-Component Industry

The provision of transport infrastructure has been a subject of main concern among economic geographers and a pertinent factor in location of new manufacturing plants as well as encouraging the formation of industry clusters [Rietveld and Bruinsma, 1998; McCann, 2013] In an international context, Woodward [1992, Pp. 690-708] highlights the importance of role of highway infrastructure in attracting plant location outside metro areas in US with respect to automotive corridors. Furthermore, Klier and McMillen [2008, Pp. 245-267] estimated the location patterns of motor auto-component plants in U.S. They found that the observed location choices are well explained by a small number of factors and two of the factors point toward transportation costs of goods (good highway access and shorter distance to assemblers' plants). One of the main forces of agglomeration in the European auto supplier industry is highway access, [Klier and McMillen, 2015, Pp. 558-573]. In the Indian context, rapid growth of the automotive industry may involve a big increase in the movement of physical goods, (e.g., raw materials, components, assemblies, and finished vehicles) both within the country and across the sea ports. For this reason, various National and State policies related to the automobile sector [SIAM, 2016; MIDC, 2019; GOTN, 2014; GOH, 2020] have focused on development of transport infrastructure in States and clusters. Ghani, Goswami, and Kerr [2015, Pp. 317-357] highlight this effect in terms of clustering of manufacturing operations and the resulting increased productivity, in line with Economic Geography insights. The authors discover that increased entrance rates for manufacturing companies close to better highways were brought on by a significant intercity road investment in India (the "Golden Quadrilateral"), and that the labour productivity and total factor productivity of these companies were higher. According to Datta [2012, Pp. 46-57], businesses near highway improvements are actually run more effectively (by storing inventories for less time).

In this Section, we show this particular characteristic of auto-component industry clusters which is the importance of highway access (state/national highways) for location of auto-component plants in the clusters. For this, we calculate the shortest Euclidean/linear distance between the location coordinates of plants to nearest coordinates on state/ national highways. These distances (km) from highways are calculated for 1501 auto-component plants located in sixteen identified clusters.

We find that the average distance of the plants to nearest highway for all the clusters is only 2.17 km, with average minimum and maximum distance being 0.03 km and 9.48 km respectively (Table 4).

Also, for all the sixteen clusters, average highway access from plants' location varies between 1 to 6.16 km. For instance, in the Gurgaon Cluster majority of plants are located on and around National Highway 8 through Gurgaon, Rewari and Alwar, and NH 48 through Faridabad. Also, For the Pune-Mumbai cluster, many auto-component plants are established on the Pune-Ahmednagar state highway and Pune-Bangalore national highway which passes through Satara and Kolhapur and goes further south. Furthermore, all clusters have significant highway coverage within them. Average length of highways in all clusters area is 435 kms. It should be noted from Table 4 that southern region clusters have more highway coverage (1.4 times) than the northern region clusters. Overall, the analysis shows that Indian auto component plants are located in proximity to highways. Thus, access and coverage of highways are crucial factors in their locations.

Table 4. Highway Access to Clusters

Cluster No.	Clusters Name	Average Distance of cluster plants to nearest highway (km)	Maximum Distance of cluster plants to nearest highway (km)	Minimum Distance of cluster plants to nearest highway (km)	Total length of highways within cluster (km)
(1)	(2)	(3)	(4)	(5)	(6)
1	Gurgaon Cluster	2.44	17.55	0.00	1583.49
2	Udham Singh Nagar Cluster	1.96	12.97	0.00	79.59
3	Ludhiana Cluster	1.14	2.49	0.00	279.39
4	Jamshedpur Cluster	1.73	8.87	0.00	70.08
5	Haridwar Cluster	2.90	6.18	0.02	92.82
6	Ahmedabad Cluster	6.16	29.60	0.00	457.35
7	Solan Cluster	1.18	2.89	0.00	560.19
8	Dhar Cluster	3.09	6.32	0.00	157.29
9	Kolkata Cluster	2.18	3.54	0.17	42.75
10	Pune-Mumbai Cluster	1.48	8.48	0.00	1749.79
11	Chennai Cluster	2.87	21.13	0.00	582.30
12	Bangalore Cluster	2.70	17.95	0.00	663.48
13	Aurangabad Cluster	1.66	4.51	0.00	114.88
14	Nashik Cluster	1.10	2.50	0.04	186.53
15	Coimbatore Cluster	1.03	4.31	0.00	259.36
16	Satara Cluster	1.07	2.42	0.28	74.09
	North Clusters	**2.53**	**10.05**	**0.02**	**369.22**
	South Clusters	**1.70**	**8.76**	**0.05**	**518.63**
	Total Clusters	**2.17**	**9.48**	**0.03**	**434.59**

Source: Authors' Calculations from ACMA data.

This highway-based analysis of plant location shows that from Gurgaon in the north to Coimbatore in the south, all clusters have formed around major transport infrastructure. In fact, it can be claimed following Ghani, Goswami, and Kerr [2015] and Datta [2012] that transport infrastructure has been preferentially provided to link these high performing industrialised western and southern regions, so that movement of goods and labour can be facilitated. Therefore, these highways were constructed to facilitate the agglomerations. With more assemblers in the south than the north, this road mobility becomes important to achieve high productivity on one hand, and a smooth supply of components and finished vehicles, especially to the north region on the other. It is difficult to say which process has preceded this development; urbanisation or industrialisation. It is clear, however, that they have reinforced each other, and the States and the Centre governments have aided this process with provision of critical transport infrastructure.

5. Implications for Urban Employment in Auto Component Sector and COVID:

In this section, we make a brief comment about the employment generation capacity of the auto sector especially in the context of the pandemic years. The COVID 19 crisis and associated slowdown in the demand for vehicles had deep implications for urban employment.

The auto sector shows a peculiar characteristic in its employment pattern. We calculate total employment per 1000 vehicles produced in India, using employment data from MOSP [GOI, several years], and annual vehicle production data from Society of Indian Automobile Manufacturers (SIAM). We find that over the period 2003 to 2019, this number has been nearly constant (Figure 4). In India, it takes about 210 workers to produce 1000 vehicles, including both vendor and assembler employment. Of these 160 are auto component workers and 50 workers in assemblers. Fluctuations around these numbers over the years from maximum 8 per cent to minimum - 11 per cent. It turns out that this is not a peculiarity of the Indian auto industry. This is true of the German auto sector as well, giving a more or less constant 140 workers to produce 1000 vehicles over the same period (Kraftfahrt-Bundesamt (KBA)) This shows that as demand for vehicles increases, employment in the auto industry increases at a nearly fixed proportion. Figure 4 shows that since 2003-04, the number of workers per vehicle has remained approximately constant in both total auto workers and auto component workers. From Figure 5 it is also clear that as production and sales have increased, employment has steadily risen. However, sales declined for the first time in nearly 20 years in 2019-20, and the COVID pandemic hit the sector in this low period.

We have used the average employment per vehicle factor to estimate what the loss of employment was likely to be during 2019 - 2022.

Total auto sector workers (Est) = 0.21 * Total Vehicle Production
Total auto component workers (Est) = 0.16 * Total Vehicle Production

The dotted lines in Figure 5 for the years 2018-19 to 2020-21 show the estimated loss of employment in automobile sector in times of Covid 19 due to decrease in production. The estimated loss in auto sector employment was 67.58% in 2020-21 compared to 2017-18, to levels last seen in 2010-11.

Figure 4. Employment per Vehicle in Automobile Industry

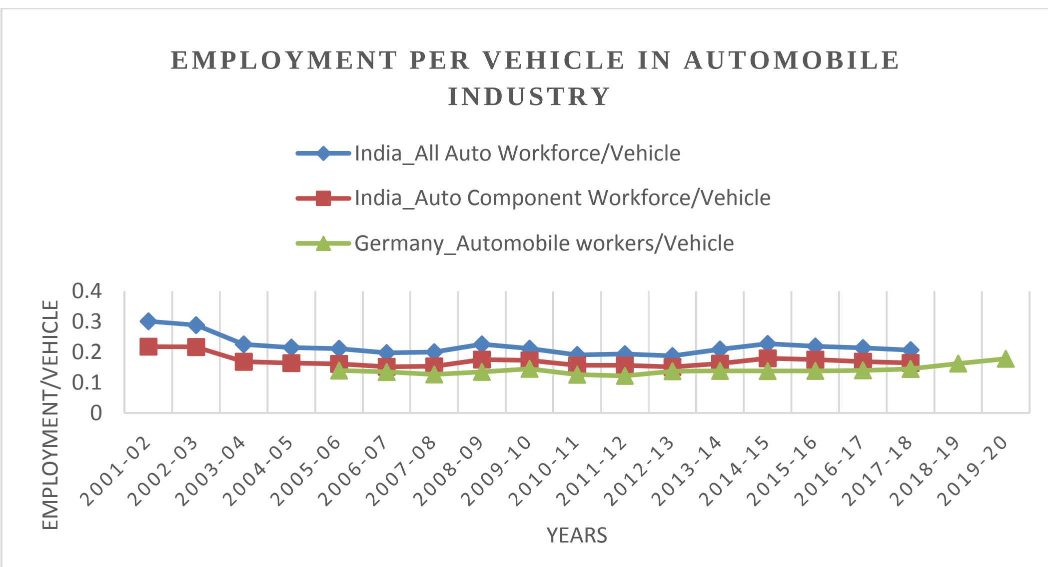
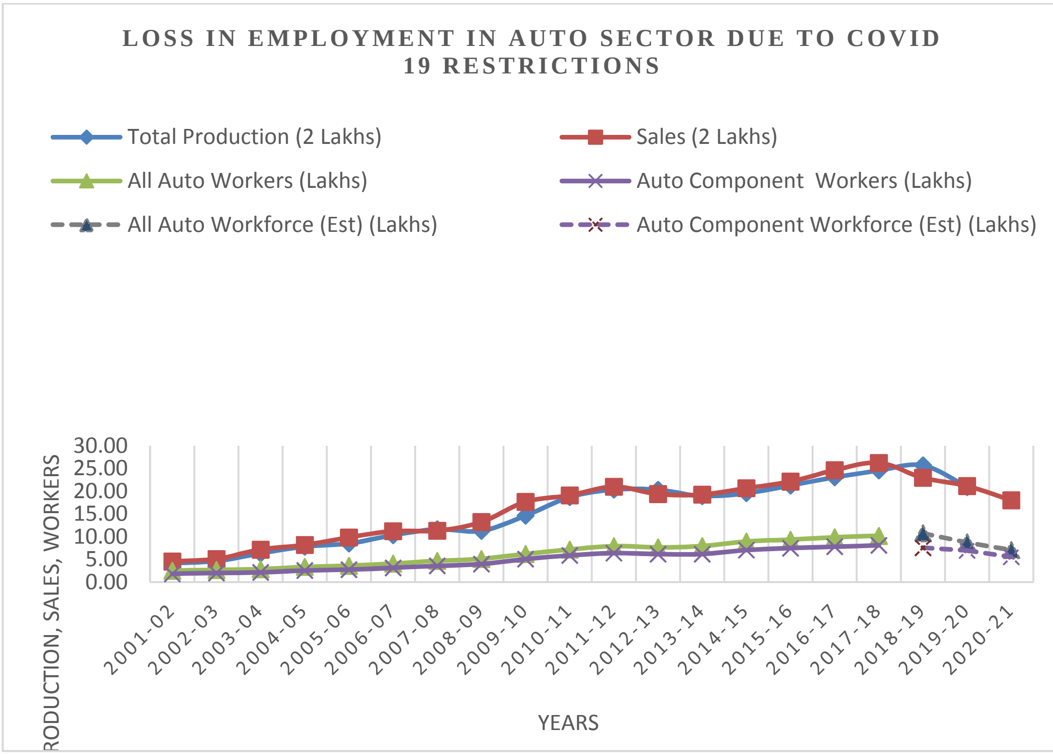


Figure 5. Loss in Employment in Auto Sector due to COVID 19 restrictions



Production and sales in Figure 5 are given in 2 Lakhs, while actual and estimated workers are given in Lakhs.

5. Summary and Concluding Remarks

Industrial clusters seem to require urban economies to sustain them. The auto component sector in India is a particularly good example of this mutually beneficial relationship. The development of the northern region is dominated by a single large cluster near Gurgaon whereas the south region has three big clusters namely, Pune-Mumbai, Chennai and Bangalore. Geographically, the southern region tends to be larger in terms of cluster development and more dispersed with larger peripheries than the northern region. The greater urbanization in the south of India explains many aspects of this location pattern, including provision of infrastructure. A larger presence of vehicle assembly plants, greater depth in ancillarisation, and far greater geographical dispersion of the industry, means that the south can be effectively considered a large agglomeration that sees the north of India as an export market, where absolute demand is much higher. This relationship is sustained by infrastructure provision that reinforces these tendencies and aids further cluster formation in the preferred regions of the country. The production and sale of motor vehicles are highly correlated with employment in their manufacture.

REFERENCES

ACMA, 2016; Buyers Guide, Auto Component Manufacturers Association of India, New Delhi.

Becattini, G., 2006; The industrial district and development economics in Raffaellit. T, Becattini, G. and Dardi, M. (eds), The Elgar Companion to Alfred Marshall, U.K., Edward Elgar.

Brown, K. and J. Burgess, M. Festing, and S. Royer, 2013; Resources and Competitive Advantage in Clusters: An Introduction, in K. Brown, J. Burgess, M. Festing & S. Royer (eds), Resources and Competitive Advantage in Clusters, München: Rainer Hampp Verlag.

Chang, L., et. al., 2016; Data Mining Applied to Oil Well Using K-means and DBSCAN, 7th International Conference on Cloud Computing and Big Data, Macau, China, November.

Datta, S., 2012; The impact of improved highways on Indian firms, *Journal of Development Economics,* Vol. 99, No. 1, doi:10.1016/j.jdeveco.2011.08.005

DBEISUK, 2017; Density Based Spatial Clustering-Identifying the Industrial Clusters in the UK, Methodology Report, Department of Business, Energy & Industrial Strategy, London, United Kingdom.

Diebolt, C. Mishra, T. and Parhi, M. (2016), A Comparative Regional Analysis of Indian Auto Components Industry, book- Dynamics of Distribution and Diffusion of New Technology: A Contribution to the Historical, Economic and Social Route of a Developing Economy, Springer Publications, Pages 85-102.

EPWRF, 2016; Database on the Annual Survey of Industries (ASI), 2015-16, Economic and Political Weekly Research Foundation, Mumbai.

Ester, M. et. al., 1996; A density-based algorithm for discovering clusters in large spatial databases with noise, KDD: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining.

Ghani, E., A. G. Goswami, & W. R. Kerr, 2015; Highway to success: The impact of the golden quadrilateral project for the location and performance of Indian manufacturing, *The Economic Journal,* Vol. 126, (591), doi:10.1111/ecoj.12207.

GOH, 2020; Haryana Enterprise Promotion Policy, Industries Department, Government of Haryana, Chandigarh.

GOI, 1948; Factories Act, Ministry of Labour and Employment, Government of India, New Delhi.

GOI, 2001 and 2011; Population Census of India, Director General and Census Commissioner, India, Ministry of Home Affairs, Government of India. New Delhi.

GOI, 2003; Scheme of Package for Special Category States, Department of Promotion of Industry and Internal Trade, Ministry of Commerce and Industry, Government of India, New Delhi.

GOI, 2004-05 to 2014-15; Road Transport Year Book, Ministry of Road Transport and Highways, Government of India, New Delhi.

GOI, 2008; National Industrial Classification, Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

GOI, 2016; Annual Survey of Industries, Central Statistics Office, Industrial Statistics Wing, Ministry of Statistics & Programme Implementation, Government of India, New Delhi.

GOTN, 2014; Tamil Nadu Automobile and auto-component Policy, Industries Department, Government of Tamil Nadu, Chennai.

Jacobs, J., 1969; The Economy of Cities, New York: Random House.

Klier, T. and D. McMillen, 2015; Plant Location Patterns in the European Automobile Supplier Industry, *Growth and Change,* Vol. 46, No. 4.

Klier, T. and D. P. McMillen, 2008; Evolving Agglomeration in the U.S. Auto Supplier Industry, *Journal of Regional Science,* Vol. 48, No. 1.

Markusen, A., 1996; Sticky Places in Slippery Space: A Typology of Industrial Districts, *Economic Geography,* Vol. 72, No. 3.

Marshall, A., 1890; Principles of Economics, London: Macmillan.

Martin, R. and P. Sunley, 2003: Deconstructing clusters: chaotic concept or policy panacea? *Journal of Economic Geography* 3.

McCann, P., 2013; The Economics of Industrial Location: A Logistics Costs Approach, Berlin: Springer Verlag.

Meine, P.V.D and A. Sverrison, 2003; Enterprise clusters in developing countries: Mechanisms of transition and stagnation, *Entrepreneurship and Regional Development,* Vol. 15, No. 3, July-September.

MIDC, 2019; Maharashtra Industrial Policy, Maharashtra Industrial Development Corporation, Government of Maharashtra, Mumbai.

Narayanan, K., 1998; Technology acquisition, deregulation and competitiveness: A study of Indian automobile industry, Research Policy, Vol. 27, No. 2.

OECD, 1999; OECD Economic Review No: 29, Organisation for Economic Cooperation and Development, Paris.

Okada A., N. S. Siddharthan, 2008; Automobile Clusters in India: Evidence from Chennai and the National Capital Region. In: Kuchiki A., Tsuji M. (eds.) The Flowchart Approach to Industrial Cluster Policy. IDE-JETRO, Palgrave Macmillan, London.
https://doi.org/10.1057/9780230589520_5.

Porter, M., 2000; Clusters and the New Economics of Competition, *Harvard Business Review,* Vol. 76, No. 6.

Q. Rong, J. Yan, and G. Guo, 2004; "Research and implimentation of clustering algorithm based on DBSCAN", Comput. Appl., Vol. 4.

Rahmah, N., I.S. Sitanggang, 2016; Determination of Optimal Epsilon (Eps) Value on DBSCAN Algorithm to Clustering Data on Peatland Hotspots in Sumatra. In: IOP Conference Series: Earth and Environmental Science.

Ranawat, M. and R. Tiwari, 2009; Influence of Government Policies on Industry Development: The Case of India's Automotive Industry. Working Paper No. 57, Working Paper Series, Technology and Innovation Management, Hamburg University, Hamburg, Germany.

Rath, S., 2022; Temporal Analysis of Industrial Location in India: Concentrations and Dispersal, *Journal of Indian School of Political Economy,* Vol. 34, Nos. 3&4.

Rietveld, P., F. Bruinsma, 1998; Is Transport Infrastructure Effective? Transport Infrastructure and Accessibility: Impacts on the Space Economy Springer, Berlin.

SIAM, 2016; Automobile mission Policy 2016-26, *Society of Indian Automobile Manufactures,* New Delhi.

SIAM, 2016; Statistics on production of Vehicles, *Society of Indian Automobile Manufactures,* New Delhi.

Uchikawa and S. Roy, 2013; The Development of Auto Component Industry in India. Mimeograph.

Wan, S., Y.P. Wang, 2020; The comparison of density-based clustering approach among different machine learning models on paddy rice image classification of multispectral and hyperspectral image data, *Agriculture*, 10.

Woodward, D., 1992; Locational determinants of Japanese manufacturing start-ups in the United States, *Southern Economic Journal*, Vol. 58.

CLUSTER PROFILES OF THE AUTO COMPONENT INDUSTRY IN INDIA

Ashish Andhale and Sharadini Rath

The Auto-Component Industry in India is highly clustered. The industry clusters are spread unequally across the northern and southern regions of the country. In this study, we have attempted to understand the detailed geographic and structural profiles of India's auto component industry clusters. The geographic profiles examined the geographic coverage and accessibility of assemblers in the context of these clusters. The study of structural properties was based on the auto component industry's product type and firm size. The study found that the southern region tends to be geographically larger (in terms of spread) while, at the same time, being more dispersed with larger peripheries when compared to the northern region. Also, it highlighted the close proximity among assemblers and component plants as an important factor for the location of the component industry. Also, an examination of structural properties showed that a cluster is a heterogeneous mixture of firms with small to large size firms. Each cluster has representation by way of most product types but the degree of this representation varies from cluster to cluster.

Keywords: Auto-Component Industry, Industrial Clusters.

1. Introduction:

Industry clustering has been discussed in the literature for a very long time. The literature identifies critical benefits of clustering such as increasing returns to scale, labour market pooling and knowledge spillover [Marshall, 1890; Jacobs, 1969; Porter, 2000, Pp. 15-34; and Becattini, 2006, pp. 664-672]. Further, Jacobs [1969] highlighted the role of cities in the development of clusters (by way of benefit of urbanisation economies). But it must be noted that Porter's [2000] definition of a cluster is most often used in cluster studies. According to Porter "clusters are geographic concentrations of interconnected companies, specialised suppliers, service providers, firms in related industries, and associated institutions, (e.g., universities, standards agencies, trade associations) in a particular field that compete, but also cooperate (p. 16). Porter's model also

Ashish Andhale is PhD student at Gokhale Institute of Politics and Economics, Pune.
Sharadini Rath is Affiliate Fellow at Indian School of Political Economy, Pune, India.
Acknowledgements: Authors would like to thank the Automobile Component Manufacturers Association (ACMA), India for making available their membership directory, without which this analysis would not have been possible.
They are gratefully acknowledge Mr Abhay Tilak, Dr. Nilakantha Rath and Dr. Naresh Bodkhe for their useful comments and suggestions on the manuscript. Also, they are thankful to Mr Ujaval Gandhi for his help in the use of geographical tools.

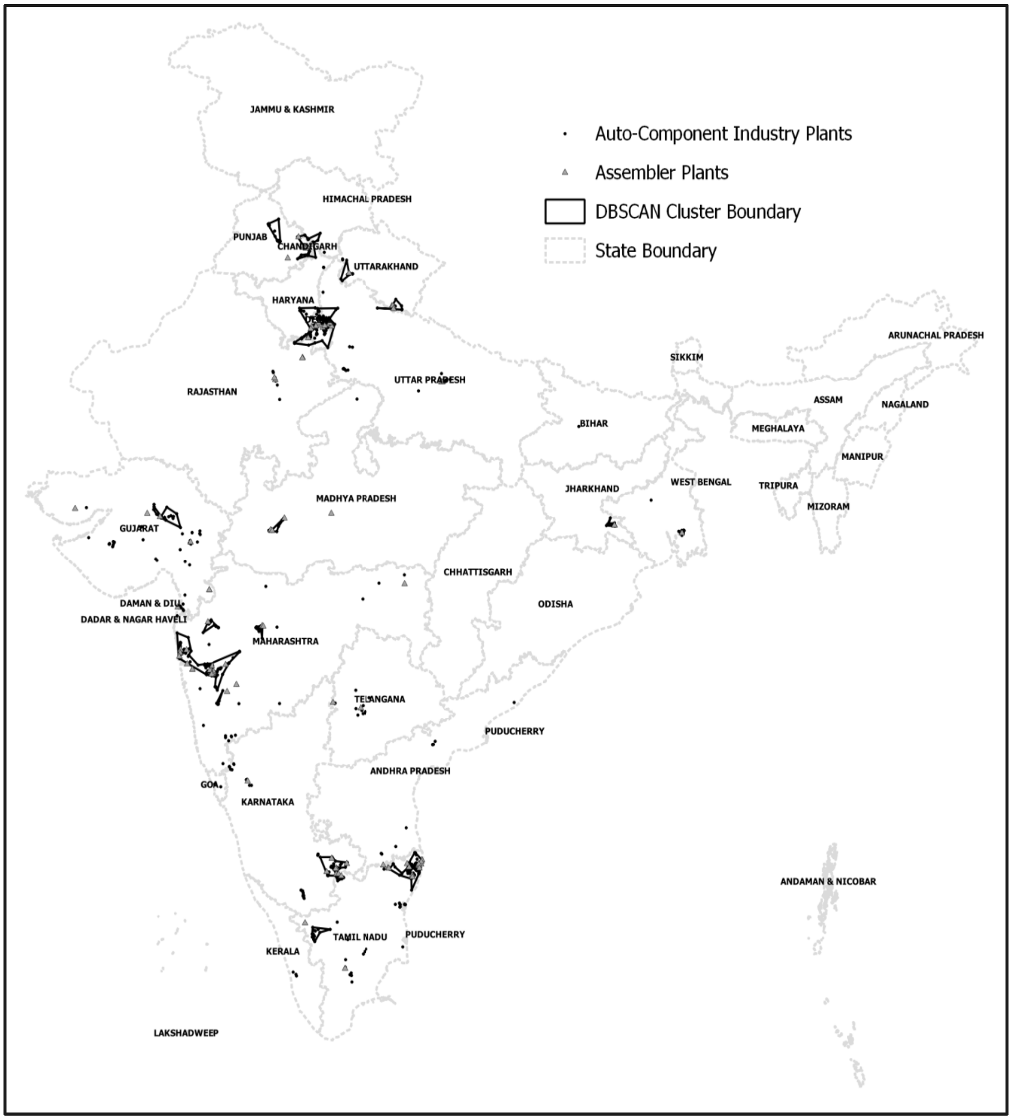
states that the geographical concentration that occurs in clusters enhance the interaction between four factors - factor input conditions, demand conditions, related and supporting industries while setting the context for firm strategy and rivalry. Therefore, clusters are considered an efficient tool in transforming policy into economic growth, competitiveness and play an increasingly prominent role in business development and industrial policy [Brown, et. al., 2013]. Here briefly give the structure of this paper.

2. A brief review of earlier papers

In the context of the auto-component industry, in an earlier paper [Andhale and Rath, 2023a], using DBSCAN methodology found that sixteen auto-component clusters in India cut across administrative boundaries; nine clusters spread across the northern region and seven clusters in the southern region (See, Map 1 below). A single large cluster, Gurgaon in the northern region, dominates the development of clusters in the northern region whereas in the southern region there are three dominant clusters: Pune-Mumbai, Chennai, and Bangalore. The clusters overlap with the pattern of urbanisation thereby highlighting the requirement of urban economies to sustain them. Also, a regional comparison undertaken in the study showed that the south has a larger presence of vehicle assembly plants, greater depth in ancillarisation in terms of employment size of plants, and far greater geographical dispersion of the industry which meant that the south can be effectively considered a large agglomeration that sees the north of India as an export market where absolute demand is much higher. This relationship is sustained by the infrastructure provision that reinforces these tendencies and aids further cluster formation in the preferred regions of the country.

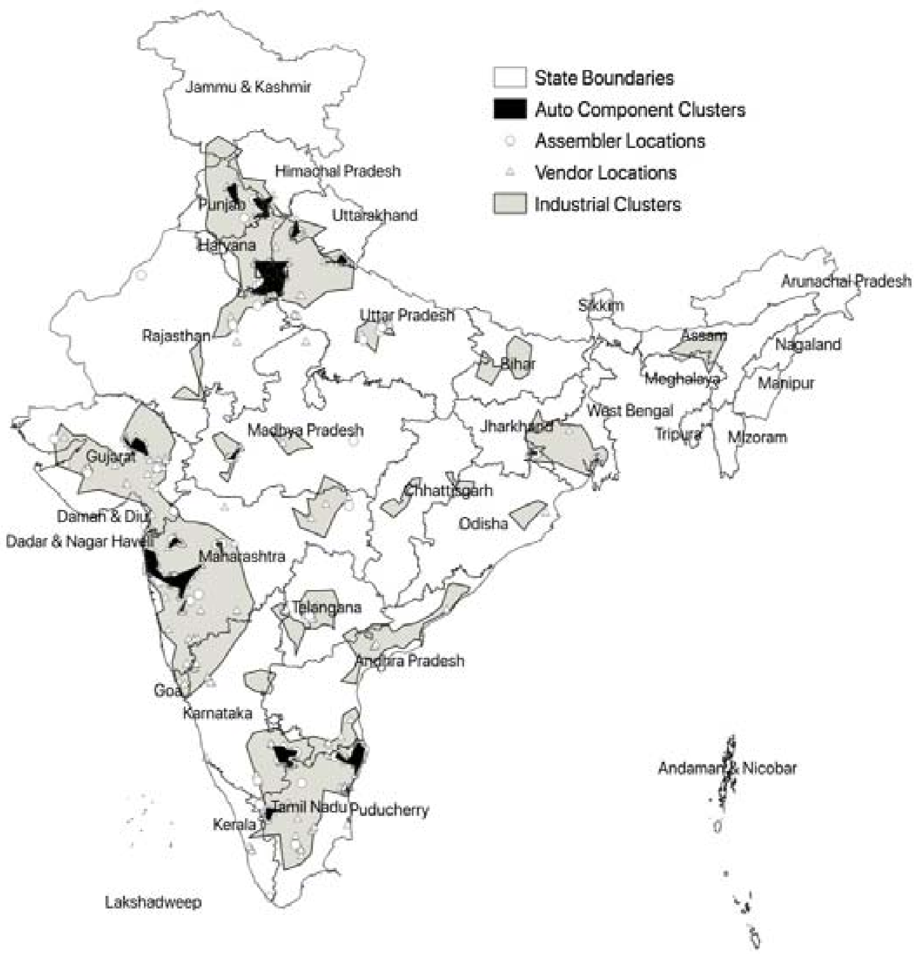
In a study undertaken to identify the spatial-temporal evolution of industrial clusters in India, the major areas of location concentration were identified using similar geographical analysis techniques of Density-Based Spatial Clustering of Applications with Noise that were used in our study. Map II in Andhale and Rath [2023a] which is reproduced below shows the overlap between the DBSCAN clusters of multi-industry types and the identified auto-component clusters identified.

Map 1: DBSCAN Clustering: Sixteen Auto-component Clusters in India



It should be noted that the database used in the [Rath, 2023a] is unrelated to the one used here but all the same uses plant level locations across the country over the period 2002-2018. It is clear that auto sector clusters are a geographical part of the larger industrial clusters in India (shown in Map II below) . This shows that single industry clusters, such as the auto-component clusters, exist within large multi-industry clusters, in effect, giving rise to clusters within clusters. The advantages of being within both, single industry, as well we multi-industry clusters are clear, especially for labour pooling and technology transfer.

Map 2: DBSCAN Clustering: Sixteen auto-component clusters and multi-industry clusters in India



Source: Rath [2022]

Given this larger perspective of industrial clustering in India, and the place of the auto-component industry within it, we now move to show the detailed profile of the auto component industry clusters that have been identified. In this, we first understand the clusters' geographical profiles in terms of geographical coverage and accessibility of assemblers for auto component plants in the clusters. Secondly, we discuss other characteristics of the profiles of dominant and smaller clusters. Further, we explore the structural properties of clusters in terms of product type and firm size. We use Auto Component Manufacturers Association (ACMA) membership directory [ACMA, 2016] for the analysis. Information from the 2016 edition of this publication comprises the basic data in the specific format used to create a database for this study. In addition, we present qualitative analyses drawing on an extensive interview conducted with the regional deputy executive of ACMA.

3. Profile of Auto-Component Industry Clusters

3.1 Differences in geographic coverage across clusters and between regions

An analysis of identification of clusters shows that they exist in different sizes, shapes and spread unequally in the northern and southern regions of India [Andhale and Rath, 2023a]. Larger the geographic size of a cluster, greater the spread in a region with larger peripheries thereby benefitting from larger local economies in terms of employment generation and also a better chance of matching workers with the right jobs as a result of which firms can maximise their productivity. If the cluster is tight and the firms in the cluster are distributed in a relatively small area, the reach of employment generation potential gets confined to a limited region of the local economy. Thus, this analysis highlights the importance of evenly spread geography of the industry for the development of regional economies.

To analyze a cluster's geographical coverage, we estimate the distances (in Km) between plants' locations (points) by calculating the Distance Matrix between them. Then, we find out the maximum distance and average distance for the sixteen clusters (See Table 1).

The estimation of the maximum distance (in Km.) between two plants in the cluster and the share of the number of plants in the cluster effectively indicate the geographical size of the cluster. On the other hand, the average distance between plants in the clusters shows the tightness of the cluster. The lesser the average, the tighter the cluster, indicating that a high proportion of plants in the cluster operate in a limited area. On the other hand, the higher the average, the lesser the cluster's tightness and the plants are well spread out in the clusters.

Table 1. Geographic coverage of the Clusters:-
Table (1) Cluster Wise Distribution of Plants and Geographic Size of the Clusters

No.	Cluster	Region	No. of Plants	Share of each Cluster in Total (%)	Maximum Distance between Plants in Cluster (km)	Average Distance between Plants in Cluster (km)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1.	Gurgaon	North	525	31.16	184.57	40.46
2.	Udham Singh Nagar	North	66	3.92	84.00	12.22
3.	Ludhiana	North	44	2.61	26.39	22.44
4.	Jamshedpur	North	37	2.20	30.04	7.23
5.	Haridwar	North	32	1.90	66.69	11.58
6.	Solan	North	30	1.78	112.20	35.42
7.	Ahmedabad	North	29	1.72	116.00	34.61
8.	Dhar	North	22	1.31	64.01	25.61
9.	Kolkata	North	18	1.07	18.51	7.93
	Northern Clusters	North Total	803	47.66	78.05	19.10
10.	Pune-Mumbai	South	256	15.19	228.55	55.57
11.	Chennai	South	187	11.10	119.34	30.71
12.	Bangalore	South	123	7.30	102.94	28.15
13.	Aurangabad	South	55	3.26	40.39	7.67
14.	Nashik	South	39	2.31	56.88	12.93
15.	Coimbatore	South	22	1.31	65.82	20.09
16.	Satara	South	16	0.95	46.69	8.35
	Southern Clusters	South Total	698	41.42	94.37	23.35
	Total Cluster Plants		1501	89.08	85.19	22.56
	Total Non-Cluster Plants		184	10.92	NA	NA
	Total Plants		1685	100.00	NA	NA

Source: Author's Calculations based on ACMA data.

Table 1 gives these figures for all 16 auto-component clusters. It becomes clear that, on an average, the southern clusters have larger maximum distances between plants as well as average distance between plants even though they have a smaller proportion of total plants than the northern region. This means that clusters in the south tend to be geographically more spread out with larger peripheries. The northern clusters are knit relatively much tighter. The southern region (105) also has more non-cluster plants than the north (79), showing greater dispersal of industrial location. This fits in with the overall industrial clustering shown in Map 1, in which the southern region is shown to have far more wide spread industrial location than the north. In some sense, the specialised clusters, within the multi-industry clusters are likely to have similar average geographical features as the larger clusters. The auto-component industry analysis here, illustrates this clearly. According to Okada & Sidharthan [2008] Indian industrial clusters are largely concentrated in the three clustered regions: NCR, Mumbai-Pune, and Chennai-Bangalore, across different manufacturing sectors. But, in this study, the geographical scale of 'clusters' not defined.

Geographic Profile of the Gurgaon Auto-Component Cluster

In the case of the Gurgaon Cluster, Maruti Udyog Ltd., (MUL) played a significant role in developing the cluster. The Company set up its two plants in Gurgaon in Haryana in 1982 and 1992. MUL's experience is unique in India, particularly in terms of its system of procurements and enormous efforts to develop first-tier suppliers in its proximity [Okada, 2008]. Therefore, initially, many auto-component plants got established in the vicinity of the MUL plants. Later, other assemblers also set up their plants in this cluster which further boosted the auto-component industry in adjoining regions of the States of Uttar Pradesh, Haryana and Rajasthan. For example, Eicher Polaris Pvt. Ltd., Escorts Ltd., General Motors India Pvt. Ltd., HMT Ltd., SAS Motors established their plants in the State of Haryana. Honda Cars India Ltd., New Holland Tractors (India) Ltd., Piaggio Vehicles Pvt. Ltd., in Uttar Pradesh. Ashok Leyland Ltd., Eicher Polaris Pvt. Ltd., Mahindra & Mahindra Ltd., and Honda Cars India Ltd., were set up in the State of Rajasthan [ACMA, 2016]. As a result, the spread of the geographical cluster grew in other neighbouring regions of Delhi. On the other hand, Delhi became an inconvenient location for larger factories because space was not readily available. Also, the

State Government of Haryana encouraged policies that promoted the transfer of factories from Delhi to the suburbs and surrounding districts [Tomozawa K, 2016, Pp. 11-27]. This led to further advantages in terms of location in the adjoining districts of surrounding states.

The Gurgaon cluster has 525 auto-component plants. This constitutes 31.16 per cent share of the total clusters in India and 65.38 per cent share of northern region's clusters, which shows that this is the primary cluster in the northern region (Table 1). The cluster's geographical coverage is relatively large compared to the average south and north clusters but is smaller than the Mumbai-Pune cluster in the south. This is seen from the maximum distance between plants in this cluster which is 184.57 km. being higher than the average maximum distance of the clusters in the north (78.04 km) and in south (94.37 km) regions. On the other hand, the Gurgaon cluster's average distance spread is only 40.46 km, and most plants are located in Haryana State (367 Plants), reflecting that Gurgaon is a very tightly knit cluster. Also, Haryana State has only 3.8% of non-cluster plants in the northern region. This means that most of the plants in this cluster are centred around the core region of Gurgaon, and dispersion outside the cluster is negligible.

This analysis also reveals that the Gurgaon Cluster dominates in the northern region with the highest geographical coverage in terms of maximum distance between plants and percentage share of plants. However, it is a very tight cluster seen from the smaller average distance between plants. The impact of this cluster in terms of localisation economies is confined to limited areas resulting in employment generation in the primarily urban areas surrounding the National Capital Region (NCR) of Delhi. In order to quantify the place of the metropolis in the location preferences of industry, Rath [2023a] attempted a geographical calculation for the category of Other Workers (all non-agricultural workers) for figures given in the Population Census 2011 [GOI, 2011]. The Euclidean distance of the centroid of the metropolis to the centroids of districts around it was measured. The percentage share of district level Other Workers in the respective State total was assigned to these centroids. The cumulative share of Other Workers as one moves away from the metropolis in widening circles was plotted (This means that, in Haryana, taking Delhi as the metropolis, about 60 per cent share of Other Workers were located within 100 kms or 45 per cent of the total distance within Haryana from Delhi. Similarly, in Uttar Pradesh, it was found that 40 per cent share of Other Workers were located within 200 kms or

25 per cent of total distance within UP from Delhi. Given the large industrial cluster around Delhi seen in Map 1, this shows that the employment impact is highly concentrated in the immediately adjoining areas of the cluster. The impact of the metropolis in location choice of industry and resulting employment patterns becomes clear from this analysis.

3.1.1 Geographic Profile of the Pune-Mumbai Auto-Component Cluster

The Pune-Mumbai auto-component cluster in Maharashtra is a location choice for many top assemblers. The cluster consists of 22 large assemblers' plants in Maharashtra, including Tata Motors Ltd, Ashok Leyland Ltd, General Motors India Pvt. Ltd., Mahindra & Mahindra Ltd., Mercedes-Benz India Pvt. Ltd., Piaggio Vehicles Pvt. Ltd., Skoda Auto India Pvt. Ltd., and Volkswagen India Pvt. Ltd., etc., which are in various parts of the cluster [ACMA, 2016].

The Pune-Mumbai Cluster with 256 plants is spread across industrially advanced districts and urban agglomerations namely Pune (205 plants), Mumbai (29) and Thane (12) districts of the State of Maharashtra. The cluster constitutes a 15.19 per cent share of total plants in India and 36.68 per cent share of the plants in the entire southern region (Table 1). The cluster is largely centred around Pune which has a long history of automotive industry of different kinds, including 2 and 3-wheeler production. For instance, Bharat Forge, one of the oldest and largest auto component manufacturers in India, is located in Pune, and is part of the historical legacy of this cluster. They are credited with creating standards of quality and a unique culture within this cluster.

With the maximum distance between plants of 228.55 km, Pune-Mumbai cluster is the largest geographic cluster in India (Table 2). The average distance between plants (55.57 km) between plants is higher than in all other clusters from south and north, reflecting good dispersal of the auto component industry in periphery regions in the cluster (Table 1). Furthermore, Maharashtra has 20 per cent of the southern region's total of non-cluster plants which is far more than for the total for northern clusters like Gurgaon. This shows significant dispersion of plants in the state, which might be because it is located within the larger western industrial cluster as shown in Map 1. Also, Rath [2023a] showed that 61 per cent share of the total Other Workers for the state was located within 200 kms or 23 per cent of the total distance from Mumbai revealing that employment generation is significantly influenced by the presence of these clusters.

3.1.2 Geographic Profile of the Chennai Auto-Component Cluster

Chennai Cluster with 187 plants are spread across mainly three districts of the State of Tamil Nadu - Chennai (92), Kancheepuram (72) and Tiruvallur (19). [Okada & Sidharthan, 2008] point out three historical factors which led to the emergence of one of the oldest -Chennai clusters. Firstly, the Tariff Commission [GOI, 1957] recommended discouragement of imports (including automobiles and auto components) while, at the same time, favoured domestic production of automobile & auto component. At that time several Tamil Nadu (then Madras) based firms like the TVS group and Standard Motors were already at the forefront and active in industrial production. Chennai had the benefit of proximity to seaports when the auto-component industry heavily depended on imports in the 1960s. Secondly, the Government of Tamil Nadu actively supported these local firms between 1950s and 1960s. Thirdly, several Tamil Nadu based industries like the TVS group, MRF, Standard Motors, Ashok Leyland and the Rane group played an active and notable role in forming the cluster in Chennai.

The Chennai cluster has 11.10 per cent share of total plants in India and 26.79 per cent share of total southern region (Table 1). The cluster is also the hub of 19 plants of major assemblers including Ashok Leyland Ltd., Ford India Pvt. Ltd., Hyundai Motor India Ltd., Nissan Motor India Pvt. Ltd., Kamaz Vectra Motors Ltd., BMW India Pvt. Ltd., and Renault India Pvt. Ltd., [ACMA, 2016]. The major auto-component industries here are the Rane and TVS group of industries.

In terms of geographical coverage of the distribution of 11.10 per cent share of total plants in the cluster given the maximum distance between plants of 119.34 km, this cluster is one of the larger clusters after Pune and Gurgaon clusters. On the other hand, the average distance between plants in the cluster is only 30.71 km which is less than in regard to Gurgaon and Pune clusters which indicates that the cluster is very tightly concentrated with less dispersion of plants within the cluster. Tamil Nadu has about 16 per cent of southern non-cluster plants showing reasonable dispersion of location. This is likely to be due to the wide spread urbanisation and accompanying industrial location in Tamil Nadu [Andhale and Rath, 2022].

3.1.3 Geographic profile of other small northern region clusters

Apart from the main Gurgaon cluster, Udham Singh Nagar and Haridwar are pertinent clusters in the northern hilly region of Uttarakhand State. The centrally sponsored hilly region policy [Scheme of Package for Special Category States, GOI, 2003] played a significant role in developing these clusters. Under that policy, fiscal incentives such as exemption from excise taxes & corporate tax and capital investment subsidies were given to attract the manufacturing industry. These incentives helped to attract the auto-component industry. As a result, the clusters have 3.92 per cent and 1.90 per cent share of total plants along with geographical size of 84 km and 66 km in Udham Singh Nagar and Haridwar clusters, respectively. Mahindra & Mahindra, Tata Motors, Ashok Leyland, Bajaj Auto have established branch plants in Uttarakhand. Auto component units from the two clusters cater to these assemblers. Large auto-component manufacturers such as, Minda Industries Ltd., Munjal Auto Industries Ltd., ABC Bearing Ltd., and Breaks India Pvt Ltd are also some of the companies that set up plants in this cluster.

Ludhiana is the other cluster in the northern region with 44 plants and a 2.61 per cent share of total plants and 5.48 per cent share of the northern region plants. These auto component plants mainly cater to nine assemblers in the Punjab region, including some tractor manufacturers - Preet Tractors Pvt. Ltd., Mahindra Swaraj Tractors Ltd., and International Tractors Ltd. Also, JCBL Ltd., International Cars & Motors Ltd., and Swaraj Mazda Ltd. The major auto component manufacturers here are Chopra Autotech Pvt Ltd., Emkay Automotive Industries Ltd., and the GNA and New Swan groups. This is a small cluster with maximum distance of 26.39 kms between plants. The average distance between plants of 22.44 km presents a tightly knit cluster.

Ahmedabad is an auto-component cluster in the northern region, which developed in Gujarat and Daman & Diu region. It has 29 plants and 1.72 per cent share of total plants with the geographical size of 116 km area. The industry caters to four assemblers, namely, AMW Motor, Ford India, Mahindra Gujarat Tractor and Tata Motors Ltd. The average distance between plants is 34 km.

Similarly, Jamshedpur, another vital cluster in the northern region, consists of a 2.20 per cent share of total plants and is spread relatively in a small area with a geographical size of 30.04 km. The cluster developed in the state of Jharkhand and caters to mainly Tata Motors Ltd., and Tata Cummins Ltd. Kolkata and Dhar are two more small clusters in the northern region with a total plant share of less than 2 per cent.

3.1.4 Geographic Profile of other small southern region clusters

In the southern region, Bangalore is another prominent cluster with a share of 7.30 per cent of the total plants and the maximum distance between plants is 103 km, showing that the geographical size of the Bangalore cluster is almost equal to the Chennai. Karnataka has 22 non-cluster plants, 22 per cent of the southern total, showing that there is good dispersal of this industry within the state, even outside the main cluster. Similarly, Aurangabad, Nashik and Satara are the other essential clusters in the southern region with the total share of plants 3.26, 2.31 and 1.06 per cent, respectively. They are well spread out in the southern region, with geographical coverage of 40.39 km, 56.88 km and 46.69 km, respectively. Also, Coimbatore is another small cluster in the southern region with a share of relatively fewer plants (1.31 %) but expanded in extensive geographical coverage compared to other small clusters in the southern region.

Overall, the northern and southern region clusters constitute 89.08 percentage share of plants in clusters areas, whereas the remaining 10.92 per cent plants operate in the non-cluster area. We see that the geographical coverage of the southern region cluster (94.37 km) is 16.32 km more than the northern region cluster (78.05 km) reflecting an average spatial spread of southern clusters that is more than the northern counterpart. The average distance between plants in both clusters is almost similar (Table 1). For non-cluster areas, southern states have more plants than northern states. These statistics suggest that South tends to have locations that more dispersed with larger peripheries than the north region. A single large cluster dominates the development of the northern region near Gurgaon, whereas the southern region has three significant clusters namely, Pune-Mumbai, Chennai and Bangalore with smaller clusters creating greater dispersal.

Apart from size of clusters, they differ from each other in terms of their business culture. A mix of Indian and Japanese business culture dominates the Gurgaon cluster and Maruti Suzuki developed a culture of actively partnering with the vendors and quality improvement. North has the peculiarity of more family-owned auto-component business than the south.

Chennai cluster is an old orthodox cluster with more focus on consciousness of the product quality than the Gurgaon and Pune-Mumbai clusters. Sundaram fasteners from Chennai are the biggest exporters of fasteners with the best quality products (Interview with regional deputy executive of ACMA).

4. Accessibility of auto component plants for an assembler in cluster

The global automotive industry is dominated by assembler centric location of the auto-component industry. In case of the United States, a study [Smith and Florida, 1994, Pp. 23-41] highlighted the role of particular type of agglomeration and the co-location of backward and forward linked plants in the process of location of industries. Spatial co-location represents the subsets of spatial events whose instances are often located in close geographic proximity. The study shows close proximity between assemblers and component plants to be an essential location factor. Another study of the location pattern of auto-component plants in the U.S found that "shorter distance to assembler plants" is an important variable for the location [Klier and McMillen, 2008, Pp. 245-267]. Various studies on different locations, including India, are consistent with the tight linkages between assemblers and suppliers observed in the context of *just in time production*, which requires most production to be performed in reasonably close proximity to the assembly location [Klier and McMillen, 2015, Pp. 558-573; Okada and Siddharthan, 2008]. In the Indian context, it becomes essential to highlight the above characteristics by examining the geographical relationship between assemblers and auto-component vendors.

The previous section points out the location of various motor vehicles assembler plants in every cluster in India. This Section analyses the accessibility of auto component plants for an assembler located in prominent clusters. For this analysis, we find the Distance Matrix to measure the distance between each 94 assembler plants and 1659 auto component plants (vendors) using geographical tools such as Open-Source Desktop Geographic Information System software (QGIS 3.24.2) to understand the accessibility of vendor plants for assemblers. Accessibility highlights the role of proximity between plants & assemblers. It also ultimately reveals the degree of concentration of plants.

The Table 2 shows the accessibility of vendor plants for an assembler up to 500 km with five distance range categories across cluster and non-cluster regions in northern and southern parts. In the southern and northern regions, more than 30 per cent of vendor plants are accessible for an assembler within 100 km. In the northern region up to 300 km. 78.82 per cent cluster plants and 16 per cent non-cluster plants are accessible for an assembler. On the other hand, in the southern region, 76 per cent cluster plants and 50.51 per cent non-cluster plants are accessible for an assembler up to 300 km. This analysis highlights two things - firstly, the prominent role of 'close proximity' (100 km. distance) in industry location and development of the cluster. Secondly, it implies that the accessibility of vendors for an assembler in northern clusters shows a high degree of concentration of vendors in close proximity of assemblers. On the other hand, the southern region shows degree of concentration is lesser when compared to the north and greater spread out of plants from assemblers, with a greater proportion of non-cluster plants within easy reach of assemblers. Interview with regional deputy executive of ACMA also highlighted that the location of vendors is primarily dependent on two factors: 'location of assemblers' and 'quantity of sales' of the product. Assemblers like to have vendors supplying products with high sales volumes close to them, for ease of inventory maintenance and quick supply. Vendors supplying products with low sales volumes can be further away.

The pattern of accessibility of non-cluster vendors for assemblers shows a significant difference between the north and south regions as seen in Table 2. In the north, non-cluster vendors are quite far away from assemblers, with 85 percent being more than 300 kms away. However, in the south, non-cluster vendors are more evenly distributed over the range of 100 to 500 kms from assemblers. This shows that the nature of industrial location in the south is quite different from the north, with greater willingness to locate outside clusters or further away from other similar industries. This probably indicates that land, labour and other inputs are more easily accessible outside clusters in the south than in the north.

Table 2. Accessibility of Vendors for an Assembler for Five Distance Range

	North Region			South Region			Grand Total
Distance Range (KM)	Cluster	Non-cluster	North Total	Cluster	Non-cluster	South Total	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0 to 100	32.68%	6.00%	32.59%	31.04%	4.96%	30.93%	31.51%
100 to 200	15.34%	2.00%	15.30%	14.94%	20.66%	14.97%	15.08%
200 to 300	30.80%	8.00%	30.72%	30.01%	24.79%	29.99%	30.24%
300 to 400	9.05%	38.00%	9.15%	11.33%	20.66%	11.37%	10.60%
400 to 500	12.13%	46.00%	12.24%	12.67%	28.93%	12.74%	12.57%
Total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Source: Author’s calculations

5. The Structural Properties of the clusters

This Section explores the two structural characteristics of the clusters, namely, firms size variation across clusters then product type variation across clusters. These characteristics help us to understand the structural nature of clustering.

5.1 Share of Firm Size in employment in Clusters

Analysis of firms’ size variation across clusters indicates the concentration of firms’ employment categories in clusters. The high share of large firms can show high technological development and capital usage. In addition, the large firms have more backward linkages which have the capacity to generate more employment in the clusters.

We have defined the firm size according to the number of employees in the firms [Diebolt et al., 2016, Pp. 85-102]:

Small firms: less than 100 employees,
Medium firms: 100-250 employees,
Large firms:more than 250 employees

Information about employee numbers in the ACMA data was at the firm level but not at the plant level. To use this data for cluster level analysis we cannot use employee data at firm level for those firms which have plants in more than one cluster, since they cannot be separated at plant level. Therefore, we have used firm level data on employment for only those firms who have all plants in the same cluster, so that the employee numbers can be aggregated at cluster level. We have used this method to analyse the product profiles of clusters as well below in

Section 5.2.

Comparing firm size by employment across our DBSCAN clusters indicated earlier, we observe that the overall share of large firms in the total is higher than the small and medium firms (Table 3). It becomes clear that there is a difference between large and small clusters in their firm size profile. Large clusters tend to have fewer large firms and a greater proportion of medium and small firms (Gurgaon, Pune-Mumbai and Chennai clusters). Smaller clusters appear to have more large firms, and a smaller proportion of medium and small firms (Coimbatore, Satara in the south and Ahmedabad, Solan, Ludhinana in the north).

Table 3. Comparison of firm size across Cluster.

Clusters \ Firm size	Large	Medium	Small	Grand Total
(1)	(2)	(3)	(4)	(5)
North	31%	46%	23%	100%
Ahmadabad Cluster	57%	43%	0%	100%
Dhar Cluster	100%	0%	0%	100%
Gurgaon Cluster	29%	51%	19%	100%
Haridwar Cluster	0%	100%	0%	100%
Jamshedpur Cluster	6%	63%	31%	100%
Kolkata Cluster	29%	29%	43%	100%
Ludhiana Cluster	41%	14%	45%	100%
Solan Cluster	50%	38%	13%	100%
Udham Singh Nagar Cluster	0%	0%	100%	100%
Non Cluster	28%	36%	36%	100%
South	34%	41%	26%	100%
Aurangabad MH Cluster	38%	50%	13%	100%
Bangalore Cluster	30%	37%	33%	100%
Chennai Cluster	30%	45%	26%	100%
Coimbatore Cluster	60%	20%	20%	100%
Nashik Cluster	38%	38%	23%	100%
Pune- Mumbai Cluster	26%	45%	30%	100%
Satara Cluster	67%	33%	0%	100%
Non Cluster	65%	24%	12%	100%
Grand Total	32%	44%	24%	100%

Source: Author’s Calculations from ACMA data

In large clusters, there is a greater scope for a multiplier effect, with assemblers and large vendors encouraging a depth in the supply chain, with medium and small firms proliferating. In small clusters, it is possible that this scope exists only to a small extent or maybe does not exist.

There is a difference in the profile of the non-cluster plants in the north and south. Most of the non-cluster firms in the south are large, with a small proportion being medium and small. In the north, this is different with a vast majority of non-cluster firms being medium and small. So location choices of large firms are quite different between north and south, with large firms possibly spearheading dispersal of industrial location in the south.

It is interesting to refer back to Andhale and Rath (2023a) here to observe that the for overall multiplier that was calculated using ASI plant for the year 2015-16 , assuming ACMA plants to be Tier-I, the south had a much larger multiplier from Tier-I to all those below, than the north. This is clearly reflected in this analysis of firm size. In fact, in the year 2015-16, ASI workers for the auto-component sector are larger in the south (334607) than in the north (245170). The south gets the advantage in two ways, namely, large firms creating a greater depth of supply chain, also more dispersed location choice.

5.2 Product Profile of Clusters

The following analysis shows the share of product composition in individual clusters revealing the specialisation of clusters in terms of product type. ACMA has divided product types into six categories: Body Parts, Drive Transmission & Steering Parts, Electrical Equipments, Electrical Parts, Engine Parts, Others Parts, and Suspension and Braking Parts. The information about product type is given at the firm level. Therefore, as in the previous section given data limitations, we only consider single-plant companies and multi-plant companies having their plants in the same cluster (Table 4).

Some notable and similar properties emerge concerning the distribution of product type in clusters. Firstly, the largest percentage of firms produce Other Parts and Engine Parts in the sixteen clusters (Table 5). Engine parts segment constitutes 142 different components [ACMA, 2016]. These are recognised as high-value components in the industry. As shown in Figure (2) and (3), with 30 per cent of plants, the southern region has a greater share of these firms than the north at 20 per cent.

Table 4. Distribution of firms in clusters as per product type

Type of Products	Multi-Plant Companies having their plants in the same Cluster	Single Plant Companies	Grand Total
(1)	(2)	(3)	(4)
Body Parts	21	38	59
Drive Transmission & Steering Parts	29	54	83
Electrical Equipments	7	15	22
Electrical Parts	6	15	21
Engine Parts	40	91	131
Others	44	104	148
Suspension and Braking Parts	11	25	36
Grand Total	158	342	500

Source: Author's Calculations from ACMA data

Firms producing 'Other Parts' also have a large share in clusters. Other parts consist of the production of bolts, nuts, rivets, screws and washers, latch assemblies, sheet assemblers, different types of bearings, etc. In industry, these are considered as low-value. Aurangabad, Bangalore, Dhar, Kolkata, Haridwar, Jamshedpur, Ludhiana clusters and Non-cluster regions have more than 30% firms producing these other parts. The majority of the clusters have more than 50% share of firms producing Engines and Other Parts.

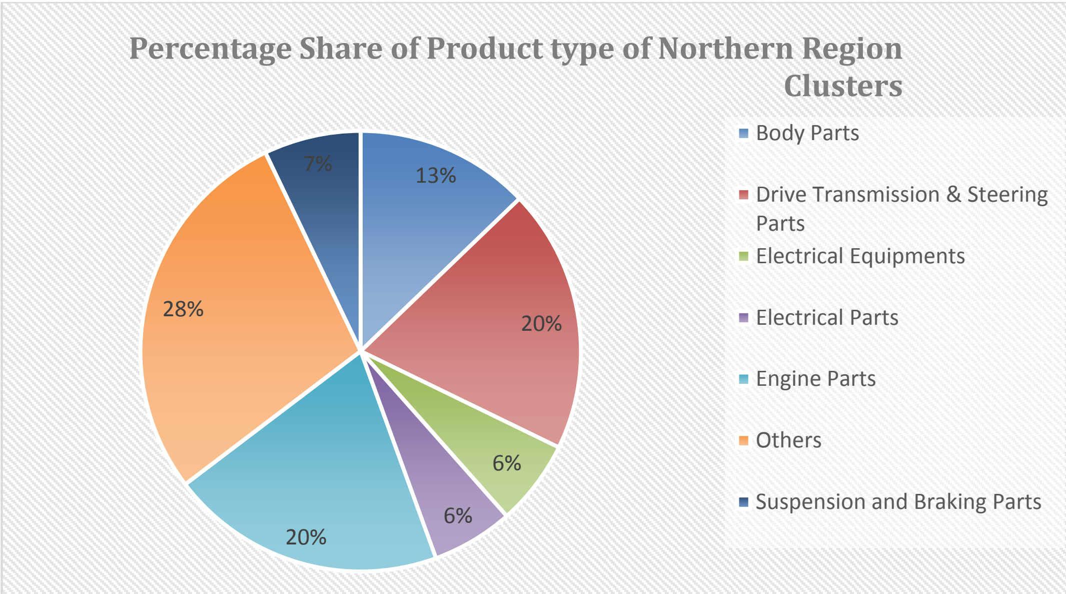
Table (5) shows that Drive Transmission & Steering Parts firms have a high presence of plants in all clusters except for small cluster like Haridwar and Satara. Notably, Northern clusters like Ludhiana and Gurgaon specialise in this product type with 36 and 17.23 per cent of all product types, respectively (Table 5). Overall, each cluster has representation in most of the product types. However, the degree of this representation varies from cluster to cluster. The lowest share of plants is for Body Parts, electrical equipments and electrical parts categories for clusters. According to officials of the ACMA, electrical equipment and parts suppliers tend to be generic across many industries, not specialised. It is possible that this is the reason for their low share within these clusters.

Table 5. Product type-wise distribution in clusters.

All clusters	Product type wise distribution in clusters								
	N/S	Body Parts (%)	Drive Trans mission & Steering Parts (%)	Electrical Equipmen ts (%)	Electrical Parts (%)	Engine Parts (%)	Others (%)	Suspension and Braking Parts (%)	Grand Total (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Gurgaon Cluster	N	13.11	17.23	7.87	7.12	20.97	26.59	7.12	100
Ludhiana	N	4.00	36.00	0.00	4.00	16.00	32.00	8.00	100
Jamshedpur Cluster	N	16.67	16.67	0.00	0.00	5.56	44.44	16.67	100
Kolkata	N	9.09	18.18	0.00	0.00	9.09	54.55	9.09	100
Solan	N	9.09	36.36	9.09	9.09	36.36	0.00	0.00	100
Ahmadabad	N	14.29	28.57	0.00	0.00	42.86	14.29	0.00	100
Dhar	N	16.67	16.67	0.00	0.00	16.67	50.00	0.00	100
Udham Singh Nagar	N	25.00	25.00	0.00	0.00	25.00	25.00	0.00	100
Haridwar Cluster	N	50.00	0.00	0.00	0.00	0.00	50.00	0.00	100
Pune- Mumbai Cluster	S	14.53	11.11	5.98	2.56	27.35	29.91	8.55	100
Chennai Cluster	S	4.41	14.71	2.94	5.88	36.76	22.06	13.24	100
Bangalore	S	8.16	16.33	2.04	0.00	30.61	38.78	4.08	100
Aurangabad MH	S	5.56	16.67	0.00	5.56	22.22	33.33	16.67	100
Nashik	S	27.78	16.67	0.00	0.00	27.78	22.22	5.56	100
Coimbatore	S	0.00	14.29	14.29	0.00	28.57	28.57	14.29	100
Satara	S	33.33	0.00	0.00	0.00	66.67	0.00	0.00	100
North Total	S	11.07	13.57	3.93	2.86	30.36	28.93	9.29	100
South Total	S	12.82	19.37	6.27	5.98	20.23	28.21	7.12	100
NON CLUSTERS		7.94	14.29	1.59	3.17	36.51	30.16	6.35	100
Grand Total		11.67	16.57	4.90	4.47	25.79	28.67	7.93	100

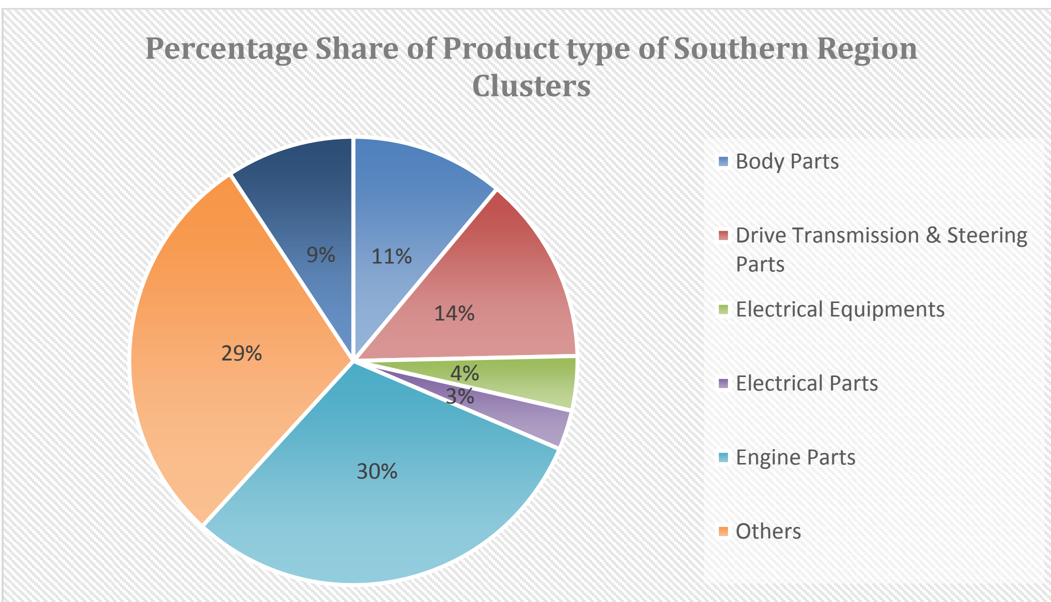
Source: Author’s Calculations from ACMA data.

Figure 2. Percentage Share of Product type of Northern Region Clusters.



Source: Author's Calculations from ACMA data.

Figure 3. Percentage Share of Product type of Southern Region Clusters



Source: Author's Calculations from ACMA data.

6. Summary and Concluding Remarks

This study sets out the geographical and structural profiles of auto-component clusters identified from plant level data. On the average, southern clusters have larger maximum distance between plants, as well as average distance between plants, even though they have a smaller proportion of total plants than the northern region. This means that clusters in the south tend to be geographically more spread out, with larger peripheries. The northern clusters are relatively much tighter. The southern region also has more non-cluster plants than the north, showing greater dispersal of industrial location. This fits in with the overall map of industrial concentrations in India, in which the southern region is shown to have far more wide spread industrial location than the north. In some sense, the specialised clusters within the multi-industry concentrations are likely to have similar average geographical features as the larger clusters. The auto-component industry analysis here illustrates this clearly.

We have also showed that there is a difference between large and small clusters in their profile of firm size by employment. Large clusters tend to have fewer large firms, and a greater proportion of medium and small firms (Gurgaon, Pune-Mumbai and Chennai). Smaller clusters appear to have more large firms, and a smaller proportion of medium and small firms (Coimbatore and Satara in the south and Ahmedabad, Solan, and Ludhinana in the north). Most of the non-cluster firms in the south are large, with a small proportion being medium and small. In the north this is different with a vast majority of non-cluster firms being medium and small. So, location choices of large firms are quite different between north and south, with large firms possibly spearheading dispersal of industrial location in the south. The south gets the advantage of both, large firms creating a greater depth of ancillaries, and a more dispersed location choice.

In terms of profile of clusters in terms of the products they manufacture, a combination of high value engine parts and low value other miscellaneous parts make up about half the share of products in most clusters. Both these vendors have a larger presence in the southern region. The northern region has a larger share of vendors of body parts, drive-transmission-steering parts and electrical parts. There are cluster-wise differences in product profile of firms, but they are small. Most clusters have some proportion of all product vendors.

REFERENCES

ACMA, 2016; Membership directory, Auto Component Manufacturers Association (ACMA), New Delhi.

Andhale, A. and S. Rath, 2022; Cluster Formation in the Auto Component Industry in India, *Journal of Indian School of Political Economy*, Vol. 34, Nos. 3&4.

Becattini, G., 2006; The industrial district and development economics in Raffaellit. T, Becattini, G. and Dardi, M. (eds.) The Elgar Companion to Alfred Marshall, U.K.: Edward Elgar.

Diebolt, C. et. al., 2016; A Comparative Regional Analysis of Indian Auto Components Industry, book-Dynamics of Distribution and Diffusion of New Technology: A Contribution to the Historical, Economic and Social Route of a Developing Economy, Springer Publications.

GOI, 1957; Tariff Commission Reports, Government of India, New Delhi.

GOI, 2001 and 2011; Population Census of India, Director General and Census Commissioner, India, Ministry of Home Affairs, Government of India. New Delhi.

GOI, 2003; Scheme of Package for Special Category States, Department of Promotion of Industry and Internal Trade, Ministry of Commerce and Industry, Government of India, New Delhi.

Jacobs, J., 1969; The Economy of Cities. Random House, New York, NY.

Klier, T., and D.P. McMillen, 2008; Evolving agglomeration in the U.S. auto supplier industry, *Journal of Regional Science*, Vol. 48, No. 1.

Klier, T. and D. McMillen, 2015; "Plant Location Patterns in the European Automobile Supplier Industry," *Growth and Change, Wiley Blackwell*, Vol. 46, No. 4, December.

Marshall, A., 1890; Principles of Economics, Macmillan, London.

Okada, A. and N.S. Siddharthan, 2008; Automobile Clusters in India: Evidence from Chennai and the National Capital Region. In: Kuchiki A., Tsuji M. (eds) The Flowchart Approach to Industrial Cluster Policy. IDE-JETRO. Palgrave Macmillan, London.
https://doi.org/10.1057/9780230589520_5.

Porter, M.E., 2000; Location, Competition and Economic Development: Local Clusters in a Global Economy, *Economic Development Quarterly*, Vol. 14, No. 1.

Rath, S., 2022; Temporal Analysis of Industrial Location in India: Concentrations and Dispersal, *Journal of Indian School of Political Economy*, Vol. 34, Nos. 3&4.

Smith, and Florida, 1994; Agglomeration and industrial location: An econometric analysis of Japanese-affiliated manufacturing establishments in automotive-related industries. *Journal of Urban Economics*, 36.

Tomozawa, K., 2016; Development and Spatial Dynamism of the Automobile Component Industry in India, *Journal of Urban and Regional Studies on Contemporary India*, Vol. 3, No. 1.

Software Availability: Open Source Desktop Geographic Information System software (QGIS 3.24.2)

PERFORMANCE ASSESSMENT OF THE AUTO-COMPONENT INDUSTRY CLUSTERS IN INDIA

Ashish Andhale*

This study attempts to analyze the performance of the automobile-component industry in India. This is undertaken by analyzing the industry at the all- India level in terms of critical financial and structural parameters. Given the geographical spread of the industry across different state clusters and the growing dominance of the industry within the overall automobile sector, this study focuses on its performance across 'states clusters' and across 'north - south regions' using key financial and structural parameters. Analysis of ratios and aggregates calculated on the basis of unit-specific information in Annual Survey of Industries (ASI) datasets aims to derive the following measures: overall partial productivity and resource use efficiency, different input intensities, factories and workers, debt rate and profit rate, partial factor productivity and total factor productivity. We find that the auto-component industry performed well in the period 2004-05 to 2016-17 in terms of critical parameters at the all-India level. However, an in-depth analysis using State groups to represent clusters shows inter-cluster differences in many parameters. Regional level analysis of the industry shows that the southern region performs better than its northern counterpart.

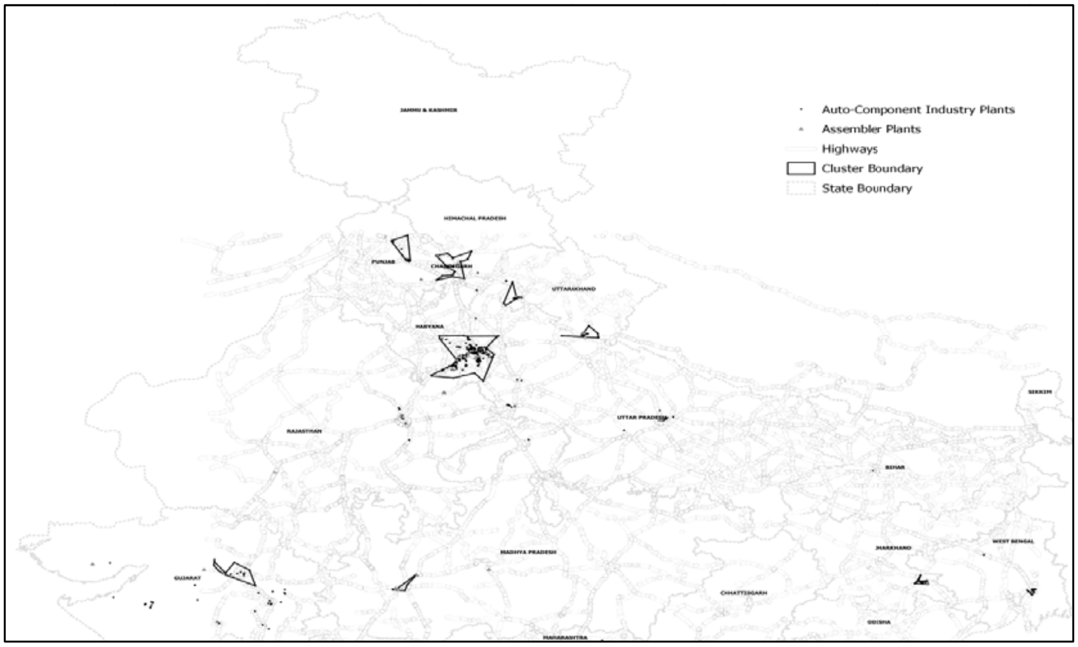
Keywords: Auto-Component Industry, Industrial Clusters.

1. Introduction

In a previous paper [Andhale and Rath, 2022a], we analysed the clustering of auto-component industry based on data given in the Automobile Component Manufactures Association (ACMA) directory, and identified sixteen auto-component clusters in India using DBSCAN methodology. The analysis revealed that nine clusters were spread across the northern region and the remaining seven clusters spread across the southern region of India. The development in the northern region has been dominated by a single large cluster near Gurgaon (Map 1 given below) whereas the southern region has three big clusters, namely, Pune-Mumbai, Chennai and Bangalore (Map 2 given below). It was

* Ashish Andhale is Ph.D. student at Gokhale Institute of Politics and Economics.
Acknowledgements: The author would like to thank the Automobile Component Manufacturers Association, India for making available their membership directory, without which this analysis would not have been possible. Author expresses his sincere gratitude to Dr Sharadini Rath for her support and valuable & constructive suggestions during this research work. He gratefully acknowledge Mr Abhay Tilak, Dr Nilakantha Rath and Dr Naresh Bodkhe for their useful comments & suggestions on the manuscript. He also thankful to Mr Ujaval Gandhi for his help in the use of geographical tools.

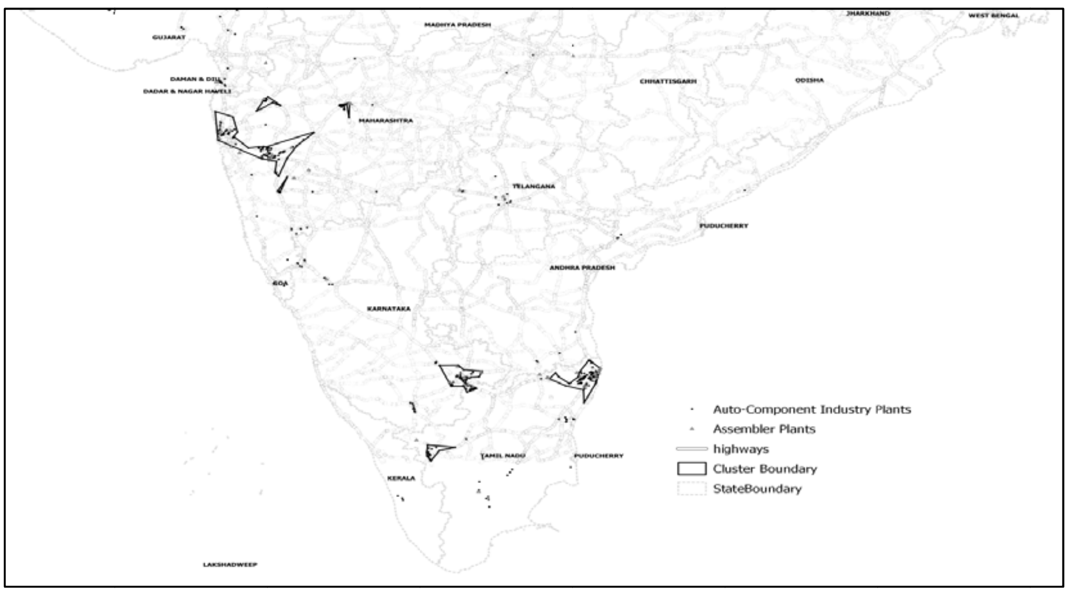
Map 1. Auto-Component Industry Clusters in the Northern Region of India



Source: Author's calculations.

also shown in [Andhale and Rath, 2022b] that clusters in the southern region tend to be geographically more dispersed with larger peripheries than in the northern region. We also showed that there is a difference between large and small clusters in their profile of firm size in terms of employment generation. Large clusters tend to have fewer large firms and a greater proportion of medium and small firms (Gurgaon, Pune-Mumbai and Chennai). Smaller clusters appear to have more large firms and a smaller proportion of medium and small firms (Coimbatore and Satara in the south and Ahmedabad, Solan, and Ludhiana in the north). Most of the non-cluster firms in the south are large with a small proportion being medium and small ones. In the north this feature is reversed, with a vast majority of non-cluster firms being medium and small. So, location choices of large firms are, thus, quite different in the north and south, with large firms possibly spearheading the dispersal of industrial location in the south. The south has the advantage of large firms creating a greater depth of ancillaries and also a more dispersed location choice. In terms of profile of clusters and in terms of the products they manufacture, a combination of high value engine parts and low value other miscellaneous parts make up about half the share of products in most clusters. Both these vendors have a larger presence in the south. The north has a larger share of vendors of body parts, drive-transmission-steering parts and electrical parts. There are cluster-wise differences in product profile of firms but they are small. Most clusters have some proportion of all product vendors.

Map 2. Auto-Component Industry Clusters in the Southern Region of India



Source: Author's calculations.

Given this background of these two studies regarding the auto component industry clusters, we now attempt to understand the performance of these clusters. This study is, thus, an extension of these earlier studies.

In order to understand the performance of these clusters, Annual Survey of Industries (ASI) data on employment, gross value added (GVA) and productivity parameters can be used. As pointed out in Andhale and Rath, 2022a there is a hierarchy of plants in the automobile vendor sector. Factories that supply parts directly to automobile assemblers are classified as being part of Tier-I while those in Tier-II and III are plants that supply intermediate inputs to Tier-I and Tier-II plants respectively. ACMA [2016] directory stated that all listed firms in their directory supply their output to at least one assembler. Therefore, we classify and assume the firms listed in the ACMA directory and their plants as Tier-I suppliers. The ASI data framework includes all those registered establishments which employ 10 or more workers with electricity connection or 20 or more without electricity without connection in accordance with the Section 2m (i) and 2m (ii) of the Factories Act, 1948 [GOI, 1948]. ASI data is, thus, available for plants in all Tiers and at the State level too.

For the purpose of the analysis here, we assumed that all information regarding this sector in the ASI for a particular state be applied to the clusters in the state. So, the ASI data for Tamil Nadu will be applied to the Chennai and Coimbatore clusters together. In short, we assess cluster performance by identifying clusters in a particular State. Table 1 below gives the details. The sixteen clusters are therefore grouped under eleven states for the analysis with ASI data.

Table 1 shows that the largest state cluster is in Tamil Nadu, with 1384 factories. The next is in Maharashtra, with 1331 factories. Both these are single state clusters. The Delhi cluster is in 3rd position, with 1142 factories across 4 states, UP, Delhi, Haryana and Rajasthan. For the ACMA Tier I plants, as shown in Andhale and Rath [2022b], this order was different. The 4 states of Delhi, Haryana, UP and Rajasthan together had 569 plants, the largest number. Maharashtra had 389 plants and Tamil Nadu 249. We also showed in that paper that the south region had a larger multiplier between ACMA Tier I plants and ASI factories of Tier II & III. In the south this multiplier was 3.20, while in the north it was 1.63. This effect can be seen in the changed ordering of the state clusters in Table 1, with Tamil Nadu the largest, Maharashtra 2nd and the Delhi state cluster 3rd.

Table 1. Conversion of DBSCAN Clusters to State Clusters

Conversion of DBSCAN Clusters to State Clusters							
State Cluster No	Name of State Cluster	Combination of DBSCAN Clusters	States	Cluster Plants	Non-cluster Plants	ACMA Plant Total	ASI Factories 2016-17
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	Delhi	Gurgaon	Uttar Pradesh, Delhi, Haryana, Rajasthan	528	33	561	1142
2	Uttarakhand	Udham Singh Nagar, Haridwar	Uttarakhand	99	2	101	298
3	Punjab	Ludhiana	Punjab	55	0	55	240
4	Himachal Pradesh	Solan	Himachal Pradesh	15	1	16	83
5	Jharkhand	Jamshedpur	Jharkhand	37	0	37	147
6	West Bengal	Kolkata	West Bengal	18	1	19	43
7	Gujarat	Ahmadabad	Gujarat, Daman and Diu, Dadra Nagar Haveli	28	48	76	201
8	Madhya Pradesh	Dhar	Madhya Pradesh	20	2	22	156
9	Maharashtra	Pune, Auran gabad, Nashik, Satara	Maharashtra	368	21	389	1331
10	Karnataka	Bangalore	Karnataka	98	22	120	498
11	Tamil Nadu	Chennai, Coimbatore	Tamil Nadu	234	15	249	1384

Source: Author’s calculations

Analysis of critical financial-economic ratios and aggregates calculated on the basis of unit-specific information in ASI data sets. The objectives of the study are: a) To Measure overall partial productivity and resource use efficiency of the auto component industry, b) To Measure different input intensities of the auto component industry, c) To measure auto component sector’s financial parameters such as debt rate and profit rate, and d) To measure partial and total factor productivity of the auto component industry.

Such an analysis was carried out for the years 2004-05 and 2016-17, at various levels, i.e., (i) All India level sectoral analysis of the auto component industry and (ii) "State Clusters" analysis of auto component industry, wherein state-level data is aggregated into data for eleven state clusters (Table 1), iii) the regional level analysis, with these eleven state clusters divided into Northern and Southern regions.

We have used Annual Surveys of Industry data, which has been sorted and formatted by the EPW Research Foundation [EPWRF, 2004-05 to 2016-17]. We selected this database as it offers a systematic and consistent dataset after carrying out the concordance for the different National Industrial Classifications (NICs) that have been used during the study period 2004-05 to 2016-17. Specific data in ASI is available by state and union territories based on two and three-digit NIC codes, including variables like the number of factories, number of workers, output and input, etc.

In the 2008 ASI survey [GOI, 2008], the auto component industry is included in the broad two-digit code 29 which refers to 'manufacture of motor vehicles, trailers and semi-trailers'. Within this, we use the three-digit sub-classification codes 292, which refers to 'manufacture of bodies (coachwork) for motor vehicles, manufacture of trailers and semi-trailers' and 293 which refers to 'manufacture of parts and accessories for motor vehicles' both directly related to the auto component industry. All industrial parameters with monetary values are deflated using the wholesale price index of manufacturing products GOI [2008-09 with 2004-05 as the base year].

After this elaborate introduction, the paper is further organized as follows-. Section 2 presents a brief analysis of the growth of auto-component industry at all India level, followed by a performance assessment of the sector by State Clusters in Section 3. Section 4 presents a summary and our concluding remarks.

2. Analysis of the growth of the Auto-Component Industry in India between 2004-05 and 2016-17

The automobile industry which is classified as 'The Manufacture of Motor Vehicles, Trailers and Semi-Trailers' in ASI survey witnessed the highest percentage rise of workers (193.54%) from 2004-05 to 2016-17. Also, this industry has added. the greatest number of formal sector manufacturing workers, next only to manufacture of chemical and chemical products (NIC-24) and the manufacture of Wearing Apparel Dressing and Dying of Fur (NIC-17) in the study period (see Table 2 below).

Table 2. Sector wise Distribution of Workers and Net Entry of Workers in Registered Manufacturing Industry in India

Two digit No	Industry	2004-05 (In 000's)	2016-17 (In 000's)	Net Entry of Workers: 2016-17 Over 2004-05	2004-05	2016-17	Percentage rise of workers: 2016-17 Over 2004-05 (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
17	Manufacture of Wearing Apparel Dressing and Dying of Fur	1258	1913	655	15.51	13.40	52.04
34	Manufacture of Motor Vehicles, Trailers and Semi- Trailers	336	986	650	4.14	6.91	193.54
26	Manufacture of Other Non-Metallic Mineral Products	516	1066	550	6.36	7.47	106.74
15	Manufacture of Food Products and Beverages	1326	1838	512	16.35	12.88	38.61
29	Manufacture of Machinery and Equipments N.E.C	434	846	412	5.35	5.92	94.85
27	Manufacture of Basic Metals	575	984	410	7.08	6.90	71.35
25	Manufacture of Rubber and Plastic Products	301	697	397	3.71	4.88	131.81
28	Manufacture of Fabricated Metal Products, Except Machinery and Equipments	315	705	390	3.89	4.94	123.67
18	Manufacture of Textiles	449	777	328	5.53	5.44	73.20
31	Manufacture of Electrical Machinery and Apparatus N.E.C.	236	490	254	2.91	3.43	107.67
19	Tanning and Dressing of Leather Manufacture of Luggage, Handbags, Saddlery, Harness and Footwear	149	392	243	1.83	2.74	163.55
36	Manufacture of Furniture; Manufacturing N.E.C.	175	360	185	2.16	2.52	106.07
35	Manufacture of Other Transport Equipment	184	321	137	2.27	2.25	74.30
21	Manufacture of Paper and Paper Products	176	251	75	2.17	1.76	42.52
23	Manufacture of Coke, Refined Petroleum Products and Nuclear Fuel	79	154	75	0.98	1.08	94.52
22	Publishing, Printing and Reproduction of Recorded Media	114	184	70	1.40	1.29	61.77
33	Manufacture of Medical, Precision and Optical Instruments, Watches and Clocks	62	117	56	0.76	0.82	90.36
32	Manufacture of Radio, Television and Communication Equipments and Apparatus	102	149	47	1.25	1.04	46.14
20	Manufacture of Wood and Products of Wood and Cork, Except Furniture, Manufacture of Articles of Straw and Plating Materials	49	90	41	0.60	0.63	84.92
15	Manufacture of Tobacco Products	471	466	-4	5.80	3.27	-0.91
30	Manufacture of Office, Accounting and Computing Machinery	26	20	-6	0.32	0.14	-21.50
	Total	8113	14276	6164	100	100	76

Source: EPWRF, [2004-05 to 2016-17].

Hence, the industry played a vital role in employment generation and is considered as a sunrise manufacturing sector for future growth. At three-digit level, we find that the share of workers in the automobile sector as a whole contributed by the auto component industry increased from 76.26 per cent in 2004-05 to 79.02 percent in 2016-17. Motor vehicle manufacturers (Assemblers) contributed only 21 percent share in 2016-17. Workers in the component industry are 3.76 times more than the assembler manufacturers (Table 3).

This shows that a growth in the large supply chain of the automobile industry and development of clusters are primarily driven by the auto-component industry in India. The industry's employment increases due mainly to growth in demand for automobiles.

Table 3. Distribution of Workers and net Entry of Workers in Automobile Industry in India

Three digit No	Three digit Automobile Industry	2004-05	2016-17	Net Entry of Workers: 2016-17 Over 2004-05	2004-05	2016-17	Percentage rise of workers: 2016-17 Over 2004-05 (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
341	341 - Manufacture of Motor vehicles	79746	206946	127200	23.74	20.98	159.51
342	342 - Manufacture of Bodies (coach work) for motor vehicles; manufacture of trailers and semi-trailers	22453	59517	37064	6.68	6.04	165.07
343	343 - Manufacture of Parts and accessories for motor vehicles and their engines [brakes, gear boxes, axles, road wheels, suspension shock absorbers, radiators, silencers, exhaust pipes, steering wheels, steering columns and steering boxes and other parts and accessories n.e.c.]	233764	719722	485958	69.58	72.98	207.88
	Auto Component_ Total (342+343)	256217	779239	523022	76.26	79.02	204.13
	Automobile Industry Total	335963	986185	650222	100.00	100.00	193.54

Source: EPWRF, [2004-05 to 2016-17]

The auto-component industry at the all-India level witnessed rapid positive growth as total output in real terms for this sector increased by nearly 110 percent from Rs. 1,344 Lakh in 2004-05 to Rs. 2,847 Lakh in 2016-17, registering an average annual growth rate of 6.45 percent. Gross value added (GVA) doubled from Rs. 308 Lakh in 2004-05 to Rs. 616 Lakh in 2016-17, growing at an annual growth rate of 5.95 percent. Also, the number of registered factories nearly doubled from 2,969 in 2004-05 to 5,775 in 2016-17 and the number of employees nearly tripled during 2004-05 and 2016-17. This buoyant growth in the component sector happened due to expansion of the domestic market and policy promotion of exports to external replacement market [Uchikawa and Roy, 2013]. On the other hand, GVA to output ratio of the industry decreased marginally from 22.97 percent in 2004-05 to 21.69 percent in 2016-17, thus reflecting concern about the overall productivity of the auto component industry. Resource efficiency also decreased marginally with GVA to input ratio decreased from 29.81 percent in 2004-05 to 27.70 percent in 2016-17 (Table 4).

With regards to factor payments, the average (real) wage per worker in the organized auto component industry grew at an annual growth rate of 4.88 per cent. Debt rate (as defined in Table 5 below) registered a marginal decrease from 28.62 per cent to 25.27 per cent thereby implying a lower dependency on borrowed capital for financing the investment over the years. The competition in the auto component industry became severe after the 2000s with new small enterprises facing difficulty to compete with existing companies [Uchikawa and Roy, 2013], resulting in an overall significant decrease in profit rate at the all-India level from 22.24 per cent in 2004-05 to 14.61 per cent in 2016-17 with an average annual negative growth rate of -3.44 per cent.

Table 4. Critical Aggregates and Derived Ratios of Auto Component Industry in India (2004-05 to 2016-17)

(At 2004-05 prices)

Variables	2004 - 2005	2016 - 2017	CAGR
(1)	(2)	(3)	(4)
Number of Factories	2969	5775	5.70
Number of Employees	25621	7779239	9.71
Fixed Capital / No of factories	406.92	957.56	7.39
Working Capital /No of factories (Rs. Lakhs)	125.95	272.16	6.63
Physical Working Capital /No of factories (Rs. Lakhs)	156.86	340.06	6.66
Invested Capital /No of factories (Rs. Lakhs)	563.78	1297.63	7.19
Gross Value Added / No of Factories (Rs. Lakhs)	308.67	617.57	5.95
Output/ No of Factories (Rs. Lakhs)	1344.03	2847.06	6.45
Total Input / No of Factories (Rs. Lakhs)	1035.37	2229.49	6.60
Profits / No of Factories (Rs. Lakhs)	121.76	167.40	2.69
Wages / No of Factories (Rs. Lakhs)	44.04	117.15	8.49
Gross Capital Formation / No of Factories (Rs. Lakhs)	156.24	180.36	1.20
GVA/Input (%)	29.81	27.70	-0.61
GVA/Output (%)	22.97	21.69	-0.47
Ratio to NVA			
Total Emoluments	28.52	39.47	2.75
Rent paid	2.00	1.80	-0.92
Interest paid	5.24	7.46	2.99
Profits	28.28	17.03	-4.14
Factor Remuneration rates			
Wages/Workers (In Lakhs)	1.02	1.81	4.88
Debt Rate (%)	28.62	25.27	-1.03
Profit rate (%)	22.24	14.61	-3.44
Factor Intensity			
Workers/Output	0.06	0.05	-2.50
Workers/Plant	86.30	134.93	3.80
Total Capital/ Output	0.42	0.46	0.69
Fixed capital /Output	0.30	0.34	0.88
Fixed capital /Workers	4.72	7.10	3.47
Fixed Capital/ Total Capital (%)	72.18	73.79	0.18
Fuel Consumed/ Workers	0.54	0.59	0.71
Fuel Consumed/ Fixed Capital (%)	11.49	8.32	-2.66
Partial Factor Productivity			
GVA/ Workers	3.58	4.58	2.08
Output/ Worker	15.57	21.10	2.56
GVA/ Fixed Capital	0.76	0.64	-1.34
Output/Fixed Capital	3.30	2.97	-0.87

Source: EPWRF, [2004-05 to 2016-17].

Table 5. Definition of Debt Rate and Profit rates

1) DEBT RATE = OUTSTANDING LOANS / (INVESTED CAPITAL + CURRENT ASSETS)

Where **OUTSTANDING LOANS** represent all loans whether short term or long term, whether interest bearing or not, outstanding according to the books of the factory as on the closing day of the accounting year;

INVESTED CAPITAL is the total of fixed capital and physical working capital;

CURRENT ASSETS is the sum total of raw materials & components, packaging materials, fuels and lubricants, spares, stores and others, semi-finished goods/ work in progress, finished goods, cash in hand and at bank, sundry debtors and other current assets.

2) PROFIT RATE= PROFITS / (INVESTED CAPITAL + CURRENT ASSETS - OUTSTANDING LOANS)

The industry also witnessed a marginal decline in partial capital productivity (measured in output per unit capital and GVA per unit capital). On the other hand, workers' partial productivity measured in output per unit of workers and GVA per unit of workers grew with the average annual growth rate of 2.08 percent and 2.56 percent, respectively. (Table 3 above).

3. Performance assessment of Auto-Component Industry by State Clusters

In this Section, we report the results of an analysis of critical financial-economic ratios and aggregates calculated on the basis of unit specific information in ASI data sets. The objective is to derive the following:

a. Measures of overall partial productivity and resource use efficiency.
b. Measures of different input intensities
c. Measures of basic key economic parameters, namely factories and workers
d. Measures of sector's financial parameters such as Debt rate and Profit rate
e. Measures of partial factor productivity and total factor productivity
f. Some other ratios which provide a summary measure of the overall business environment and sectoral efficiency.

The analysis for these measures was carried out for the years 2004-05 and 2016-17 at (i) All India level, i.e., sectoral analysis of the auto component industry and (ii) analysis of "State Clusters" of the auto component industry wherein state-level data is aggregated into data for eleven state clusters (Table 1) and iii) analysis at the regional level analysis. The eleven state clusters were divided into those in the northern region and those in the southern region.

3.a Cluster wise analysis of Key indicators of industry

At the all-India level, the number of registered factories in clusters doubled from 2,814 in 2004-05 to 5,523 in 2016-17, with an average annual growth rate of 5.78 percent. However, this addition in factories differs significantly across state clusters. Tamil Nadu cluster from the southern region attracted the maximum number of factories and workers in this period (Table 6), 820 of the total addition 2,709 in 2004-05 to 2016-17 (30.26 percent) in the country. Its share of factories to total factories increased from 20 per cent in 2004-05 to 25 per cent in 2016-17. Next to this, State Cluster 9 (Maharashtra) added 555 factories of total addition of 2,709 factories (20.48 percent). Even though Maharashtra's share of factories to total factories decreases from 27.50 percent in 2004-05 to 24 percent in 2016-17, these two southern clusters (Maharashtra and Tamil Nadu) continue to contribute 50 percent of total factories in this period.

On the other hand, in the northern region, State Cluster 1 (Delhi) is one of the prominent and larger clusters which added only 445 factories during this period (16.42 percent). Its share of factories in the all clusters total decreased by 4 percent by 2016-17. The other small clusters added some factories Uttarakhand (10.89%), Punjab (2.47%), Himachal Pradesh (2.58%), Gujarat (3.32%), and Madhya Pradesh (3.14%) during the period 2004-05 to 2016-17 (Table 6). Only Uttarakhand state cluster managed to improve its net share of factories significantly going from 3 to 298 factories over this period. In contrast, all other northern clusters lost their net share of factories in 2016-17 except the marginal increase in Himachal Pradesh and Madhya Pradesh State clusters. We also see the number of factories in the West Bengal cluster decrease during this period (Table 3).

Similarly, in the case of employment, all clusters in the northern region lost share of its workers to total workers in all clusters in 2016-17 compared to 2004-05, except for Uttarakhand state cluster, going from 100 to 20102 workers. On

Table 6. Cluster wise Analysis of Factories and Workers in Auto Component Industry

State Cluster No	State Clusters	Factories	% Share of total Factories	Factories	% Share of total Factories	Workers	% Share of total Workers	Workers	% Share of total Workers	Net Addition of factories 2016- 17 over 2004-05	Net Addition of workers 2016- 17 over 2004-05	Net change in share of factories: 2016-17 over 2004-05 (%)	Net change share of workers: 2016-17 over 2004-05 (%)
		2004-05		2016-17		2004-05		2016-17					
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1	Delhi	697	24.77	1142	20.68	56817	28.07	133098	22.96	445	76281	-4.09	-5.11
2	Uttarakhand	3	0.11	298	5.40	100	0.05	20102	3.47	295	20002	5.29	3.42
3	Punjab	173	6.15	240	4.35	8147	4.02	17974	3.10	67	9827	-1.80	-0.92
4	Himachal Pradesh	3	0.46	83	1.50	460	0.23	5650	0.97	70	5190	1.04	0.75
5	Jharkhand	97	3.45	147	2.66	6531	3.23	14245	2.46	50	7714	-0.79	-0.77
6	West Bengal	52	1.85	43	0.78	5600	2.77	761	0.13	-9	-4839	-1.07	-2.64
7	Gujarat	111	3.94	201	3.64	5763	2.85	12680	2.19	90	6917	-0.31	-0.66
8	Madhya Pradesh	71	2.52	156	2.82	6436	3.18	13233	2.28	85	6797	0.30	-0.90
9	Maharashtra	776	27.58	1331	24.10	38351	18.95	116899	20.16	555	78548	-3.48	1.22
10	Karnataka	257	9.13	498	9.02	19323	9.55	50457	8.70	241	31134	-0.12	-0.84
11	Tamil Nadu	564	20.04	1384	25.06	46413	22.93	160891	27.75	820	114478	5.02	4.82
12	All India	2814	100	5523	100	202414	100.00	579777	100.00	2709	377363	0.00	0.00
13	Northern Clusters	1217	43.25	2310	41.83	89854	44.39	217743	37.56	1093	127889	-1.42	-6.83
14	Southern Clusters	1597	56.75	3213	58.17	112560	55.61	362034	62.44	1616	249474	1.42	6.83

Source: EPWRF, [2004-05 and 2016-17]

the other hand, Tamil Nadu and Maharashtra state clusters showed a positive net change in share of workers in 2016-17 compared to 2004-05. Together both the clusters employed 48 percent share of total workers in 2016-17. Also, the net share of workers of the southern clusters increased by 6.83 percent in the study period (Table 3). The geographical analysis of clusters in Paper-1 shows that southern clusters are larger in geographic size compared to the northern region. This analysis shows that the southern region contributes more factories and more employment generation than the northern region. It is possible that the new factories located more in the southern region than in the north, because south consists of old clusters like Pune-Mumbai and Chennai that have more domestic and foreign assemblers, more tier two and tier three vendors and more technical and infrastructural facilities than the northern regions.

3.b Cluster-wise analysis of Gross value added (GVA)

At all India level, the real GVA estimate for the industry increased from Rs. 11,01,627 lakh in 2004-05 to Rs. 35,25,530 lakh in 2016-17, registering an average annual growth rate of 10.18 percent. This buoyant growth rate of the industry contributes from all the state clusters. In addition, the cluster-wise contribution of GVA is highly correlated with its share in factories. Among the northern region, Delhi cluster which accounts 20.67 percent share of total factories contributed 23.78 percent share of GVA in 2016-17 but showed the lowest average annual GVA growth rate of 5.13 percent from 2004-05 to 2016-17. Uttarakhand accounts for only 5.40 percent share of total factories, but shows the highest average annual GVA growth rate among all clusters with 80.80 percent.

In Uttarakhand, this significant growth rate in GVA and factories and employment results from the central government's policy provision in the state. In 2003, the Department of Industrial Policy & Promotion (DIPP) of the Central Government of India implemented the "Scheme of Package for Special Category States" in the hilly regions of Uttarakhand and Himachal Pradesh [GOI, 2003]. Hundred per cent excise duty exemption on outright basis to new industrial units and the existing units on substantial expansion, and 100% income tax exemption for five years and thereafter 30% for companies and 25% for other than companies for a further period of 5 years were provided under the scheme. As a result, many of the assemblers and auto-component units shifted to the small industrial areas mainly in

Udham Singh Nagar district of Uttarakhand. Himachal Pradesh has not seen such growth in the auto-component sector.

All other northern clusters, including Punjab, Himachal Pradesh, Jharkhand, West Bengal, Gujarat and Madhya Pradesh, show low CAGR of GVA compared to Uttarakhand, but the lowest was in the Delhi cluster (Table 7). In the southern region, Maharashtra, Tamil Nadu, Karnataka showed impressive CAGR of GVA with 13.34, 11.57 and 10.82 percent respectively. In 2016-17, the contribution of southern region clusters to GVA is 1.61 times that of northern region clusters.

Table 7. Cluster-wise Gross Value Added in Auto-Component Industry (2004-05-2016-17)

State Cluster No	State Clusters	GVA (Rs. Lakhs) 2004-05	GVA (Rs. Lakhs) 2016-17	CAGR (%)	GVA /Output (%) 2004-05	GVA /Output (%) 2016-17	CAGR (%)	GVA/ Input (%) 2004-05	GVA/ Input (%) 2016-17	CAGR (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1	Delhi	459917	838431	5.13	19.23	21.29	0.85	23.80	27.05	1.07
2	Uttarakhand	175	213579	80.80	25.25	22.19	-1.07	33.78	28.51	-1.40
3	Punjab	22669	86825	11.84	24.21	24.04	-0.06	31.95	31.64	-0.08
4	Himachal Pradesh	3130	26867	19.62	39.85	24.10	-4.10	66.24	31.75	-5.94
5	Jharkhand	31526	82553	8.35	32.78	18.44	-4.68	48.77	22.62	-6.20
6	West Bengal	2379	5086	6.54	6.24	26.15	12.69	6.65	35.40	14.95
7	Gujarat	14140	56539	12.24	19.82	22.98	1.24	24.72	29.83	1.58
8	Madhya Pradesh	19493	40217	6.22	21.38	23.18	0.68	27.19	30.17	0.87
9	Maharashtra	210916	948085	13.34	21.10	24.81	1.36	26.75	32.99	1.76
10	Karnataka	94801	325352	10.82	26.44	23.33	-1.04	35.94	30.42	-1.38
11	Tamil Nadu	242481	901996	11.57	25.45	18.42	-2.66	34.14	22.57	-3.39
	ALL INDIA (Clusters)	1101627	3525530	10.18	21.59	21.53	-0.02	27.54	27.44	-0.03
	Southern Clusters	548198	2175432	12.17	23.72	21.51	-0.81	31.10	27.40	-1.05
	Northern Clusters	553429	1350097	7.71	19.83	21.57	0.70	24.73	27.50	0.89

Source: EPWRF, [2004-05 and 2016-17]

In 2004-05, except for West Bengal, all other state clusters showed impressive double-digit figures for GVA per unit of output (in %) and GVA per unit of input (in %). There is not much difference across clusters in these two indicators. This picture remains the same over the years with a very CAGR in all clusters (Table 4). This shows that clusters have impressive overall partial productivity and resource efficiency but the growth rate of these indicators across all the clusters remained low over the years.

3.c Cluster wise analysis of Input Cost

At all India level, the sector's average annual wages per worker increased from Rs. 1.11 Lakh in 2004-05 to Rs. 1.78 lakh in 2016-17, registering an average annual growth rate of 4.03 percent only. Similarly, the average wage per worker increased marginally in all clusters under consideration. Northern region clusters witness the most increase in CAGR of wages per worker, Uttarakhand jumped up by 12.68 % and Himachal Pradesh by 6.55% per annum (Table 8 given below). While Delhi shows the lowest CAGR of wages per worker at 2.87%, other clusters witnessed a marginal increase in the range of 3-6 percent. In the industry, there is a reliance on employing capital than labour, which explains the low wages and its low growth in the industry.

The debt rate is an indicator of the level of indebtedness of the cluster. It is defined as the ratio of outstanding loans to the sum of the invested capital and current assets [Singhal and Saksena, 2017, Pp. 15-40]. Lower the debt rate, the better it is for an organization implying lower dependence on borrowed capital for financing the investment. Interestingly, from 2004-05 to 2016-17 the debt rate has increased in all the clusters except the marginal decline in Punjab and West Bengal (Table 5). At the all-India level, there is a small decline in the debt rate. However, that analysis includes other states that are not part of the state clusters. It is possible that medium and small auto-component industry in these states has contributed to this difference in debt rates in these two cases.

Profit rate is the ratio of profits to the sum of invested capital and current assets less outstanding loans [Singhal and Saksena, 2017]. The profit rate declined over the years in all clusters except for the major increase in Uttarakhand with 14 percent CAGR (Table 5). Also, Maharashtra has improved its profit rate from 34.75 to 38.34 percent with a marginal CAGR 0.82 percent. But, profit rate in clusters like Himachal Pradesh and West Bengal has fallen drastically (Table 5). The annual growth rate of profit is negative across clusters, similar to the all-India level, indicating that the return on owned capital decreased over the years.

Table 8. Cluster wise Analysis of Returns to Factor Inputs in Auto-Component Industry (2004-05-2016-17)

State Cluster No	State Clusters	Wages/ worker (Rs) 2004-05	Wages/ worker (Rs) 2016-17	CAGR %	Debt Rate 2004-05	Debt Rate 2016-17	CAGR %	Profit Rate 2004-05	Profit Rate 2016-17	CAGR %
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1	Delhi	1.22	1.71	2.87	18.68	29.55	3.90	34.80	15.62	-6.46
2	Uttarakhand	0.26	1.10	12.68	2.22	20.46	20.32	10.85	52.19	13.99
3	Punjab	0.89	1.51	4.47	42.59	34.00	-1.86	17.21	15.51	-0.86
4	Himachal Pradesh	0.60	1.29	6.55	3.74	174.19	37.72	90.64	-27.85	
5	Jharkhand	0.98	1.87	5.50	20.02	30.80	3.66	34.38	5.93	-13.62
6	West Bengal	0.83	1.66	5.97	297.68	112.80	-7.77	35.02	-74.43	
7	Gujarat	0.60	1.26	6.36	32.04	63.36	5.85	31.93	27.76	-1.16
8	Madhya Pradesh	0.80	1.12	2.93	45.34	57.53	2.00	21.19	17.02	-1.81
9	Maharashtra	1.06	2.05	5.63	31.94	50.04	3.81	34.75	38.34	0.82
10	Karnataka	1.36	2.24	4.28	21.43	36.94	4.64	20.21	19.36	-0.36
11	Tamil Nadu	1.12	1.79	4.04	25.09	42.00	4.39	29.51	13.94	-6.06
12	ALL INDIA (Clusters)	1.11	1.78	4.03	26.06	40.64	3.77	31.69	21.97	-3.01
	Southern Clusters	1.14	1.96	4.59	27.11	44.57	4.23	29.67	23.48	-1.93
	Northern Clusters	1.07	1.54	3.11	25.00	34.80	2.80	33.68	20.07	-4.22

Source: EPWRF, [2004-05 and 2016-17]

The debt rate is an indicator of the level of indebtedness of the cluster. It is defined as the ratio of outstanding loans to the sum of the invested capital and current assets [Singhal and Saksena, 2017, Pp. 15-40]. Lower the debt rate, the better it is for an organization implying lower dependence on borrowed capital for financing the investment. Interestingly, from 2004-05 to 2016-17 the debt rate has increased in all the clusters except the marginal decline in Punjab and

West Bengal (Table 5). At the all-India level, there is a small decline in the debt rate. However, that analysis includes other states that are not part of the state clusters. It is possible that medium and small auto-component industry in these states has contributed to this difference in debt rates in these two cases.

Profit rate is the ratio of profits to the sum of invested capital and current assets less outstanding loans [Singhal and Saksena, 2017]. The profit rate declined over the years in all clusters except for the major increase in Uttarakhand with 14 percent CAGR (Table 5). Also, Maharashtra has improved its profit rate from 34.75 to 38.34 percent with a marginal CAGR 0.82 percent. But, profit rate in clusters like Himachal Pradesh and West Bengal has fallen drastically (Table 5). The annual growth rate of profit is negative across clusters, similar to the all-India level, indicating that the return on owned capital decreased over the years.

Heavy requirement of capital for survival in a competitive market and the dominance of a few large tier one auto-component firms and assemblers seem to be factors that explain the deterioration in profit rate across clusters. There is evidence based on an interview with a regional deputy executive of the ACMA (30th November 2019) that vendors are part of the pan-India supply chains Major vendor firms supply parts to assemblers far outside their own clusters. So, there is likely to be intense competition resulting in increasing debt and declining profit for the sector as a whole.

3.d Cluster wise Analysis of Partial Factor Productivity

In the following analysis, measures of 'partial factor productivity' are defined in terms of labour (measured in output or GVA per worker) and capital (measured in terms of output or GVA per unit of capital). Over the years, figures of partial productivity of labour show an increasing trend in all the clusters except a marginal decline in Delhi and Himachal Pradesh (Table 9), which results in its overall decline in northern region. On the other hand, southern clusters show a positive CAGR in both indicators of labour productivity. Presence of more technical training and research institutions, along with more Tier-I and II factories explain higher labour productivity in the south. In 2004-05, indicators of capital productivity show phenomenal figures across all the clusters, but fail to maintain this trend by 2016-17 in all clusters except Uttarakhand. Uttarakhand indicates a high CAGR of 8.11 percent of GVA per fixed capital from 2004-05 to 2016-17. This is because the cluster in this state is very new, with factories increasing from 3 to 298 over this period. A large amount of capital expenditure has taken place over this period, with a very fast process of industrial production.

7DEOH&OXVWHUZLVH$QDO\VLVRI)DFWRU3URGXFWLY

6W Cluster No	6WDWH&	GVA/Worker (Rs. Lakhs) 2004-05	GVA/Worker (Rs. Lakhs) 2016-17	CAGR %	Output/ Worker (Rs. Lakhs) 2004-05	Output/ Worker (Rs. Lakhs) 2016-17	CAGR %	GVA/ Fixed Capital 2004-05	GVA/ Fixed Capital 2016-17	CAGR %	Output/ Fixed Capital 2004-05	Output/ Fixed Capital 2016-17	CAGR %
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
9 S	Maharashtra	4.07	6.09	3.42	19.29	24.56	2.03	78.73	47.17	-4.18	373.07	190.16	-5.46
6 N	West Bengal	0.30	5.03	26.33	4.88	19.24	12.11	30.24	42.32	2.84	484.96	161.84	-8.74
10 S	Karnataka	3.68	4.81	2.26	13.91	20.62	3.34	75.68	50.42	-3.33	286.23	216.14	-2.31
1 N	Delhi	6.59	4.60	-2.95	34.28	21.62	-3.77	83.13	38.14	-6.29	432.36	179.14	-7.08
11 S	Tamil Nadu	4.07	4.23	0.31	16.00	22.95	3.05	78.32	38.50	-5.75	307.74	209.08	-3.17
2 N	Uttarakhand	1.28	3.90	9.75	5.06	17.58	10.94	27.65	70.47	8.11	109.48	317.65	9.28
3 N	Punjab	2.14	3.90	5.12	8.85	16.23	5.18	72.84	36.56	-5.58	300.85	152.11	-5.53
5 N	Jharkhand	3.70	3.79	0.18	11.30	20.53	5.17	6.92	25.90	-8.67	234.62	140.41	-4.19
4 N	Himachal Pradesh	4.94	3.76	-2.23	12.39	15.62	1.95	200.38	47.58	-11.29	502.88	197.43	-7.50
7 N	Gujarat	2.01	3.35	4.38	10.12	14.60	3.10	108.83	34.99	-9.02	549.00	152.25	-10.14
8 N	Madhya Pradesh	2.26	2.51	0.87	10.57	10.82	0.19	67.73	27.06	-7.36	316.79	116.73	-7.98
13 N	Southern Clusters	4.00	4.98	1.85	16.85	23.16	2.69	78.00	43.53	-4.74	328.81	202.38	-3.96
14 S	Northern Clusters	4.89	4.19	-1.28	24.67	19.44	-1.96	81.72	39.29	-5.92	412.15	182.17	-6.58
12 T	ALL INDIA (Clusters)	4.40	4.65	0.45	20.38	21.58	0.48	79.83	41.80	-5.25	369.71	194.15	-5.23

Sorted by GVA/Worker (Rs. Lakhs) (2016-17)
Data Source: ASI, [2004-05 and 2016-17] Via EPWRF

Overall, the results show that there is no significant variation in labour and capital productivity trends across all the clusters. This shows that the industry is following similar patterns of deployment of capital and labour across clusters and regions for output targets.

3.e Cluster wise Analysis of Total Factor Productivity

Total Factor Productivity (TFP) is measured by a simple growth accounting method across clusters. This is to understand the differences among overall productivity in clusters. In the literature, productivity is measured by a production function which with capital and labor in the Cobb-Douglas [Cobb and Douglas, 1928, Pp. 139-165] form is formulated as,

$$Y = A + K^{\alpha} + L^{\beta}$$

The equation represents total output (Y) as a function of total-factor productivity (A), capital input (K), labor input (L), and the two inputs respective shares of output (α and β are the share of contribution for K and L, respectively). In an analysis by Solow [1956, Pp. 65-94], A refers to technological improvement. However, another form of this equation was proposed by Goldar, B. [1986; 2006], Goldar and Kumari [2003], Singh [2016, Pp. 108-120] which used the growth accounting method to calculate TFP using industry data. Singh [2016; 2017, Pp. 821-1833] proposed a simple technique to calculate TFP in the growth accounting method. As the name suggests, growth accounting refers to growth in total factor productivity over time, and the equation is:

$$\%\Delta Y = \%\Delta A + aK \cdot \%\Delta K + aN \cdot \%\Delta N$$

In the equation $\%\Delta Y$ represents growth in output/year, $\%\Delta K$ is the growth in the capital stock/year, $\%\Delta N$ is termed as growth in the labor supply/year, $\%\Delta A$ is termed as growth in total factor productivity/year, aK is represented as $\%\Delta Y/\%\Delta K$ or elasticity of output with respect to capital (holding A and N fixed) and aN is termed as $\%\Delta Y/\%\Delta N$ or elasticity of output with respect to labor (holding A and K fixed). Our paper used the growth accounting technique (Solow Residual) because the objective is to estimate how much output, on average, is acquired from a set of inputs by an industry-focused data set. Table 10 summarizes the variables and data used to find out the TFP of auto-component clusters.

At the all-India level, total factor productivity calculated by this growth accounting method, shows -12.8% average annual negative growth from 2004-05 to 2016-17.

Table 10. TFP variables, definitions and data source

Variable	Definition	Deflator	Data Source
(1)	(2)	(3)	(4)
Output (Q)	Gross value added (GVA) to the firm	Deflated by industry-specific Wholesale Price indices (WPI)	1) GVA obtained from the Annual Survey of Industries (EPW) 2) WPI obtained from the Ministry of Commerce & Industry of India
Labor (N)	Total Workers		Total workers obtained Annual Survey of Industries (EPW)
Capital (K)	Fixed capital stock series constructed by perpetual inventory method.	Deflator is derived from the data on gross fixed capital formation in registered manufacturing at current and constant prices given in NAS	1) Net fixed capital formation obtained from Annual Survey of Industries (EPW). 2) Deflator obtained from Macroeconomic Aggregates.

Source: Singh (2017).

Table 11 shows mean growth rates of Inputs, Output and TFP from 2004-05 to 2016-17 for state clusters. While there is phenomenal growth in inputs and output in all the clusters (Tables 3, 4), productivity is not up to the mark. Total factor productivity in prominent clusters in the north like Delhi, Uttarakhand, Himachal Pradesh, Gujarat and Madhya Pradesh shows negative growth in TFP for the given period. Only Punjab and Jharkhand show low positive growth in TFP. Similarly, large southern clusters like Maharashtra, Chennai and Karnataka also show negative TFP growth for 2004-05 to 2016-17. The above analysis of TFP and partial factor productivity shows that overall, the industry successfully attracted resources and investment in all the clusters but failed on the productivity front. It is clear that in spite of reaching full production in a short time period, the Uttarakhand cluster has not been able to achieve significant growth in TFP. Policy driven cluster formation of this type creates employment, but remains low on the productivity front, compared to old and more mature clusters. It should be mentioned here that, with the introduction of the Goods and Services Tax (GST), the policy instrument for hilly regions has lapsed. It will be interesting to watch how industry reacts to this change and how the future trajectory of this cluster takes place.

Table 11. Average annual growth rate of Inputs, Output and TFP from 2004-05 to 2016-17

State Cluster No/ N-S	Clusters	Inputs Number of Workers (AAGR)	Stock NFCF (AAGR)	Output Gross Value Added (AAGR)	Total Factor Productivity AAGR
(1)	(2)	(3)	(4)	(5)	(6)
3 N	Punjab	7.91%	32.04%	29.62%	4.29%
5 N	Jharkhand	10.35%	59.38%	40.10%	2.49%
6 N	West Bengal	47.28%	51.38%	48.84%	1.20%
8 N	Madhya Pradesh	0.36%	25.38%	8.56%	-2.11%
11 S	Tamil Nadu	10.10%	26.39%	14.41%	-7.28%
10 S	Karnataka	3.23%	31.54%	6.03%	-10.69%
9 S	Maharashtra	1.93%	33.42%	6.79%	-11.56%
1 N	Delhi	8.14%	31.05%	10.55%	-13.38%
7 N	Gujarat	14.39%	65.45%	28.33%	-17.44%
4 N	Himachal Pradesh	27.09%	90.07%	45.66%	-18.59%
2 N	Uttarakhand	122.71%	296.75%	119.47%	-81.09%
12 N	North	7.16%	33.31%	9.61%	-12.76%
13 S	South	3.50%	30.59%	6.65%	-10.83%
14 T	All Clusters	4.88%	31.67%	6.62%	-12.82%

Source: Author's Calculations

3.f Analysis of Difference in Inter-regional (North & South) variation

We conduct t-tests to detect if there are "statistically significant" differences in the average performance of regions on certain critical ratios. A t-test can be used when comparing the means of two groups (a.k.a. pairwise comparison). The formula for the two-sample t test (the Student's t-test) [Student, 1908, Pp. 1-25] is shown below.

$$t = \frac{x_1 - x_2}{\sqrt{\left(s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)}}$$

In this formula, t is the t value, x_1 and x_2 are the means of the two groups being compared, $s2$ is the pooled standard error of the two groups, and n_1 and n_2 are the number of observations in each of the groups.

These tests were undertaken for the

year 2016-17 data. Results (Table 12) show that the basic indicators of successful industrialization such as number of factories and GVA vary significantly across the South and North region. The southern region has larger mean factories and mean GVA than north. Even though the differences in mean workers across regions are insignificant, mean partial productivity of labour varies greatly when measured in terms of output per unit of worker. This can be the result of better environment for workers in the southern region in terms of training and educational facilities and availability of urban amenities, with better infrastructure than the north. High labour productivity in south also gives higher returns to labour there. Variations in the mean returns to labour is significant across the regions when measured in terms of wages per unit of worker. The south performs better than the north.

Table 12. Analysis of Difference in Inter-regional (North & South) variation

Difference in the Mean of	Measure Of	Test Statistic (T-test)	p-value	Significant or not
(1)	(2)	(3)	(4)	(5)
1. Factories	No of Factories	2.499	0.044*	Yes
2. Workers	No of Workers	2.311	0.05*	Yes
3. GVA	GVA	2.493	0.044*	Yes
4. GVA /Output	Overall Partial Productivity	-0.292	0.396	No
5. GVA/ Input	Resource use efficiency	-0.282	0.398	No
6. GVA/ Fixed Capital	Partial Productivity of Capital	0.813	0.220	No
Output/Fixed Capital		1.207	0.131	No
7. GVA / Worker	Partial Productivity of Labour	1.986	0.050*	Yes
Output/ Worker		3.372	0.006**	Yes
8. Profit Rate	Returns to Owned Capital	1.280	0.116	No
9. Wages / Worker	Returns to Labour	3.588	0.008**	Yes
10. Debt Rate	Ability to Repay Loans	-1.169	0.138	No

*Significant at 5%; **Significant at 1%
Source: Author's Calculations

4. Summary and Concluding remarks

We have used the Annual Surveys of Industry data to link these clusters to the state level data for the sector, to study their performance, and to create regional insights. We have created groups of States that represent each of the clusters so that ASI data can be linked to their performance. We find that the auto-component industry performs well in period 2004-05 to 2016-17 in terms of critical parameters at the all-India level. However, an in-depth analysis using state groups to represent clusters shows inter-cluster differences in many parameters. Employment, number of factories and GVA doubled or more at the all-India level. However, the share of the north in employment and workers declined, and its GVA also grew at a slower pace than the south. Surprisingly, the variation in the mean values of some performance parameters for individual clusters, such as labour and capital productivity, wages per worker, profit and debt rate is not significant. This shows that there is industry standard of deployment of labour, capital for an expected output. All clusters operate to this basic model. On the other hand, regional level analysis shows that taken together, southern region performs better than the north in most of these. Overall, growth in total factor productivity was negative for most clusters, showing that the industry successfully attracted resources and investments in all clusters, but failed on the productivity front.

In 2003, Government of India introduced the Scheme of Package for Special Category States in the Hilly Regions [GOI, 2003]. It gave huge exemptions on excise duty and income tax for industries locating in those regions. This led directly to the creation of the clusters in Uttarakhand and Himachal Pradesh. This is the only instance of policy driven cluster formation that we could see in the course of our study. Clusters confer many advantages on those locating within them. However, given the poor productivity figures for clusters in this study (Uttarakhand also showed the lowest total factor productivity growth of all states/clusters), it is not clear that we would advocate a purely cluster -based policy induced initiatives to drive industrial growth.

REFERENCES

ACMA, 2016; Buyers Guide, Auto Component Manufacturers Association of India, New Delhi.

Andhale, A. and S. Rath, 2022a; Cluster Formation in the Auto Component Industry in India: Links to Urbanisation, *Journal of Indian School of Political Economy,* Vol. 34, Nos. 3&4.

Andhale, A. and S. Rath, 2022b; Cluster Profiles of Auto Component Industry in India, *Journal of Indian School of Political Economy,* Vol. 34, Nos. 3&4.

Cobb, C. W., P. H. Douglas, 1928; "A Theory of Production" (PDF), *American Economic Review,* 18 (Supplement).

EPWRF, 2004-05 to 2016-17; Database on the Annual Survey of Industries, *Economic and Political Weekly Research Foundation,* Mumbai.

GOI, 1948; Factories Act, 1948, Ministry of Labour and Employment, Government of India, New Delhi.

GOI, 2003; Scheme of Package for Special Category States, Department of Promotion of Industry and Internal Trade, Ministry of Commerce and Industry, Government of India, New Delhi.

GOI, 2008-09; Index Numbers of Wholesale Price in India (Base: 2004-05=100), Ministry of Commerce and Industry, Government of India, New Delhi.

GOI, 2008; *Annual Survey of Industries,* Central Statistics Office, Industrial Statistics Wing, Ministry of Statistics & Programme Implementation, Government of India, New Delhi.

Goldar, B., 1986; *Productivity Growth in Indian Industry,* Allied Publishers, New Delhi.

Goldar, B., 2006; *Productivity Growth in Indian Manufacturing in the 1980s and 1990s,* In: Tendulkar, S.D., Mitra, A., Narayanan, K. and Das, D.K., (Eds.), India: Industrialization in a Reforming Economy, Academic Publishers, New Delhi.

Goldar, B. and A. Kumari, 2003; Import Liberalization and Productivity Growth in Indian Manufacturing Industries in the 1990s, The Developing Economies.

Singh, A. P., 2016; Do Technology Spillovers Accelerate Performance of Firms? Unravelling a Puzzle from Indian Manufacturing Industry, *Economics and Applied Informatics,* 26.

Singh, A. P., 2017; "Keeping it simple: comparative analysis of TFP across manufacturing industries and major states of India", *Theoretical Economics Letters,* Vol. 7 No. 1.

Singhal, R. and S. Saksena, 2017; "Performance Assessment of the Storage and Warehousing Industry in India", *The Journal of Industrial Statistics,* Vol. 6, No. 1.

Solow, R., 1956; "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, Vol. 70, No. 1, Feb.

Student, 1908; The probable error of a mean, *Biometrika,* Vol. 6, No. 1.

Uchikawa, S. and S. Roy, 2013; ‘The development of auto component industry in India’, *Mimeograph.*

DOCUMENTATION

The purpose of this section is to make available to the readers official documents such as reports of committees, commissions, working groups, task forces, etc., appointed by various ministries, departments, agencies of central and state governments and international organisations, which are not readily accessible either because they are old, or because of the usual problems of acquiring governmental publications, or because they were printed but not published, or because they were not printed and remained in mimeographed form. We also present in this section, official documents compiled from scattered electronic and/or other sources for ready reference of the readers. It will be difficult and probably not worthwhile to publish the documents entirely. We shall publish only such parts of them as we think will interest our readers. The readers are requested to send their suggestions regarding official documents or parts thereof for inclusion in this section.

We are also keen to publish Papers, Notes or Comments based on the material included in this section. We invite the readers to contribute the same to our journal, which we shall consider for publication in subsequent issues of the journal, after the usual refereeing process.

In the present section, we publish:

1. Prof. V. M. Dandekar's articles published in 1965-1970.

2. Toward Modelling Poor Cities: A Review of Urban Economic and Planning Models, World Bank Staff Working Paper No. 232, April 1976

MINIMUM SUPPORT PRICES FOR FOODGRAINS: GUIDELINES FOR A POLICY AND A PROGRAMME*

V. M. DANDEKAR

1. The demand for a minimum support price for foodgrains comes from the producers of foodgrains. Naturally, in the minds of many, the demand for and the concept of minimum support price are closely associated with the notion of 'cost of production' of foodgrains. One of the reasons, possibly a minor one, for the failure to evolve a public policy and a programme in this field in the past has been the difficulty of knowing what the 'cost of production' meant in this context and an imperfect understanding of how it might be related, in concept and in operation, to a minimum support price. In this paper, we shall argue that 'cost of production' is only remotely connected to the notion of a minimum support price and that therefore an operationally meaningful policy and programme in this field can be evolved without an active reference to the notion of 'cost of production'.

2. We shall begin by distinguishing two rather different sets of circumstances under which the demand for a minimum support price is made. For convenience we shall refer to them as the short-term and the long-term considerations, respectively. In the short-term considerations, the inputs or costs incurred by a producer in the process of production are supposed to be given. Nevertheless, the output is variable due to fortuitous circumstances such as weather. The price which the produce fetches in the market affects immediately the well being of the producer. In the long-term considerations, the inputs or costs that a producer incurs in the process of production are variable and are determined by conscious production decisions by the producer in which his price expectations play an important part. The output is significantly related to the inputs. Therefore, the price that the out put fetches in any year influences the future production decisions of the producer and hence also the production. Both sets of circumstances lead to a demand for a minimum support price. However, the two are different and need to be kept distinct. We shall first examine the demand for a minimum support price under the shot-term considerations.

3. Tn a free market economy, the prices are determined by the conditions of supply and demand and the function of the market price is to equate the two. In

* Paper read at the Seminar on Fixation of Support Prices for Foodgrains at Vallabh Vidyanagar on 15-17 October 1965 organized by Sardar Vallabhbhai Vidyapeeth and Charutar Vidya Mandir.
Reprinted from: 'Artha Vijnana', Vol. 7, No. 4, December 1965, Gokhale Institute of Politics and Economics, Poona.

the short-run, which in relation to agricultural commodities is often called the 'stock-period' that is the period after the harvest has arrived on the market, the supply is given. The market price then is so determined that all the supply is taken up and no more is demanded. Thus during a year, the market price of food-grains is determined predominantly by the size of the harvest and the price elasticities of consumer demand for foodgrains. Because of generally low price elasticities of demand for foodgrains, it needs a disproportionately large rise in price in order to restrict the demand to a somewhat smaller harvest; and it needs a disproportionately large fall in price, to expand the demand to a somewhat large harvest. The market price of foodgrains is thus liable to large fluctuations from year to year, firstly because of the fluctuations in the size of the harvest and secondly because of the generally low price elasticities of demand for foodgrains.

4. From the standpoint of the producers, the fluctuations in the harvest are entirely fortuitous and given acreage sown under the drops and barring exceptionally bad seasons, the inputs or costs incurred by the producer are more or less independent of the harvest. This means that the costs of production per ton of the foodgrains are low in a year of good harvest and high in a year of bad harvest. Therefore, in the short-term considerations high prices in a year of bad harvest and low prices in a year of good harvest are not entirely contrary to the interests of the producers.

5. However, on account of the generally low price elasticities of consumer demand for foodgrains, the fluctuations in prices are likely to be more pronounced than warranted by the fluctuations in the harvest. Thus in a year of bad harvest, the rise in prices might more than compensate the fall in yield and the producers might actually reap a higher total revenue. On the other hand, in a year of good harvest, the prices might fall so low that in spite of the high yields, the producers might end up with a smaller total revenue. The short-term case for minimum support prices rests on this phenomenon and consequent need to support the prices in a year of good harvest.

6. Let us examine the point of view of the consumer on this question. Obviously there is little reason why in a year of good harvest, a consumer should pay a price higher than the one demanded by the market unless this helps him, in a year of bad harvest, to pay a price lower than that demanded by the market In fact, it does not help the consumer to have a low price and high consumption in a year of good harvest and a high price and low consumption in a year of bad harvest.

Therefore, he would prefer the price to be supported in a year of good harvest and thus not be allowed to fall below a certain minimum provided a reverse obligation is accepted and in a year of bad harvest the price is not allowed to rise above a certain maximum. Thus in the view of the consumer, a maximum ceiling price is a necessary corollary to a minimum support price.

7. As we shall presently see, even, operationally, a maximum ceiling price is a necessary corollary to a minimum support price. Let us very briefly see how a minimum support price may be made operative in practice. A minimum support price begins to operate when, in an year of good harvest, the market price tends to fall below the support price. It follows that if in such a situation the support price were made effective, all the supplies would not be taken up by the consumers and that there would result an excess of supplies in the market. The support operation consists in buying up these excess supplies and thus prevent them from depressing the market price below the minimum support price.

8. Once supplies are bought and stored in a year of good harvest, their disposal becomes an operational necessity. The most opportune time for their disposal is when, in a year of bad harvest, supplies are short and prices are rising. If used purposefully and effectively, the disposal can be so operated as to prevent the market price from rising above a certain maximum. Thus, even operationally, a maximum ceiling price becomes a necessary corollary to a minimum support price.

9. As mentioned above, the producers require only the support price. However, they must recognize that a ceiling price is a necessary corollary both because consumer interest demands it and because it would be impossible operationally to offer support without enforcing a ceiling. The two together reduce the range within which the market price may fluctuate. This is good for the consumer and the smaller the range within which the market price is allowed to fluctuate, the better it is for him. The relatively stable prices are also in the interests of the producers in the long run. However, in the short-run, it must be recognized that stable prices with fluctuating harvests result in fluctuating incomes for the producers which may lead to wasteful expenditure in years of good harvest and hardship in years of bad harvest. Therefore, in the interest of the producers, any measures of stabilization of prices such as through operating support and ceiling prices, should be accompanied by measures of income stabilization through appropriate credit and insurance policies.

10. The above describes the nature of the demand for support prices under short-term considerations and also the manner in which such support prices may be operated. Let us now ask the question as to how such support prices may be fixed and in particular whether in fixing the support prices, the "cost of production" can be and need be explicitly taken into account. In the above we have emphasized how a minimum support price and a maximum ceiling price are necessary complements of the proposed operation. In the following, we shall point out that the minimum support price and the maximum ceiling price must stand in a certain mutual relationship. An essential condition of the price support that can be offered on the kind of short-term considerations as we have described above, is that the support operation must not lead to a permanent accumulation of stocks with the agency responsible for the operation and must also not involve the agency into financial losses in its buying and selling operations. The first condition means that over a period of years, the agency must be able to sell the stocks it acquires in good years. This requires a certain relation, pragmatically arrived at, between the buying and the selling prices of the agency. The second condition means, neglecting for the sake of simplicity the costs of transport, storage, and handling which the agency must incur, that the selling price of the agency must not be below its buying price. In fact, it is on this supposition that we call its selling price the ‘ceiling’ and its buying price the ‘support’. The agency agrees to buy all supplies offered to it at the ‘support’ price. It undertakes to release on the market stocks at the ‘ceiling price’ whenever the market price crosses the ceiling and so long as it remains above the ceiling. For it to fulfil its second obligation, it must have sufficient stocks which in turn depend upon the support price. It is thus that the support and the ceiling prices at which the agency buys and sells must bear a certain mutual relationship. The higher the support price, the larger will be the stocks that the agency will be obliged to buy and the lower will be the ceiling it will be able to enforce. The support must not be placed so high that the ceiling will have to be below the support. The lower the support, the smaller will be the stocks that the agency may acquire and the higher will be the ceiling, it may be able to enforce. The support must not be so low and the consequent ceiling must not be so high that both become meaningless in practice. These considerations provide the broad limits within which the support price may be fixed.

11. While fixing the support price somewhere between these broad limits, the following considerations will appear relevant. The higher the support price the

larger will be the stocks that the agency will have to acquire and deal with. This involves buying, transport, storage, rotating and, finally selling in a manner that the price for the consumer will remain effectively below the ceiling. This requires knowledge, organization, expertise and competence. All these can be built only through experience. It will therefore be entirely legitimate to begin with a low support price and to raise it gradually.

12. In principle, one may raise the support price so high that the difference between the ceiling and the support may become very small. In deciding how far one may go in this direction, the following consideration is relevant. The minimum support price and the maximum ceiling price determine the limits within which the normal market is allowed to function. The higher the support price, the lower will have to be the ceiling price and thus smaller will be the margin left for normal market to operate within. Decisions in this matter must therefore be largely governed by considerations of expediency and advisability of how large a field maybe left to the normal market.

13. It would be obvious that throughout these several considerations, the "cost of production" of the producer did not appear anywhere explicitly. Therefore it seems that it should be possible to evolve a policy and a programme in this field without any active reference to "cost of production". This should be true of a price support policy at least to the extent demanded by what we have called the short-term considerations.

14. Before we proceed to examine the demand for price support under what we have called, the long-term considerations, let us once again emphasize what we have called the essential condition of price-support under short-term considerations. Briefly stated, this condition is that the agency responsible for operating the support programme should be able to sell, over a period of years, all the stocks it acquires at the support price, at prices not below the support price. If the condition is satisfied, it indeed means that over a period of years, the programme does not give any price support over and above what the market offers. Its operative purpose is merely to equalize the supplies and consequently also the prices over the years. It creates an agency to undertake the operations necessary to equalize the supplies over the years. The agency may use prices as indicators of quantities coming to the market and withdraw or reimburse the supplies as occasion demands. But the resulting prices on the market are such as the market determines them. It still is a free

market and the programme is a part of it. It offers no price support over and above the market price.

15. Let us now examine the long-term considerations of the producers. These are basically different from the short-term considerations. In the short-term considerations, we presumed that the size of the harvest was more or less independent of the inputs or costs incurred by the producers. This is not so in the long run. In the long-run, the size of the production is positively related to the inputs or costs incurred. The producer knows that he can expand production by increased inputs or costs. But he would naturally not do it unless it paid him to do so. Thus the inputs or the costs that the producer incurs get related to the price that he expects. Under these conditions, the demand for a minimum support price arises on two slightly different grounds. One is that the producer wants an assured minimum price which he may count upon and which may provide a basis for production decisions regarding inputs and costs to be incurred. The underlying consideration here is closely akin to those we have termed short-term considerations. However, there is a difference. In the short-term considerations, we have supposed that the size of the harvest is more or less independent of the inputs or costs incurred in the sense that it is determined by entirely fortuitous circumstances beyond the control of the producer. In the long-term considerations as we are now examining them, this need not be so. Here we suppose that the size of production depends upon the inputs or costs incurred and that the producer is aware of the possibility of expanding production by means of incurring higher inputs or costs. Under these conditions, production may expand and may bring down the market price below the economic level and thus ruin the producer. The producer needs an assurance against this eventuality. The second consideration is related. In order to continue production at a given level, the producer needs an assured minimum price. In order to expand production beyond the given level, the producer wants a higher minimum. He is willing to and able to incur higher inputs and costs, if and only if a higher minimum is assured.

16. Thus it seems that the minimum support price to be assured has to have some relation to what might be called the "cost of production" or the "supply price" of foodgrains. The two are not quite the same and in view of the fact that "cost of production" is not uniquely defined or understood, it may be more convenient and meaningful to talk in terms of the "supply price". The last mentioned consideration then only means that the supply schedule of foodgrains is of the usual

kind with a positive slope and thus requiring a higher supply price for a larger market supply. Nevertheless, certain doubts are often expressed on this point and it may be useful to clarify certain issues involved.

17. For instance, it is asked whether, with expansion of production, we should expect the unit cost of production to fall or rise. The argument runs as follows: Expansion of production will come about through adoption of improved technology. By definition, improved technology must be cost-saving. Hence, with expansion of production, unit cost of production and hence also the supply price should fall. Evidently, the argument suffers from over-simplification. In the first instance, it is not true that at every stage, expansion of production will come about only through the adoption of a new technology. It is quite conceivable that at each stage, a certain expansion of production may take place, within a given technology through increased inputs even beyond the point where diminishing returns and increased average costs set in. Secondly, it is not necessary that an improved technology must be cost reducing. A technology is an improved technology because it enables a larger volume of foodgrains to be produced out of the same area of land; it is an improved technology because it converts something which was not available for human consumption into something that is available for human consumption. It does not have to be necessarily cost-reducing. Whether or not it will be adopted in practice will of course depend upon considerations of costs. However, even then it is not necessary that an improved technology would be accepted only if it were cost-reducing. It would be perfectly legitimate to accept an improved technology if it made possible a higher production, though at higher unit cost, simply because there was no other method of securing a higher production. Thirdly, it may be readily conceded that many technological innovations may be cost-reducing. However, it is in the nature of an innovation that its adoption requires higher incentives and higher profit margins. There are considerable costs of research, and experimentation even at the individual farmer level and there is need for compensating the risk associated with an innovation until a farmer finally adopts it. It seems that the progress in agriculture will come about mainly through increased inputs within a given technology even beyond the point of diminishing returns, punctuated by adoption of innovations which may be cost-reducing but the acceptance of which nevertheless requires higher incentives and profit-margins. Finally, human labour, both of the farmer's family and of hired workers, constitutes and will continue to constitute for quite some

time, an important input in the production of foodgrains. Hired labour is valued naturally at the existing wage rates and labour of the farmer's family is often not counted at all in certain definitions of 'costs of production'. With agricultural development leading among other things, to increased production of foodgrains, there is no reason to suppose that the wages of hired labour in agriculture will remain the same or that expectations of the farmer and his family will remain the same. Whether or not one includes such increased expectations of human labour engaged in the process of production as an element in the "cost of production", they certainly enter the supply price. Therefore, broadly speaking, one must conclude that the supply schedule of foodgrains will be of the usual kind with a positive slope and that the producer will ask for a higher unit price for a larger market supply.

18. Let us now examine the nature of price supports under such long-term considerations. In the first instance, we should be clear that, the basic mechanism of offering the support would remain the same, namely, an agency which would buy whatever supplies are offered at the minimum support price. Secondly, the support price required to meet the long-term considerations will have to be higher than the one needed under the short-term considerations. The critical difference between the two is the following: Under the short-term considerations, we said that the support price must not be so high that, over a period of years, it becomes impossible to dispose of the stocks acquired at the support price, except at prices below the support price. The long-term considerations demand that the support price, will be at least as high. It means that stocks acquired at the support price will accumulate with the agency concerned. Long-term price support consists in buying up all supplies offered at a support price with full knowledge that over a period of years, it will be impossible to dispose of the stocks except at prices below the price at which they were purchased. One of the principal problems of operating price supports at such a level is how to dispose of the stocks acquired at that price.

19. There is another and a closely related problem. As mentioned above, with high support prices offering long-term support, stocks accumulate which cannot be disposed of except at below the support price. The reasons are two-fold. In the first instance, because of high support prices, the production and market supplies expand. Secondly, because of high prices, consumption is restricted; a class of consumers is unable to buy all its foodgrains requirements at the support price. Thus accumulated stocks in the hands of the public marketing agency and

unsatisfied demand of a class of consumers are found side by side. An obvious solution therefore is to distribute the acquired stocks at subsidized prices that is at prices below the support price but taking due care that such subsidized distribution does not depress the market price below the support price.

20. We shall not discuss here the problems of subsidized distribution but shall emphasize two relevant points. The first is that subsidized distribution is a necessary corollary of a long-term price support. The higher price provides a necessary condition for expanding production. This price has to be supported because the market fails to provide the support for want of necessary purchasing power. This means that while conditions are being created for expanding production, steps must be taken simultaneously to promote consumption to the same extent as the production may expand. This is the purpose of subsidized distribution. While discussing a minimum support price under short-term considerations, we had observed that a maximum ceiling is a necessary corollary to a minimum support price. In the same manner it must be emphasized that subsidized distribution is a necessary corollary to a price support under long-term considerations.

21. The-second point to be emphasized in this connection is the requirement that the subsidized distribution must not depress the market below the support price. This is obvious, otherwise the subsidized distribution would defeat its purpose. This means that the subsidized distribution must be arranged outside the free market and must be effectively sealed off from the free market. This is in sharp contrast to the situation obtaining in the context of price support under short-term considerations. There, as we have earlier observed, the operation does not indeed offer any support over and above that given by the market and hence the operations of the agency responsible for operating the support prices, together with the ceiling prices as a necessary corollary, may all fall within the framework of a free market. This is no longer so if the price support is to be given under long-term considerations. Here the intention is to offer support over and above what the market gives. This cannot be done by operating within the limits of a free market. Therefore the agency responsible for operating such price supports has to engage in extra-market activities. Subsidized distribution sealed off from the free market is one such activity.

22. Let us now ask the question as to how does one decide upon the level of support prices under such conditions and

in particular whether "cost of production" of the producer offers any explicit guidance in the matter. We may begin with an important consideration which is obvious. The higher the support price, the larger will be the stocks that may have to be acquired at that price and later distributed at below the purchase price. This means that the higher the support price, the larger will be the subsidy that will have to be provided to promote consumption. Hence, financial resources available for the purpose clearly set a limit to how high to set the support price.

23. Though the decisions in this matter will be ultimately governed by financial considerations, it is worthwhile being clear about the basic issues involved. As pointed out above, higher support prices lead to larger stocks on account of two reasons: (a) expanded production and (b) restricted consumption. The second is obvious. So long as the first is also assured, namely that so long as it is reasonably certain and so long as continued evidence becomes available to the effect, that higher support prices lead to higher production and larger supplies, there is no reason why the support prices should not be kept so high as would lead to production and supplies large enough to satisfy the physical needs of all the people. That all the supplies which may be forthcoming at high support prices may not be taken up by the consumers, even when they are needed, is an irony of the way the incomes are distributed. It is possible that in the process of economic development of which agricultural development and expansion of food production are a part, may bring about a levelling up of incomes in the lowest income class of consumers partly through increased national income and partly through a better distribution of the same, so that the market itself may support the high minimum necessary to enable the producers to produce enough for everybody. However, until that happens, a better distribution of the incomes at the lowest level must be brought about through subsidized distribution of food. As explained above, subsidized distribution of food operated as a necessary corollary of long-term support price, helps achieve simultaneously a higher production of food and a better distribution of the same. These are evidently the most primary aims of economic development. How much financial resources may be devoted for the purpose, depends upon the strategy and the path of development that one chooses.

24. To return to our question as to how high to set the support price, one obvious answer then is to set it as high as financial resources permit. It is difficult to see that there can be any other consideration or

that in particular the "cost of production" of the producer may offer any explicit guidance in this matter.

25. Let us then sum up the main guidelines that emerge from the above discussion for evolving a policy and a programme for a minimum price support to foodgrains. They are as under:

(a) Make a beginning.

(b) Begin with establishing a programme of minimum support prices under short-term considerations, namely a programme of support prices to be operated in conjunction with ceiling prices in such a manner that it does not lead to any net accumulation of stocks over a period of years. Set up the necessary agencies and the organization.

(c) Begin with a low price support and a high ceiling. Gradually, as experience gathers, raise the support and lower the ceiling, until the difference between the two is just enough to cover the operational expenses of the programme.

(d) The stage will then be reached where the price support may assume the long-term character. Gradually raise the support and simultaneously organize subsidized distribution of all stocks that cannot be disposed of on the market above the support price.

(e) So long as there is continuing evidence to suppose that further raising of the support price may help further expansion of production, continue to raise the support and continue the subsidized distribution until there are enough staple foodgrains for everybody and every one gets them in adequate quantities.

(f) Finally, MAKE A BEGINNING.

Summary

The paper suggests that a programme for setting up and operating minimum support prices for foodgrains is possible without an active reference to ‘cost of production’. The basic operation consists of an agency which would buy whatever supplies are offered at the minimum support price and sell the same in a manner which would not depress the market below the support price. It is suggested that a programme for this purpose may be set up and evolved in two stages. In the first stage, the support price should be kept low enough so that the agency would be able to sell, over a period of years, all the stocks it acquires at prices not below the support prices.

Hence at this stage, ceiling prices become a necessary corollary to support prices. One may begin with a low support and a high ceiling and then gradually raise the support and lower the ceiling until the difference between the two is just enough to cover the operational expenses of the programme. Supporting the price above this limit may be justified only if there is evidence that higher support will lead to increased production. A higher support will also mean that the operating agency may not be able to dispose of the stocks except at prices below the price at which they are purchased. This will require subsidized distribution sealed off from the rest of the market. The level of the support prices at this stage must depend upon the evidence that it will lead to increased production and also upon the financial resources available for subsidized distribution.

GOPAL KRISHNA GOKHALE AS AN ECONOMIST

Among the Indian political leaders of the last generation few showed an understanding of the economic and fiscal problems then facing the country and of the need for, and the direction of reform in this field, more complete than did Gopal Krishna Gokhale. His record in the Imperial Legislative Council of which he was a member from 1902 until his death in 1915, shows that he was undoubtedly the greatest authority on public finance of his time in India.

Gokhale became a member of the Imperial Legislative Council in 1902, when Sir Ferozshah Mehta retired from that body. That was a period when the Indian budgets showed substantial surpluses year after year. The revised estimates for the year 1901-02 revealed a surplus larger than was anticipated and Sir Edward Law, who was the then finance member remarked: "The realization of so large a surplus cannot but be a matter of congratulation to India and following the relatively large surplus which we were able to announce last year it confirms the hope I then ventured to express that if no new ill-fortune overtakes us, we might, as regards finance, look forward to a period of increasing prosperity". It was in the face of this able and efficient bureaucracy in a complacent mood wanting to demonstrate an increasing prosperity of India under the British rule that Gokhale rose to make his maiden speech in the Imperial Legislative Council on 26th March 1902. Speaking of the budget surplus Gokhale observed: "Coming as it does on the top of a series of similar surpluses realized when the country has been admittedly passing through very trying times it illustrates ...... the utter absence of a due correspondence between the condition of the people and the condition of the finances of the country ..... Indeed these surpluses constitute a double wrong to the community. They are wrong in the first instance in that they exist at all - that Government should take so much more from the people than is needed in times of a serious depression and sufferings; and they are also wrong because they lend themselves to an easy misinterpretation and, among other things, render possible the phenomenal optimism of the Secretary of State for India who seems to imagine that all is for the best in this best of lands."

Point by point Gokhale replied to the Finance Member's case of India's growing prosperity: "The growth under land revenue, excise, and stamps is sometimes mentioned as indicating increasing prosperity. But the growth of land revenue is a forced compulsory growth. It is a one-sided arrangement and the people have either to pay the increased demand or give up their lands and thereby part with the only resource they have. The growth of excise revenue to the extent to which it is secured by increased consumption, only shows that the operations of the Abkari department ..... are leading to increased drunkness in the land. This of course means increased misery and is thus the very reverse of the indication of increasing prosperity ..... Similarly an increase of revenue under stamps only means an increase in litigation which undoubtedly shows that the people are quarreling more; but which is no proof of their growing riches ..... The only taxes whose proceeds supply an indication of the material condition of the people are the income tax and the salt tax, the former roughly speaking for the middle and upper classes and the latter for the masses. Now the revenue under both these heads has been more or less stationary all these years and the salt revenue has not even kept pace with the normal growth of the population. They, therefore, lend no support to the contention that the people are advancing in material prosperity".

He was unrivalled in the skilful use of statistics and was quick to discover their fallacious use. In the budget discussion of the previous year Lord curzon, dealing with the question of India's prosperity, had tried to show on the basis of certain figures of per capita national income that the movement was distinctly in a forward and not in a retrograde direction. Commenting on it Gokhale said: "The attempt to determine the average income per head for a given population is useful only for the purpose of obtaining a statistical view of the economic condition of the people and from this point of view our average income, whether it works out to Rs. 18 or Rs. 20, or Rs. 27 or Rs. 30 per head, is exceedingly small and shows that we are exceedingly poor people. But when these calculations are used for taking a dynamical view of the economic situation, the method Is open to serious objection, as the necessarily conjectural character of many of the data render them of little value for such a purpose". He then proceeded to put forward an array of facts and figures covering such items as census returns, vital statistics, salt consumption, agricultural outturn, crop areas, areas wider superior crops, exports and imports and concluded: "These and similar facts taken cumulatively lead to the conclusion that the material condition of the masses of the people in India is steadily deteriorating ...... The phenomenon is the saddest in the whole range of the economic history of the world. Here is a peasantry which taken all in all is inferior to no other people in industry, frugality and patient suffering. It has enjoyed the blessings of uninterrupted peace for half a century, and at the end of the period the bulk of them are found to be in a worse plight than they have ever been in."

Returning to the budget surpluses, he pointed out: "(The surplus) does not connote . . . . any advancing material prosperity of the country ... (but) is a clear proof of the fact that the level of national taxation is kept justifiably high even when Government are in a position to lower that level." He urged that an immediate tax relief was necessary, and emphasised: "The obligation to remit taxation in years of assured surpluses goes ..... with the right to demand additional revenues from the people in times of financial embarrassment". In particular, Gokhale wanted the salt duty to be reduced. Speaking of the salt duty, he said: "I do not think any words are needed from any one to establish the increased hardship which the present rate imposes upon the poorest of the poor of our community It may be noted that the consumption of salt during the last 14 years has been almost stationary, not even keeping pace with the growth of population showing a rise of less than 6 per cent in 14 years against a rise of 18 percent in 4 years following the reduction of duty in 1882". Again next year speaking on the same subject he observed: "The manufacturer of salt in India is strictly under Government control and practically a government monopoly. And the monopoly is enforced under restrictions and in a manner which have the effect of transferring about a third of the industry to the foreign manufacture ..... We have extensive sea board and salt mines too and can manufacture every pound of salt we need and yet under the existing fiscal system about a third of our supplies comes from foreign countries .... The imports have increased 38 per cent in 10 years. I submit that in respect of such a prime necessity of life as salt, specially when we have plenty of it within the four corners of this country, we ought not to be forced to depend on foreign supplies to such an increasing extent".

The salt duty was subsequently reduced in stages and it was found that this reduction led to an increase in the consumption of salt. Attention was drawn to this fact by the Finance Member, Mr. E.N. Baker, in the financial statement of the year 1906-07. Commenting on it Gokhale said: "Time was not so lone ago when it was the fashion, both in this Council and outside, to regard the burden imposed on the masses by a high salt duty as after all only a light one and to deny that its right could seriously affect consumption ..... I hope ..... no one will again venture to contest the proposition that, in dealing with a prime necessity of life such as salt the only right policy is to raise an expanding revenue on an expanding consumption under a diminishing scale of taxation".

Another tax relief he urged was the reduction of the excise duties on cotton goods. Speaking on them during the budget session in 1903 Gokhale said: "These excise duties illustrate what John Stuert Mill has said about the Government of the people of one country by the people of another. They were levied not for revenue purposes but as a concession to the selfish agitation of Manchester..... In no other country would such a phenomenon of the Government taxing an internal industry for the benefit of a foreign competitor be possible". But the reasons why Gokhale opposed the excise duties on cotton goods were very different from those for which many others particularly the spokesmen of the mill industry opposed them and he made this quite clear. Thus when in 1911 the Hon'ble Mr. Dadabhoy moved the resolution in the Council recommending abolition of the excise duties on cotton goods manufactured in India, Gokhale while supporting the resolution said: "I rise to accord my support to the resolution which my Hon'ble friend Mr. Dadabhoy has moved though I do so on grounds somewhat different to those on which he and some of the other speakers who have followed him have based their case. I approach this question not from the standpoint of the representative of the mill industry but from that of a member of the general community ..... If it was the case that these excise duties fell on the producers and not on the consumers, I would not stand up here to support their abolition today. My friend the Hon'ble Mr. Dadabhoy complained of the extreme depression of the mill industry and several other members have also spoken in similar terms. I think, however, it is necessary to point out that before this depression came they had a spell of extraordinary prosperity. If we take an average of good and bad years, I am not quite sure that there is such a strong case to urge for the abolition of these duties from the standpoint of the industry. Coming to the larger aspect of the free trade versus protection, it is necessary to remember that there are two kinds of protections, the right kind and the wrong kind. The right kind of protection is that under which the growing industries of a country receive the necessary stimulus and encouragement and support that they require but under which care is taken that no influential combinations prejudicial to the interests of the general community, come into existence. The wrong kind of protection on the other hand is that under which powerful influences and combinations, and interests receive assistance to the prejudice of the general community. And I believe that the right kind of protection if available will do good to India. But situated as India is, I fear that there is no likelihood of that kind of protection being available to us. I do not, therefore, join in the plea that the abolition of excise duties would be a measure of protection to the Indian industry and that the Government should accede to it on that ground".

It was remarkable the way he kept the interests of the common man always in the foreground. For instance, in 1912 the Hon'ble Mr. Dadabhoy moved the resolution in the Council recommending that the minimum of income assessable to the income tax be raised to Rs. 1500 a year. Speaking on the resolution, Gokhale said: "If every tax is to be discussed solely from the standpoint of those who pay it, I do not think that there will be any tax which will escape adverse criticism. But the State has to look at it from another standpoint. The State has to look at the whole scheme of taxation, first, from the standpoint of its own necessities and, secondly, from the standpoint of the comparative ability of the different classes to pay their particular share of the total revenue raised from the community. Now, judged by this standpoint, I really do not think that the class for which my Hon'ble friend seeks a remission has any substantial grievance...... In judging the comparative ability of the different classes to pay, the point to be really considered is whether the scheme of taxation taken as a whole hits any one class harder than any other class. Now, from the standpoint I have no hesitation in saying, and I have urged this view again and again in this Council, that the poorer classes of this country bear really a larger burden than the class to which my Hon'ble friend has referred or the classes above them. The upper and the middle classes of the country contribute really much less to the exchequer than poor classes relative to their resources".

While he was thus pleading for a judicious reduction in taxation he was at the same time critically examining the manner in which the budget surpluses were spent. Even in his maiden speech in 1902, he had warned the Government of the dangers of such surpluses leading to wasteful expenditure. He said: "A succession of large surpluses is little conducive to economy and is apt to demoralize even the most conscientious Government by the temptation it offers for indulging in extravagant expenditure. This is true of all countries but it is specially true of countries like India where public revenues are administered under no sense of responsibility to the governed".

In 1903 Gokhale commented upon the military expenditure: "As things stand at present, Indian finance is virtually at the mercy of military considerations and no well sustained or vigorous effort by the State on an adequate scale for the material advancement or the moral progress of the people is possible while our revenues are liable to be appropriated in an ever-increasing proportion for military purposes ..... Military safety is no doubt a paramount consideration to which every effort must yield, but military preparedness has no definite standard and might absorb whatever resources could be made available for it practically without limit Military efficiency must, therefore, be always relative that is determined in the case of each country by a combined consideration of its needs of defence and the resources that it can fairly devote for the purpose".

The year 1904-05 was the seventh year in succession showing a surplus. Examining the manner in which the surpluses realized in the previous seven years were utilised, Gokhale said: "Thus after allowing the expenditure to increase in all directions on an unprecedented scale, after making large special grants to provincial governments from time to time, after spending annually 16 crores out of the current revenues for non-recurring charges, and after laying by $12\frac{1}{2}$ crores for purposes of the Gold Reserve Fund, the Government have still been able to devote a sum of about $36\frac{1}{2}$ crores in seven years, or a little over 5 crores a year on an average, to the reduction or avoidance of debt. I submit that such a system of finance is unsound in theory and indefensible in practice. For it involves grievous injustice to the present generation".

He particularly objected to the use of the budget surpluses for the reduction of the public debt. He argued: "The ordinary debt of India, as distinct from the public works debt which is fully covered by valuable assets is not large and there is no justification for being in such a hurry to reduce it. The utmost that the Government might do in the matter is to provide a small sinking fund, say about a million sterling a year, but beyond this it is indefensible to go especially as in the absence of a reduction of taxation, there are so many ways, all intimately connected with the well being of the people, in which the surplus revenue could be spent".

He also objected to the surplus being used for the construction of railways. He observed: "During last eight years, the surpluses: of the Government of India have amounted to not less than 85 crores of Rupees, and the whole of this money has been spent by the Government on railways, in addition to the large amounts borrowed for the purpose. Now I do not wish to say anything against the construction of railways as a commercial undertaking ..... though even bare, the claims of irrigation to a larger share of the capital raised must be recognized better than they have been in the past. But I have the strongest possible objection to our surplus being devoted to railway construction when they are urgently needed for so many other objects vitally affecting the interests of the masses ..... I submit that there should be some sense of proportion in this matter...... Are railways everything, is mass education nothing, is improved sanitation nothing, that the Finance Member should lay hands on every rupee that he could get either by borrowing or out of surpluses and devote it to the construction of railways only?" And again: "Judging from the manner in which the surpluses are applied year after year to railways construction, one would conclude that in the opinion of the Government what people needed most was a vigorous extension of railway facilities ..... We all know that by spending surpluses as capital on railways, the Government is able, in the final adjustment, to reduce by a corresponding amount, the unproductive debt of the country. And it may be contended that though the surpluses are in the first instance devoted to railway construction, they are in the end virtually utilised for the reduction of debt. My answer to this point is that our debt, by which I mean unproductive debt of the country for that is the only real debt is so small in amount that its further reduction is not an object of much importance". He pressed this point year after year. In 1900, he observed: "I have been contending again and again in the Council that as our railways earn a very fair rate of interest and as the credit of the Government is excellent, the Government should confine the outlay on railway construction only to such bias as can be raised in the market out of borrowings; and any sums available out of current revenues should be used for other more pressing objects."

In the meanwhile, the fears that Gokhale had expressed in 1902 had come true. The public expenditure in India kept mounting. In 1910, additional taxation was proposed in order to meet the mounting expenditure. Criticising this Gokhale said: "This Circumstance, namely the levying of extra taxation in a normal year, suggests that something is wrong with the financial position of the country and in any case it suggests an inquiry. ..... It has been said by some critics that the present difficulties of Government have arisen from the fact that during those fat years, Government remitted taxation which should not have been remitted. ...... I must protest strongly against this view .... The whole position is this, that in the ten years 1898 to 1908, while 6 crores were remitted in taxation, the annual civil expenditure was allowed to grow by 15 crores and about 5 crores of additional expenditure was incurred every year in connection with the army. This gives an increase of about 20 crores in civil and military expenditure in the course of ten years or an average growth of 2 crores a year. ..... I think these figures suggest ..... the necessity for an inquiry in the growth of civil and military expenditure during the last ten years". Next year in 1911, Gokhale moved a resolution calling for such an inquiry.

In 1912, Gokhale again returned to the question of the use of the budget surpluses and moved a resolution recommending the creation of special provincial reserves by means of grants from the imperial surpluses. He said, "Money is required for non-recurring expenditure in many directions in this country, specially for education, sanitation and medical relief .... These will require not ten but hundreds of crores of rupees and the problem cannot be satisfactorily dealt with unless the Government made large regular allotments for this purpose. The Government, however, is reluctant to make a large regular allotment out of the current revenues. Therefore, I propose another method which, though not equally satisfactory, will be found to answer the requirements to some extent ..... What I urge is that two-thirds of this surplus, as also all future surpluses, should be placed at the disposal of provincial governments for non-recurring expenditure on the objects I have mentioned. The present policy of making grants for a year only out of the surplus that accrues during that year is wasteful policy ..... The local governments cannot rely upon a continuance of their good fortune and therefore they cannot take in hand any large scheme which requires finances over a series of years ..... The Government of India, on the other hand, can think only of a single surplus at a time, and therefore, it is not surprising that it does not find itself to make any further grant until another surplus is realized."

These are a few illustrations of the depth of analysis and the breadth of vision that Gokhale brought to bear on the debates on the public finances of the country. Because of his ill-health Gokhale did not attend the Council during the debate on the budget in 1913. Without him the debate must have been a tame affair. While most of the Council members were making speeches showering superlative praise and congratulations on the Finance Member, the Hon'ble Sir Guy Fleetwood Wilson must have sat there missing only one man, Gokhale. Winding up the debate, Sir Wilson expressed himself in so many words: "I do not propose to detain the Council long to offer many remarks; but I should like to preface my remarks by saying how deeply I regret the absence of one who has proved himself to be a true servant of India, and that is Mr. Gokhale..... It is difficult for those who served in previous Councils to express to those who are new to this Council, what a great blank Mr. Gokhale's absence creates. It seems to be a play of Hamlet without Hamlet."

Gokhale Institute
of Politics & Economics, V. M. Dandekar.
Poona 4.
30th April 1966.

CATTLE PROBLEM

V. M. Dandekar

The current agitation for ban on cow-slaughter has focused public attention on a serious problem with which Indian agriculture is riddled. The few religious men who are sincere in their agitated sentiment have never paid attention to the material aspects of this question; or else they do not under stand them. The many politicians who are operating under the religious robe are plainly exploiting the ignorance of the mass of people. The government should know better. However, it has obviously decided to take the political line. There are frequent references to the directive principles of the Constitution. The politicians, both in the government and the opposition, apparently believe that the Constitution has enfranchised the cow. So they are looking for the votes. There is thus a little change that the problem will be examined rationally and dispassionately.

Let us all understand that this is a grave problem of the first magnitude and that by making a religious and political issue out of it, we are doing great harm to ourselves and to our children. The country has already a very large human population and a very large cattle population. The density of human population in the country is 370 per square mile of geographic area and for every 100 persons there are over 50 heads of cattle. Both these are too large for our resources to feed adequately. Moreover, both the populations are increasing at an increasingly rapid rate during the last fifteen years. This is because of the continuous improvement in medical, public health and veterinary services we are achieving through our development programmes. Previously, both the human and cattle populations in the country were kept within limits because of high mortality on account of disease. Thanks to the improvement in medicine and public health, major epidemics such as small-pox, cholera, plague and influenza no longer take a heavy toll of life. For many common illnesses, better medical treatment is now available more rapidly. This has resulted in a phenomenal reduction in mortality. Consequently, the human population has been growing at an increasingly rapid rate. The estimated rate of growth at present is about 2.5 per cent per annum. The same is true of the cattle population. Rinderpest and foot-and-mouth are now effectively controlled and veterinary services have greatly improved. This has led to a reduction in mortality among the cattle population which has also been growing at an increasingly rapid rate. During the quinquennium 1956-1960, the cattle population in the country grew by 10 per cent which gives an annual rate of growth of almost 2.0 per cent. In some of the

States, the rates of growth have been even higher. The annual growth rate in Orissa was 4.5 per cent, in Uttar Pradesh it was 2.7 per cent, in Bihar 2.3 per cent and in Madras 2.2 per cent. With improving veterinary services, these growth rates are bound to increase even further. In fact, one wonders if in Orissa, the cattle population will not soon he larger than the human population.

The rapid growth in the human and cattle populations of the country in recent years is thus due to improvement in medical, public health and veterinary services and consequent decline in mortality. However, no one will argue that we should for that reason give up our programmes in medical, public health and veterinary services and let Mortality check the growth in population. As regards the human population, this is obvious. It is one of the aims of economic development that every individual should live healthy and long. In fact, we wish that everyone should live upto 100 years. The medical and public health services must therefore be improved and mortality reduced even further. Even with the cattle population, there are several reasons why we must not cut down our veterinary services, and let a higher mortality rate check the growth of the cattle population. First, diseases and epidemics do not act selectively. They kill not necessarily the animals which are useless or unproductive. They operate indiscriminately and kill good, bad and indifferent animals alike. In fact, epidemics, if uncontrolled, may endanger the entire Stock wholesale. Secondly, diseases do not necessarily kill. Often they merely disable the animals and let them live in an unproductive and useless condition. Thus for several reasons it is essential to control diseases and death among the cattle population as much as among the human population. In other words, growth in population, whether human or cattle, cannot be checked through uncontrolled mortality.

Nevertheless, it is obvious that if the human and cattle population in the country are allowed to grow unchecked, they will jeopardize the entire process of development. Already, the burden of these populations on our resources is too heavy to support. If it grows any further, it will without doubt push the country from the present poverty, to destitution and starvation. It is therefore imperative that something is done to check the growth in these two populations. In relation to human population, we have recognized the urgency of this problem. We have also seen the logical solution, namely, that if mortality is to be reduced fertility must also be reduced. In other words, if the number of deaths as reduced, the number of births must also be reduced so that there may be no growing balance of births over deaths. After considerable debate and deliberation, we have accepted this logical necessity and launched a massive programme for controlling human population through birth control.

The same logical necessity prevails in the case of cattle population. We cannot allow the cattle population to grow indefinitely. This is a crucial point and must be understood firmly. Our agricultural resources in land and water are limited and they cannot support and sustain an indefinitely growing population whether it is human or cattle. We have agreed to limit the human population and we must agree to limit the cattle population. We must restrict the cattle population not only because it is in our interest to do so, but also, because it is in the interest of the cattle as well. Even if we were to extinguish ourselves and hand over the country to the cattle, it will be necessary to restrict the cattle population. If this is not done, increased mortality through disease and starvation will ultimately begin to operate. This is the crucial point in the understanding of this question. I hope that even the protagonists, of a ban on cow-slaughter will agree to this need, namely, that the cattle population will have to be restricted. No further discussion is possible with those who do not accept this preliminary proposition. Hence before parting company, we should try to understand their precise position on this point. I shall therefore ask the following questions:

(1) Do they desire that the human population should be limited but that the cattle population should be allowed to grow unrestricted? If they do, they may know that under the circumstances, it will not take much time for the cattle population to outgrow the human population.

(2) Or, do they desire that neither the human nor the cattle population should be restricted and that both should be allowed to grow unrestrictedly? If they do, let them clearly understand the consequences. The combined human and cattle populations will soon outstrip all our resources and man will find it difficult to live very differently from the cattle. There are then two possibilities the 'human' beings and cattle will continue to live harmoniously in which case increased mortality through disease and starvation will begin to operate in both the populations. Alternatively, the two populations may begin to compete for the limited resources and a struggle for survival and existence will ensue. Then either species must win by killing and controlling the other.

(3) One final question: do they recognize any difference between man, cattle, horse, dog, birds, insects, bacteria and several other forms of life? If they do, is the difference religious or economic? If they do not recognize any difference between several forms of life, do they advocate that all forms of life should be allowed to grow unrestrictedly? In that case, they may know that it will not take much time to return to the jungle and that ultimately, the law of the jungle will prevail.

Let these questions be considered dispassionately and answered publicly. We may not agree but let us understand our respective positions. In the meantime let us move in the company of those who recognize that just as we have agreed to control and limit the human population, so also we must agree to take effective, steps to control and limit the cattle population.

The accepted method to control the human population is controlling births. How do we control the cattle population? Serious suggestions have been made that we should adopt the same method namely, birth control through contraceptives such as sterilization or ringing of the cows. Recently, the Central Council of Gosamvardhan has reportedly recommended the use of the loop for cows. These suggestions of course arise out of the desire to treat the cows exactly like our mother or like the mother of our children. Suppose for a moment that we accept this suggestion and fit loops to the cows. Let us then consider the consequences. The cows fitted with loops will of course not calve and hence the population will be controlled. But if a cow does, not calve, she will also not give any milk. Who will then feed her? Of course, if we desire to treat the cow exactly like our mother or like the mother of our children, we should agree to feed her even when she does not calve and give any milk. We feed our mother or wife even after she is fitted with a loop. Why should we hot feed a cow? We should. Unfortunately we do not. We do not feed a cow unless she promises to give us milk. This is a hard fact that, when the chips are down, we do make a difference between our mother and the cow. We need not be ashamed of this because the cow has precisely reciprocal feelings. We call her mother and some of us do behave like her sons. Nevertheless, she makes a difference between ourselves and her calves. She refuses to give milk unless she has a calf. The hard facts about our relations with the cow are: (a) a cow refuses to give milk unless she calves, and (b) we refuse to feed her unless she promises to give milk.

There is ample evidence to show that we shall not feed the cow once we fit her with the loop. Even without the loop, there are a large number of unproductive cows in the country and we merely have to examine how we are treating them. We do not feed an unproductive cow but let her loose. She must then roam around and feed herself on refuse. In rural areas, these cows roam with hunger and soon become wild. Wild cows roaming and destroying crops are a serious problem in many districts. Finally, if she cannot feed herself, she faces starvation, emaciation and death. The evidence of this can be seen in the much smaller number of cows, as compared to the number of bullocks in our adult cattle stock. Let us consider the situation in Uttar Pradesh where presumably the cow is most loved and best looked after. According to the Livestock Census of 1961, among the adult cattle stock in Uttar Pradesh there were 195 bullocks for every 100 cows. This means that in the adult stock, only one-third are cows, and two-thirds are bullocks. How does this happen? The male and female calves are born in more or less equal numbers, and they appear in more or less equal numbers in the young stock. How is it then that the number of cows is so much reduced in the adult stock? How does it happen without slaughtering? The answer is neglect and starvation. It is through the neglect and starvation that in Uttar Pradesh the number of cows is kept down. The same is true in several other States. Take Bihar and Gujarat. In Bihar, among the adult cattle stock, there are 163 bullocks to every 100 cows.

How are these numbers of cows reduced so much below the number of bullocks without slaughter? It is through neglect and starvation. There are notable exceptions where this is not true. They are, most importantly, Kerala with a predominant Christian population, Jammu and Kashmir with a predominantly Muslim population, and Rajasthan with land resources which are as yet not overburdened. In the remaining States, the ratios vary but everywhere the number of cows is smaller than the number of bullocks. Here, the cows are eliminated, through neglect and starvation. This is the general rule and these are the prospects that the cow will face if she is fitted with the loop.

It is suggested that unproductive cows should be taken care of in special, cattle camps set up for the purpose. We have no such care-taking camps for old men and women. Nevertheless, we may certainly set up care-taking camps for old and unproductive cows if we feel that the old cows should have higher priority over old parents. In fact, a few such camps called **Gosadans** have been set up. The experience generally has been that they have failed to attract private charity in sufficient measure and the public funds provided for them have not been used for the care of the cows exclusively. However, this is incidental. The main point, is that whether it is private or public charity, whether the cows are fed at home, or they are let loose to feed themselves or are taken care of in **Gosadans,** they are a claim on the limited resources of the country and we cannot afford to let their number grow unrestrictedly.

The cattle problem has often been mistaken as the problem of the old and unproductive stock. Old and unproductive stock indeed is a serious problem. However, it must be emphasized that even if we succeed in doing something to the existing stock of old and unproductive animals the basic problem remains. This is a point which must be firmly understood. The fact of the matter is that given any stock, it brings forth a certain number of calves which is much larger than is needed to replace the original stock. As a result, the size of the stock continues to grow without limit. When the stock outgrows the resources, there arises the problem of unproductive stock. The existence of unproductive stock is only a signal that the stock has out-grown the resources. The root cause of the problem is thus not old and unproductive animals, but the large number of calves that come forth and are allowed to grow. Nevertheless, we cannot reduce the number of calves coming forth, because that would render a large number of cows unproductive. Therefore we must permit the largest number of calves to come forth, but not let all of them grow indefinitely.

To be sure, today we do not permit all the calves to grow. This will be evident if we compare the number of young stock under one year and the number of adult stock above three years. A little computation shows that the mortality in the young stock is so high that fewer than 30 per cent of the calves grow to the adult age of three-years. These ratios are again different in different states. In Uttar Pradesh, less than 20 per cent of the calves grow to the adult age of three years. The same is true of Punjab. It is only in Gujarat, Himachal Pradesh and Madhya Pradesh that between 40 and 50 per cent of the calves grow to the adult age of three. In all other states, the ratios are less than 30 per cent.

How does this happen? We do not eat veal. Why do then so many calves die so young? Because almost immediately, after it is born, we pull the calf away and deny it the full share of the milk of its mother. We cheat the cow with a false calf and steal away all her milk while her young one is starved to death. If the cow had the slightest notion of the fate of her young one, she would readily walk to the slaughter-house rather than deliver her, calf in our hands to be starved to death.

These are then the two cardinal principles of the affectionate care that we bestow on the cow. First, we do not feed the-cow unless she promised milk. But we do not kill her either, we let her starve to death. Secondly, when she calves, we cheat her, steal the milk away and leave the young one alone but do not kill it; we let it starve to death. What is the religious sanction to this between starvation and slaughter. In what sense is starvation of animals more humane, more in consonance with our cultural heritage than is slaughter? These are questions for the religious men to answer. Let them ponder these questions and answer them sincerely.

If, the choice may be made on economic grounds, it is obvious that slaughter is far more economic than starvation. In the, first instance, with deliberate slaughter, one can be deliberately selective. One may select the animals one would like to keep and kill the rest. Starvation cannot be equally selective. When an animal starves, it nevertheless eats a little and denies that food to other animals. Hence all animals starve in varying degrees. Starvation is not selective also in the sense that the animals which survive the process of starvation are not necessarily those who should survive. Those who survive are often fit for nothing else except mere survival. Thirdly, a well-fed animal when deliberately slaughtered has high economic value. A starved and emaciated animal eats its own meat, and finally when it dies, it leaves behind little except bones and inferior hide. It has become impossible to consider these questions dispassionately because of importing into the discussion much argument by analogy. Let us give up this analogy between the cow and the mother. We know it is false and dishonest. We know that our relation with the cow is not that between mother and son. We know that the relation is based on solid material considerations. We know that we would not feed her if she did not give milk. Let us then call a cow, a cow. That will help establish a normal, healthy relation between ourselves and the cow.

Once this is understood, we shall find it possible to feed a cow not only because she gives milk but also because she can give us meat. We may then choose to slaughter some cows but we shall feed them well, to the last moment, rather than letting them starve to death as at present. We shall feed them well not for charity or sentiment but because it will pay us to feed them. We shall let all the young stock such at the mother because it will pay us to do so. As the stock grows, we shall have to select and weed out but we shall feed all the animals well up to the last moment. We shall then have around us cattle that are well-fed and well-looked after. Sooner or later every one may be slaughtered. But every one will be fed well while it lives.

These are the elementary principles of stock management. Feed, weed and select. Without adequate weeding through appropriate selection between sex and age and between one animal and another, there is no possibility of feeding the stock well. Reference is constantly being made to the directive principles of the Constitution. The Constitution certainly directs that the State shall endeavour to take steps for prohibiting the slaughter of cows and calves and other much and draught cattle. But the Constitution also directs that the State shall endeavour to organize agriculture and animal husbandry on modern and scientific lines. There is a clear conflict between these two directives of policy and we must decide which one of the two principles should prevail. Let us discuss and debate the issue dispassionately.

Gokhale Institute of
Politics and Economics,
Poona-4.

8th December 1966.

THE PLAN FRAME AND THE MOTIVE POWER

V. M. Dandekar

The central thesis underlying D. R. Gadgil's assessment of the present economic situation and his prescriptions for the future is that the shape and direction of economic policy are largely the result of influence of vested interests in private business and trade. He recommends a comprehensive regulatory framework in order to curtail and regulate the work and influence of this section.

However, we should recognise that in the process we may create another vested interest, namely, the bureaucracy. While the consequences of replacing completely one vested interest, private business, by another vested interest, bureaucracy, deserve to be carefully considered, this is not an immediate prospect. The immediate prospect is a much worse one, namely, that the two vested interests, private business and bureaucracy, may co-exist, join hands and undermine the entire regulatory apparatus.

This may be the real danger in setting up an all-pervasive regulatory framework.

Gadgil is probably justified in his criticism of the physical and the perspective planners who presumably believe that once a given capital investment is achieved, the economy functions automatically according to plan and leads to the development aimed at. So, as Gadgil emphasises, in order to succeed, a plan of investment must be supported by the necessary policy and regulatory framework.

However, it will be equally wrong to suppose that a policy and a regulatory framework will function automatically merely because it is logical and comprehensive. To the extent that a regulatory framework curbs ordinary economic incentives, it will be necessary to create and provide other motivating incentives which, in the circumstances, have to be political and patriotic.

Or else, we may have a frame without the motive power.

IN HIS PAPER on Indian Planning D. R. Gadgil begins with an assessment of the performance of Indian planning during the last three Plan periods and a diagnosis of its failures. His main conclusion is that the Government finds it possible to raise financial resources and make them available for financing capital investment and current inputs but is unable to direct and regulate the use of these resources once they pass into private hands. In other words, the Government succeeds best when it directly undertakes large programmes of public expenditure but fails when the planned use of resources lies in or passes into private hands. The reasons for this, in his opinion, are a total absence of (a) a policy frame and (b) a regulatory framework. In fact, he accuses the Government of deliberately avoiding constructing the necessary policy frame and regulatory framework because of the growing influence in the Government of (a) the vested interests with laissez faire bias and (b) the physical or perspective planners, in other words the technocrats, with bias towards automatism of development following investment. In consequence, in his opinion, the Plan resources have passed into the hands of the vested Interest and have been employed in market-oriented uses. The net result of all this is the failure to maintain stability in the economy which has finally jeopardised planned development.

Emphasis on Stability

Gadgil attaches great importance to maintaining stability in the economy. In his opinion, the most serious aspects of the failure to maintain stability are (1) unchecked increase in internal prices, (2) growing imbalance in the external balance of payments and (3) growing inequality in the economy. He rejects the notion, currently prevailing, that mere avoidance of deficits financing supplemented by greater financial discipline and appropriate monetary and credit policies will be able to check the rise in prices. In his opinion, in order to achieve price stability, the essential requirements are (a) restraints on domestic consumption and (b) a price policy supported by much greater control over production, trade, speculation and banking. The growing imbalance in the external balance of payments is, of course, due to growth in imports and failure to promote exports. Gadgil blames the growth in imports on (a) failure to direct agricultural production effort into required channels and to restrain and rationalise demand leading to unrestrained imports of foodgrains and cotton and (b) unplanned growth of private industry committing the economy to compulsive imports to maintain established industry. Failure to promote and expand exports has occurred in both the traditional and the new sector. The failure in the traditional sector is primarily due to the inelastic world demand for our traditional exports, inelastic supply of agricultural production and rising internal demand for the exportable commodities. Gadgil suggests that to counter the latter two factors, it was essential to restrain the domestic consumption of exportable commodities. The failure to promote exports in the new sector, in his opinion, is primarily due to quantitative, restrictions on imports unaccompanied by their detailed allocation and control over their prices and consequent failure to eliminate extra and illicit gains to private parties in the business. To do this, again, would require comprehensive regulation of private industry. Gadgil attributes, the growing inequality in the economy to the so-called production-production-before-distribution policy, failure to achieve transmission and diffusion of the forces and effects of development beyond the initial points of injection, and neglect of the employment aspects of development planning. In his judgment, the growing inequality in the economy is causing not only economic instability but social and political unrest as well.

Following this assessment and diagnosis of the situation, Gadgil, proceeds to prescribe what he calls a corrective approach. The primary objective of his prescription is to achieve and maintain stability. He points that not only is there no conflict between stability and development but that stability is an essential condition for continued and sustained development. The essential elements, in his prescription are (a) an integrated policy frame and (b) a comprehensive regulatory framework.

As Gadgil has spelt it out, the policy for stability has two aspects: (1) a price policy and (2) an income policy. The main instruments of his price policy are (a) a comprehensive system of socialised purchase, storage and sale of all major agricultural products with a view to establishing and maintaining a stable structure of agricultural prices; (b) a widespread network of retail shops under public control for distributing all essential consumer goods and necessary procurement of essential manufactured goods in order to supply the retail shops and (c) control of prices of intermediate and capital goods mainly through detailed allocation of imports and effective elimination of all extra gains which quantitative restrictions on imports make possible. The essential elements of his income policy are (a) stabilisation of agricultural incomes through stabilisation of agricultural prices; (b) avoidance of technological unemployment through deliberate use of existing equipment and skills; (c) giving a deliberate employment orientation to the plan of new investment and (d) checking rising incomes at higher levels through appropriate fiscal measures.

Gadgil advocates a very comprehensive regulatory framework to regulate private trade and industry. The main operative instruments of regulation are: (a) fiscal, monetary and credit policies aimed at reducing the importance and scope of the private capital market, limitation of bank credit and its allocation to different industries; (b) closer scrutiny of the industrial sector, in relation to its import commitment with a view to achieving increasing self-reliance; and, (c) scrutiny and control of the financial affairs of the private sector through a system of public audit.

One might dispute Gadgil's assessment of the past performance of our planning and his diagnosis of its failures. One might also disagree with his policy prescription for the future. In the following, I wish to do neither. Instead, I wish to spell out in some detail certain implications of Gadgil's policy prescription which I believe he has deliberately chosen not to elaborate upon.

Overall Policy Frame

The key element in Gadgil's prescription is an overall policy and regulatory framework which affects all persons both as consumers and producers. The policy aims at restraining consumption and directing production in appropriate channels. Its principal objective is to achieve and maintain stability in the economy. Let us first consider the restraint on consumption. This Is to be achieved through (a) quantitative restrictions in the case of all essential consumer goods of mass consumption and also of intermediate and capital goods in crucial sectors; and, (b) through high excise duties and high prices in the case of all other commodities. It must be emphasised that the restraints on consumption are necessary not only in respect of those commodities of which domestic supplies are short of domestic demand but also where we have a surplus. In the latter case, we need restraints on consumption in order to procure quantities for export. If there are any commodities where there is a domestic surplus but no export market, we shall have to cut down their production. In short, we can not have unrestrained consumption of anything whatever. This is obvious if we must maximise investment.

This is one side of the proposition. The other side is what are called economic incentives. If we are to restrain consumption all round, what are the economic incentives for additional production effort, it is asked. The promise of a brighter future is all right; however, people also need something tangible in the immediate future, it is argued. This is a serious matter and we should examine the implications carefully.

We might divide the question into two parts and first consider, it in relation to the great mass of people. What are the economic incentives that we could provide to these people so that they might put in additional production effort? The consumption of these people would be confined, in any event, by and large to the essential consumer goods of mass consumption. If the overall restraints on the consumption of these goods are accompanied by a more equitable distribution of the given supplies to a majority of these people it might in fact mean higher living standards than at present. This is what Gadgil suggests should be done through proper price and income policies and a comprehensive retail distribution system under public control. If this is successfully achieved this will provide, let us say, the necessary economic incentives to the great mass of people.

Let us next consider the residual minority, a small class but nevertheless important for economic development, namely, the middle and upper middle classes and above all the elite. An equitable distribution of the essential consumer goods does not offer much attraction to these people; they can acquire necessary supplies in any event. On the other hand, a section of this class thrives when the prices and distribution of the essential commodities are not controlled. Therefore, such controls and regulation result in direct disincentives to at least a section of these people. Of greater concern to the class as a whole are the restraints on consumption of non-essential consumer goods. These commodities are non-essential to the mass of consumers but nevertheless constitute essential economic incentives to the class under consideration. Gadgil suggests that the consumption of these commodities should be restrained through appropriate price and income policies supported by necessary fiscal measures. The purpose of restraining consumption of these commodities is not only to make large quantities of them available for export, or to make larger resources available for investment, but also to prevent, what Gadgil has called "warping of the production structure" whereby production of non-essential commodities prospers at the expense of the production of essential commodities of mass consumption.

Vigilance over Perquisites

These are obviously important considerations. Let us, therefore, see what are the necessary steps and what are the implications if consumption of the non-essential commodities is to be effectively restrained. In the first instance, this will require strict scrutiny and vigilance over the perquisites and fringe benefits which private industry and business provide to their employees. This must specifically cover travel abroad on business account. Secondly, the curbs on consumption must extend to foreign businessmen, experts, technicians, diplomatic personnel, visitors, tourists, in short all foreigners. This is absolutely essential if we have to prevent infiltration through hospitality and drink and dinner diplomacy generally. Thirdly, once we effectively curb domestic consumption of the non-essential commodities, we shall have to take steps to prevent the flight of this class and especially of the scientific, technical, managerial and administrative personnel, to other countries. If we must do this effectively, for all practical purposes we shall have to close our doors. There are several good reasons why we should in any case do this immediately. Once we do this, we must turn attention to the question of incentives and conditions of work more conducive to greater effort on the part of these persons. It is obvious that in the circumstance these incentives cannot be economic in the narrow sense of the term. The incentives will have to be political or patriotic.

The need to curb non-essential and ostentatious consumption appears paramount from yet another and wider viewpoint. The central thesis underlying Gadgil's assessment of the present situation and his prescription for the future is that "the present shape and direction of economic policy in India are largely the result of influence of vested interest" and that this needs to be immediately corrected. In this, probably, Gadgil had in mind mainly the vested interests in private business and industry. But undoubtedly he would, readily recognise other vested interests as well. Landowning interest and the bureaucracy are obviously two other important vested interests. We should also recognise the middle and upper-middle classes as an other important vested interest. I may add yet another category, namely, the one represented by the mother cow and her dear sons. Gadgil has not devoted specific attention to the role of landowning vested interests in the present economic situation. Therefore, I shall not here consider the problems that this category presents. I shall also not examine the problems that the mother cow and her sons raise. In any case, these two categories of vested interests fall in a class apart from the other three categories mentioned above, namely, private business and industry, the bureaucracy and the middle and upper middle classes whom we might refer to as the intellectuals. These three categories of vested interests together constitute what we may call the urban vested interests. There can be no doubt that at present economic policy in this country is being shaped and directed by these interests.

The nature of vested interest in private business and industry is well known and well understood. Gadgil recommends a comprehensive regulatory framework in order to curtail and regulate the power and influence of this section. In this framework, the main instruments of operation are: a fiscal policy to reduce the importance of private capital market, a credit policy to limit and allocate bank credit, allocation of imports together with control over their prices, requisition of essential consumer goods of industrial production for retail distribution, procurement of production for export, price control over intermediate and capital goods in crucial sectors, high excise duties on non-essential goods to be released for domestic consumption, high income taxes for curbing incomes at higher levels and public audit of the financial affairs of private business and industry. In short, what Gadgil suggests is that the fiscal policy should ensure that there are little or no savings in the private sector so that it would be possible to direct production through allocation of bank credit. Production should be directed into production of standardised essential consumption goods and when they are produced they should be procured for retail distribution under public control. The detailed allocation of imports and procurement for export imply that the import and export trade should be in the public sector. Finally, to eliminate all illicit gains arising out of regulation, Gadgil suggests public audit of the financial affairs of private business and Industry.

A New Vested Interest

This is a formidable apparatus and there is little doubt that if it could be operated successfully, it would effectively curb the power and influence of vested interests in private business and industry. However, we should recognise that in the process we may create another vested interest, namely, the bureaucracy. We should, therefore, beware of the consequences. Here I am not concerned about the consequences of replacing completely one vested interest, namely, private business, by another vested interest, namely, the bureaucracy. That is not an immediate prospect. I am concerned with a prospect which is worse and more immediate, namely, that the two vested interests private business and bureaucracy, may co-exist, join hands and undermine the entire regulatory apparatus. Many consider this to be the real danger in setting up an all-pervasive regulatory framework. Gadgil has blamed our failure in the past on what he calls the total absence of a policy frame and a regulatory framework. He might be right. However, there are others who believe that there was in fact a policy and also a regulatory framework though undoubtedly not as comprehensive as Gadgil suggests. In their opinion, our failure has been mainly in implementing the policy and operating the regulatory framework and that the collusion between the bureaucracy and the private business was the main reason. Whether or not this has in fact happened in the past, the danger is undoubtedly real.

In their collusion, the bureaucracy and private business are aided and abetted by the middle and upper middle classes or what we may broadly term the intellectuals. A large section of what we superficially recognise as bureaucracy or private business indeed belongs to this category. Besides, it includes all liberal professions, doctors, lawyers, teachers, writers, artists and above all the politicians. There is no need to emphasise the importance of this class in influencing and shaping crucial policy decisions. Nevertheless, this class is easily the most vulnerable from the economic pressures which the hard core of the other two categories can bring to bear. When this class succumbs, it is the signal for loosening of the policy frame. The bureaucrat then becomes bold and joins hands with the private business and undermines the entire regulatory framework. It is thus that the private business and industry ultimately wins and we have all the consequences that Gadgil has so well emphasised.

The middle and upper middle classes thus have a crucial role to play in upholding the policy frame and in ensuring that the bureaucrat operates the regulatory framework effectively. It is, therefore, essential that these classes are adequately protected or insulated from the economic pressures that private business and industry can exert. The vested interest of these classes is in a comfortable life and that is their undoing. It also alienates these classes from the general mass of people and places them within easy influence of private business and industry. It is in this context that the importance of curbing non-essential and ostentatious consumption becomes paramount. If this is done, it will effectively reduce the vulnerability of the middle and upper middle classes and the regulatory framework, such as the one suggested by Gadgil, will have a better chance to function successfully.

Is Stability Possible?

Let me now turn to the second important element in Gadgil's prescription, namely, the emphasis he lays on stability. Gadgil lays primary stress on achieving and maintaining economic stability-stability in internal prices, stability in external balance of payments and stability, if not progressive reduction, in inequality. He thinks that stability is an essential condition for continuous and sustained development. This is a matter which may be debated at various levels of sophistication, mathematical, statistical or philosophical. However, even if one agrees with Gadgil regarding the importance of stability, one might wonder whether even with a comprehensive policy frame and a regulatory framework of his prescription, it will be possible to achieve and maintain stability of the kind he desires. There are several elements in the situation which lie outside even the most comprehensive framework of policy and regulation and which must, therefore, be treated as exogenous. First and foremost of all, agricultural production is liable to great seasonal fluctuations and these do not affect all crops and all regions uniformly. These are bound to affect income distribution between rural and urban areas, between different regions and between different categories of agricultural producers. These will inevitably affect the price and distribution systems. Some of the fluctuations will also affect industry and some others will affect directly the imports and exports. The export market on which we must depend in order to achieve and maintain stability in our external balance of payments lies entirely outside our regulatory framework and any changes in the export market situation will have to be inevitably absorbed within our economy. The extent and nature of external assistance that we may receive depends on an international political situation which is extremely fluid and hence liable to great fluctuations. Gadgil realises this and, rightly suggests that our strategy and route of industrialisation must not be put in a single, rigid frame but must provide for a series of possible situations. As he has suggested, it must aim to reach self-reliance early or late, with large or no external assistance and must define a correspondingly narrow or wide basis. These are the compulsions of circumstances and are bound to make the task of maintaining stability a difficult one. It is, therefore, difficult to agree with Gadgil when he says that in a developing economy, policy should be geared to steady maintenance of stable conditions and not to dealing with constant ups and downs.

In this context the analogy he draws between economies of poor developing countries and the war economies of industrial countries is relevant. The purpose of the policy and regulatory framework in the war economies may, in one sense, be described as steady maintenance of stable conditions. However, it seems to me that a more appropriate description of their purpose would be to deal with ups and downs of a constantly changing situation. The primary function and purpose of a comprehensive and effective regulatory framework in a poor developing economy might, be described in the same manner, namely to meet the ups and downs originating in exogenous factors beyond our regulatory framework-to make the best use of unexpected favourable circumstances and to absorb the shocks of unforeseen adverse developments.

The analogy with the war economies of the west is, however, not complete and it seems to me that there is missing an important element in our economy which must be supplied for the analogy to hold good. The war economies of the western countries functioned under comprehensive and effective regulatory framework and hence these economies certainly provide us with a model to adapt and adopt. However, their success was due not only to a comprehensive regulatory framework but also, and at least in equal measure, to the incentives which the war provided. To put the matter somewhat differently, the regulatory framework functioned effectively not only because it was comprehensive but also because the war provided the necessary motivating incentives. These incentives were dot economic in the narrow sense of the term. They were essentially political and patriotic. To make complete the analogy between our economy and the war economies of the western countries, we shall, therefore, need not duly an integrated policy and a comprehensive regulatory framework but also the necessary political and patriotic incentive. Under peacetime conclusions, the nature of these incentives and the manner in which they may be supplied are such that it may be more appropriate to seek an analogy with communist economies of the developing countries rather than with the war economies of the western industrial countries.

Other Incentives Necessary

Gadgil is probably justified in his criticism of the physical and the perspective planners or the technocrats who presumably believe that once a given capital investment is achieved, the economy junctions automatically according to plan and leads to the development aimed at. As Gadgil has emphasised, in order to succeed, a plan of investment must be supplied with the necessary policy and regulatory framework. However, it will be equally wrong to suppose that a policy and a regulatory framework will function automatically merely because it is logical and comprehensive. To the extent that a regulatory framework curbs ordinary economic incentives, it will be necessary to create and provide other motivating incentives which, in the circumstances, have to be political and patriotic; or else, we may have a frame without the motive power. Gadgil has not emphasised this point because, I suppose, he wanted to keep his observations within the limits of a professional economist.

DEMOCRATIC DECENTRALIZATION

V.M. Dandekar

I propose this evening to speak on certain aspects of democratic decentralization. As you are aware, the concept of democratic decentralization, also known as the panchayati raj, was established ten years ago with the publication of the pioneering report of the Study Team under the distinguished chairmanship of Shri Balwantrai Mehta. During the ten years, since the publication of this report, considerable progress has been made broadly along the lines recommended by the Study Team. There are at present established and functioning panchayati raj institutions in twelve states, namely, Andhra Pradesh, Assam, Bihar, Gujarat Madras, Maharashtra, Mysore, Orissa, Punjab, Rajasthan, Uttar Pradesh and West Bengal. There are now established in the country in all 250 district level bodies and about 3500 intermediate level bodies between the district and the village. I do not propose to survey the constitution or functioning of these bodies or to assess their success or failure in the tasks assigned to them. Instead I intend to raise a few rather broad issues of principles and policy in the context of this subject. Let me recapitulate briefly its background.

The origins of the concept of democratic decentralization lay in the need to reorganize the administration in the districts. The structure of administration developed by the British was based upon the district as the principal unit with the District Collector as the government's principal representative in touch with the people. Besides being in control of the administration of law and order and revenue in the district, the Collector had a co-ordinating responsibility for the activities of all the departmental agencies within the district. In the hierarchy of administration, he enjoyed status and much delegated power which gave him considerable influence over the local population. For the limited functions with which he was especially responsible, there was no need for a greater degree of popular co-operation than the Collector could thus command.

With the country becoming independent and with the emphasis shifting from law and order to economic development, it became evident that the official agencies by themselves could not carry even a modest programme of rural development and that therefore greater initiative and participation on the part of the people in the development programmes was needed. This was recognized in the First Five Year Plan. However, little progress was made along these lines during the First Plan period. The Second Five Year Plan again emphasized the need to create, a well organized democratic structure of administration within the district and suggested that 'the functions of the popular body should come to include the entire general administration and development of the area other than such functions as law and order, administration of justice and certain functions pertaining to revenue administration.' It was against this background that the Committee on Plan Projects of the National Development Council appointed in January 1957, a study team under the chairmanship of Shri Balwantrai Mehta to study and report on the working of the Community Development Projects and the National Extension Service and also to examine the question of reorganizing the district administration. The Study Team submitted its report by the end of 1957. After reviewing the working of the several local bodies and the development committees then existing in the districts, the Study Team recommended that all these bodies be replaced by a single representative and vigorous democratic institution to take charge of all aspects of development work in the rural areas.

The issues of principles and policy which I propose to raise and discuss centre around this central recommendation, namely, to establish in each rural area 'a single representative democratic institution to take charge of all aspects of development work' in the area. I have two issues to discuss. Firstly, I shall question the advisability of setting up a single institution to take charge of all aspects of development work. I shall argue that this leads to a kind of concentration or authority and power which is not desirable. Secondly, I shall raise questions regarding the precise nature and content of representative democracy desired in panchayati raj institutions. I shall argue that the parliamentary form of democracy in these institutions leads to an undue and unjustifiable concentration of power in the hands of a single political party which is not desirable.

The proposal to establish in each area a single body to take charge of all aspects of development work in the area arises naturally because our concept of decentralization is essentially a concept of geographical or regional decentralization - from the state level to the district level, from the district level to the block level and from the block level to the village level. We have, therefore, discussed whether decentralization should be carried upto the district level or whether we should carry it further to the block level. But once it is decided that it should be carried to the district or the block level, all appear agreed that all junctions at that level should be concentrated in the hands of a single body. This achieves geographical or regional decentralization Nat inevitably leads to functional concentration. I shall presently argue that this is undesirable and suggest that our notion of democratic decentralization should extend to both regional as well as functional decentralization.

However, before I do this, I wish to clear a small point which has received much attention namely the regional level to which decentralization should proceed. The Balwantrai Mehta Study Team had suggested that decentralization should be carried to a level below the district, namely the development block, and that real authority should be devolved on bodies called Panchayat Samitis to be created at the block level. They had also recommended the creation of Zilla Parishads at the district level but their functions were to be largely advisory. There has been considerable discussion on this point and naturally different state governments have taken different views. For instance, the Government of Maharashtra took the opposite view and decided to devolve real authority on Zilla Parishads established at the district level while caking the Panchayat Samitis mere regional sub-committees of that body. I suppose it will be readily agreed that if we, wish to devolve authority at a lower level, we shall be able to devolve only smaller authority. On the other hand, if we wish to devolve greater authority it is obvious that this will have to be done at a higher level. The reason is simple. At lower levels, such as the Panchayat Samiti level, the average calibre of the elected elements as also of the technical services cannot be high enough presently to enable greater devolution of power at that level. Thus, if the Government of Maharashtra did not decentralize teal authority at the block level and instead restricted it to the district level, I believe, that it has devolved much greater authority than was envisaged by the Balwantrai Mehta Study Team. Personally I am in favour of decentralization of greater authority and, hence approve the position taken by the Government of. Maharashtra on this point.

However, from what I have earlier said, it follows that if we devolve greater authority at a higher level, it results in a greater concentration of authority and power at that level. This is best illustrated by the situation in Maharashtra. As a consequence of its decision to devolve real authority at the district level and to devolve on the district bodies all functions, other than law and order and administration of justice, which legitimately fall within the jurisdiction of a district administration, the Zilla Parishads in Maharashtra are fairly strong bodies with considerable financial and personnel resources. A Zilla Parishad has an annual budget of about 2 to 3 crores of rupees and service personnel of about 2 to 3 thousand. Practically all district-level functions in the fields of education, health, public works, agriculture and co-operation are devolved on these bodies, There can be little doubt that this is genuine democratic decentralization. However, this has inevitably led to an excessive concentration of power and authority within the district. This happens in both the political and the administrative wings of the Zilla Parishad. On the one hand, there is an excessive concentration of power in the hands of the President of the Zilla Parishad which leads to potentially undesirable political activity which all such concentration of political power leads to. There is similar and parallel concentration of authority in the hands of the Chief Executive Officer of the Zilla Parishad. He belongs to the senior cadre of the Indian 4dministrative Service and all executive power of the Zilla Parishad vests in him. All service personnel in all departments including officers belonging to Class I of the State service working with the Zilla Parishad are subordinated to him. This has many undesirable consequences the most serious of which is the snapping of the connections between the technical staff working in the districts and the senior officers in the technical departments functioning at the divisional or state levels. Concentration of political power and administrative authority in such measure at a local level is undesirable.

Thus we appear to be faced with a dilemma. If we decentralize authority to a lower level, we cannot effectively decentralize much authority. On the other hand, if we wish to 4ecentralize considerable authority, we have to do it at a higher level and that leads to an excessive concentration of power and authority at that level. The dilemma arises because, as I have earlier said, we have limited our notion of democratic decentralization to regional decentralization. The dilemma can be resolved and a more desirable pattern of decentralization can be evolved if we combine with regional decentralization the concept of functional decentralization. Let me illustrate again with reference to tin functioning of Zilla Parishads in Maharashtra.

The Zilla Parishads in Maharashtra function through a set of seven committees, namely, the Standing. Committee and six subject committees. The six subject committees are chaired by three chairmen and they have fairly defined functions. And powers in relation to their subjects. Nevertheless, both politically and within the hierarchy of the Zilla Parishad, they have a status junior to that of President. This leads to effective concentration of authority in the hands of the President. In the same manner, the heads of respective technical departments in the district act as secretaries to the subject committees and are answerable to those committees. Nevertheless, in the administration they are subordinated to the Chief Executive Officer and that leads to excessive concentration of authority in the hands of the Chief Executive Officer. What is, therefore, needed is a certain degree of functional decentralization so that the chairmen of the subject committees may not be subordinated to the President of the Zilla Parishad nor may the heads of technical departments in the district be subordinated to the Chief Executive Officer of the Parishad.

This may be done by suitably federating the district level subject committees into a divisional level board for districts within a division. Consider for instance, the education committee of a Zilla Parishad. At present it functions within the framework of the Zilla Parishad and hence is subordinated to the President on the one hand and the Chief Executive Officer on the other. Suppose now that we constitute for a group of six or eight districts, a divisional education board consisting of the chairmen of the district education committees and the district education officers who incidentally are the secretaries of the district education committees. The elected members of this board may elect a President to preside over the board and the divisional officer of the Education Department may function as its secretary. The board may then be entrusted with the responsibility of implementing the educational policy and programme in the division with the district subject committees functioning as its subsidiaries. Similar divisional boards may be established in other subjects such as public health, public works, agriculture, co-operation, etc. The Presidents of all such divisional boards may enjoy status equal to that of the Presidents of the Zilla Parishads. At the same time, all district heads of technical departments together with all junior cadres in these departments should be placed firmly within the respective technical departments and not subjected to the control of the. Chief Executive Officer of the Zilla Parishad.

Programmes of economic development are becoming, increasingly technical and it is necessary that the popular representatives at the local level are sufficiently acquainted and informed on these aspects along with questions of broad social-economic policy in these matters. The divisional boards in different subjects will offer the necessary forum for the purpose. Firstly, the divisional board in a subject will offer a forum where the chairmen of the district subject committees may meet and get acquainted with larger issues of development policies and programmes in their respective subject fields. Secondly, the divisional board will offer a forum where the district heads of the department, under the guidance of the divisional head, may meet the elected popular representatives on level and may inform them on the technical aspects of the programmes. At present, this does not happen in sufficient measure. It is true that the district head of a department is the secretary of the relevant district subject committee. However, in that committee he is in the minority of one and has to face the political pressure of the rest of the committee. He is also isolated from his equals and seniors in the technical department and naturally submits to political pressure rather than squarely placing before the committee the technical aspects of the question. In the divisional board, the district heads of the department in all the districts will sit together and will in addition have the benefit of advice of their senior namely the divisional head. They will have, therefore, no difficulty in putting before the board the technical aspects of the question under discussion. Even though they may have no voting power in the board, it will be important that the decisions made by the board will be fully informed. Finally, the proposal will re-establish the link between the district level service cadres and the divisional and state level senior officers in the department which at present is very much snapped. This is necessary for more than one reason. The most important reason is that a direct link with the department is regarded a necessary element of the satisfactory service conditions by the service personnel. We have paid much attention, to creating desirable political conditions in the local level administration but too little attention to creating right service conditions for the public servants. This has led to much heart-burning, resentment and frustration in the service personnel in the districts so much so that in general they have not been willing partners in the, process of democratic decentralization.

It may be asked as to what will be the function of the Presidents of the Zilla Parishads under such a reorganization. There is little doubt that in the proposed reorganization, the political power of the presidents and the administrative authority of the Chief Executive Officers of the Zilla Parishads will be greatly reduced. In fact, that is the intention. It seems that the functions of the Chief Executive Officers may be reduced so greatly that it may not be necessary to have IAS officers to man these ranks, But the Presidents of the Zilla Parishads will still have important functions to perform. Their chief function will be as chairmen of the district standing committees. On the district standing committee sit, besides other elected members, all chairmen of district committees. The standing committee, therefore, provides the natural forum to review the progress of implementation of various policies and programmes in the district. It should be the principal function of the Presidents of the Zilla Parishads to do this and bring to the notice of the appropriate divisional subjects boards any shortcomings discovered. The Presidents may also be held responsible for the maintenance of the financial accounts of the Zilla Parishads. They may also be responsible for providing the general establishment for the functioning of the several subject committees in the district and should be in charge of the same. In short, though the Presidents of the Zilla Parishads may be very much reduced in their status and political power, they will still have useful coordinating functions to perform.

Let me sum up. I do not approve of the idea of establishing for each area a single body to be in charge of all development functions. Besides leading to excessive concentration of political power and administrative authority in the area, such a body tends to become political in one of its wings and bureaucratic in the other, tails to bring together the two wings in harmonious relations and generally neglects technical aspects, of development programmes. I therefore plead for a functional decentralization at a higher level whereby several specialized agencies may operate in each larger area. In the context of the constitution of Zilla. Parishads in Maharashtra, I suggest that divisional boards may be established in each subject which will have the real authority in relation to that subject. The existing subject committees in the districts will continue to function as at present except that they will be answerable to the divisional boards rather than the Zilla Parishad. The principal function of the Zilla Parishads should be to co-ordinate the work of the several subject committees in the district. This leaves unaffected the Panchayat Samitis and they will continue to function as at present. These are matters of detail and I have illustrated them with reference to the constitution of Zilla Parishads in Maharashtra simply because I am familiar with their working. Once the principle of functional decentralization is accepted, the details will have to be adapted to the structure of panchayati raj institution in each state.

Let me now turn to the second of the two issues which I said I would discuss, namely, the precise nature and content of representative democracy to be desired in the panchayati raj institutions. Because of our notion that democratic decentralization means regional decentralization, we have tended to look at the new institutions as local governments, that is governments at the district or block level, and in fact called them panchayati raj institutions. Consequently, we have implicitly understood that the nature and content of representative democracy in these institutions should be, as far as possible, the same as at the state level. Therefore, ideally, these bodies should be constituted on exactly the same lines as those on which the state assemblies are constituted. For instance, in Maharashtra, the Zilla Parishad in a district consists of between 40 and 60 councilors chosen by direct election from single member constituencies determined so that there is one councilor for not more than 35,000 of the population. In other states the election to the district body or the intermediate level body is not direct. However, whether the election is direct or indirect, the principle of the rule of the majority is everywhere accepted. Some people believe that this is inevitable in any democratically constituted body and point out that after all the governments at the centre and in the states also function on the same principle. I shall argue that the analogy between the governments at the centre and in the states and the panchayati raj institutions is not correct and that therefore the content of representative democracy in the panchayati raj institutions has to be different from the simple rule of the majority.

The difference between the governments at the centre and in the states on the one hand and the panchayati raj institutions on the other will become clear if we examine closely the functions of the panchayati raj institutions. These functions are of course a part of the functions of the state governments which these governments are pleased to devolve on the panchayati raj institutions. The state governments have two major functions to perform, namely, legislative and executive, besides of course administration of justice. The legislative function is performed in the legislative assemblies and here the majority rules supreme. The executive functions are performed by the permanent civil service under the direction of the government of the day but which nevertheless is expected to function above the party lines. The cabinet provides the link between the legislature and the civil service and has a duel role to play. When it functions in the legislature, it seeks legislative approval to a certain policy and programme. It is for this reason that it must constitute a politically homogeneous group and must command the confidence of the majority in the legislature. In other words, it should be chosen from the majority party or a majority coalition. In this role, the cabinet avowedly functions in the interests of the majority party and tries to promote its policies and programme. In its functions as the executive at the head of the civil service, its purpose is to give effect to the policy and programmes approved by the legislature. In this function, it is expected that it will not promote the interests of the party and that instead it will function on non-party lines. In other words, it is understood that though the policy and the programmes may be of the majority party, their implementation shall not be affected by party interests.

It is a part of some of these functions that the state governments have powers to devolve on the panchayati raj institutions and we should be clear about which these functions are. At the outset we should note that no part if the legislative functions of the state governments are devolved on the panchayati raj institutions and that therefore whether they are elected directly or indirectly, the Zilla Parishads or the Panchayat Samitis are not legislative bodies at the local level. The only functions which are devolved on the panchayati raj institutions are a part of the executive functions of the state governments. Therefore, though democratic, the panchayati raj institutions are no more than executive organs of the state governments charged with the responsibility of implementing policies and programmes of the state governments in certain specified fields. It follows that the panchayati raj institutions must not conduct their affairs on party lines and that therefore the content of representative democracy here has to be different from the parliamentary democracy operating at the centre or in the states.

Something of this principle has been admitted in the constitution of the Zilla Parishads in Maharashtra. I am not aware, if it has also been accepted in the constitution of panchayati raj institutions in any other states. Let me therefore illustrate my point with reference to the constitution of Zilla Parishads in Maharashtra, I have already mentioned that in Maharashtra, the Zilla Parishad in a district consists of between 40 and 60 councilors chosen by direct election from single member constituencies. The Zilla Parishad is presided over by a president who is elected by the elected councilors. However, the Zilla Parishad does not function on the basis of the cabinet system; that is to say, the President does not select his own team or cabinet, as does the Chief Minister in the state government, to run the affairs of the Zilla Parishad. Instead the Zilla Parishad functions by means of a number of committees, namely, Standing Committee, Finance Committee, Workers Committee, Agriculture Committee, Co-operatives Committee, Education Committee and Health Committee. If we neglect minor details, each subject committee consists of seven councilors elected by the Zilla Parishad from amongst its councilors and what is important, the elections to the committees are held in accordance with the system of proportional representation by means of a single transfer able vote. This ensures that on all the committees, all parties and groups in the Zilla Parishad are given representation in proportion to their strength. The Standing Committee is presided over by the President of the Zilla Parishad. The remaining six committees are presided over by three chairmen, each chairman presiding over two committees. As already mentioned, the three chairmen are not selected, by the President and hence do not constitute his cabinet. They are directly elected by the Parishad from among its councilors.

Let us now see to what extent this system ensures functioning of the Zilla Parishads on non-party lines as we desire it to be. From this standpoint, I suppose, it will be agreed that the committee system is to be preferred to the cabinet system, especially if the committees are elected on the principle of proportional representation. Therefore, undoubtedly this is a commendable feature of the constitution of the Zilla Parishads in Maharashtra. However, once the principle is accepted that the affairs of the Zilla Parishads should be conducted on non-party lines, it seems that the principle of proportional representation could be extended further and beyond the elections of the subject committees. Though the business of the Zilla Parishads is conducted through the subject committees the executive and financial powers of these committees naturally vest in their chairmen. Consequently the President and the three chairmen of the committees wield the real executive authority of the Zilla Parishads. In fact they are the four full-time elected office-bearers of the Zilla Parishad and they receive suitable honoraries and other benefits on that account. It is therefore important that the chairmen are not selected by the President and that they do not constitute his cabinet. As already mentioned, the three chairmen, like the President himself, are all elected by the Zilla Parishad. This is good as far as it goes. But it does not ensure that all parties and groups in the Zilla Parishads share its authority. In fact, because the, President and the three chairmen are elected by the Zilla Parishad in separate elections, almost invariably all of them belong to the majority party or the coalition. Consequently, the President and the three chairmen constitute in effect a patty cabinet.

This is undesirable. As already mentioned, the primary purpose of the panchayati raj institutions is to implement the policy and programmes of the state government at the local level. With all the office-bearers belonging to the majority party, the implementation of the programmes are unavoidably affected and influenced by party considerations. This is contrary to the principles of a healthy democracy. Secondly, the majority party gets undue credit for all the development programmes and activities in the district whereby it may entrench itself at the local levels. This is undesirable for the functioning of a healthy democracy.

It seems to me therefore that it is necessary to ensure that the authority of the Zilla Parishads is appropriately shared by all parties and groups in the Zilla Parishad. In the context of the constitution of the Zilla Parishads in Maharashtra, this may for instance be achieved by changing the manner in which the President and the chairmen of subject committees are elected. As already mentioned, the committees are elected on the principle of proportional representation but the President and the three chairmen are elected in separate elections, What is necessary and desirable is that the President and the three chairmen are elected in a single election on the principle of proportional representation. This will ensure that even a party or a group claiming no more than one-fourth of the total strength in the Zilla Parishad, will get one of the offices and that the authority of the Zilla Parishads will be appropriately shared between all parties and groups. This is as much as can be done to place the functioning of a democratic body, such as the Zilla Parishad, on non-party lines.

It may be felt that if the office-bearers of the Zilla Parishad are elected on the principle of proportional representation, they will not form a homogeneous team and hence that the administration of the Zilla Parishad will be unstable. It seems to me that these fears are based on a false analogy between the Zilla Parishads and the governments at the centre or in the states. As I have emphasized, Zilla Parishads are not legislative or policy-making bodies. Their function essentially is to implement the policy and programmes of the state government. For this reason, the unanimity and homogeneity of the office-bearers is much less important here than in the governments at the centre and the states. Moreover, once it is understood by all concerned that the office-bearers of the Zilla Parishads shall not belong to a single party or a coalition, the political activities in Zilla Parishads will take a different orientation. There will be no more horse-trading and changes in the party affiliations in search of the office. In consequence, administration in the Zilla Parishads will be more and not less stable. Further, if the principles is accepted that the office bearers of the Zilla Parishads shall not belong to a single party which may be the ruling party in the state governments or may be one in opposition there, the relations between the Zilla Parishads and the state governments will be placed on a healthier footing and the Zilla Parishads will be able to assert their rightful autonomy.

I have illustrated my point with reference to the constitution of Zilla Parishads in Maharashtra. However, the principle I wish to enunciate is of general applicability and the details should be adapted to the needs of the panchayati raj institutions in each state. What is needed is to ensure that the functioning of the Zilla Parishads or the Panchayat Samitis on whom real authority has been devolved, should be placed on a democratic but at the same time non-party footing.

In most states, real executive authority has been devolved not on the Zilla Parishad at the district level but on the Panchayat Samiti at a lower level. Members of the Panchayat Samiti are also usually elected not by direct election but by indirect election. There is also often not more than one full-time elected office-bearer. Let me therefore indicate how the principle of proportional representation may be applied to such cases. I may again illustrate the point with reference to the structure of Panchayat Samitis in Maharashtra the members of which are also elected by indirect election and which has only one full-time elected office-bearer. In Maharashtra, the Panchayat Samitis are joint committees of the Zilla Parishad and the village panchayats in an area. Their function is to implement and execute all programmes of the Zilla Parishad in that area. On a Panchayat Samiti, the Zilla Parishad is represented by all its councilors elected from the Panchayat Samiti area. The village panchayats in that area are represented by village sarpanchas whose number is fixed at twice the number of the councilors of the Zilla Parishad sitting on the Panchayat Samiti. Thus if there are five Zilla Parishad councilors sitting on the Panchayat Samiti, the village panchayats are represented by ten sarpanchas. This is the normal size of a Panchayat Samiti. The sarpanchas sitting on the Panchayat Samiti are elected from among all the sarpanchas in the area. For the purpose of this election, the members of all the village panchayats in the area constitute the electoral college. The village panchayats in the area are divided into single member constituencies and the members of the village panchayats in each constituency elect one of their sarpanchas to the Panchayat Samiti. The Panchayat Samiti is presided over by a chairman who is elected by the Panchayat Samiti. There is also an elected vice-chairman but he has no real functions and he may be neglected for the purpose of the present discussion.

Let us now examine to what extent and in what manner the functioning of the Panchayat Samiti may be placed on a democratic but nevertheless non-party lines. The Panchayat Samiti as above constituted has only one office-bearer namely the chairman. We cannot therefore improve upon the system by which the Panchayat Samiti elects him by a majority vote. It is also not possible and desirable to have more elected office-bearers for such a small body as the Panchayat Samiti charged with the responsibility of administering programmes in a small area. Therefore, what needs to be examined is the membership of the Panchayat Samiti and see whether we could ensure on it appropriate representation to all parties and interest groups in the area.

As I have mentioned, the Panchayat Samitis are conceived as joint committees of the Zilla Parishad and the village Panchayats. On the Panchayat Samiti, the Zilla Parishad is represented by its councilors elected from the area. These councilors are elected by direct election from single-member constituencies. It is well-known that in this system of elections, the majority party usually secures a more than proportionate representation in the elected body. However, this is unavoidable because in spite of this fault, the single-member constituencies have obvious advantages in such a large operation as a direct election involves, The village panchayats are represented by sarpanchas elected by an indirect election. Their number in the Panchayat Samiti is double that of the Zilla Parishad councilors sitting there. Therefore, if we can improve the system of indirect election by the members of the village panchayats and ensure a better representation to all parties and interest-groups appearing in village panchayats, that might cause a substantial improvement in the character of the Panchayat Samitis. With this in view, I suggest the following two measures.

My first proposal is simple. At present, the election to the Panchayat Samiti is confined to the sarpanchas. This is unnecessary and undesirable. As a class, the sarpanchas represent overwhelmingly a specific class of the rural society and represent very inadequately the other classes. Therefore, all members of the village panchayats who constitute the Electoral College should be considered eligible for the election to the Panchayat Samiti. This is a simple proposal but may nevertheless change the present composition of the Panchayat Samitis considerably. However, to cause the full change that is necessary, a more radical change in the system of indirect election will have to be accepted. At present the village panchayats in the area are divided into single-member constituencies and the members of village panchayats in each constituency elect one member to the Panchayat Samiti. This naturally secures a more than proportionate representation to the majority party. This needs to be corrected, It may be clone by treating all the village panchayats in the area as a single multi-member constituency and requiring their members to elect the requisite number of members to the Panchayat Samiti on the principle of proportional representation. Thus, if the village panchayats are to send 10 members to the Panchayat Samiti, all the members of the village panchayats as a single constituency should elect to members together on the principle of proportional representation. If this is accepted, parties and interest groups having not more than even 10 per cent of the strength in the village panchayats will be able to send a representative to the Panchayat Samiti. If this appears unmanageable, the village panchayats may be grouped into two multi-member constituencies each electing say five members. This should certainly be manageable. I am aware that this system of election is more complicated than the system of single-member constituencies. However, I do not think that this is impracticable in a system of indirect election where the electoral college does not comprise more than a few hundred voters. If the system is adopted, it will also take away much of the stigma attaching to the indirect elections.

The burden of my theme is that in order to achieve a democratic decentralization of the executive authority of the state government, it will not do to establish at the local levels democratic bodies which operate on the principle of the rule of the majority party because that will mean that the implementation and execution of the policies and programmes of the state governments will be allowed to be influenced and affected by party considerations. What is necessary is to establish democratic bodies but to ensure that in their constitution they represent as far as possible all parties and interest groups in proportion to their strengths and that the authority of these bodies is appropriately shared between them.

Let me summarise the two proposals I have made. Firstly, I suggest that the concept of democratic decentralization should not remain confined to regional decentralization hut should extend to include functional decentralization. I have suggested that if we combine suitably the two concepts of regional and functional decentralization, we shall be able to avoid under concentration of political power and administrative authority in single institutions and in single persons. My second proposal is that we should introduce in these institutions the principle of proportional representation wherever possible so that the authority of the panchayati raj institutions will he shared by all political parties and interest groups in that area. I believe that these two proposals will achieve the democratic decentralization more fully and more meaningfully.

(From 'Democratic Decentralization' a lecture delivered by Prof. V.M. Dandekar at Harold Laski Institute of Political Science, Ahmedabad on December 3, 1967)

STATISTICAL REFORM OF THE SYSTEM OF EVALUATING RESULTS OF MASS-CONDUCTED EXAMINATIONS*

V. M. Dandekar

The author examines the recommendations made by the University Grants Commission's Committee on Examination Reform as regards the marking of examination scripts, and particularly the statistical methods employed by the University of Gauhati. He then puts forward a comprehensive proposal for a statistical reform of the examination system.

The unsatisfactory character of mass-conducted examinations is generally recognized. It is also recognized that so long as the examinations continue to be mass-conducted, any reform of the system will have to be based ultimately on certain statistical principles. Numerous proposals has been made from time to time towards this end. The Committee on Examination Reform appointed by the University Grants Commission has also made certain recommendations along the same lines.[1] Since the publication of the report of this Committee, a certain amount of experience has also become available on the use of these principles, through their application in the University of Gauhati (Taylor, 1963; Tluanga arid Taylor, 1964). The purpose of this paper is to examine the recommendations made by the UGC Committee and the specific statistical principles employed by the University of Gauhati and to put forward a comprehensive proposal for a statistical reform of the examination system.

The UGC recommendations

The relevant recommendations made by the UGC Committee are the following:

> No. 7. The present methods of marking examination scripts and of combining and tabulating marks in university examinations without reference to recognized statistical procedures are not satisfactory. The procedures will have to be developed to make marking and combining of marks more objective.
>
> No. 9. In universities where more than one media of examination exists, the examiners in the different media (at least head examiners) should meet and define the standards and spread of marks to be adopted in the evaluation work undertaken by them, in order to avoid variation in the marking of scripts.

* Reprinted from Indian Educational Review Volume 3, Number 1, January 1968

1. Report on Examination Reform. University Grants Commission. New Delhi, 1962, p. 45.

There are thus three distinct issues: (i) the marking of scripts in each individual subject; (ii) ensuring comparability between the marking of scripts in the same subject but answered in different media; and (iii) the combining of marks secured by a candidate in the different subjects offered by him so as to obtain a 'Grand Total'. We should first understand what the problems in each of these issues are.

There are fundamental questions regarding the evaluation or grading of the performance of different candidates in a given test. These problems exist even if the test is being administered by a single examiner and only to a small number of candidates. Here we shall not concern ourselves with these problems. But there are special problems in this respect in a mass-conducted examination in which several hundred or, in fact, several thousand candidates are involved. The problems arise from the fact that the answer scripts of all the candidates cannot be examined by one and the same examiner but that they have to be distributed over a large number of examiners. Hence, in marking scripts submitted by different candidates in the same subject in a mass-conducted examination, the problem is one of ensuring comparability between different examiners. Ensuring comparability between the marking of scripts in the same subject, but answered in different media, is only a special case of the same problem aid, if a workable solution is found for (i), the same may be extended to (ii). The problems of combining marks secured by a candidate in different subjects so as to give a Grand Total are of a different nature. The purpose of the Grand Total is to serve as a measure of aggregate performance. The basic problem here is one of determining the relative importance or 'weight' to be attached to performance in different subjects. Once this is determined or decided upon, the marks secured in different subjects can be combined into a single measure of aggregate performance. Let us now examine the three sets of problems, one by one.

Non-comparability between marking by different examiners

The basic factors resulting in non-comparability between marking by different examiners are the subjective judgments and biases of individual examiners. There are several ways in which such subjective elements affect the marking of scripts. One is the imperfect and varying understanding on the part of examiners of the questions set in the question paper and of what is expected of the candidates. This factor is sought to be kept to a minimum by specifying in elaborate detail the questions set and the answers to be expected. This, of course, helps but it does not work equally well with all examiners and in all subjects. Therefore, there arise frequent cases when an examiner has either uniformly overrated or underrated the scripts examined by him. There also occur cases of quite erratic marking. However, these may be either because of lack of understanding or due to sheer irresponsibility on the part of the examiner. Unfortunately, irresponsible behaviour is not uncommon. This also results in gross marking whereby marking is confined to only a few typical marks, or in safe marking whereby all borderline cases are pushed into the Pass class. The second important factor which lets the subjective biases of the examiner affect his marking of scripts is the conditions and the environment in which he works. The only care being taken in this respect at present is to give each examiner only a reasonable load of work and an adequate remuneration. However, by and large, the examiners examine the scripts at home where the working conditions are far from ideal. There is little doubt that if the examiners are required to examine the scripts only during regular working hours and in officially provided working rooms or halls, a great improvement in the standard of marking will take place.

There is yet another way in which the subjective judgment of the examiner affects the marking of particular scripts. It arises out the average quality of the bunch of scripts which an examiner is required to examine. It is well recognized that in spite of the clearest instructions and understanding about the answers to be expected, the expectations of the examiner are affected and influenced by the quality of the scripts he is examining, If the average quality of the bunch of scripts is good, a sub-average candidate is often judged too harshly. On the other hand, if the average quality of the bunch of scripts is not so good, an above-average candidate is judged too liberally. In fact, it is not difficult to demonstrate that the same script, placed in different bunches of a different average quality, may be judged differently even if examined by the same examiner. The only means to correct this situation is to ensure that the bunches of scripts being allocated to different examiners are more or less of equal average quality.

The need to allocate bunches of equal average quality to different examiners

There is another and probably more fundamental reason why the bunches of scripts allocated to different examiners should be of equal average quality. Suppose that the bunches of scripts given to different examiners are not of equal average quality or that it is not known that they are of equal average quality. In the circumstances, we cannot reasonably expect that the marking of these bunches by different examiners should be nearly equal. Even if there are pronounced differences between the marking of the different bunches, we cannot be certain that they do not reflect real quality differences between the bunches. Therefore, we dare not adjust the marks given by an examiner even if we suspect that they are affected by the strong personal biases of the examiner. All that we may do is to require a moderator to re-examine a few scripts, probably some borderline cases, and revise the marks given by the examiner if it is found necessary-but only in those few cases actually re-examined. However, even if a whole bunch seems to have been overmarked, or undermarked, or erratically marked, or uniformly marked, or grossly marked, or safely marked, there is nothing that we can do except to ask the moderator to examine the whole lot. Except for re-examination and reassessment by the moderator, the marks given by the examiner are thus sacrosanct and cannot be touched. Because of the large number of scripts to be examined and assessed and the extremely short time within which the final results have to be prepared, the scope for re-examination and reassessment by the moderator is indeed small, and in practice, for a large majority of the scripts submitted, the marks given by the examiner stay untouched. Thus, so long as the bunches of scripts allocated to different examiners are not known to be nearly equal in their average quality, there is no way of comparing the behaviour of different examiners and correcting the one who is found to behave rather peculiarly. In order to be able to do this, the bunches of scripts allocated to different examiners must be of equal average quality.

The principle of randomization

We must therefore look for a method of allocating bunches of scripts to different examiners so that their average quality would be more or less equal. There is of course no way of knowing the quality of a script until it is examined. Nevertheless, the allocation of scripts to different examiners must be done before the scripts are examined, and this has to be done in such a way that the average quality of the bunches given to different examiners is approximately the same. There is only one way of doing this, and, surprisingly, it is the process of random allocation. It is a basic statistical principle and it can be easily demonstrated that if the scripts to be given to different examiners are allocated on the basis of random selection by means of a process equivalent to the drawing of lots, then the lots given to different examiners will be approximately of the same average quality. The lots, when they are so allocated, will not of course be of exactly the same quality. There will naturally be some differences between them. However, the differences between such lots are governed by three important statistical properties:

1. Firstly, the differences between the lots are smaller, the larger the lots. For instance, lots of 400 scripts each will be more nearly equal in their average quality than lots of, say, 100 scripts each.

2. Secondly, and this is even more important, given the size of the lots, the range of likely differences between them can be judged, in advance. For instance, with lots of, say, between 250 and 300 scripts each, it can be demonstrated that the mean or the median marks of different lots should all lie within a narrow range of two or three marks.

3. Thirdly, the randomized lots are more or less alike not only in their average quality as judged by the mean or the median marks, but also in their internal variation or the spread of marks over different scripts in each lot. Further, whatever differences may legitimately exist in such respects can also be judged in advance.

Consequently, once lots of scripts are allocated to different examiners on the principle of randomization, any lot howing a very different level or pattern of marking may be suspected to have been affected by rather pronounced subjective factors of the expect and hence calling for revision. In other words, once the different examiners are given random lots of scripts, it will be reasonable to expect a common pattern of marking and, if the different lots show unreasonably different patterns of marking, it will be justifiable to reduce them to a common pattern. Subsequent procedures are concerned with how to reduce the markings by different examiners to a common pattern.

Procedures adopted by Gauhati University

The several problems that arise are best illustrated by reference to the procedures adopted by Gauhati University. In a report on this subject by Dr. H. J. Taylor, Vice-Chancellor of the University, (Taylor, 1963) the following is observed:

> Ideally markes should be scaled so that all sets of marks have ; (a) the same mean or median, and (b) the same standard deviation, Of these (b) presents the more difficult problem, which needs further studo before a solution can be supplied with confidence. In the present work only median scaling has been carried out, but this ensures that at -least the major errors arising from the variation of standards are removed. (Taylor, 1963.)

The median scaling, as adopted by Gauhati University, consists of the following:

i) For the scripts examined by each examiner, the median mark is found out. The median mark is the mark above and below which half the candidates lie.

ii) The average of the median marks of different examiners is then worked out, and treating this as the norm in those cases where the median mark of an examiner differs from the norm beyond a certain range, the marks in the whole lot examined by that examiner are adjusted, that is to say raised or lowered until the median comes to value chose as the norm. There are certain complications at both the extremes, so that candidates ò are not given marks either below zero or above 100, but that is a matter of detail which need not concern us here.

Appreciating fully that Gauhati University has only recently made a beginning in this new direction trusting that the procedures employed will be improven as more experience is gathered, we offer a few comments with the sole purpose of raising certain importment issues involved. In the first instance, there is the fact that in the procedure at present adopted marks are adjusted for the differences in the median marks only and not in the spread of marks. The reason is that the latter adjustment is the more difficult of the two to carry out. We shall later examine the question as to how difficult it is to adjust for the differences in the spread of marks. Our immediate concern is to emphasize the implications of adjusting for the median marks only and not for the spread at the same time. Adjusting for the median marks only presumes that the nature and the magnitude of the subjective bias of the examiner are the same at the median level as also at the high and the low levels. Thus, if the median marks given by an examiner are above the norm by 5 marks it is presumed that the examiner has overmarked all scripts uniformly by just 5 marks. For all we know, this may not be true. It is not only the magnitude of the bias that is likely to be different at the two ends of the scale, but sometimes also its direction. It is quite conceivable, for instance, that an examiner is considerate at the lower end but more exacting at the higher end or that he may be more hard at the lower end and generous at the higher end. Hence, a mere adjustment for the median mark by means of a straight raising or lowering of all marks might merely exaggerate his biases at either of the end rather than correcting them. The Gauhati report in fact points out a phenomenon of this nature, which it terms the ‘J-effect’, namely, the large clustering at the Pass mark of 35. It will be clearly disastrous to lower these marks along with the median mark. On the other hand, if they are raised along with the median, we could be giving a few additional marks to a group of candidates who have already received grace marks at the hands of the examiner. In the procedure adopted by Gauhati University, these borderline cases are presumably left untouched. The case however illustrates the defects of adjusting for the median without at the same time adjusting for the spread.

The second important aspect of the procedure adopted by Gauhati University is the choice of the norm to which the median marks are adjusted. As already pointed out, the average of the median marks given by the different examiners is taken as the norm. This means that the actual median marking by the several examiners in a year is itself taken as the norm for adjusting the median marking of different examiners. It follows that the norm to be adopted for adjustment of marks may fluctuate from year to year. Therefore, it is necessary to ask the question as to what extent such year-to-year fluctuations in the performance of candidates, as judged by the examiners, are justified. The annual fluctuations may be due to two sets of circumstances. In the first instance, the groups of candidates appearing in different years may differ in their quality, in the standards of teaching they were subjected to, and generally in the level of their preparation for the examination. This is not impossible. However, considering the we are concerned with groups of several thousand candidates, taught in several hundred schools by several hundred teachers, it is almost certain that in their quality, in the standards of teaching they are subjected to, and generally in their preparation for the examination, these groups cannot show wide fluctuations from year to year. Changes in these matters are bound to take place but such changes will be slow, gradual and evident only over a period of time. From one year to the next the groups of candidates will remain almost stable in their preparation for the examination. Consequently, the year-to-year fluctuations can almost certainly be due to only the second set of circumstances, namely, the standard of the question paper set, the standard of marking adopted and a host of other circumstances incidental to the conduct of the examination. Such fluctuation are probably unavoidable. However, it is necessary for that reason to protect the candidates from their effects. It is clearly necessary therefore to adopt norms for each subject, such that they will not vary from year to year, at the same time keeping open the possibility of review ing them periodically and adjusting them in the light of the changes taking place in the quality of the candidates and their preparation for the examination.

Adjusting the marking by different examiners to a common pattern

As has already been said, the purpose of these critical remarks on the procedures adopted by the Gauhati University is to raise and illustrate certain relevant issues. It is now proposed to put forward a procedure for adjusting the marks given by different examiners in a subject which we trust is free from these shortcomings and which we hope will also be simpler to operate.

The basis of this proposal is exactly the same as that which the Gauhati report mentions to be the ideal procedure; it says: "ideally, marks should be scaled so that all sets of marks have (a) the same mean or median, and (b) the same standard deviation". The mean or the median are particular measures of the ò average level and the standard deviation is a particular measure of the spread of marks. Therefore, if two sets of marks are so adjusted that they have the same mean or median and also the same standard deviation, we may expect that the two sets of marks would have the same average level and the same pattern of distribution of marks. However, this is not entirely true. Two sets of marks having the same mean or median and also the same standard deviation may differ in several important respects. For instance, the following results are based on the analysis of marks obtained by a sample of 1000 candidates appearing for the March 1959 examination of the Maharashtra S.S.C. Board.

Subject	Mean marks	Standard deviation	Percentage of passes
Language (Higher level)	41.50	11.09	84.17
Language (Lower level)	39.03	11.65	76.01

Thus two sets of marks, very nearly equal in the mean and also in the standard deviation, differ greatly, by as much as 8 per cent, in the percentage of passes. Several such examples may be given. The reason why two sets of marks with the same mean and standard deviation do not agree in several important respects is simple. As pointed out above, the mean and the standard deviation are particular measures of the average level and of the spread of marks. These measures would have special significance only if the distribution of marks, as given by the examiners, were perfectly ‘normal’. The term ‘normal’ here does not mean more than a particular form of statistical distribution. If the distribution of marks were perfectly 'normal' in this sense, it could be shown that two sets of marks having the same mean and the same standard deviation would agree in all other respects. In point of fact, sets of marks are known to be distributed not quite normally. It is for this reason that two sets of marks having the same mean and the standard deviation may differ in quite important respects. Therefore, adjusting marks for mean and standard deviation is not entirely satisfactory in that even after such adjustment, two sets of marks may still differ in quite important respects. More over, as the Gauhati report points out, adjustment for mean or median is relatively simple but that for standard deviation is fairly complicated. In fact, for that reason the procedures adopted by the Gauhati University permit adjustment only for the median. This is all the more unsatisfactory. In the first instance, a flat adjustment for median might introduce serious errors. Secondly, two sets of marks adjusted only for the median may, and in fact almost invariably do, differ in several major respects. Thus, it will appear that adjustment of marks for median and standard deviation is neither simple to carry out nor does it adequately satisfy its purpose, namely, to bring the different sets of marks to a common pattern of marking. In what follows, it will be argued that if the purpose of adjusting marks is to reduce the different sets of marks to a common pattern of marking, then this can be effected more directly and simply without reference to either the mean or the median and the standard deviation.

Pattern of distribution of marks

As mentioned above, the mean or the median are particular measures of the average level of marks obtained by a set of candidates and the standard deviation is a measure of the spread of marks within the set. These measures are often referred to as the measures of 'central tendency' and of 'dispersion' of a distribution. When, the distribution of marks is 'normal', which is a particular form of statistical distribution, the arithmetical mean and the standard deviation as measures of 'central tendency' and 'dispersion' are of particular significance because two 'normal' distributions, agreeing in their mean and standard deviation, agree in all other respects. When the distribution of marks if not 'normal', these measures are of no particular significance and they offer only inadequate description of the pattern of distribution. Two 'non-normal' distributions agreeing in their mean and standard deviation may differ in many important respects.

The distributions of marks as given by examiners are usually not normal, and often they are far from being normal. Under these conditions, it may be advisable to adopt certain measures which are more direct descriptions of the pattern of distribution. The simplest and the most direct description of a pattern of distribution is to indicate the number or the proportion of candidates who have secured marks between certain ranges. For instance, we may indicate the proportion of candidates who have secured 60 or more marks; between 45 and 59 marks; between 35 and 44 marks; between 25 and 34 marks; and 24 or fewer marks. These ranges of marks are the same as those which are usually valued as first class, second class, pass, condonable failure and non-condonable failure. Let us for convenience denote them by A, B, C, D and B classes. Two sets of marks which agree in respect of the proportion of candidates falling in each of these classes will have patterns of distribution which are similar for all practical purposes. If a closer agreement is desired, we may split up some of the five classes mentioned above and obtain, say, seven or even nine classes. Two sets of marks agreeing in respect of the proportion of candidates falling in each of seven classes will evidently agree in their pattern of distribution even more closely, and if they agree in respect of nine such classes, the agreement will be all the more close. It seems that specifying five classes will be found to be quite satisfactory; specifying seven classes will be found to be quite adequate; and specifying nine classes will provide an almost complete description of the pattern of distribution. In the following, for purposes of illustration, we shall adopt a specification of nine classes as under:

Range of marks	Class
90 and above	A_1
75-89	A_3
60-74	A_3
45-59	B
35-44	C
25-34	D
15-24	E_1
5-14	E_2
4 or less	E_3

If we specify the proportion of candidates which should all in these classes, we would have specified the pattern of distribution of marks almost completely.

The problem of norms

Let us recapitulate for a moment the position we have reached. It is this. If different examiners in a subject are allotted random lots of answer scripts, then, provided the examiners are free from subjective biases, we should expect the set of marks given by each examiner to follow the same pattern, which means that the proportion of candidates given marks falling between various ranges of marks such as those indicated in the system of five, seven or nine classes mentioned above, should be the same for each examiner. If different examiners do not show the same pattern of marking, we should suspect the markings as being affected by the individual biases of the examiners, and hence we should proceed to adjust the marks given by an examiner whose marking departs from the common pattern, so as to bring it in line with the common pattern. The question naturally arises: what is the common pattern to which the marking by each examiner should conform or else be adjusted? In the following, two alternative solutions are put forward.

In the first instance, we may adopt the pattern of marking indicated by all the examiners if their markings were pooled together. Thus, after all the examiners have marked their lots of answer scripts, we should obtain a combined distribution of marks. Let us suppose, the proportion of candidates in the various classes in the combined distribution to be as follows:

Range of marks	Per cent of candidates
90 and above	1
75-89	4
60-74	10
45-59	15
35-44	20
25-34	20
15-24	15
5-14	10
4 or less	5
	100

This is, of course, the pattern of marking in the combined distribution. The pattern of marking of each examiner within the set of scripts examined by him will of course differ from the 'common' pattern by varying degrees. As the first alternative, we suggest that the marks given by each examiner should be adjusted so as to bring the pattern of his marking in conformity with the 'common' pattern. This is adequate for the immediate purpose in hand though it has two shortcomings which we should note. In the first place, as we have earlier noted, it implies that the common pattern may vary from year to year-which variation may be wholly unjustifiable. Secondly, this procedure means that the process of adjusting marks given by individual examiners cannot begin until the lists of marks are received from all the examiners, pooled together and norms established. In practice, this may cause operational delays. However, subject to these two shortcomings, the procedure meets the main purpose, namely, to bring the level and pattern of marking of different examiners to a common norm. This procedure has also the advantage that it does not require any externally established norms. The marks given by different examiners are brought in conformity with their own combined average.

As mentioned above, the procedure implies that the norms for a subject may fluctuate from year to year. We have earlier argued that such year-to-year fluctuations in the norms may be wholly unjustifiable. If this view is accepted, it will necessitate establishing norms which will not change from year to year but which may be revised in the light of any long-term changes taking place in the quality of students, or in the quality of their preparation for the examination. The simplest method to do this is to establish the norms on the basis of recent past experience. With this purpose in view, we should examine the actual marks given by all the examiners in the last two or three years, work out what proportions of students were given marks falling between the various ranges such as 90 and above, between 75 and 89, etc. These proportions would then establish the norms which might remain unchanged for the next two, three or five years, or which may be changed as and when evidence of any long-term changes taking place becomes available.

It should be clear that, whether we accept the average level and pattern of marking of all the examiners in the current year as the norm for adjustment or the average level and pattern of marking of all the examiners during the past two or three years as the norm for adjustment, the subsequent procedures would be exactly the same. The distinction between the two methods of establishing norms is minor but nevertheless important. The distinction mainly lies in that while the first method implies that norms may fluctuate from year to year, the second implies that they should not, unless positive evidence becomes available for long-term changes taking place. The distinction is a matter of operational convenience. The first method requires the mark lists from all the examiners to come in before norms can be established and the process of adjustment begun. In the second case, the norms are already established and hence the marks given by each examiner can be immediately compared with the norms and adjusted if found necessary. Subsequent procedures of adjustment are, however, exactly the same in the two cases.

Procedures for adjusting marks given by an examiner

Let us now turn to the procedures for adjusting the marks given by an examiner if they are found to deviate from the norms (whether established on the basis of the current year's performance or the past few years' performance). To begin with, we must emphasize that the norms that we are now talking about are not just mean or median, or even mean or median with standard deviation. The norms consist of the entire statistical distribution of marks specified by proportions of students falling in various ranges of marks. As we mentioned earlier, specifying the distribution by means of five class ranges would be satisfactory; specifying by means of seven class intervals would be quite adequate; and specification by means of nine class intervals would be almost complete. As an illustration, let us suppose that the norms are specified in terms of proportions of students falling in nine classes. To make the example correct, let us suppose that the norms correspond to the distribution given on page 10. For the purpose of the subsequent discussion, it makes no difference whether the norms are established the basis of the current year's performance or the past few years' performance.

Referring to the distribution as given on page 10 and noting that the norms require 1 per cent of all students to secure 90 or more marks, 4 per cent to secure between 75 and 89 marks, and so on, the obvious thing to do would be to rank all the students in the descending order of the marks given to them by an examiner and then suppose that the top 1 per cent have, in fact, earned 90 or more marks that the next 4 per cent have, in fact, earned between 75 and 89 marks, and so on. Thus the process of adjustment involves the following:

1. Ranking all scripts examined by a single examiner, in the descending order of the marks;
2. Supposing that the top 1 per cent, for instance, have in fact earned 90 or more marks; that the next 4 per cent, for instance, have in fact earned between 75 and 89 marks; and so on.

It is thus that we could bring the distribution of marks given by the examiner in complete agreement with the distribution comprising the norms. Let us now ask the next question. What do we do to the marks given on each script? Do we change them in accordance with what we would want them to be? It seems that this will have to be done. If this is to be done, the procedure may be as follows. Suppose that the top 1 per cent should have 90 or more marks, but that, in fact, the last of these is given only, say, 86 marks. We may then raise the marks of the top 1 per cent of the scripts by 4 marks, subject to the condition that no one should have more than 100 marks. In the same manner, if we need that the next 4 per cent should have between 75 and 89 marks but find that the last of these scripts is given 72 marks; then we may raise the marks on all these scripts by 3 marks, subject to the condition that none in that group should get 90 or more marks. We may proceed thus, class by class, throughout the distribution. Evidently, in some cases we may have to reduce the marks given on certain scripts in order to bring them within specified limits. Clear and unambiguous rules may have to be framed for the purpose but there should be no difficulty, in principle, to doing this. It is obvious, nevertheless, that the procedure would be cumbersome and, operationally, too complicated. In what follows, proposals are put forward, which will obviate all the difficulties accompanying any changes to be made in the marks given by an examiner.

Let us for a moment return to the first stage of the adjustment process. As has been mentioned earlier, it requires the ranking of the scripts examined by a single examiner in the descending order of the marks given to them. Having done this, we have to identify the first 1 per cent, the next 4 per cent, and so on, of the scripts which the specification of norms requires should fall within certain class ranges. This is what we have called the first stage in the process of adjustment. The second stage consists of changing the marks given by the examiner. It is proposed that at this stage, instead of changing the marks given by the examiner, the numerical marks should be converted into letter grades. Concretely, in terms of our illustration, it is proposed that having identified the top 1 per cent scripts, instead of changing the marks on them so as to bring them within the specified range of 90 to 100, we should merely place the grade A_1 on these scripts. In the same manner, having identified the next 4 per cent scripts, instead of changing the marks on them so as to bring them within the specified range of 75 to 89 marks, we should place the letter grade A_2 on these scripts. In brief, taking the distribution given on page 10 as specifying the norms and taking the letter grade nomenclature suggested on page 9, both of which are purely illustrative, the proposal amounts to the follows: After ranking all the scripts examined by each examiner, in the descending order of the marks given to them, the top 1 per cent should be given the letter grade A_1; the next 4 per cent should be given grade A_2; the next 10 per cent, grade A_3; the next 15 per cent, grade B; the next 20 per cent, grade C; the next

20 per cent, grade 13; the next 15 per cent, grade B; the next 10 per cent, grade E_2; and, finally, the last 5 per cent should be given the grade E_3; To avoid any misunderstanding, it is repeated that the grade nomenclature and the proportions attached to the various grades in the above example are purely illustrative.

Implications of the proposals

Let us now examine the implications of these proposals. The first and the foremost implication is the limited importance and meaning attached to the actual marks given by an examiner. We accept that a script which is given a higher mark than another is better than the other. In other words, we accept the rank order indicated by the marks given by the examiner. This is as it should be; for, it must be understood, nothing short of re-examination of the scripts can alter the rank order indicated by an examiner. So we accept his rank-ordering. But we accept no more. We do not accept the absolute distances his marking indicates as between different scripts. In other words, we do not attach any absolute importance to the numerical marks given by the examiner. For the same reason, we do not care to change the marks given by him. We accept his rank-ordering of the scripts and we accord, say the first 1 per cent, grade A_1; the next 4 per cent, grade A_2, etc. This implies that a certain proportion of scripts from the lot examined by each examiner will be accorded a certain grade such as A_1, A_2, A_3, B, C, D, E_1, E_2 or E_3. The justification for this is the fact that each examiner is given a random lot of scripts to examine.

The converting of numerical marks given by different examiners into letter grades constitutes a major step forward in reforming the present examination system in India. Operationally speaking, it is immensely more simple as compared to any system of changing marks given by an examiner. Combined with the principle of allocating to each examiner a random lot of scripts to examine, it achieves comparability between evaluations by different examiners without doing explicit violence to the marks given by them. At the same time, it rids the examination system of the absolute omniscience of the parks given by an examiner. It should be understood that we cannot get rid of the numerical marking by the examiners altogether. For instance, it will not do if we ask the individual examiners to assign grades rather than numerical marks. In that case, there will be no guarantee that the grades assigned by different examiners are comparable. It is a necessary step in the above procedure that each examiner assigns numerical marks. It is only by rank-ordering the scripts on the basis of the marks assigned by each examiner that grades may be assigned which will ensure comparability of evaluation as between different examiners. Thus, though we finally get rid of the numerical marks assigned by the examiners, it is important to realize that the numerical marks perform an essential function, namely, that of rank-ordering the scripts examined by each examiner.

The problem of combining grades in different subjects

Once numerical marks are converted into grades, there arises the problem of combining the grades obtained in different subjects. Numerical marks in different subjects can be added together but the letter grades cannot be so added. However there are well-known statistical procedures for converting letter grades into numerical scores, normalized or otherwise, and after such scores have been added together, for reconverting the total score into an over-all grade. Any one of these procedures may be adopted. Of course, the conversion of grades into scores has to be done for purposes of arriving at an over-all grade. Scores imply a system of weights to be attached to different grades before they are combined to give a single over-all grade. Scores have no further use and they need not appear explicitly on any statement of the examination result to be given to a candidate. Thus the final result of a candidate will not consist of a statement of marks as at present. It will consists only of the grades secured by him in different subjects and a combined over-all grade.

While combining the grades secured in different subjects, yet another problem arises, namely, the relative importance to be attached to different subjects. Is English as important as, say, mathematics or a regional language or geography? It is to be clearly understood that these questions must be answered on pedagogic and not statistical considerations. Once a number of subjects are regarded as of equal importance, the simplest method to accord equal importance to them is to assign to them equal marks in the grand total. Thus, if each of seven subjects is accorded 100 marks in a grand total of 700, it means that they are all considered as being of equal importance. On the other hand, if it is intended to attach greater importance to any subject, the simplest method is to raise its marks in the grand total. For instance, if it is intended to attach double importance to English, it should be assigned 200 marks while the other six subjects would have 100 marks each in a grand total of 800 marks. It follows logically that if English is to be given double importance, the examination in English should be of double load or duration and in fact the curriculum in English should also be double in comparison to that in other subjects. In other words, and using common parlance, there should be two papers in English. This seems reasonable. However, from the standpoint of our present discussion, this is not necessary. It will be perfectly legitimate to keep the curriculum and examination load in any subject such as English the same as in other subjects and nevertheless give it double importance by assigning 200 marks to it in a grand total in which the other subjects are assigned only 100 marks each. The more important point for the purpose of the present discussion is that the relative importance to be attached to different subjects that a candidate offers is a matter which must be settled entirely by pedagogic considerations, and that, once this is settled, no problems arise while combining the marks in the different subjects. In the event, the marks in each subject are converted into letter grades as we have proposed above; and while combining the grade scores in different subjects, the intended weights must be attached. There are no further problems in combining marks or grades in different subjects.

Certain problems in the establishment of norms

This completes the discussion of the issues raised by the recommendations made by the UGC. Before concluding, a few extensions of these issues may be briefly indicated. The central issue, as will be realized, is how to ensure comparability between the evaluations by different examiners of answer scripts in a subject. This requires the establishment of norms. Our procedure requires that the norms should comprise specification of certain grades in descending order of merit and specification of certain proportions, of candidates which should be accorded these grades. Earlier, we have raised the question as to the basis on which such norms may be established. One method is to base the norms on the average judgment of all the examiners in the current year. This method has the operational inconvenience that the norms are not established until all the examiners return their scripts and a certain amount of tabulation of the total results is done. A more serious objection to the procedure is that it may result in the norms fluctuating from year, to year, and these fluctuations may be entirely unjustifiable. As an alternative, the norms may therefore be established on the basis of the past three or five years' average evaluation by the examiners. This may keep the norms more steady and, at the same time, permit a gradual change in them from year to year. The choice between the one or the other method will depend upon what view we take regarding the year-to-year fluctuations in the results.

There is another analogous problem. Should the norms be the same for different subjects or may they be different? If we base our norms on the average evaluation by the examiners in each subject, whether in the current year or in the past three or five years, it is obvious that the norms will be very different in different subjects, in the sense that the proportions of students who will be accorded certain grades will be different in the different subjects. The wide range of variation in the standard of evaluating scripts in the different subjects, as it exists today, will be evident from the following percentages of passes in different subjects at the March 1965 S.S.C. Examination in Western Maharashtra.

Subject	Percentage of passes
Marathi (Lower level)	89.10
Gujarati (Lower level)	93.83
Hindi (Lower level)	57.21
English (Lower level)	43.58

Thus the pass percentages vary from 43.58 per cent in English to 93.83 per cent in Gujarati. Languages are usually offered at the lower level when they are not the mother tongue. Therefore, Marathi and Gujarati in the above were presumably as foreign to the students offering them as were Hindi and English. Why then this large variation in the level of passes? It may be suggested that possibly Marathi and Gujarati were mother tongues and Hindi and English were not. However, even supposing this to be so, why the difference between the pass percentages in Marathi and Gujarati, or in Hindi and English, which though smaller is by no means inconsiderable. The differences in the pass percentages are equally striking if we consider the languages at the higher level in which case presumably they are all mother tongues of the students concerned. The pass percentage in the languages at the higher level were as follows

Subject	Percentage of passes
English (Higher level)	75.87
Marathi (Higher level)	75.93
Hindi (Higher level)	81.17
Kannad (Higher level)	90.84
Gujarati (Higher level)	92.56
Urdu (Higher level)	93.01
Sindhi (Higher level)	96.86

The above requires little comment. The same is true in respect of classical and modern European languages which are again available as equivalent options. The pass percentages in these were as follows:

Subject	Percentage of passes
Sanskrit	60.38
Ardhamagadhi	56.97
Persian	82.12
French	87.98

Such variations in the pass percentages are not, of course, confined to languages. They are equally evident in the social sciences, physical science and mathematics. The following are the pass percentages in social science subjects:

Subject	Percentage of passes
Social studies	78.24
History	81.89
Geography	65.73
Indian administration and civics	72.82

The difference between the pass percentages in history and geography is noteworthy. The following are the pass percentages in natural science subjects:

Subject	Percentage of passes
General science	70.93
Physics and chemistry	61.54
Physiology and hygiene	81.71

One wonders why it should be so much easier to pass in physiology and hygiene than in physics and chemistry. Or, finally, consider the mathematics subjects:

Subject	Percentage of passes
Elementary mathematics	58.21
Algebra and geometry	81.90
Arithmetic	45.78

Here again, the difference between the pass percentages in algebra and geometry and in arithmetic in striking.

It seems that, in general, it will be extremely difficult to justify the differences in the pass percentages as they exist today. However, this raises several questions which are beyond the scope of the present discussion. For our immediate purpose, what is relevant is the following If on any pedagogic considerations, the existing differences between the pass percentages in different subjects are thought to be justifiable, it is obvious that the norms on the basis of which different grades may be accorded in the different subjects will have to be different. In that case, no further problem arises. However, if it is felt that such large differences between the pass percentages are not justifiable, it is obvious that something will have to be done to narrow down the existing differences. One method to achieve this is to reduce the curricular burden in those subjects in which the pass percentages are low and/or to increase it in those subjects in which the pass percentages are high. Simultaneously, instructions may have to be given to the examiners and moderators to liberalize marking in those subjects where the pass percentages are low, and/or to tighten it in those subjects in which pass percentages are high. By such means, it may be possible gradually to reduce the existing disparities in the pass percentages in different subjects, though the process will by no means be simple and smooth. We should also remember that once a movement is started in this direction it will be hard to resist the logical conclusion that pass percentages should be more or less the same in all subjects. Once this is achieved, it will, amount to adopting uniform norms for accrediting grades in all subjects.

However, if the ultimate intention is to establish uniform norms for adjudging performance in all subjects, as we shall presently see, it is quite unnecessary to do this in the manner suggested above, namely, by: (i) adjusting curricular loads in different subjects, and (ii) instructing examiners and moderators to aim at certain pass percentages. In the present instance, let us be clear that the question of the curricular load in each subject must be decided entirely on pedagogic considerations add not to suit particular marking habits evolved by the examination system. If, therefore, we leave this question out, and if we desire that in each subject, a certain proportion of students should pass, there are two choices open to us. One is that suggested above, namely, to instruct the examiners and moderators to aim at the specified pass percentage and to liberalize marking if the pass percentage in the subject is too low and to tighten the marking if the percentage is too high. The other alternative is to vary the passing mark in order that a given percentage of students acquire the passing mark; that is, to lower the passing mark if the pass percentage is too low, and to raise the passing mark if the percentage is too high. This is obviously the simpler thing to do. But we would not do it because of the sanctity which the particular pass mark, namely 35, has acquired in our examination system.

As has earlier been mentioned, the absolute omniscience at present attaching to the marks given by an examiner to a script is the main hurdle in the reforming of our examination system. The notion is so deep seated in our thinking that it will require resolute steps to rid our examination system of the same. Let us therefore be quite clear on this point. In the first place, let us repeat, the assigning of numerical marks to a script by the examiner is a necessary step in the process and therefore it is not suggested that the system of assigning numerical marks should be abolished. All the elaborate instructions that are at present prepared for the guidance of the examiners in their evaluation work are also necessary. In fact, in any mass-conducted examination, these are absolutely essential in order to achieve the maximum possible objectivity in the evaluation procedures. However, having said this, we must be clear regarding the meaning and importance to be attached to the numerical marks given by the examiner. As we have earlier mentioned, the numerical marks given by an examiner to a number of scripts examined by him indicate a certain rank-ordering of the scripts by merit.

This is the essential function of the numerical marks given by the examiner and no violence to the implied rank-ordering can be done except by re-examination of one or more scripts by a moderator. But rank-ordering must also be the sole function of the numerical marks given by an examiner and no further absolute importance should be attached to them. This is the crucial point and its implications are:

1. In spite of all the instructions given to the examiners and in spite of all the steps taken to achieve the maximum objectivity in the marking system, marks given by two different examiners to two different scripts are not comparable. They are not comparable in the sense that a mark such as 40 given by one examiner to a script examined by him does not necessarily mean the same thing as mark 40 given to another script examined by another examiner. This is the starting point, and if this is not accepted, no reform in the examination system of the kind under discussion is called for. However, once this point is granted, the other two follow logically.

2. The marks secured by a certain candidate in a subject in a particular year are not comparable with the marks secured by another candidate in the same subject but in another year. This is obvious because not only the examiners are likely to be different in the two cases, but also, the candidates would be answering two different question papers, though in the same subject.

3. The marks secured in one subject are not comparable with the marks secured in another subject, even in the same year and by one and the same candidate.
   This, too, is obvious, for in this case not only the examiners would be different but they would be evaluating scripts in two different subjects under, naturally, very different marking instructions.

Thus the numerical mark assigned by any single examiner to a script has no further validity than as an indicator for rank-ordering the scripts evaluated by him. Once this is accepted, the rule for making the evaluation by different examiners comparable, is simple. Firstly, allocate to each examiner lots of scripts drawn on a random basis. Secondly, after rank-ordering the scripts examined by each examiner, certain specified proportions of them taken from the top to the bottom, successively, should be assigned certain letter grades. If these proportions are predetermined, as we suggest they should be, the work of assigning letter grades after rank-ordering the scripts on the basis of the marks given may indeed be left to each examiner. Further, if it is intended that the standards of evaluation, that is to say the standards of assigning different grades, should be uniform for all subjects, the same may be achieved very simply. The rule is to use for all the subjects the same norms, specifying certain percentages for assigning certain grades in each lot examined by a single examiner. This implies that the norms will have to be determined by quite independent considerations and not on the basis of the actual marking by several examiners in the current year or in the past few years. In the final analysis, the norms would be based on an arbitrary judgment as to what proportion of candidates should be given various grades; but the arbitrariness would be no greater than that involved in the existing pass mark of 35, second class mark of 45 or 50 and first class mark of 50 or 70. Finally, in order to work out the over-all average grade for each candidate, the grades in different subjects should be converted into scores, averaged simply or weighted appropriately, and reconverted into an over-all grade. This would constitute the average grade for the candidate for the examination as a whole. The decision as to whether he may be deemed to have passed the examination will have to be taken on the basis of this average grade together with the grades secured in single subjects, by means of suitable rules. There would be nothing new in such procedures.

REFERENCES

Report on Examination Reform. University Grants Commission, New Delhi, 1962.

Taylor, H.J., *Operation Mass Mark: An Account of the Methods Used in the Matriculation Examination of 1963.* Gauhati: University of Gauhati, July 1963.

Tluanga, L. N. and Taylor H. J., *The Influence of Scaling on Examination Results: A Report on a Sample Analysis of the Matriculation Examination Results of 1963.* Gauhati Examination Research Unit, University of Gauhati, December 1964.

V.M. Dandekar is Director of the Gokhale Institute of Politics and Economics, Poona.

ASIAN DRAMA - AN INQUIRY INTO THE POVERTY OF NATIONS*

V. M. Dandekar

"The title of the book ***Asian Drama*** was chosen in order to express the conception of events in South Asia held by the author at the beginning of his work and fortified in the course of study" (p. 34). The book is primarily addressed to the Western economists, social scientists and generally the makers of Western policy towards South Asian countries. "It is intended to be undiplomatic" aimed at "truth and blunt truth-speaking" (p. 23). That truth seems to be that the conception of the European colonial masters about South Asia was closer to reality than the illusions harboured or, for diplomatic reasons, given expression to by western academicians and politicians during the past twenty years. (pp. 20-21). The essence of this concept is the following: "The obstacles to rapid economic expansion (in South Asia) ... are rooted in the inefficiency, rigidity and inequality of the established institutions and attitudes and in the economic and social power relations embodied in this framework of institutions and attitudes" (p. 47).

The sub-title "An inquiry into the Poverty of Nations" is evidently aimed at immortality. Unfortunately, what it inquires into is more topical than fundamental. It inquires into the current difficulties encountered by the South Asian countries in lifting themselves out of poverty and not into the more fundamental question why these countries fell into poverty in precisely the same period when western countries rose to affluence. Consequently the survey of "Political Problems" in Part Two (pp. 129-409) and of "Economic Realities" in Part Three (pp. 411-705) is confined to the post Second World War developments. A fuller background of the political and economic history of the colonial period is needed for a better appreciation of even the current political problems and economic realities in these countries. The two Parts thus strengthened would have constituted a useful, self-contained single volume.

Part One (pp. 37-125) is devoted to a discussion of the values, interests and ideals, summed up as 'Modernization Ideals', held by the peoples of the countries concerned. Major elements in this system of values are national consolidation and independence, political democracy, rationality, social and economic equalization, improved institutions and attitudes, development

* Asian Drama by Gunnar Myrdal. 3 vols. p. xxx + 2284. Pantheon. New York

and democratic planning for the purpose. In Part Four (pp. 707-958) the rationale of many prevailing ideological elements such as equality, democracy, socialism and economic planning is examined in the light of the values discussed in Part One. Much pedantic repetition could have been avoided by combining the two Parts. Many of the appendices such as 1-5 (pp. 1839-2040) and 8-9 (pp. 2077-2120) offer more pedantic elaboration of the same discussion. The Prologue titled the ***Beam in Our Eyes*** (pp. 5-35), Parts One and Four, and the relevant appendices put in a single volume might have constituted a more homogeneous reading material.

It seems that "in all the countries of South Asia the ideal that social and economic stratification should be changed . . . is commonly accepted" and that "in general, it is held that social and economic institutions and attitudes should be changed ... that such a revolution is necessary for development" (pp. 59-60). This is also what the book prescribes. What then is the problem? Apparently, there are - two difficulties. Firstly, "the attitudes that are thought to need changing are a function of the low levels of living and culture, and that these levels can only slowly be elevated" (p. 63). This is the viciously circular proposition: "a country is poor because it is poor" (p. 1844). In spite of much air to the contrary, the book offers little analytical penetration into and much less political-economic strategy to achieve a break through this vicious circle. The second difficulty is the following: "... the success of planning for development requires . . . rigorous enforcement of obligation, in which compulsion plays a strategic role ... Under present South Asian conditions development cannot be achieved without much more social discipline than the prevailing interpretation of democracy in the region permits." (p. 67). But the book offers little guidance on the alternative ‘democratic’ forms which the countries must adopt.

From the standpoint of economic analysis and prescription, Parts Five, Six and Seven (pp. 959-1828) together with appendices 6 and 7 (pp. 2041-2075) form the core of this study. They discuss the problems of labour utilization, population size and population quality, respectively. "In the mainstream of economic theory, unemployment and under-employment on a vast scale are regarded as a primary cause of poverty in the South Asian countries. At the same time, the large volume of unutilized labour possessed by these countries is thought to have a productive potential . . . that can be used to eliminate poverty. The supreme task of planning is thus to drain this labour reservoir ... into productive work." (p. 962). This theory is dismissed as being "but a species of a genus, namely, aggregative model-building of the western type" (p. 963). More mercifully, the inadequacy of this approach in South Asia is discussed (pp. 994-1012) and a "More Realistic Approach" is outlined (pp. 1012-1027). In the first section (pp. 1012-1016), after some definitions, we come to the conclusion that the level of actual 1abour utilization can be expressed as the product of three ratios, 'participation' (working members/ labour force), duration (man-hours/working members) and labour efficiency (output/man-hours). In the next section, (pp 1017-1019) we learn that the three ratios are interrelated. The following two sections (pp. 1019-1027) which complete the out-line of the "More Realistic Approach" are devoted to discussing "The Inappropriateness of Western Conceptual Categories for the compilation of Statistical Data" (pp. 1019-1026) and "The Quality of Statistics on Labour Utilization" (pp. 1026-1027). This is typical of the treatment in many parts of the book and partly explains its volume. What is more disconcerting is that questions are raised which are left unanswered without acknowledging that they have not been answered. For instance, in the present case, we do not know whether, in the light of the "More Realistic Approach" unemployment is a primary cause of poverty, whether the unutilized labour is a productive potential that can be used to eliminate poverty and whether planning can be directed to this purpose. One wades vainly through the successive chapters seeking answers to these initial questions.

In the discussion of the policy choices, the book has managed to achieve a distinctly Asian flair and flavour. Consider, for instance, the chapter on Agricultural Policy (pp. 1241-1384) which comes at the end of Part Five on Problems of Labour Utilization. At the beginning of this chapter we are told: "The goal must be ***to increase the utilization of the at present greatly underutilized agricultural labour force"*** (P. 1244). Towards the end of the chapter, the Policy Alternatives are presented. (pp. 1366-1384). The Basic Fact, we are told, are the following: Perhaps the most conspicuous result of post-war policies ... has been the strengthening of the upper strata in the villages and a corresponding reduction in the, position of share-croppers and landless labourers in the lower strata of rural society ... The political consequences .. are far reaching. The evidence suggests that the opportune moment for a radical reshaping of the agrarian structure has passed" (p. 1367). Nevertheless, it is emphasized that in South Asian situation "the promotion of social and economic equality is a precondition for attaining substantial long-term increases in production" ... and "A frontal assault on the equality issue requires a much more radical approach to agricultural policy than has yet been seriously attempted" (p. 1369).

But Communism, even in theory, does not seem to provide a viable treatment for the ills of South Asian agriculture" aud "neither of the two extreme measures for agrarian re-organization - radical redistribution or radical land consolidation - appears to provide a feasible approach to the problem of agriculture in the region" (p. 1377). Hence "it will be useful to begin with the realities of the existing situation" (p. 1378) and the conclusion is that .... it may be preferable to make a deliberate policy choice in favour of capitalist farming allowing and encouraging the progressive cultivator to reap the full rewards of his enterprise and labour" (p. 1380). The author, who is generally well informed on the facts, and the situation, does not appear to know in this case that this choice has long been made by most Asian governments and leaders of opinion and that they came to this conclusion by precisely the same steps he has taken. Unaware of this, and wanting to sell his 'radical' idea to the Asian leaders and knowing them as well as he does, the author assures them that this policy "is to be recognized as a quite radical 'land reform', though of a different species from the schemes usually discussed" (p. 1382). Not feeling certain that the idea would still be accepted, he adds: "High priority in this scheme should be accorded to a programme ***to give a small plot of land - and with it a dignity and a fresh outlook on life as well as a minor independent source of income - to member of the landless lower strata"*** (p. 1382). (all italics original). This is the familiar Indian rope-trick taken straight from the Indian Five Year Plans. This is not surprising. The longer one dwells on Asian problems, the more readily and inevitably one seems to arrive at the Asian solutions, namely, to try and forget the problems.

There can be no denying however that this is a major "study of major economic problems in South Asia" and "one that maintains a constant awareness of the broader setting of these issues" (p. 43). The Western readers to whom it is addressed will find the book, in spite of its pervasive presumption, a most valuable addition to the growing western literature on South Asia. The Asian readers will find in it, in spite of its equally pervasive condescension, the most sympathetic treatment and summarization of the problems which their countries have muddled through in the past twenty years. And in spite of its length, the general reader will find it most readable though a firmer editing would have proved immensely helpful. Judging by its size, one might suppose that all that was written has been put in print, and almost in that order. This is not true. Much that was written remains unpublished. We are informed that "all outlines, drafts, proofs, staff memoranda, correspondence and other documents relating to the study are deposited with the Royal Library in Stockholm, at its invitation, where they are open to the public" (p. xviii). It seems certain that if Professor Myrdal had known earlier that the Royal Library would provide this excellent facility, he would have deposited the present manuscript there and prepared another, much shorter and much more firmly edited version for the press.

June 7, 1968.

COW DUNG MODELS (REVIEW)

V M Dandekar

Investment in Livestock in Agrarian Economies: An Analysis of Some Issues Concerning "Sacred Cows" and "Surplus Cattle" by K N Raj; Centre for Advanced Studies, Department of Economics, Delhi School of Economics, University of Delhi; 1969; Pp 33; Rs 2.00.

THIS is a reprint of an article publish ed in Indian Economic Review, Vol. IV, No 1, April 1969. Its purpose is to examine the "widespread assumption about the existence of large surpluses of livestock" in India. The article may be divided broadly in two parts. The first part (pp. 2-17, Sections II-V) offers what may be called the 'reasoning'. The second part (pp. 17-31 Sections V-VIII offers the 'data'. There is an introductory Section I and a concluding Section IX. Two Sections carry the same number V, apparently to supply some link between the 'reasoning' and the 'data'.

The 'reasoning' consists of setting up two models which we may call the 'poor-country model' and the 'rich-country model' and trying to demonstrate that the livestock situation in India and USA broadly accords with the two models, respectively. In fact, without the author's realising it, the 'poor-country model' totally fails in its application to the Indian situation.

In the second part, the model is left behind and several fragmentary data are presented to throw light on the 'economic' bases of India's cattle population leading to the conclusion that religious sentiment does not have much effect on its size and composition.

BREEDING FUNCTION IGNORED

The 'poor-country model' (Section II) postulates a community which consumes only grain, milk and beef, per capita per annum consumption being 300 lbs of grain, 30 gallons of milk, and 10 lbs of beef. Valued at unit prices of Rs 0.50 per lb of grain, Rs 2.00 per gallon of milk and Rs 1.00 per lb of beef, the per capita consumption amounts to Rs 220. Other assumptions of the model are: per acre grain output = 500 lbs; per cow milk output = 300 gallons; beef yield per cow/bull = 500 lbs; per acre draught power requirement = 0.25 bulls; per cow/bull feed consumption - 800 lbs; per calf (defined as below 3 years of age) feed consumption = 400. lbs. Finally, "we assume that no calves are, killed and that, if less than 5 per cent of the 'adult' cattle are slaughtered each year, the cattle population can be maintained constant by keeping calves equal to 30 per cent of the adult cattle in number". We do not know if any animals die outside the slaughter-house and at what ages. From one passage, it seems that no animal indeed dies outside the slaughter-house except at very ripe old age and that the few animals which are slaughtered to meet the demand for beef are picked just when they are about to drop dead. But presumably this is not essential for the working of the model. Equally, we do not know the number of calves which, say, 100 cows give annually. The essential assumption of the model is that no calves are slaughtered and that whatever the number born annually and whatever the mortality at various ages, the stock of adult cattle is maintained provided the young stock (below 3 years of age) forms 30 per cent of the number of adult stock. Starting with these assumptions, it can be shown that a human population of 100 will require 65 heads of cattle (38 bulls, 12 cows and 15 calves) and 152 acres of land.

The 'rich-country model' (Section III) postulates a community with per capita consumption twice as much as in the 'poor-country model', namely, Rs 440. This comprises 480 lbs of grain, 75 gallons of milk and 50 lbs of beef. Assumptions regarding per acre grain output, per cow milk output, per cow/bull beef yield, per cow/bull feed consumption and per calf feed consumption remain the same as in the previous model. There is however one difference: animal draught power is replaced by tractors. Fuel for the tractors is imported by exporting grain, the grain equivalent of fuel needed to cultivate one acre of land being 130 lbs. It will be remembered that the grain equivalent of draught power needed to cultivate one acre in the 'poor-country model' was 200 lbs. However, while in the 'poor-country model' bulls are reared by feeding the necessary young-stock, in the 'rich-country model, tractors are neither manufactured nor imported; presumably they are a gift from the World Bank.

In view of the larger consumption of beef, the assumption regarding the ratio of young stock to adult stock needed to maintain the necessary stock has also to be modified. "The extent of the modification needed will depend not only on the exact increase in death rates caused by slaughter but also on the age group affected. Without going further into this question, we shall simply assume that, in the second situation, calves equal in number to 60 per cent of the number of adult cattle will have to be kept in order to maintain the total livestock population constant." Thus here again no explicit note is taken of the breeding function of the cows.

Starting with these assumptions, it can be shown that a human population of 100 will require 56 head of cattle (25 cows, 10 cows or bulls and 21 calves) and

228.10 acres of land. This may be compared with the corresponding requirements in the 'poor-country model', namely, 65 head of cattle and 152 acres of land. Incidentally, instead of comparing the two model situations in this manner, Raj seems to believe that the 'poor-country model' with 100 human population is comparable to the 'rich-country model' with only 50 human population. For instance, the 65 head of cattle needed in the first model are compared with 28 head of cattle (not 56 needed by a human population of 100) in the second model. Why? Because, the second model is derived from the first by imagining that 50 members of the first community have migrated elsewhere!

HUMAN AND LAND RATIOS

Section V Opens with the following statement: "A glance at the livestock data for India and the United States would show that the size and pattern of the livestock held in these countries in the early sixties relatively to their population and cropped area are not very different from what one should expect on the reasoning outlined above". Let us first consider the size of the livestock. Table 6 shows that in USA, the number of bovine animals was 54.42 per 100 of human population and 53.95 per 100 hectares of arable and cropped land. In the 'rich-country model' these numbers are 56 and 60.67, respectively, and may be regarded as not very different. But the same table shows that in India, the number of bovine animals was 51.35 per 100 human population and 139.88 per 100 hectares of arable and cropped land. In the 'poor-country model' these numbers are 65 and 105.67, respectively and Raj believes that these too are not very different! In particular, the important point is completely lost, namely, that in comparison with the 'poor-country model', India has 21.0 per cent fewer bovine animals per 100 human population and 32.4 per cent more bovine animals per 100 hectares of arable and cropped land.

RESCUE OF MODEL

Let us next consider the age and sex composition of the livestock population. Raj begins with an assertion: "The sex and age composition of the live stock population in the two countries also display characteristics suggested by our exercise". Let us first compare the 'poor-country model' with the Indian situation. As we have seen, in the 'poor-country model' the livestock requirement per 100 human population is 65 head of cattle comprising 38 bulls, 12 cows and 15 calves. If we divide the calves equally between the two sexes, as we should because the model does not permit slaughter of calves, the sex ratio works out to be 70 per cent male and 30 per cent female. Raj observes as follows:

"According to the Livestock Census of 1956 for India", (one wonders why the 1956 census when 1961 census figures are equally well available) "about 55 per cent of the total cattle population was male, only the remaining 45 per cent being female". Thus clearly, India has proportionately a much larger female stock than that indicated by the 'poor-country model'. The situation is even worse because as Raj observes: "The inclusion of the buffalo population shifts this numerical balance in favour of the female animals, as male buffaloes are not used in India for draught purposes except in regions with very heavy rainfall while buffaloes are highly preferred as milch animals in many parts of the country". If due allowance were made in the 'poor-country model' for this preference for buffaloes as milch animals, the number of cows needed (mainly for the supply of milk) would appear to be smaller than 12 per 100 human population as appearing in the present model. This would make the proportion of female animals in the cattle population in the revised model smaller than 30 per cent as in the present model and it would be evident that India has proportionately a much larger female stock than is justified for supplying milk. In fact the whole model collapses in its application to India.

Raj proceeds heroically to rescue his model. He pleads: "However, if one leaves out of the calculation (as in the model) the adult cows and she-buffaloes not in milk and compares their number with those of bullocks and he-buffaloes actually 'at work' (again as in the model) the proportions that emerge approximate more closely to those indicated by the illustrative exercise relating to situation I". But Raj is making a bad mistake in believing that the 'model' leaves out of calculation adult cows not in milk and bullocks not at work. The model assumes that the output of milk per cow per annum is 300 gallons, This does not mean that every cow gives 300 gallons per annum. It also does not mean that there are no cows not in-milk. The output of 300 gallons per annum is an average for all the cows that there are. Similarly, the assumption that the number of bulls required for ploughing an acre is 0.25 does not mean that there are no bullocks not-at-work. Indeed, with slaughter rate of no more, than 5 per cent, it is inevitable that there are cows not-in-milk and bulls not-at-work. Not only such animals exist in the model but they also eat and the feed requirement assumed to be 800 lbs per cow or bull per annum is again an average for all productive and unproductive animals. Hence if, as Raj says, these animals are left out of the model, we should know how they are left out. For instance, are they left in a *pinjra pole or a gosadan?*

And where does this juggling lead us to? According to Raj, the adult productive cattle and buffalo population in India (1956) comprises 31.8 per cent female stock and 68.2 per cent male stock. According to the figures quoted in the first unnumbered table following Table 6, these proportions in the 'model' are 21.2 per cent and 78.8 per cent, respectively, Is the correspondence supposed to be satisfactory? Incidentally, the correspondence would improve a little if the proportions for the 'model' appearing in this table, were correctly quoted; the correct proportions appear to be 24.0 (not 21.2) per cent female stock and 76.0 (not 78.8) per cent male stock.

Let us next consider the age composition. Raj says that according to the Livestock Census of 1956 for India, the animals 3 years and below in age accounted for less than two-fifths of the adult cattle population, He does not give the actual proportion. But it does not matter. In the model, this proportion is assumed to be 30 per cent. To the extent it agrees with the actual, the agreement is achieved by assuming the appropriate proportion slave no other reason was given for assuming the particular proportion. To the extent, it does not agree with the actual, one is left wondering why a more appropriate proportion was not assumed.

The correspondence between the 'rich-country model' and USA situation appears better. Raj follows this comparison with a series of *obiter dicta* leading to the sound recommendation: "There are therefore a number of circumstances specific to each case which should be taken into account in relating our model to any particular situation".

In relation to India, Raj discovers one such specific circumstance. He observes: "it is evident that in India the grain equivalent of the feed given to livestock per head is not as high relatively to the per capita grain consumption of the human population as we have assumed in our numerical exercise ... It follows that, though some of the aggregated percentages yielded by our illustrative exercise for the 'initial' situation and the data for India bear some resemblance to each other, there are important differences which might even make this broad conformity *just on accident.* For, in our model it is the relatively high feed requirement of the livestock that make such a large livestock population necessary.' (italics added.)

NO RESEMBLANCE

In point of fact, there is no resemblance, even by accident, between Raj's 'poor-country model' and the Indian sitution. It was by assuming a relatively high feed requirement of live stock that the model led to a large cattle population; and it was by assuming that the cow was the only milch animal that the model gave a proportion of 30 per cent cows in the cattle population. Both these assumptions are untrue so far as the Indian situation is concerned. Nevertheless, in relation to the land resources, India has a larger cattle population and larger proportion of cows than indicated by the model. Obviously Raj's 'poor-country model' is a poor model to explain even the size and sex composition of India's cattle population. He does not seem to realise this. But he drops the model discreetly and turns to an examination of the Indian data and the situation, unfettered by the model but guided by little else.

He begins with the relevant question "whether, given the amount of food that can be spared for maintaining animals in the situation, a higher out put is secured by giving this food to a smaller number and raising their productivity or by distributing it over a larger number" but immediately offers the apology: "A firm answer to this question cannot be offered without much deeper investigation into the relevant production functions". Fortunately, he admits that "the productivity of livestock as much animals is highly responsive to the quantity and quality of the feed" and even puts forward "such evidence as is readily available". Evidence readily available to him is the one produced by Mellor and de Ponteves. "From data available from 13 rural and urban centres in India (covering the period 1954-60). Mellor and de Ponteves have worked out the relationship between concentrates fed per day to milch animals and milk yield per year." Raj is aware of the several reasons for exercising caution in the use of such results. Nevertheless, he thinks that "they do help to bring out the broad characteristics of the production response to feed input in the form of concentrates". Consequently, he agrees that "one should expect the holders of milch animals to be more concerned about their feed requirements and, ... choose perhaps to hold a smaller number of animals and feed them well than keep a large number in a serious undernourishment by distributing the food among them"; and adds: "The number of milch animals kept will depend of course on the demand for milk and on the relative price of milk and of the feed required".

FARMERS' HOME ECONOMICS

Very well, Does the situation in India accord with these economic principles?

No. Raj has no difficulty in understanding "why Indian farmers choose to feed their cattle only as much as they do". He explains: "Consider the large number of relatively undernourished cows maintained in the rural areas. When large markets for milk are not open to the holders of this livestock ... they have only their own demand to satisfy. The smaller the holdings of land and the lower the standard of consumption of the farmers the smaller is likely to be their own demand and the higher the cost of concentrates to them. But it may still pay to have a milch animal around - underfed by standards elsewhere as long as they have some minimum requirements of milk for family consumption or if there is a small market for milk within the village". So much for the small land-holder. As for the large land-holder, Raj observes: "Since farmers with larger holdings, of land are likely to keep milch cows for the same reason as the small farmers, the size of the village markets is likely to be on the whole very limited. There would therefore also be an upper limit to the number of animals that is worth maintaining both in individual households and generally in rural areas far away from the towns,"

One wonders what Raj is trying to say. Is he explaining why there are many more and undernourished cows when the same quantity of milk could be produced more economically by maintaining fewer and better-fed cows? Or is be explaining why there are no more cows than there are in the rural areas? And if, as Raj says, "farmers with larger holdings of land are likely to keep milch cows for, the same reason as the small farmers" what shall we do with the "much deeper investigations into the relevant production functions"? Whatever is his point, Raj thinks that "this becomes still clearer when one examines closely the number of milch cattle held in rural areas according to size of the land held by households" and presents "data yielded by a farm management study in two selected districts of North India (to be precise, in the State of Uttar Pradesh ...)". A welcome emphasis on precision indeed!

These data show the classification of 596 farms according to their size and according to the number of milch cattle held in them. Evidently, the farms vary in size ranging from under 2.5 acres to above 25 acres; some of them have no cows, some only one cow, some two cows and the others have three or more cows each, the larger farms generally having more cows, Nothing becomes clearer than this from the data as they are presented and apparently Raj does not know what to look for in them. In particular, he does not notice the one point which is relevant to his discussion. It is this: The 596 farms have on an average 1.6 cows per farm. This gives a total of about 950 cows. Out of these, 274 cows, which account for less than 30 per cent of the total, are held in farms which have only one cow each. We may explain their existence on the ground that "it might still pay to have a milch animal around -

underfed by standards elsewhere" in order to meet the home demand of the farmer. But what about the remaining 70 per cent of the cows which are held in farms with two or more cows each? Should we not expect the holders of these animals to "choose perhaps to hold a smaller number of animals and feed them well than keep a large number in a state of serious under-nourishment"? We might, as another concession to farmers' home economics, leave out even the cows held in farms with only two cows each and justify their existence on the ground that reduction from two cows to one cow is not possible. There are 296 such cows which is a little more than 30 per cent of the total. But still 40 per cent of the cows are held in farms with three or more cows each. Should we not expect these farmers to be more economic in their behaviour and to choose to hold a smaller number of betterfed cows? Or, shall we suppose that these farmers too are waiting for the results of "much deeper investigations into the relevant production functions"?

BUFFALO IGNORED

Throughout his discussion, Raj has concentrated on the cow as the milch animal though he is aware that this is not generally true in India and that buffalo is a serious and superior rival milch animal. The predominant function of cow in India is the breeding of bulls needed for draught power in agriculture. If Raj had taken this, into account, he could have examined the size of the cow population in India from this point as well. In a subsequent Table, he presents data regarding the number of draught cattle held by the same 596 farms. The average number of draught cattle per farm is 2.5 which gives a total of about 1,490 draught animals. Let us suppose, as Raj argues, that under the existing technology and institutional set-up in Indian agriculture, this number is irreducible. There remains nevertheless the relevant question: Are 950 cows needed to maintain a stock of 1,490 draught animals? Raj has totally neglected this aspect of the problem.

Finally, Raj puts forward an amazing argument purported to demonstrate that religious sentiment does not have much effect on the size and composition of the livestock held in India. He observes that "it is in the Iudo-Gangetic valley where Hindu orthodoxy is deeply entrenched and the sentiment against the killing of cows is strongest, that the pressure of human and bovine population on resources makes it most necessary to get rid of cows in preference for bulls for traction purposes and she-buffaloes for milk". He asks: "How does the table get turned so dramatically against the cows in Bihar and Uttar Pradesh?"; and answers: "Obviously, killing must be taking place, but perhaps the main technique adopted for getting rid of the cows is infanticide and deliberate starvation". Apparently, this does not concern Raj as an economist. Does he suggest that infanticide and starvation are exact substitutes for slaughter and result in exactly the same size and age-sex composition of the cattle population? Even supposing that they do, does he suggest that infanticide and starvation are economic substitutes for slaughter? Does he not recognise that the choice between infanticide-starvation and selective slaughter affects materially the relevant production functions in the livestock economy?

THE ECONOMIC QUESTION

In the end, Raj asks: "Does all this mean that there are no surplus cattle even in a country like India?"; and concludes that "the generalisations that have been made hitherto on this, subject are too simple-minded and superficial to bear scrutiny". His present paper is no improvement. "Is there surplus cattle in India?" is a superficial question to ask. The right question is: Are the size, age-sex composition and quality of Indian cattle the most economic in relation to India's resources and needs? If not, is the Hindu orthodox sentiment about the cow responsible? To answer these questions, let Raj try another exercise. He might as well begin with his first model. There is only one assumption of this model which is relevant: No slaughter of calves and less than 5 per cent slaughter of adult stack. Let him follow the consequences of this single assumption. Let the cattle population growing under this assumption be given the fodder and feed resources that the country has and let it be brought in contact with the human population which must compete for survival. If possible, let the superior rival milch animal, buffalo, be brought into the picture. Raj will inevitably meet the Indian cattle population with all its characteristic features. If he fails, it is time for him to shift to the Govardhan Peeth of the Shankaracharya of Puri.

INDIA'S SACRED CATTLE AND CULTURAL ECOLOGY

V. M. Dandekar

An important contribution to the debate on India's sacred cattle is Marvin Harris's article "The Cultural Ecology of India's Sacred Cattle" (Current Anthropology, Vol. VII, No a, 1966).

Harris contends that "the irrational, non-economic and exotic aspects of the Indian cattle-complex are greatly overemphasized 'at the expense' of rational economic and mundane interpretations" and 'urges that "explanation of taboos' customs, and rituals associated with the management of Indian cattle be sought in the "positive-functioned" and 'adaptive' processes of the ecological system of which they are a part rather than in the influence of Hindu theology".

This article, however, argues that, in Marvin Harris's thesis is summarised the entire defense of Hindu cow-worship garbed in pseudo-science.

It is suggested here that a cattle population smaller in number but better fed than at present will make a better contribution in milk, traction and dung to human survival and well-being in India. Even if, because of prevailing, technology and institutional arrangements in agriculture, it is not possible to reduce the number of draught, animals, the large number of cows is wholly unjustified. A smaller number of better-fed cows will contribute more milk and dung and will give enough calves to replace the needed stock drought animals.

I
Ecological Hypotheses

IN the current debate on India's sacred cattle, an important contribution has been one by Marvin Harris, namely his article "The Cultural Ecology of India's Sacred Cattle" (*Current Anthropology,* Vol VII, No. a, 1966). It presents a most sympathetic view of the Indian cattle complex and - almost justifies the present situation. Naturally it has made much impression on many Indian thinkers on the subject. For instance, K N Raj commends the article in the following words: "An excellent paper which views this problem of surplus cattle from a sociological angle and presents a point of view very similar to the one presented here is that by Marvin Harris" ("Investment in Livestock in Agrarian Economies", *Indian Economic Review,* Vol. IV, No. 1, April 1969). The article was also strongly

Reprinted from Economic and Political Weekly, Vol. IV, No 39, September 27. 1969.

recommended by the editor of *Samaj Prabodhan Patrika*, a Marathi two-monthly journal (March-April, 1969). Not being a normal reader of *Current Anthropology,* I had not seen the article until it as thus brought to my notice. Now that I have seen it, I am compelled to offer my comments because, I think, here is summarised the entire defence of Hindu Cow-worship garbed in pseudo-science.

Harris begins his article with an apology. He says: "I have never seen a sacred cow, nor been to India. I hope this will not deprive me of that expert advice and informed criticism which alone can justify so rude an invasion of unfamiliar territory." I suppose I should begin with a return apology for I seem to be about to invade, equally rudely, an equally unfamiliar territory. I have never met, knowingly, a cultural ecologist. Moreover, after reading Harris's article, I realise that I am thoroughly unfamiliar with the logical method employed in Cultural Ecology. Harris observes significantly: "Failure to separate emics from etics contributes greatly to confusion surrounding the Indian cattle question." I confess I do not quite know what emics and etics are, nor how to separate them. Harris's original article on this question to which he makes a reference was not easily accessible to me. A reference to dictionary did not help much. But I suspect that it is precisely this failure to separate emics from etics which is causing difficulty for Harris to come out of the confusion surrounding the Indian cattle question. Indeed, he seems to he making significant contribution to the same.

Harris believes that "the irrational, non-economic and exotic aspects of the Indian cattle complex are, greatly over-emphasised at the expense of rational, economic and mundane interpretations" and urges that "explanation of taboos, customs and rituals associated with management of Indian cattle be sought in "positive-functioned" and 'adaptive' processes of the ecological system of which they are a part rather than in the influence of Hindu theology". He suggests that this is possible if "the cattle complex is viewed as part of an eco-system rather than as a sector of national price market".

More specifically, Harris puts forward a series of hypotheses. They are, firstly: "Ecologically, it is doubtful that any complement of the cattle complex is 'useless', i.e., the number, type and condition of the Indian bovines do not *per se* impair the ability of the human population to survive and reproduce." I suppose that this hypothesis will be readily accepted. Human species has survived on the Indian subcontinent probably longer than in most other parts of the earth and recently has been reproducing itself faster here than in most other places. Therefore, whatever be the number, type and condition of the Indian bovines, Indian humans cannot legitimately complain that the bovines are impairing the human ability to reproduce and survive; surely not yet. Whatever tone might speculate about the future, this is simply not true of the present. I do not think any more evidence is needed to support this particular hypothesis.

Moreover, even supposing that it turns out that the Indian bovines, because, of their number, type and condition, are impairing the ability of the Indian humans to reproduce and survive, would it be right, on that account, to regard the Indian bovines to be ‘useless’? Would it not be possible rather to view them as performing a distinctly positive function in the total eco-system of which they are an essential part?

Harris’s second hypothesis is the following: "The relationship between bovines and humans is symbiotic instead of competitive." The meaning of the term ‘symbiotic’ is explained in a footnote as, follows; "According to Zeuner, ‘Symbiosis includes all conditions, of the living together of two different species, provided both derive advantages there from. Cases in which both partners benefit equally are rare’. In the symbiosis under consideration, men, benefit more than cattle." Here again, it seems that the hypothesis, in large parts is self-evident. For instance it is obvious; that the relationship between bovines and humans is symbiotic, in Zeuner’s sense, for otherwise the bovines would not be and could not be domesticated. What is not so self-evident is Harris’s assertion at the end, namely that ‘in the symbiosis under consideration, men benefit more than cattle’. We shall see what evidence, if any, Harris puts forward to support this part of the hypothesis. Of course, even if it turns out that cattle benefit more than men do, the relationship between bovines and humans would nevertheless be symbiotic in Zeuner’s sense.

Finally, Harris’s third, hypothesis is as follows: "It [the present relationship between bovines and humans in India] probably represents the outcome of Intense Darwinian pressures acting upon human and bovine population cultigens wild, flora and fauna, and social structure and ideology. Moreover, presumably the degree of observance of taboos against bovine slaughter and beef-eating reflect the power of these ecological pressures rather than *ahimsa* in other words, *ahimsa* itself derives power and sustenance from the material rewards it confers upon both, men and animals."

A little clarification seems necessary in the context of the present discussion, what is relevant is not the, doctrine of *ahimsa* but the sentiment attaching to all large animals which are maintained primarily for milk and, work. These include cattle and buffaloes and also horses, camels and donkeys This is a personal sentiment of the owner towards the particular animal which has served him during its working life. I am not sure that this is peculiar to India; it may be that it is a little stronger here than elsewhere. In India, this general sentiments greatly strengthened and fortified in the case of the cow because of the sacredness Hindu theology attaches to this particular animal. In the context of the present discussion, what is relevant is this general sentiment toward all large animals fortified by the religious feeling attaching to the cow. In contrast *ahimsa* is a much wider doctrine. It makes no distinction between the large and the small animal, and would be offended equally by the slaughter of sheep, goats, pigs and poultry which is quite common in India. Therefore, in what follows, I shall avoid the term *ahimsa.* Instead, for brevity, I shall use the term "cow-worship" and it should be taken to mean the general sentiment towards all large animals fortified by religious feeling attaching to the cow.

With this clarification, it seems to me that the first part of the hypothesis is self-evident or that it must be regarded to be so. The present relationship between bovines and humans in India is the outcome of years of co-existence of the two species in the midst of cultigens, wild flora and fauna and conditioned by a certain social structure and ideology. Therefore, it is only reasonable to suppose that it is the outcome of the Darwinian process; that is to say, if we, accept the Darwinian hypothesis, we should accept that the relationship between bovines and humans in India, as, I suppose, anywhere, else, is the outcome of the process postulated by that hypothesis.

But I should have supposed that we should regard the doctrine of cow-worship to, be a part of that ideology which is exogenously determined, such as by, Hindu theology, but which conditions the Darwinian process. Evidently, Harris views it differently and that constitutes the second part of his hypothesis under consideration. He thinks that the doctrine of cow-worship "reflects the power of the ecological pressure" and "itself derives power and sustenance from the material rewards it confers upon both men and animals." In other words, he views the doctrine of cow-worship not as an exogenously determined factor conditioning the Darwinian process but as itself, being an outcome of the Darwinian process. This, I presume, is the crux of Harris's argument. We shall examine what evidence he puts forward to support it.

II
Size of Cattle Population in Relation to Its Contribution

As Harris enumerates them, the principal positive functioned or useful contributions of India's sacred cattle to human survival and well-being are milk, traction, dung, beef and hides. We shall first examine milk, traction and dung, namely the contributions which cattle make during their lifetime. Harris observes as follows:

Milk: "In India, the average yield of milk per cow is, 413 pounds compared with the 5,000 pound average in Europe and the USA. However, this does not mean that the cows are useless or uneconomic, since milk production is a minor aspect of the sacred cow's contribution to the eco-system. It is more relevant to note that, despite the marginal status of milking in the symbiotic syndrome, 46.7 per cent of India's dairy products come from cow's milk".

Traction: "The principal positive ecological effect of India's bovine cattle is in their contribution to production of grain crops from which about 80 per cent of the human calorie ration comes."

Dung: In India, cattle dung is the main source of domestic cooking fuel. Many observers stress the slow burning qualities of dung and its special appropriateness for preparation of ghee and deployment of woman-power in the household. As manure, dung enters the energy system in another vital fashion. So vital is this, that old, decrepit and dry animals, might have a use after all, especially when the dung they manufacture employs raw materials lost to the culture-energy system unless processed by cattle."

These observations are well taken though, I think, it is necessary to discount the rather romantic appreciation of dung as fuel. Farmers use dung as fuel because no other, fuel is available and, in spite of "its special appropriateness for preparation of ghee and deployment of woman-power in the house hold", there is little doubt that, if other, fuel were available, farmers would cease using dung as fuel because they understand its alternative use as manure. But let us return to the main point.

Milk, traction and dung are all admittedly useful contributions of cattle to human survival, and well-being. But then two questions arise. The first is the question of costs. Harris recognises that the final evaluation of the contributions of cattle "must involve assessment of energy costs in terms of resources and human labour input which might be more efficiently expended in other activities". But evidently he does not know what these are. His only observation on this point suggests that he probably believes that there are no costs at all. He says: "Direct and indirect evidence suggests that in India men and bovine cattle do not compete for existence. On the contrary, the bulk of foods (straws and crop residues) that are ploughed into the soil in other countries are [here] converted into milk."

It seems to me that this is too superficial an assessment of the "energy costs in terms of resources, and human labour" and too naive, a conclusion to draw that "in India, men and bovine cattle do not compete for existence". Firstly, with the extension draw that "in India, men and bovine cattle do not compete for existence". Firstly, with the extension of irrigation and cultivation, the present cattle grazing is bound to be reduced. Secondly, the cattle consume besides straws and crop residues, a certain amount of other feeds, known as concentrates, all of which have alternative uses for human consumption or for manure as, for instance, is the case with oilseedcake. Thirdly, the straws and crop residues are mainly by-products of grain crops and with the introduction of high-yielding varieties of grain-crops, it is becoming evident that there is a competition between grain and atraw or crop-residue, Finally, once India begins to produce sufficient grain crops proportionately larger areas will be diverted to other crops such as sugarcane, cotton, jute, tobacco which have relatively little crop-residue ft for animal consumption. Hence, it seems not quite true that "in India, men and bovine do not compete for existence".

But of immediate relevance is another consideration, which Harris recognises, namely, whether "the energy costs in terms of, resources and human labour", which are presently expended on cattle, "might be more efficiently expended in other activities". Among these other activities is the possibility of maintaining a smaller cattle population but at a better level of nutrition. We should recognise this as another activity because it is qualitatively quite, different from India's present cattle industry, This leads us to the second question: Whether, with available resources, more milk, traction and dung could not be produced with fewer but better-fed cattle than at present. Apparently, Harris admits of this possibility. But he argues: "Much has been made of the fact that one large animal is more efficient dung, milk and traction machine than two small ones. But once again, there is another way of looking at the evidence. Is it not possible that India's cattle are under-sized precisely because other breeds never could survive the atrocious conditions they experience most of the year? I find it difficult to believe that breeds better adapted to the present Indian eco-system exist elsewhere."

Here, I think, Harris is not facing the original question squarely. That question is: Could more milk, traction and dung be produce with fewer but better-fed animals than at present? I submit that the question does not imply a necessary replacing of India's cattle by some exotic breeds from abroad. Therefore the question is not answered by asserting that breeds more productive and better adapted to the present Indian eco-system do not exist elsewhere. To meet the question squarely, Harris must make the following propositions: The atrocious conditions which the Indian cattle at present suffer most of the year are not the result, even partly, of the unduly large size of the cattle population that India possesses in relation to its land resources. Hence, there is no change, with a smaller cattle population, of relieving even partly these atrocious conditions and thus feeding the fewer animals somewhat better, If this is attempted, even the Indian cattle will not survive because the Indian cattle are not adapted to conditions of better feeding. Finally, if fewer but better-fed animals at all survive, they will not produce the same amount of milk, traction and dung as is presently being produced. For Harris, this would be a logically correct position to take, whatever the facts may be.

Expert appraisal of India's cattle population is that India has some fine breeds of cattle, some good for milk others good for traction and, presumably, all good for dung; that they are well adapted to Indian conditions and that with better feeding their productivity can be more than proportionately increased; that, therefore, with given resources, a smaller stock of such animals can give more milk, traction and dung than the much larger stock of nondescript animals at present gives. Harris is, of course, free to dispute this export appraisal or, once again, to view the situation differently.

III
Would Not a Smaller Number of Cows Be More Economic?

Before pursuing this point further, it will be convenient to get a certain detail out of the way. The contributions of cattle which we are presently considering are milk, traction and dung. Of these, milk is supplied by the female stock, traction, is supplied, generally, by the male stock and dung is supplied by both. With regard to traction, Harris suggests, along with many others, that with the existing technology and institutional arrangements prevailing in Indian agriculture, it is not possible to reduce the number of draught animals; that fewer draught animals, even if better-fed, would not do. As 'technology' and 'institutional arrangements' are sufficiently omnibus terms, it will simplify argument if we accept this proposition. This means that India has no surplus of draught animals.

But Harris, along with many others, proceeds to assert that in India, there is not only no surplus but that there is actually a serious shortage Of draught animals. In evidence, he Says that the existing stock is "insufficient to permit a large portion, perhaps as many as 1/3, of India's farmers to begin their agricultural cycle under conditions appropriate to their techno-environmental system". This is quite true. He further says: "Sharing of draught animals on a co-operative basis might reduce the need for additional animals [But] under existing property arrangements, there is perfectly good economic explanation of why bullocks are not shared among adjacent household". This is also quite true. But why do as many as 1/3 of India's farmers not have draught animals adequate to their techno-environmental system? The reason is not that not enough male calves are born in the country but that these farmers, with theft inadequate land resources, cannot afford to raise and maintain the draught animals they need. The plain fact is that, these farmers, functioning under a given techno-environmental system, need more draught animals but, because of their small land-holdings, cannot afford them. Therefore, a solution to their problem is to be sought not in a larger number of draught animals but in a change in the techno-environmental system, including possibly the existing property arrangements, prevailing at least in the relevant sector of Indian agriculture.

Let us therefore suppose that India needs all the draught animals that there are in the country but cannot afford any more. In a discussion of the right size of the cattle population, we thus isolate the cows. They produce milk, dung and calves to replace the existing stock of draught animals. The relevant question now is: Whether a smaller number of better-fed cows would not produce more milk and dung and also calves sufficient to replace the present stock of draught animals.

Harris now puts forward what he considers a 'subtle' argument to support his contention that a larger stock of under-nourished cows is to be preferred to a smaller stock of better-fed cows. He emphasises that "the principal function of bovine cows is not their milk-producing but their bullock-producing abilities" and argues that a large herd of under-nourished cows re-produces faster than a smaller herd of better-fed cows for the simple reason that there are more cows in the larger herd then in the smaller one.

This is indeed a subtle argument but it overlooks two points. One is the poor breeding performance of the under-nourished cows. Harris is aware of this but apparently thinks that it adds to the subtlety of his argument. He says: "In India, since cows suffer from malnutrition through restriction to marginal pasture, they conceive and deliver in, unpredictable fashion. The chronic starvation of the inter-mon-soon period makes the cow an irregular breeder. To the agricultural specialists with knowledge of what a healthy dairy stock looks like the hot weather herds of walking skeletons must indeed seem without economic potential. Many of them, in fact, will not make it through to the next monsoon. However, among the survivors are an unknown number still physically capable of having progeny. Evidently, neither the farmer nor the specialist knows which will conceive and when." So the farmers must keep them all.

May we ask: Are not the malnutrition and chronic starvation loft of the inter-monsoon period, which the Indian cows suffer, the consequences, at least partly, of the unduly large number of cows which India maintains in relation to its fodder and feed resources? Is it not that by holding the number of cows down, it may be possible to feed them better, improve their breeding performance and produce the same number of male calves as the hot weather herds of walking skeletons presently produce? This is the same question which we raised earlier: Are not the atrocious conditions which the Indian cattle at present suffer most of the year a result, at least partly, of the unduly large size of the cattle population which India possesses in relation to it land resources. Is it not that with a smaller cattle population, it may Be possible to feed the animals better and to improve their productivity sufficiently to compensate their smaller number?

Harris has brushed aside these questions by means of the following profound observation "Much has been made of India's having 115 heads of cattle per square mile, compared with 28 per square mile for the USA and 3 per square mile for Canada. But what actually may be most characteristic of the size of India's herd is the low ratio of cattle to people. Thus, India has 44 cattle per 100 persons, while in the USA the ratio is 58 per 100 and in Canada, 90. Given the increasing growth rate of India's human population, the critical importance of cattle in the eco-system, and the absence of fundamental technical and environmental changes, a substantial increase in cattle seems necessary". Harris does not see that what is indeed most characteristic of India's herd is the low ratio of cattle to people coupled with the high ratio of cattle to land and that therefore what is needed is a substantial increase of land so that the cattle-to-people ratio may improve without deteriorating the cattle-to-land ratio. So long as this is not possible, the fact will persist India's human population needs a much larger herd of cattle but it cannot afford even the existing one.

The second point completely over looked by Harris is the following. We said that India needs all the draught animals that there are; that, with the existing technology and institutional arrangements, their number cannot be reduced Very well. Does India need all the cow that there are in order to reproduce the present stock of draught animals? More specifically, must all the calves that are born today be born in order to replace the present stock of draught animals? The answer to this question is a straight and simple 'No'. It can be easily demonstrated that many more calves are born today than are needed and that the excess young stock is severely culled through a process of deliberate starvation so much so that barely one-third, of the calves are allowed to reach the adult age of 3.

We should therefore conclude that there is a distinct possibility, worth exploring, that a cattle population smaller in number but better fed than at present will make a better contribution in milk, traction and dung to human survival and well-being in India. Even if we suppose that, because of prevailing technology and institutional arrangements in Indian agriculture, it is not possible to reduce the number of draught animals very much, the large number of cows appears wholly unjustified. It seems that a smaller number of better-fed cows will contribute more milk and dung and will give enough calves to replace the needed stock of draught animals. Harris's cultural ecological view to the contrary appears unconvincing and illogical.

IV
Ecological of Starvation

Let us now consider beef and hides being the contributions which cattle Make to human survival and well-being on their death. As Harris says, in spite of the doctrine of cow-worship, India's cattle do not fail to make this contribution. He notes: "Few, if any, Hindu farmers kill their cattle by slaughter. But it is, not true that they do not kill their animals when it is economically important for them to do so. People will admit that they 'neglect' their animals, but will not openly accept responsibility for the *etic* effects, i.e., the more or less rapid death which ensues. There is much evidence that the Hindu farmer calculates carefully which animals deserve more food and attention. He culls his stock by starving unwanted animals and also, under duress, sells them directly or indirectly to butchers. The Indian peasant is nothing if he is not practical."

This is quite true. But, if Harris knows so much, one wonders why in another place he asks rhetorically: "How then do peasants tolerate the widely reported herds of useless animals?" and answers: "Perhaps, 'useless' means one thing to the peasant and quite another to the price-market - oriented agronomist". In fact there is no conflict between the practical calculating farmer and the price-market oriented agronomist. As Harris say, the Indian farmer "calculates carefully which animals deserve more food and attention" and when he discovers that some of his animals do not deserve any more food and attention, he "culls his stock by starving unwanted animals". What does it mean? It means plainly that the farmer finds that all the useful contribution some of his animals may make in milk, traction, dung and reproduction is not worth the food and attention he may have to expend on them to keep them alive. This is the simple meaning of the animals being "useless" and there is no conflict between the farmer and the agronomist on this point. Harris may of course, accuse the practical calculating farther, along with the agronomist, of being 'price-market-oriented'. The Indian farmer may indeed welcome such accusation coming from an American cultural-ecologist, because he has been far too much accused by the American agronomist of being exactly the opposite.

There is no conflict between the farmer and the agronomist. The conflict is between the practical calculation of the farmer and the Hindu doctrine of cow-worship. Caught between this conflict, the farmer resorts to the only alternative which is open to him, namely, 'neglect' of what he considers 'useless' animals. He neglects them and refuses to give them any more food and attention. The animals on their part, deserted by main begin to fend for themselves and to live on anything from garbage to standing crops which they can put their mouths to. The more aggressive and virile among them soon turn wild, and roam the countryside in marauding herds destroying crops sown by man. Most other, sooner or later, succumb to malnutrition, consequent disease and dire starvation. But, while they live, Harris wonders "how do the peasants tolerate the useless animals" and suggests that they may not be useless after all. This, I suppose, is ecologic.

As an aside, we may note how cultural ecology accounts for and justifies stray cattle or even marauding herds of cattle. Harris is aware of these phenomena and argues: "Many who protest most against destructiveness of marauding herds of beasts may perceive the situation from very special vantage points in the social hierarchy. But to have one's cow to eat other people's crops may be a very fine solution to the subsistence problem of those with no crops of their own." The socialistic fervour of this argument should surely put to shame the, advocates of socialistic pattern of society in India.

Harris seems to believe that the Indian farmer does not send his useless animals to the butcher because "the low price of beef, caused by the beef-eating taboo, makes it economically infeasible to send animals to slaughter." This is not correct. The beef-eating taboo probably keeps the price of beef low. But it would nevertheless pay to send the animal to the butcher rather than allow it to starve. This should be obvious. An animal before it is neglected and starved has more and better beef in it and also a better skin on it. Naturally, it would fetch the farmer a better price if he sold it to the butcher rather than allowed it to starve, and tall to the ground. A Hindu farmer does not send his animal to the slaughter not because it is "economically infeasible" but because it directly, offends the doctrine of cow-worship. If this doctrine were to be dropped, there is little doubt, that the farmer would sell his useless animals to the butcher rather that neglect and starve them to death as at present. If this were not so, what does Harris mean when, he says that "under duress" the farmer, sells his useless animals "directly or indirectly to butchers"? It only means that when the farmer is too hard-pressed to afford the luxury of cow ship, he contrives to send his animals to the butcher.

But, ordinarily, the Hindu farmer patiently suffers the luxury of cow worship and rarely sends his animals to slaughter. Consequently, the majority of Indian cattle at present die of 'natural' causes, primarily malnutrition and deliberate starvation and only few by slaughter. For information, only about 5 per cent of all deaths of cattle in India are by slaughter. But this does not prevent the cattle making their final contribution of beef and hides to human survival and well-being. Harris notes: "Although most Hindus probably, do not consume beef the eco-system under consideration is not confined to Hindus. Much of the flesh on the 25 million annually dead cattle and buffaloes probably gets consumed by human beings whether or not the cattle die naturally." Also, "The slaughter taboo does not prevent, depressed castes from utilising skin, horns and hoofs of the dead beasts. The quality of India's huge leather industry leaves much to be desired but the problem is primarily, of outmoded tanning techniques and lack of capital, not ahimsa." All this is substantially correct. But, from this, Harris concludes that if animals are slaughtered rather than neglected and starved to death, it will have "deleterious consequences to the rural tanning and carrion-eating castes." I confess I do not see how.

So far as I can see, there need be no deleterious consequences to the tanning castes. If animals are slaughtered, there will be more hides becoming available for tanning because as is known, a number of fallen hides, estimated to be about a third of the total, are presently not recovered because the tanner does not reach them in time. Moreover the hides will be in a better condition if animals are slaughtered in good time rather than starved and allowed to fall to the ground. The utilisation of bones, horns and hoofs of the fallen cattle is even more incomplete; presently, probably no more than 15 per cent are utilised because, presumably, their collection from a wide area is not economical. The fat from the bones is also almost totally lost before they are so collected. Other products such as blood, casings and glands of great pharmaceutical importance cannot be recovered at all from the fallen carcasses; these can be recovered only from, scientific slaughter.

So far as the carrion-eating castes are concerned, we should note that some of them have reportedly given up carrion-eating because untouchability goes with carrion-eating and to get untouchability removed, carrion-eating has to be abandoned. The flesh on the fallen animals is therefore probably no longer being utilised as fully as in earlier days. In any event I suppose, the carrion-eating castes may not resent a change from carrion to slaughter beef.

But Harris raises the relevant question whether the carrion-eating casts can afford slaughter beef. He says: "Indeed, could it be that without the orthodox Hindu, beef-eating taboo, many marginal and depressed castes would be deprived of an occasional, but nutritionally critical, source of animal protein?" It seems that Harris is equating cow-worship with the beef-eating taboo. This is not necessary. Of course, beef-eating taboo follows cow-worship. But giving up cow-worship does not necessarily imply giving up beef-eating taboo. Many Hindus do not eat the meat of buffaloes, horses, camels, donkeys and pigs; but they would have much less objection, to the slaughter of these animals than that of cattle. Hence, it the doctrine of cow-worship is given up a farmer will begin to sell his 'useless' animals to the butcher; but he may not necessarily accept beef in his diet. In fact, it seems certain that beef-eating taboo will last much longer than cow-worship.

Consequently, if more animals are sent to slaughter the price of beef will fall but, presumably, never so low as to make starvation of animals to death, more economical than sending them to slaughter. The lower price of beef will enable some of the carrion-eating people to shift to slaughter beef. Of course, some of them may not be able to buy enough beef even at the lower price; but they will lean to do without. They will not be the only people who will not have enough animal protein. There will be non-carrion-eaters too who may not be able to buy enough slaughter beef. Besides, there will be non-beef-eaters who may not be able to buy enough of other meats because their prices may be relatively high. Carrion-eating is surely not the fulcrum of India's cultural eco-system and the system will not collapse if carrion-eating vanishes.

V
Slaughter vs Starvation

Let us return to the central issue will slaughter, rather than starvation killing, improve the supply of animal protein from cattle both in the form of milk and beef including carrion? The answer to this question is a simple 'Yes'. Nevertheless, it will be convenient to consider it in two stages. In the first stage let us suppose that cow-worship is given up but beef-eating taboo persists in large measure. As we have seen this keeps the price of beef low but not so low that the farmers would rather starve their animals to death than send them to slaughter! However, because the price of beef is low the farmers will send only their 'useless' animals to slaughter. Nevertheless, they will send them to slaughter in good time namely before they are neglected and starved. The reason, as we have noted, is that there is more and better beef in the animals at this stage than after they are neglected and starved. It follows that slaughtering of animals rather than starving them to death will increase and improve the supply of beef.

Slaughter has another advantage over neglect and starvation. Animals when they are merely neglected, continue to subsist, on something until they inevitably succumb to starvation. Hence, though they are useless they are not entirely costless. Slaughters by removing such animals altogether, makes such resources available to the productive animals and consequently may improve their productivity. In particular, if 'useless' cows are sent to slaughter rather than allowed to live in a state of neglect and starvation, the productive cows may have a little more to eat and may yield some more milk.

Hence, it seems that slaughter rather than neglect and starvation of 'useless' animals will improve the supply of animal protein from both milk and beef. These are important gains from slaughter of 'useless' animal's. However, they will remain only marginal so long as the beef-eating taboo persists.

Real fundamental changes in the cattle economy can come about only when The beef-eating taboo breaks down substantially. As mentioned earlier, this is bound to take time, Beef-eating taboo may break down gradually with the development of more rational attitudes towards these matters. But more surely and directly, the taboo will break down under the pressure of low price of beef in comparison with prices of other meats such as mutton. Under pressure of continued price differential between beef and other meats, part of the demand will inevitably shift from other meats to beef. This will support the price of beef and the price of beef will begin to rise.

When this happens a fundamental change will occur in the attitude of the farmer towards his cattle. When the price of beef is very low, a farmer will not send his animal to slaughter until he is convinced that the contribution which the living animal may make in milk traction, dung and reproduction is not worth the food and attention he has to expend on it to keep it alive. But once the price of beef rises substantially, farmer will face a quite different situation. He will begin to compare the value of milk, traction dung and reproduction which the living animal makes, not with the cost of keeping it alive, but with the price it may fetch if sent to slaughter short. Consequently, he will be inclined to send the animal to slaughter earlier than otherwise and hence with more and better beef in it. Indeed, if the price of beef is high enough, he will find it worth his while feeding some animals, not so much for milk, traction dung reproduction as for the beef they accumulate on themselves. He will feed them so long as the additional beef they produce is worth the cost of their feeding and then send them to slaughter. Thus, he will find it possible to fatten some animals before sending them to slaughter. In other words, a proper beef producing industry will emerge. In short, under incentives provided by the high price of beef, the production of beef will of beef will increase.

It needs to be pointed out that the beef industry does not necessarily deplete the stock animals continuously. Many Indians, not being familiar with the nature of the beef industry, harbour this misconception. The beef industry in, the first place enables the stock to be maintained at the desired, level, by culling the excess stock Secondly, it enables the stock to be culled selectively and not indiscriminately as happens with the process of starvation; witness the large number of starvation deaths among young stock while old decrepit animals live. Thus, the beef industry does not deplete the stock continuously and indiscriminately it rotates the stock more frequently by replacing the less efficient by the more efficient animals. This improves the production not only of beef but also of milk because the beef industry also helps keep the dairy stock rationalized Indeed the beef industry is economically the most efficient method of regulating the number and quality of the Stock of animals so that a stock of optimum efficiency is constantly maintained. It is thus that the entire structure of the cattle economy gets fundamentally altered. It advantages are too obvious to need more explanation.

But again, Harris has his peculiarly ecological view. He says: "Since the least efficient way to convert solar energy into comestibles is to impose an animal converter between plant and man it should be obvious that without major technical innovations or drastic population cuts, India could not tolerate a large beef-producing industry. This suggests that insofar as the beef-eating, taboo helps discourage growth of beef-producing industries, it is part of an ecological adjustment which maximises rather than minimizes the calorie and protein output of the productive process." In short Harris think that breakdown of the beef-eating taboo and consequent emergence of a beef-producing industry will reduce rather than increase the calorie and protein output of the productive process.

It seems to me that Harris has now turned full circle. Consider first the proposition that "the least efficient way to convert solar energy into comestibles is to impose an animal converter between plant and man". 'We should note that Harris makes a distinction between the calorie and protein content of the comestibles. He also makes s distinction between vegetable and animal Proteins and believes that animal proteins are nutritionally critical (*vide,* "Indeed, could it be that without the orthodox Hindu beef-eating taboo, many marginal and depressed castes would be deprived of an occasional, but nutritionally critical source of animal protein?"). Hence if we suppose that the comestibles for human beings should have enough calories, sufficient proteins and at least the critical amount of animal proteins, one wonders whether imposing an animal converter between plant and man is in fact the least efficient way to convert solar energy into comestibles.

But let us take the proposition to be true. Should not its application be extended to other meats such as mutton and poultry? As Harris knows, in India, there is no taboo against these meats comparable with the beef-eating taboo and India has a large population of goats, sheep and poultry imposed between plant and man converting solar energy into comestibles. I suppose Harris will view this as a failure of the ecological process of adjustment. Evidently, beef-eating taboo does not go far enough in the ecological process of adjustment which Harris is envisaging; it needs hundred per cent vegetarianism.

Indeed, even hundred per cent vegetarianism would not do because, though vegetarianism would not recognise it as such, milk too is an animal product and, in India, some cattle are kept for milk. I suppose Harris will recognise this as yet another failure of the ecological process of adjustment. To be logically consistent, Harris should argue that a milk-drinking taboo added to total vegetarianism would have maximised further the calorie and protein output of the productive process in India.

As we have noted. Harris began by enumerating milk, traction, dung, beef and hides as the principal positive functioned contributions of India’s sacred cattle to human survival and well being. It seems we should now drop both beef and milk out of the list. We are then left with traction, dung and hides. But are not traction and dung also contributions, of cattle by means of which they indirectly convert the solar energy into comestibles? As Harris says: "The principal positive ecological effect of India’s bovine cattle is in their contribution to production of grain-crops from which about 80 per cent of the human calorie ration comes." This they do by providing traction for agriculture. On dung, Harris says: "Since grain-crops, cannot be digested unless boiled or baked, cooking is indispensable" and "In India cattle dung is the main source of domestic cooking fuel". Further: "As manure, dung enters the energy system in another vital fashion ... The sedentary, intensive rainfall agriculture of most of the subcontinent is heavily dependent upon manuring". It is thus that, by supplying traction and dung, cattle help indirectly to convert solar energy into comestibles.

Hence, if we accept the proposition that imposing an animal converter between plant and man is the least efficient way to convert solar energy into comestibles, we should also drop traction and dung from the list of positive functioned contributions which India's sacred cattle make to human survival and, well-being. How then do we fit in the symbiotic syndrome so large a cattle population imposed between India's meagre plant life and large human population?

It seems to me that Harris has moved full circle through the confusion surrounding the Sacred Cow, touching her piously at both ends and in the middle, as a good Hindu does, and is now back where he started, standing face to face with the Sacred Cow, unable to separate emics from etics. He may rest assured that many, before him, have done precisely the same, but without assistance from Cultural Ecology.

The truth seems to be that, considering the kind of comestibles that man needs, imposing an animal between plant and man may not necessarily be the least efficient way to convert solar energy into comestibles. But the number, kind and quality of the animals to be so imposed must surely depend upon the plant life available for such conversion. Whatever be the most appropriate number and quality of animals needed for this purpose, they cannot be secured except by firm, regulation and control of animal population; surely not by the automatic adaptive processes of Cultural Ecology.

TOWARD MODELLING POOR CITIES: A REVIEW OF URBAN ECONOMIC AND PLANNING MODELS

World Bank Staff Working Paper No. 232

THE WORLD BANK

Bank Staff Working Paper No. 232
April 1976

TOWARD MODELLING POOR CITIES.

A REVIEW OF URBAN ECONOMIC AND PLANNING MODELS

As the cities in developing countries and the attendant problems have been growing rapidly, interest in urban analysis has also increased markedly. This paper provides a survey of the main varieties of urban models which have been developed so far, with a view to finding approaches useful for improving our understanding of cities in IDC's. In the process it reflects upon the complexity of urban phenomena which intrinsically cause difficulties in formulating urban economic models. This review should prove of particular interest to several groups of readers, including: (1) researchers experienced in other fields who want to become familiar with the context of urban research; (2) urban researchers more familiar with other aspects of the urban field who want to know more about the state of formal urban economic modelling; and (3) urban policy analysts who wish to evaluate the potential effectiveness of recent urban modelling efforts in assisting policy formulation. A summary of this relatively lengthy paper on a generally unfamiliar subject is provided under cover.

Prepared by:
Rakesh Mohan
Department of Economics
Princeton University
Urban and Regional Economics Division
Development Economics Department
Development Policy Staff

PREFACE

This paper surveys the main varieties of urban models with a view to finding approaches that can be useful for understanding cities in less developed countries. Some comments are also offered on the complexity of urban phenomena which intrinsically cause difficulties in formulating urban economic models.

Notation is always a problem when comparing a large number of models. Here we have tried to maintain a consistent notation so that the same letter stands for similar variables in different models. One has to say "similar" since each model has its own variation in the definition of essentially the same variable. The notation is explained as it is introduced, and every effort is made to maintain reasonable comparability with the original articles.

I am indebted to the following who read the first draft with care and graciously provided useful comments: Orville Grimes, Jr., Peter Watson, Edward Holland, J. Ben-Bouanah, Sven Sandstrom, Chris Turner, Raymond Struyk, James Ohls and Edwin Mills. The final form of the paper owes much to the diligence and care with which Bertrand Renaud and Douglas Keare have read and re-read it - always providing helpful critical comments and not letting me get away with much.

TABLE OF CONTENTS

SUMMARY

This paper explores the possibilities of applying current urban modelling techniques to cities in less developed countries. Some problems intrinsic to the modelling of urban areas are discussed first. Uncertainty about the present, unforeseen changes in the future, indivisibilities and economies of scale are posited as the reasons for the existence of cities, and these are the phenomena which are most difficult to codify and to include in models. The peculiarities of cities in LDCs are then described to explain why modelling their activities is an even more complex task than modelling Western cities. These peculiarities stem from unprecedentedly high population growth rates in LDC cities, the co-existence of high and low levels of technology, along with predominantly low incomes and the decline in relative prices of transport and communications as compared with a century ago. Before a review of models is begun the criteria for model evaluation are set out.

Urban models are divided into two classes: analytic models and operational or policy oriented models. The former are mainly based on economic theory and are usually aggregate type models attempting to offer basic insights into urban form. Almost all of these models investigate optimal residential location -- either from the point of view of the household itself or for maximization of social welfare as expressed in a welfare function. Particular attention is paid to the land/transport trade-off; the uniqueness of location; the effects of transport congestion on city form; and the consequences for egalitarian welfare functions. Though these models are at a sufficiently theoretical plane, they should be regarded as conceptual building blocks towards more realistic models.

Operational models characteristically require the use of a computer for their solution. There are two basic strands in such urban modelling: the social physics variety now based on entropy maximization techniques and the behavioral variety mainly drawing on the economic analytic models. The social physics variety of models is introduced with an explanation of the increasingly used entropy maximization techniques. The objective to distribute activities spatially in a city. A of these models is typical output of a model would be allocation of residential and employment in the city -- often disaggregated by socio-economic types location by zones of households, types of residential structures, types of employment, etc. These models reproduce current city structures fairly accurately, after careful calibration. However,

without behavioral underpinnings, their methodological basis raises doubts as to their usefulness as predictive or planning tools, given the rapidly changing conditions of IDC cities. The behavioral models reviewed have many pleasingly realistic features; however, they are to unwieldy to operate, are highly data intensive and do not easily permit evaluation of alternative courses of action within a realistic decision-making context. Thus many of the operational models tried so far are found to be of limited practical use in LDC cities.

Some productive approaches to urban modelling in LDC cities are suggested on the basis of three models: each quite distinct in approach. The first is the multi-level approach to modelling of the Master Planning can Commission of Stockholm. Their model be operated at various levels of aggregation and is particularly suited for dialogue with policy-makers. Moreover, it is an evaluative model. The second is the Urban Institute Housing Model which models the behavior of households, owners, a building industry and government by positing 'model' representatives of each group. Thus the model is small in size and flexible in data requirements, though mathematically complex. It is particularly well suited for modelling the direct and indirect effects of different governmental policies. Finally, Apps' model of housing demand at a disaggregated level is presented as an example of a useful explanatory model to explore behavioral relationships and parameters which can then be used in operational models.

It is the major conclusion of this paper that analytic and operational models for LDC cities have to be developed simultaneously: operational models should be of small, sketch-planning types while analytic models can be more disaggregated.

I. INTRODUCTION

1.1 Why Model Cities at All?

> "But as for those who posit the ideas as causes, firstly, in seeking to grasp the causes of the things around us, they introduce others equal in number to these, as if a man who wanted to count things thought he would not be able to do it while they were few, but tried to count them when he had added to their number".

Aristotle (Metaphysics Book 1, Ch.9)

Such was Aristotle's criticism of Plato's Theory of Forms. Plato sought to comprehend reality around him by defining general 'Forms' and interpreting things similar to them as their particular manifestations. The objective was tc' reduce the size of the problem by having to comprehend only classes of things rather than each by itself. The difficulty with this was, however, a question like, "what is the essence of a table that makes it a table?". Thus ultimately the number of Forms precisely equals the number of 'things around us' and our quest for comprehension of the universe brings us back to our starting point.

Such is also the predicament of the urban model builder. One seeks to build a model of an urban environment with the object of reducing the complexity of the observed world to the coherent and rigorous language of mathematical relationships. (Lowry, 1965). When this language becomes complicated or the size of the 'simplified' model becomes so large that it assumes a complexity of its own, one begins to question the usefulness of such an exercise. Thus when a model of an urban environment is sought it is crucial to keep in mind the objectives of the exercise.

Planners and policy makers in the urban sphere have become prisoners of the idea that in the city everything affects everything else. (Lowry, 1965). If, indeed, everything is interrelated then every public decision needs to be an informed one if it is to achieve its aims. The urban policy makers must be aware not only of the immediate effects of their decisions but of the indirect effects as well. For example, the provision of a new speedy travel mode has the direct effect of reducing travel time for its users. However, its indirect effects on industry location and consequently on employment and residential location could be far greater in magnitude. The policy-maker, knowing the existence of such interrelations, demands knowledge of their magnitudes. Hence, the demand for models. Models should therefore help the policy maker in understanding the underlying determinants of spatial location within a city; in analyzing the causes of city growth and decay; and in predicting future land uses in specific parts of urban areas. If a model can help to show that commercial development A will take up X hectares of city center land, cause y percent more traffic congestion, and generate, z percent less tax income and jobs in ancillary industries than industrial development B, then a policy maker will be interested in that model.

These, however, are a daunting set of demands which social science can scarcely supply even if some of its practitioners pretend that they can. Recent developments in computer technology and in the use of mathematical techniques in the social sciences have raised people's expectations of what can be predicted about the future. Social scientists themselves have been instrumental in raising the level of demands that policy makers make of them. This results in disappointment more often than not.

In the context of urban systems the demand for models is particularly great because of the large number of variables, large number of available policy parameters, the complex interrelationships among them and the long term consequences of public decisions. The sheer size of the number of variables would not be much of a problem if they were not thought to be correlated. If they were independent, problems could be solved partially or sequentially and there would be no need for models. Planners, therefore, need models for the following purposes:

(i) prediction and projection
(ii) impact analyses of alternative strategies
(iii) plan design
(iv) educating planners
(v) controlling and directing urban change.

A comprehensive model would be one which met all these needs. Such a model has not yet been formulated. Different models meet different needs. The following kinds of models can be distinguished:

(i) theoretical
(ii) policy making
(iii) data manipulating
(iv) educational
(v) measurement devices.

Theoretical models are at a high level of abstraction seeking to gain basic insights into urban structure. Such models can be useful in educating planners as well as in the design of more operational models. Policy making models can be merely predictive, or optimizing or of the impact analysis type: they all seek to help planners and policy makers. Data manipulating models, (e.g., input-output models) are good devices for checking consistency of data while they can also throw light on the structure of some interrelationship. Educational models can be game playing devices with which planners can be educated. While they may not be strictly operational policy makers can use them to test possible effects of policies.

It is evident from this classification that it is not a mutually exclusive nor an exhaustive one. It is presented to illustrate the point that different models do different things which attempt to satisfy the different needs mentioned earlier. Users of models must be as clear about what they cannot do as about what they can. If, for example, a model is merely predictive (in the positive rather than normative sense) it must not be taken to imply what ought to be.

In conclusion, policy makers must be clear about why they want to model cities before they set about financing model builders' dreams. They then have some chance of achieving their objectives.

1.2 Some Reflections on City Phenomena

Having concluded that there is some need for models of the urban environment, it is appropriate to reflect on the intrinsic characteristics of cities which cause problems for model builders. City phenomena are replete which analytic inconveniences such as increasing returns to scale, indivisibilities, interdependencies and minimum size thresholds. Some of these can be captured by models and some cannot. Whatever the case, they are the phenomena that need to be studied and understood.

To comprehend and, perhaps, manipulate the structure of cities we need some understanding of why they exist in the first place. Cities have existed for a long time. Jane Jacobs (1970) has offered some conjectures on the reasons for their existence and their role in economic development over the ages. She is of the view that cities are the primary economic organs and that most (if not all) economic advances of note have taken place in cities. She even argues that agriculture (as we know it) started in cities and only later spread to the countryside. Furthermore, most agricultural innovations have been city-based and agricultural productivity is therefore a derivative of city productivity. Whatever the merits of her historiography and chain of temporal reasoning some lessons can be drawn from her work. The main one is the idea of cities as centers of innovation. One immediately wonders 'why are cities innovative?' Words such as agglomeration, proximity and minimum threshold size come to mind but more precise explanations do not. The ideas of agglomeration and proximity are at the root of the existence of cities and immediately point to the importance of the spatial structure of cities.

Many cities evolved originally as market or trading centers. This was necessary for product markets to expand and consequently for economic growth. Roland Artle (1972) has suggested that the characteristics of the income elasticity of demand for goods provide us with clues as to why cities invariably seem to go together with economic growth. It is the essentially simple idea that the income elasticity of demand for primary goods like food is low and declining with income while that for urban type goods and services is high and increasing with income. The question then is, 'can these products be manufactured and provided in rural areas?' The answer appears to be negative. Even craft products need markets and capital. Only if a craftsman can expect to sell this product can he afford to invest in capital. Risks are reduced if he has proximity to his customers and more so if the potential number of customers is large. The production process does not need economies of scale for the size of expected market to be a factor. Such a consideration for production is difficult to capture in a mathematical model.

As markets expand and production increases concern over availability of inputs also becomes important. The likelihood of obtaining labour, more importantly labour of the right type, at a given time is clearly greater in a concentrated population than in one that is dispersed. Other inputs like raw materials and manufactured goods are also easier to obtain. We thus have the notion of interdependencies between products. We must distinguish between two kinds of interdependence. One is the kind represented by an input-output matrix. A dense matrix, i.e., one with many non-zero entries represents a high level of interdependence. The process of production is then an intricate mesh: most products require as inputs many other products. This can be termed technological interdependence. Such interdependence does not necessarily produce a city if transport costs are not high. As Moses and Williamson (1967) have attempted to show, cities that grew in the nineteenth century have the structure they do because the cost of moving goods within cities was high relative to (a) moving people within cities and (b) moving goods between cities (by train or water transportation). Industries were thus located very densely and near major transportation modes. The advent of the truck made intracity transportation of goods cheaper and this is posited as one cause of decentralization. The other kind of interdependence arises from technological interdependence but is more related to 'uncertainty'. Physical proximity of establishments is then of some economic benefit. Vernon (1959) found this to be a major characteristic in relation to the existence of certain types of industries in New York City. Activities which are dependent on changing output demand have to change their own input demands in response. The necessity to do this in a short time span makes proximity necessary. Risk is spread: a supplier who loses one client can easily switch to another before deciding to change his product line. To summarize, technological interdependence is important for spatial concerns because of transport costs while the other kind of interdependence is important because of information costs and uncertainties.

Moving on to the wider aspects of modern economic activity the interdependence of production with such services as banking, insurance, marketing, etc. has become more important in the agglomeration economies of modern cities. Some of these economies can be dealt with by thinking of indivisibilities (or minimum threshold size) in the size of such activities and can be represented by constraints to a production function. If, however, such activities are regarded as inputs into the production process their representation is not so easy. Agglomeration is necessary for these activities because of the need for face to face contact. One would have expected the telephone and other communication advances to make face to face contact unnecessary, but this does not appear to be so. It is, however, difficult to put economic values on the benefits of such contact and therefore to subsume it in economic analysis.

The existence of economies of scale in certain industries is the next most important reason for the existence of cities. Even if a few industries exhibit economies of scale their effect through backward and forward linkages will be much greater. We thus see whole cities like Detroit, which are based primarily on a single industry which has economies of scale. This can, of course, be represented easily by production functions e.g. by a Cobb-Douglas production function whose sum of exponents is greater than one. The problem created by such functions is that they lose many of the 'nice' properties of constant returns or diminishing returns to scale production functions.

Having dealt with the production side it is also import-ant to recognize the indivisibilities and economies of scale on the consumption side. There are many collective services whose consumption is characteristically joint. Learning and research are activities which are done jointly; moreover, they are tightly correlated and further produce technical change. They thrive on agglomeration. The provision of health services, transportation and recreation also tends to be collective. The more densely an area is settled (within limits) the greater are the potentials for the use of such services. Since such services are collective the private market does not operate very well in handling them. Thus some kind of public authority is necessary to provide them. Consequently, policy makers come into play and ask for guidelines to help in taking decisions.

Finally, there are 'bads' associated with agglomeration. On the production side there are negative externalities like pollution and monopoly resulting from economies of scale. On the consumption side, over-crowding can make the provision of collective services difficult.

In addition to the problems mentioned above, although in part because of them, cities change in rather unpredictable ways. The growth process is of two kinds: one is the multiplication of existing facilities and the other is the initiation of new activities. The stability of a city depends on its ability to cope with change. Most industries have long term cycles: they are created, go through a boom period and then decline. If a city is based on one industry the city itself goes through the same cycle. If, however, the city is based on a variety of industries it is unlikely that all their cycles will be in phase and the city will then be seen to be more flexible and able to cope with change. At any one time, then, there will be both efficient and inefficient industries. This is, perhaps, what Jane Jacobs means when she says that cities need to have inefficiencies and impracticalities built into them to cope with the uncertainties of chaning technology and the economic environment. This is not unimportant, since urban infrastructure usually lasts about fifty years. Apart from single industry bases, cities built with single transport modes are more vulnerable than others. European cities built when no motorized transport was available have adapted well to the many changes in the modes of transport that have taken place over the past century. In contrast, North American cities built in the last seventy to hundred years are so dependent on automobile transport that they will find the effects of oil price rises more difficult to handle. Models can scarcely capture such characteristics of cities. If a model is an optimizing one and is calibrated on contemporary data its prescriptions may well produce an efficient city for the present but a disastrous one for the future.

Uncertainty about the present, unforeseen changes in the future, indivisibilities, and economies of scale have been posited as the reasons for the existence of cities. These are precisely the phenomena most difficult to codify and to include in models. At best approximations can be tried. Non-linearities can be built in but they make solution and handling of the model more difficult. Complex interrelationships between variables and the durability of structures also makes the impact of decisions more difficult to analyze. Thus model builders have to be modest about what they are trying to accomplish and policy makers less sanguine about what they can expect from models.

1.3 Cities in Less Developed Countries

Cities in the less developed countries are in many ways even more difficult to handle than those in the Western world. The main reason for these difficulties is that unlike most Western cities the former have grown suddenly and explosively in the present century and have attained sizes comparable to those elsewhere but at lower levels of income. Analysis is more complex because of the coexistence of technology from different When analyzing transport systems, for example, we have a limited number of modes to worry about in Western cities. In an Asian city, the other hand, the electric train coexists with the hand-pulled rickshaw, bicycles, scooters, automobiles and, of course, walking.

Even during the period of rapid urbanization in Europe, rates of population growth in cities were on the order of about 0.5 percent per year (IBRD, 1975); whereas the populations of cities in the LDCs are presently growing at the rate of 3 to 7 percent per year. With a growth rate of 0.5 percent European cities had time to adapt and evolve as they grew in size. Diffusion of innovations was slow and economic, social and political institutions emerged to regulate patterns of growth and to govern cities as they grew. This can be seen as an equilibrium process, i.e., one which is roughly at equilibrium at every stage.

The current LDC experience is quite different. A primary factor is that the population growth rate is truly explosive as compared with historical experience. Even with the best of intentions, controlling this growth rate is not easy, especially in the short run, for it has a number of implications. First, the growth rate of cities is high without migration. Second, technical change in agriculture which raises farm productivity has to provide for the growing rural population in addition to the burgeoning urban population. In times when overall population growth rates were less than 0.5%, a 3% growth in agricultural productivity was tremendous. Now such a growth rate is barely enough to feed the rural population thereby leaving precious little marketable surplus and hence a sluggishly growing demand for urban products. Cities are therefore not as tightly connected with their hinterlands. Third, a high population growth rate implies that cities can become larger without significant increases in the urbanization levels of a whole country. Fourth, such a pace of increase makes the structural growth of the city even more unpredictable and consequently more difficult to manage.

The second major difference from the European experience is the coexistence of 'high' and 'low' technologies in LDC cities. This makes the demand structure of the rich qualitatively different from that of the poor in a manner more pronounced than that in European cities at similar levels of incomes. Thus the demand pattern of the rich in LDCs corresponds roughly with the current Western rich. They demand, and receive, services and products which use twentieth century technology, while the poor still live in much the same way as they might have a century ago. Thus the 'technological inequality' in consumption between the poor and rich is more pronounced than the *income* inequality. The latter was probably quite similar in Western cities in their early stages to that existing in LDC cities now.

Unlike the European experience, then, technological growth, income growth and population growth are unbalanced in cities in LDCs today. High population growth and technological growth makes it possible to have cities with a multi-million population at low levels of income. This was just not possible in the last century. The sewage disposal, water supply and transport problems of a city of 8 million are qualitatively different from those of a city of half a million, which was the size of major western cities at similar income levels.

The third major difference is the decline, in relative prices of transportation and communication costs. This makes for less centralized cities, as seen in some urban densities in Latin America. Workers can live away from their place of work; information and innovations travel faster. Because of other imbalances this diffusion is uneven and urban centered thereby making primate cities more important. The international diffusion of information as well as innovation reinforces this tendency of concentration in primate cities, thereby further exacerbating technology and income inequalities. Elites are internationally mobile: they characteristically have more contact with Western cities than with their own hinterlands. Their demand structure is therefore more in tune with Western factor proportions. The obvious urban effects are those in housing where differences between rich and poor become extreme.

These major differences combine to produce effects which lead us to expect cities which have a different spatial structure from those for which models and plans have been developed. Until now, LDC cities have been lumped together and discussed as if they are all similar, and different as a group from Western cities. It is, however, worthwhile to disaggregate them. Clearly, there are systematic historical and geographical differences among LDCs. The IBRD (1975) classification, based on the urbanization experience, is interesting, but this paper is more concerned with urban form. Our classification below (suggested originally in Mohan, 1975) is a hybrid one and consequently less well ordered logically than the IBRD one:

(a) *Pre-Industrial Cities:* These are cities which existed in LDCs before colonization, or before the modern era. They grew 'naturally' with highly congested central areas reflecting their existence before the motorized age. Having grown within and along with the traditional economy they have tight links with their immediate hinterland. Their own pattern of production, commercial relationships and transportation patterns can be expected to reflect this. They often have superimposed on them modern influences usually characterized by a modern business center existing along with the old one though spatially separated. Examples of these are some cities in the Middle East and North Africa, e.g., Ibadan in Nigeria, Hyderabad and Lucknow in India. Mombasa in Kenva.

(b) *Industrial Cities:* These are usually large port cities which were the main colonial importing and exporting centers. They are characterized by a great amount of Western influence. They have links with a wider hinterland, but the links are purely commercial. They have few 'traditional' activities. Their structure is more like Western cities, more geared to motorized transportation. This category should, perhaps, be further disaggregated to reflect varieties of colonial experience. Examples are Singapore, Hong Kong, Calcutta, Bombay, Rio de Janeiro, and Buenos Aires.

(c) *Post Industrial Cities:* These should really be called administrative cities but are called post-industrial for logical neatness. They are post industrial in the sense that they have really become 'cities' in the last two or three decade. These cities are usually quite sparsely populated and much more planned (initially). They have grown primarily as capital cities with somewhat questionable 'economic bases'. Some are in the process of attracting more industry and diversified economic activity in addition to governmental activity. Examples are Brasilia and most African capital cities.

The object of giving this brief account of different types of LDC cities is to illustrate the reasons why these cities can be expected to require different kinds of analyses. They also have, large differences in income levels which have to be taken into account.

We now illustrate how the three major differences between LDC cities and Western cities combine to make the life of model builders difficult.

(i) *Rapid population growth* of these cities at low levels of income but with high and low technology produces segmented markets. Here we focus on the resulting distinctions between the 'formal' and 'informal' labor markets. The model builder is usually accustomed to looking only at the formal markets. The journey to work is then regarded as one of the most important components of urban economic models and as the major transport activity. Whether this is so in LDC cities is not clear. With less organized economic activity employment appears to be more diffuse. Firstly, the large 'informal sector' provides moving employment: hawkers, service oriented people, etc., have no fixed place of work but do have areas of operation. Such employment may well account for about 10% of all urban employment, as in Peru. Their journey to work cannot really be defined and their work location covers a large area. Secondly, small manufacturers, artisans and shopkeepers often live and work in the same location. They are mostly self-employed and could account for another 10% of urban employment. Thirdly, it is often the case that people commute from the neighboring countryside and stay in the city for part of the week and return to their homes for the other part. Thus it would be surprising if journey-to-work oriented models of urban location are appropriate for LDC cities. (However, precise information on these transportation patterns is scarce so these remarks should be taken as speculative). The informal market is itself heterogeneous and the range of incomes within it is large. Analysis is further complicated by the fact that a majority of informal sector workers are secondary income earners in the household. This makes the derivation of residential location from work location much more complicated. If poorer households characteristically have more than one income earner then this problem is a serious one for models generating residential location distributions. The existence of informal markets also causes difficulties for models which are based on the 'basic' or 'export' oriented employment concept which is then used to generate all other kinds of employment. The relationships between markets in LDC cities needs more investigation to find such multipliers.

(ii) *Transportation modes:* Analysis and model building is simplified considerably by the assumption of single modes or, at the most two modes. Existence of low incomes along with high technology produces great problems for the model builder in this area. In LDCs, the poorest walk, those slightly better off use a bicycle, then a bus or a train; and the richest own automobiles. There are also other modes like the jeepney, scooter, rickshaws and cycle-rickshaws. Some evidence from Delhi (Sarna, 1975) is worth quoting here. Excluding walking, 42% of commuters used bicycles and about 37% public transport. For low income people at least 35% of work trips were walking trips and about 25% for middle income people. Average work trip length was remarkably short: 12 minutes for fast vehicles, 16 minutes for mass transit and 11 minutes for bicycle users (no data for walkers). This evidence coupled with the spatial origin of vehicular work trips indicates dispersed employment. Those going to work by fast vehicles live closest to the city center and those not using fast vehicles the farthest. The calculation of urban transport costs and their effect on residential location is then not a simple task but one that requires ingenuity if models are to be developed and are to be of use in LDC cities.

(iii) *Housing:* High migration rates, low incomes and capital scarcity produce squatter settlements, shanty towns and slum areas. Such housing accounts for about half of total housing in many LDC cities. For model builders this produces two problems. First, data are intrinsically difficult to gather from such neighborhoods since many structures are illegal. Secondly, the variety in housing is as great or greater still than in modes of transport. The production technology for straw huts is clearly different from that for tall buildings. Even if the form of the production function is not different the factor proportions are clearly so. In Western cities this is not the case. At most one can usefully distinguish between low rise and high rise buildings. Production technology is not much different. In LDC cities, different kinds of housing require different kinds of inputs from different markets and the housing market itself is more segmented.

(iv) *Factor proportions:* The existence of a city implies higher capital-land and capital-labor ratios (although the latter is not obvious). Given that LDCs are capital scarce but have similar sized cities as the West (in population) we should expect their structures to be different. Given that LDC cities do have some skyscrapers it would appear that the mere existence of such structures distorts the allocation of urban resources. Large populations densely packed in, drive up the price of land in central cities, resulting in the use of skyscrapers as a form of capital to substitute for land. From the context of the whole economy this may not be an optimum allocation of resources. The provision of urban services: sewerage, water supply, electricity, roads, mass transportation, etc., is necessary as well as capital intensive. The factor proportions in these activities are probably nearer those in developed countries than those in the country generally. Here the problem is probably of limited technological choice. Nonetheless, it creates allocational and financial problems for LDC cities.

(v) *Choice:* Low levels of income limit the alternatives available to the poor. Here the model builder is perhaps helped in that some distributions could be over-determined. The limiting case would be one where the poor may be so constrained that their location decision has no element of choice. Usually, though, their choice is expected to be limited, but decisions still have to be made even if they are among a few bad alternatives.

Finally, we come to one aspect of LDC cities that is difficult to account for in economic models (except exogenously) but is in some sense the reason for building such models. This is the role of public authorities, in the spatial decisions in the city. Households can be regarded as utility maximizing and firms profit maximizing. Modelling the behavior of public authorities can follow no such rules. Although centralized political authority has long been one of the reasons for the existence of some cities, it has played an especially dominating role since World War II. Governments have become more active in general since then and particularly so in LDCs. The problem is further complicated by the fact that though the effects of governmental decisions are all pervasive in the formal sector they are less visible in the informal sector. We have very little information on these effects. Theories adequately capturing the interaction between governmental decisions and economic activity have yet to be developed, although governmental reaction functions and the like are now being posited. Some role has obviously to be assigned to governmental activity in urban models of LDC cities.

This has to be exogenous as well as endogenous for the effects of exogenous public decisions are easily dealt with by sensitivity analysis.

To summarize, the major differences between LDC cities and Western cities stem from:

a. The unprecedentedly high population growth rates of most LDC cities;

b. the coexistence of 'high' and 'low' technology along with predominantly low incomes; and

c. the decline in relative prices of transport and communication compared with a century ago.

These factors give rise to segmented markets, the coexistence of many transportation modes, squatter settlements, shanty towns and slum areas; while the high intensity of capital in urban areas is not commensurate with factor proportions existing elsewhere in the country. These effects make formal models of LDC cities difficult to develop. Because information on LDC urban phenomena is seriously lacking, these notions are somewhat speculative.

The next section suggests criteria by which urban models may be judged.

1.4 Criteria for Model Evaluation

In reviewing urban models the first question to be asked is the objectives of the models. This has been addressed in the first part of this section. In listing criteria for judging models we present one set for explanatory analytic models and another for policy-oriented ones because their objectives are different.

A. Explanatory Models

(i) *Model Structure:* Model structure should be easily comprehensible by those reasonably literate in the field. Since explanatory or analytic models operate at high levels of abstraction making many compromises with reality, their structure must not be obtuse. If a complex reality is being modelled we cannot always ask for the model structure to be simple. But a complex model structure can also be presented so that interrelationships within the model are clearly articulated. This is important because larger policy-oriented models are often built on the basis of these analytic models.

(ii) *Correspondence to Reality:* As mentioned earlier, it is recognized that analytic models have to make certain compromises with reality to remain explanatory or analytic models within a reasonable size. However, they are explanatory only to the extent that they explain some urban phenomena however aggregated. Thus a good model remains close to reality in its structure and output.

(iii) *Output:* The results of a good analytic model should give insight into *some* aspect of urban phenomena. For example, a model can focus on the location decision and illuminate its behavioral determinants. An explanatory model is of little value if it does not accomplish at least as much. It is at this level of modelling that the rationale behind urban structure is seen.

(iv) *Normative and Positive Results:* Models should not create confusion between what is and what ought to be. Assumptions behind a good model are made explicit. This is important because the results can then be judged on the basis of these assumptions. Moreover the role of the assumptions in the results can then be clearly seen. If a model is predicting something, a good model should make clear if it is a projection or if it is a prescription for a plan of action. If it is a prescription then the criteria for such a prescription should be spelled out. This is a problem when mere projections of current trends are regarded as targets.

Before proceeding to the criteria for policy oriented models it is useful to see some connections between the two kinds of models. Explanatory models are not necessarily small. They essentially seek to further understanding of the urban process. They cannot usually be directly used in policy planning. Policy-oriented models have to simulate reality reasonably closely to aid in decision making. They have to take account of all interrelationships that have a bearing on the problem under consideration. This usually necessitates sub-models within a larger model. The relationships in these sub-models usually operationalize the behavioral relationships found in explanatory models. An analogy from macro-economics is to use all the research on the consumption function to form the consumption equations in large econometric models. This set of equations could be disaggregated for different product markets and could comprise a set of 30-40 equations. The form of all these equations would have come from the basic consumption function studies. Similarly, the insight of declining land rents as one moves outward from a city center when built into a policy oriented model will probably be disaggregated into segmented markets for different kinds of demands - office space, residence, manufacturing, etc.

Even highly theoretical models are then not totally useless as long as they exhibit the qualities specified above. They ultimately become useful for policy design purposes through their utilization in policy models.

B. Policy Oriented Models

These models are usually more difficult to evaluate since they are usually larger and more complex. They are difficult to understand because of the many bits of inputs that have to be fed into them, the many equations which constitute their structure and the many outputs they provide. They can be judged according to the following criteria:

(i) *Objectives*: Complex as they are, policy-oriented models must be absolutely clear about their objectives. Clarity of objectives makes it easier for policy makers to use them. They can then judge if the model is accomplishing the objectives it was designed for. A predictive model is quite different in structure from a prescriptive one. A spatially disaggregated one is again different from an aggregated one, and so on.

(ii) *Data requirements:* In evaluating a model for LDC application its data requirements have to be looked at closely. As described later, half the budgets for policy-oriented models are characteristically devoted to data collection. Since data collection in LDCs is even more of a problem, data requirements of models have to be scrutinized even more carefully. Several issues are involved:

- measurability
- accessibility
- cost

It often happens that some kinds of data requirements are just not measurable. They might fit into neat conceptual slots but are not possible to measure in practice. Other kinds of data are measurable but often not easily accessible. The length of time involved in their collection is a consideration that is particularly important given that the rate of growth of LDC cities is so high. If, for example, a set of data requires two years to collect, it may well become obsolete sooner than it is collected. Resources for such data collection are scarce, as are the facilities for processing such data. Thus, the more a model uses readily available data, the cheaper they are to obtain and the more quickly the model can be put into operation.

(iii) *Inputs:* In the description of a model there is often a great amount of confusion between which variables are endogenous and which are exogenous - in other words which variables have to be externally provided. Among the exogenous variables, the distinction between control variables and other variables should be clear. Control variables are the ones that policy makers can play with and should therefore be made explicit. Their operational meanings should also be clear. The clear specification of inputs is also important because it makes it easier to understand what a model is --and is not-- simulating.

(iv) *Model Structure:* This is the core of the model on which its usefulness really depends. One believes a model to the extent that one has faith in its structure. We discuss various aspects of model structure:

-- *Silmplicity:* The structure of a good model is such that it can be comprehended. It need not be simple in the sense that it does not have complex relationships embedded within it. It should be simple from the viewpoint of explanation. An opaque structure is such that it seems like a black box which cannot be easily explained. Such explanation is facilitated if the model has clear connections with theory and is consistent with it. If it is not consistent it is useful to explain why it is not. Thus the theoretical bases of a good model are made explicit. This is particularly important if a model is to be used as a policy device. Policy makers would be more inclined to believe results if they could comprehend how these results come about.

-- *Logical Structure:* This is a variation of the issue discussed above. A logical structure aids in comprehension.

-- *Flexibility:* Models are usually made in relation to one city and then attempts are made to apply them to other cities. On the one hand, the more flexible a model the better it is. On the other hand, flexibility implies that the model is not that closely related to the original city and therefore breeds some suspicion. If a model is claimed to be flexible in this respect justifications should be given about its flexibility. Such justifications depend on the kind of relationships embedded in the model. Is the flexibility due to easily changed parameters which differ from city to city or because the relationships themselves can easily be changed. If a model is lifted from one country to another or from one culture to another, explanations should be given as to the feasibility of such a jump.

Another kind of flexibility is in the changeability of the parts of a model. Where a model is achieving a number of objectives and has a number of submodels it is useful to know how tightly interrelated these parts are. Are they modules that can be lifted or whose sequence can be changed around to fit different circumstances and objectives? In this sense, the more flexible a model is the better it is as long as the flexibility is explicable.

Role of Model Structure in Decision Making:

The structure of the model affects the outputs. Where the output is used for decision making it should be possible to decide whether certain results are due to technical quirks of the model or are believed to be inherent in the city as modelled. In other words, clarity about the model's role in decision making is desirable.

(v) *Output:* The policy maker receives the output from the model. It should, above all, be intelligible. What it represents should be clear. It should also be reasonable in the sense of correspondence to reality. If a projection is a wild one in that it is unexpected, then it must be explained. Calibration of the model is used to make its outputs reasonable. Since all outputs are really point estimates of some distribution, some indication should be given about their level of accuracy. They should therefore be accompanied with some tolerance measures. Probability estimates are useful when extrapolation is involved.

Policy makers are also interested in the robustness of outputs. They want to know how sensitive they are to policy changes. Thus a good model should be capable of providing sensitivity tests. In addition, predictive outputs should be clearly distinguished from prescriptive ones.

(vi) *Cost of Operation:* Costs are of different kinds. Firstly, there is the cost of developing the model and putting it into operation. Secondly, there are time costs, i.e., one likes to know how long it would take to put a model into operation. Thirdly, there are costs of skill. Clearly, one would like all these costs to be minimized. Since skills are in particularly short supply in LDCs this is a crucial variable for consideration. Furthermore, since it requires intimate knowledge of a city to build a good model, it is desirable to find models which can be built with available skill levels. Finally, one is interested in the costs of running the model - usually the cost of running the model on a computer.

In the reviews of models that follow comments on the quality of models are based on the above criteria. To the extent that they are not always made explicit is a shortcoming of this evaluation. It would therefore be useful to refer back to this list of criteria while reading the evaluations.

In concluding this section it is instructive to quote from Garry Brewer (1973) when models should not be used at all:

When - simpler techniques exist
- data are inadequate
- objectives are not clear
- short term deadlines exist
- problems are minor.

If any of these conditions exist the question of evaluation of models does not arise, since they should not be used at all.

II. ANALYTICAL OR EXPLANATORY MODELS

2.1 *The Classical Economic Models*

Any review of urban economic models must begin with the seminal contributions of Muth (1969), Wingo (1961), Alonso (1964) and Mills (1967). They are of the same family: utility maximizing households constrained by their budgets trying to find optimum residential locations. The city is located in a featureless plain and possesses a single Central Business District ,(CBD) where all employment is located.

Muth's households have a utility function

$$U = U(x, q) \quad (1)$$

where x represents consumption of goods other than housing and q is consumption of housing.

Their budget constraint is

$$M = x + p(u)q + T(u, M) \quad (2)$$
$$T_u > 0$$

where M is household income;
u is distance from the CBD;
p(u) is price per unit of housing, a function of distance from CBD;
T is the cost per trip and is a function of location and income.
(and subscripts denote partial derivatives, i.e., T_u is $\delta T/\delta u$).

Income includes money value of travel and leisure time. Prices of goods other than housing and transportation are the same everywhere in the city. Housing is regarded as a bundle of services yielded both by structures and the land they are built on. These services are a flow not an asset and thus the price is also of the flow, not of the asset. No distinction is made between owners and renters since they are both seen to be consuming a bundle of services. In the basic model, Muth assumes that households make a fixed number of trips to the CBD, (i.e., it is not a decision variable) and these costs are composed of money costs which vary with distance ($T_u > 0$) from the CBD and time costs which are assumed to vary with income ($T_M > 0$).

Maximizing (1) constrained by (2) yields the standard first order conditions:

$$\frac{\delta L}{\delta x} = U_x - \lambda = 0 \tag{3a}$$

$$\frac{\delta L}{\delta q} = U_q - \lambda p = 0 \tag{3b}$$

$$\frac{\delta L}{\delta u} = -\lambda(qp_u + T_u = 0 \tag{3c}$$

$$\frac{\delta L}{\delta \lambda} = M - \{x + p(u)q + T(u + M)\} \tag{3d}$$

The marginal utilities are in the proportion of the prices (equations 3a, 3b)

$$\frac{U_x}{U_q} = \frac{1}{p(u)} \tag{4}$$

and from (3c)

$$-qP_u = T_u \vee p_u = -\frac{1}{q}.T_u$$

$$(T_u > 0)$$

(5) shows that, in equilibrium, the result of a small move will not result in any savings. T_u is the marginal change in transport costs and qp_u is the change in housing expenditure occasioned by such a move. Thus a move outwards, (i.e., $\delta u > 0$) will increase transport costs by T_u which will be exactly balanced by a saving qp_u in housing expenditure. A small move inwards balances the saving in transport costs by an equivalent increase in housing expenditure $T_u > 0$ hence $p_u < 0$, i.e., housing costs per unit decline with distance.

Muth then investigates the effect of small changes in each variable and obtains the following results by totally differentiating (5) with respect to u

$$-qp_{uu} - p_u \frac{dq}{du} - T_{uu} \leq 0 \tag{6}$$

(i) $P_{uu} > 0$

The price of housing decreases at a numerically decreasing rate.

ii) $\frac{\delta u}{\delta M} \geq 0$

Optimum location is more distant from the CBD the higher the income. This is not an unambiguous result but depends on various assumptions concerning the income elasticity of the demand for housing and the elasticity of T_u with respect to income assumptions regarded as plausible by Muth. Stated more simply, this result states that the benefits of increased housing consumption outweigh the increase in transportation costs.

This is a very important implication for city structure and change in structure over time. If Muth's assumptions are correct, a general increase in incomes leads to an increase in housing consumption on the part of all households and the city spreads out.

Where transport is slow as it is in developing countries the time costs for higher income people may well be high enough to contravene this assumption. The location of high income people nearer CBDs rather than in suburbs could then be regarded as an optimal location within that framework. This is clearly something that can be tested econometrically if one had good data on housing expenditures and trip times for different income groups by location for some developing country cities.

iii) $\frac{\delta u}{\delta p_o} \leq 0$

where p_o is the price of housing services if they were located at the centre. Optimum location will be nearer the CBD if the price of housing services increases (keeping the price-distance gradient unchanged).

iv) $\frac{\delta u}{\delta T_o} \leq 0$ and $\frac{\delta u}{\delta T_u} \leq 0$

an increase in either the fixed or marginal cost of transport decreases the optimum distance from the CBD.

Muth assumes the income elasticity of demand for housing to be greater than 1 in the derivation of all these results. He supports this assumption by his own empirical work.

Muth then modifies his basic model by relaxing some initial assumptions:

a. The number of CBD trips is made a decision variable by introducing it in the utility function.

b. Similarly, preferences for location are introduced in the utility function, and

c. Uniformly distributed local employment is introduced in addition to the CBD employment.

"a" does not affect the results of the model in any substantive sense. "b" makes the derivation of qualitative results. almost impossible. "c" makes possible the derivation of a wage gradient with distance from the CBD. This obviously appears because the local workers incur no transport costs and can therefore accept lower wages to remain at the same utility level at the same location.

Muth then has a production side to his model to describe the behaviour of profit maximizing firms producing housing services:

$$\Pi = pq(L,NL) - rL - \rho(NL) \qquad (7)$$

where π is profits;
L is quantity of land inputs;
NL is quantity of non-land inputs;
r and ρ are their respective prices; and
q (L, NL) is the production function of housing services.

All the firms are identical and have the same production functions. Thus they all have the same profits irrespective of location. The prices of the inputs vary with locations so differing combinations are used according to location. Equilibrium conditions yield

$$dr^* = \frac{1}{S_L} dp^* - \frac{S_{NL}}{S_L} d\rho^* \qquad (8)$$

(where * indicates natural logarithm of the number);

S_L and S_{NL} are the shares of land and non-land in the firms revenue,

i.e., $S_L = \frac{rL}{pq}$ and $S_{NL} = \rho\frac{(NL)}{pq}$

(8) shows that land price is high where the price of housing, (e.g., due to location) is high and where the price of other inputs is low, (e.g., of raw materials).

We can rewrite (8)as

$$\frac{ru}{r} = \frac{1}{S_L}\frac{(Pu)}{p} - \frac{S_{NL}}{S_L}\frac{(P_u)}{\rho} \qquad (9)$$

This shows that the land rent gradient $\frac{(ru)}{r}$ is a multiple of the housing price gradient since $S_L < 1$, assuming the gradient of other input prices, (e.g., wages) to be negligible. For land share of 5 to 20 percent the land rent gradient can be anything from 5 to 20 times the housing price gradient. By assuming a Cobb-Douglas production function for housing, Muth then derives the housing price, land rental and population density functions as declining exponentially with distance from the CBD.

This model has been presented in some detail since it is an example of simple economic reasoning stretched to its limit. The assumptions underlying the model are highly unrealistic but that is the price of a simple manageable model. Muth does extensive empirical testing of his propositions: indeed his work is an example of the close relationship of good economic theorizing with empirical work. Many qualitative results are not possible to obtain without the assumption of robust parameter estimates - the income elasticity of the demand for housing being one example.

Wingo's concern is with the cost and programming of transportation in a city although this is embedded in a larger model of land use and transportation. Indeed, he set out to develop a transportation model but soon discovered that land use and transportation were too interrelated to be separated. Wingo was much more guided by policy considerations than Muth and tried to articulate a model which could be of operational use for policy planners. However, it is more of interest as an analytical model. His approach was guided by three considerations:

(a) A model should have explicit differentiation between policy and structural effects.

(b) It should enhance its analytical value by bringing the main elements of the problems within the framework of economic theory.

(c) It should have conditions for treating the problem of intra-urban distribution of population as a part of the urban economy as a whole.

His model concentrates on how labor services are organized in space given the characteristics of the transportation system, the spatial arrangements of production, the nature of the labor force and the institutions by which the labor force is articulated with the processes of production. The model provides the following:

1. The concept of transportation demand based on characteristics of the labor force and of the journey to work.

2. A systematic general description of the transportation function.

3. A general transportation cost function.

4. A system of location rents which result from the transportation cost function.

5. The manner in which a household demands space and how supply is equated to demand.

Here we will describe the derivation of the transportation cost function in some detail and neglect the rest of the model. This is because that is really the core of tie model. Moreover, the derivation of his cost function is of great interest since it is built up from fairly simple notions but takes account of many of the complexities of transportation. Wingo treated the problems of congestion at that early stage: others rediscovered it almost a decade later. The development of this cost function is a lesson on how a particular part of an urban model requires care and thought as well as empirical observation for a proper specification. Wingo assumes that the journey to work is the most important transportation function and then proceeds to analyze its cost.

He observes that since the journey to work is both spatially as well as temporally concentrated these peaks in demand result in saturation in a given capacity. Other demands on the transportation system can therefore be excluded and the journey to work analyzed. The costs can be broken into two components: the time costs incurred and the actual transportation costs. Both involve a prior characterization of technology of the urban transportation system. A key assumption made here is that the homogeneity of travellers as well as of carriers (in his case: the automobile). This clearly simplifies the analysis: indeed makes it manageable. The calculation of time spent in the journey to work is done by a function

$$T = T(u, v, n, c) \qquad (10)$$

where u is distance travelled;
v is velocity;
n n is number of workers; and
c c is a measure of capacity of the transportation system.

The specification of the function depends on the mode of transport used. Wingo's contribution here is in the inclusion of n and c in this function to demonstrate the interdependence between the users of a transportation system when demand (represented by n) exceeds capacity (c). The time lost because of this excess demand is because of:

a. **ingression:** the irreducible minimum of time loss because of aggregation of demand - this is a technological function; (analogous to changes in water pressure in a pipe according to the volume moving through it).

b. **congestion:** which arises because of reduction of free flow, e.g., because of bad driving. This is not necessary but arises because of human errors which are proportional to pressure of traffic.

If we define t_n as the time loss to the n_{th} unit entering the transportation system at peak time,

$$t_n = \frac{n-1}{c} \qquad (11)$$

and

$$t = \frac{n(n-1)}{2c} \qquad (12)$$

is time lost to all units in queue. This is proportional to the square of demand. Capacity of a system is also determined technologically.

$$C = C(v, v^{\beta}, \xi)^{1} \qquad (13)$$

where ξ is length of the carrier. The exact specification of the function clearly depends on mode of transport used, condition of roads, etc.

Having derived the time spent in journey to work, its value is derived from the marginal value of leisure function: the supply function of labor. Although a worker is paid according to time spent at work Wingo argues that the wage rate subsumes the time costs of the journey to work. In Figure 1, OP is the time spent at work, PQ is the journey to work. The worker needs to be compensated for OQ hours at hourly wage QJ. However, he gets paid for OP hours only, hence his wage rate may be PH, i.e., W_2W_1 is the "pure" wage rate. $W_1KH\ W_2$ gives a measure of the value of the time spent in the journey to work.

The actual transportation costs, i.e., money costs, have two components:

a. Those that vary according to distance travelled; and

b. Those that vary according to number of trips made.

The sum of these along with the time costs finally gives us the total costs of transportation for an individual.

We can now describe the rest of Wingo's model for the sake of completeness. Since all workers are paid the same wages, the differences in transport costs incurred by each account for the differences in land rental. In other words, land near the center of the city commands a higher rent, the difference being equivalent to the excess transportation costs incurred in living farther out. A household's demand for land depends on the rental value of land and the elasticity of demand is constant. The supply of land is proportional to the distance from the city center and is given exogenously. The model is closed by balancing the supply and demand of residential land. Wingo's condition for locational equilibrium is to is that the saving in transport cost equal the increase in outlay on residential land. This is not strictly correct: the condition should really be that no one can increase his utility by moving. The two are equivalent if the time costs of commuting include the disutility of commuting - which Wingo does appear to account for.

1 Wingo's expression is $C = \dfrac{v}{\rho(\alpha' V^{\beta} + r'v) + \xi}$

where ρ is a risk coefficient and β is approximately 2-2.5 for automobiles.

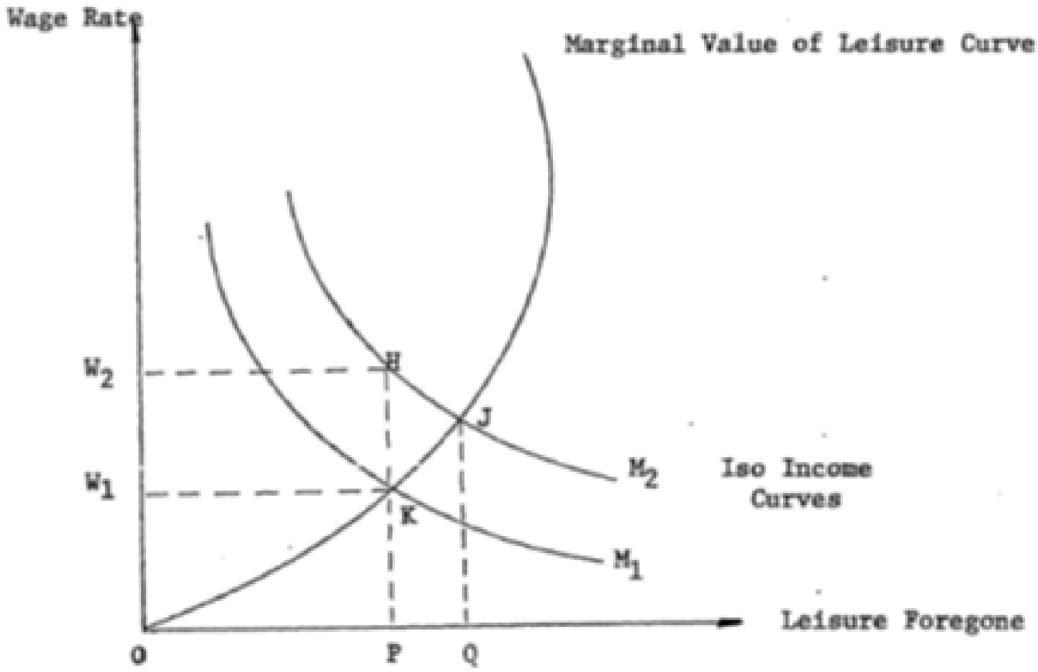


Figure 1: VALUING TIME SPENT IN JOURNEY TO WORK

The importance of Wingo's work lies in the demonstration of the complexity involved in specifying just one component of an urban economic model. The valuation of transport costs involves knowledge of the technological relations of the particular transportation system and of the workings of the labor market. While the assumptions about homogeneity of carriers and travellers are justified at this level of abstraction for developed countries, they are not for less developed countries. The modes of transport are much more mixed - from walking to bicycles to electric trains -- as are the workings of the informal labor market. Indeed, Wingo himself makes the point that valuing time with money implies fungibility and this is only valid for well operating markets. If time and money are not exchangeable a person may behave as though the scarcity of his time or his money were governing his behaviour - in either case the other is ineffectual in allocation. Constraints are probably distributed among the population in accordance with income levels. Money is more likely to be the binding constraint for low income groups, time for upper income groups. Such considerations can go a long way in the explanation of location patterns in developing country cities. Modelling this can be done with a Winge-type approach modified as suggested above.

Alonso (1964) provides a rather complete and general model of urban location and urban land markets. He starts with utility maximizing house-holds constrained by their budgets:

$$U = U\,(x, q, u) \tag{14}$$

This is similar to Muth's formulation except that q is quantity of land rather than housing and u - the distance from CBD - is introduced explicitly in the utility function with $U_u < 0$. Alonso derives a bid price function for each household from the equilibrium conditions. Each household has a bid price curve for a given level of utility. The result of adding u to the utility function is that one of the equilibrium conditions becomes:

$$-\frac{\delta r}{\delta u} = \frac{1}{q}\left(\frac{\delta T}{\delta u} - \frac{1}{\lambda}\frac{\delta U}{\delta u}\right) \tag{15}$$

where λ is the Lagrange multiplier denoting the marginal value of money. This condition says that residential rent compensates for different travel times. Since $T_u > 0$ and $U_u < 0$ the R.H.S. of (15) is positive and the rent (of land) distance function is negatively sloping. The reason Alonso does not go the route of demand curves to analyze the land market is because each location has a different demand curve. The uniqueness of locations makes the derivation of an aggregate demand curve invalid. This problem is circumvented by the use of bid-price curves. A bid-price curve represents the prices a household is willing to pay for land in each location in order to maintain a constant level of utility. A bid-price curve is therefore derived by fixing utility and then varying distance to obtain a function.

$$b = b(u) \tag{16}$$

People with the steeper curves locate nearer the center. The market also yields a price-structure curve showing the market price (rent) of land at each location, r(u). Tangency between these determines a household's location, i.e.,

$$\begin{gathered} p(u^*) = b(u^*) \\ \text{and} \\ p'\,(u^*) = b'(u^*) \end{gathered} \tag{17}$$

for equilibrium.

On the production side, firms are profit-maximizing. Profit is defined by revenue minus the sum of land and non-land production costs. Alonso uses curious revenue and costs functions:

$$\begin{aligned}&\text{Revenue (Volume of Business)} = R(u, q)\\&\text{Operating (Non-land) costs} = NL\ (R,u,q)\\&\text{Land costs} = p(u)q\end{aligned} \tag{18}$$

The firm's bid price function is derived for each level of profits; i.e., the rent a firm is willing to pay for each location in order to make the same profits. Their location is then determined by the equilibrium tangency condition as mentioned above for households. At each location, of course, profit is maximized.

Market equilibrium is achieved when each user's land bid-price is tangent to the price structure. The price structure is the envelope of all bid-price functions. Finally supply of and demand for land should be equal. Alonso analyzes the case where each user's bid-price function is a family of parallel straight lines. He concludes that users will be ranked from the city center according to the ranking of the slope of their bid price lines - steeper ones being nearer.

Alonso's approach is largely diagrammatic though he does give some mathematical analysis. A rigorous mathematical formulation of his model would be quite complex since he allows for different tastes among households; indeed that is what produces different bid-price functions. Mills (1972) has shown that Alonso's assumptions are not adequate to produce a solution to his model. The specification of an equilibrium utility level is necessary for a solution. Particular specification of the form of the utility functions is also necessary to derive stronger implications from the model.

Alonso's work is essentially an extension of Von Thunen's (1826) theory of the values of agricultural land. He has adapted it to an urban area but runs into difficulties precisely because urban land has no intrinsic productivity differences as agricultural land does because of fertility. He thus loses one determining variable and ends up with an n-person, n-firm game. A solution to such a game needs assumptions concerning strategy, permitted coalitions, etc.

Although Alonso's model is not entirely satisfactory it is useful because it once again illustrates the difficulty of modelling urban areas even at a highly simplified level. The difficulties arise from the intrinsic nature of urban areas - the uniqueness of each location which is created by a complex set of interdependencies.

Drawing on the work of Wingo, Alonso and Muth, Mills attempts to build simple general equilibrium models for urban structure. He has a family of models, all with similar bases but each with a different wrinkle. Here we will review the most complex and earliest of his models (Mills 1967) and then comment on the others.

He begins by speculating about the primary reason for the existence of cities and posits that non-homogeneity of land and non-constant returns to scale in production functions are sufficient to justify the existence of cities. If land is heterogeneous and some land is more productive than other land it will pay to concentrate production on the better land, thus producing a city. This can be represented in models in two ways . One is to introduce variables such as natural resources, topography and climate into the production function and the other is to have just one land input but to associate different efficiency parameters with different sites. Mills chooses the latter for the purposes of this model. Agglomeration economies of different kinds are all broadly interpreted to be scale economies and represented as such in an aggregative model.

The city is assumed to be a homogeneous plain and has 3 activities. The first is the production of goods. The goods production function has non-constant returns to scale. All goods production takes place in the CBD. This represents site advantages of the CBD.

$$X_{1s} = A_1 L_1^{\alpha 1} N_1^{\beta 1} K_1^{\gamma 1}, \alpha_1 + \beta_1 + \gamma_1 = H_1 \gtrless 1 \qquad (19)$$

where X_{1s} is total output of goods produced; and subscript s denotes supply.
L_1, N_1, K_1 are the land, labor and capital inputs and
H_1 <> represents non-constant returns.

$$X_1 = \int_{CBD} X_1(u)du \qquad (20)$$

where X_1 (u) refers to the amount of goods produced in a ring of width du, u miles from the center.

The other two activities are transportation and the production of housing. Transportation has only one input - land - with a fixed coefficient.

$$L_2 (u) = b \; X_{2s} (u) \qquad (21)$$

where X_{2s} is transportation produced.
$L_2(u)$ is land used at distance u.

'Housing' subsumes all goods other than those produced in the CBD. The assumption is that all goods with non-constant returns to scale will be produced in the CBD while the others will be forced by competition to locate adjacent to customers in order to avoid transportation costs. The production function for 'housing' defined thus is

$$X_{3s}(u) = A_3 L_3(u)^{\alpha 3} N_3(u)^{\beta 3} K_3(u)^{\gamma 3}$$

$$\alpha_3 + \beta_3 + \gamma_3 = 1 \qquad (22)$$

On the demand side X_1 is thought of as an export good-with an exogenously given price elasticity λ_1, i.e.,

$$X_{1D} = ap_1^{-\lambda_1}$$

A fixed proportion δ of the workers resident at each u work adjacent to their residences in the suburbs -- presumably in housing and transportation. The demand for transportation is then

$$X_{2D}(u) = (1-\delta)\int_{k_o}^{k1} N(u')du',$$

$$k_o \le u \le k_1 \qquad (24)$$

where k_1 is the radius of the city;
k_o is the radius of the CBD;
$N(u')\,du'$ is the number of people living in a ring of width du' and radius uradius u'.

Within the CBD,

$$X_{2D}(u)\int_0^{k_0} N_1(u')du'$$

$$0 \le u \le k_0 \qquad (25)$$

where $N_1(u')\,du'$ is the number of workers working in a ring of width du' and radius u'.

The demand for housing is constant per worker

$$X_{3D}(u) = N(U)x_3 \qquad (26)$$

Market Conditions

All factor markets are competitive and w the wage rate and p the rental rate for capital are given exogenously. Rental rate for land r(u) is exogenous.

In industry 1 we have

$$w = \frac{\delta(P_1X_1)}{\delta N_1}, p = \frac{\delta(P_1X_1)}{\delta K_1} \quad \text{and}$$

$$r(k_o) = \frac{\delta(P_1X_1)}{\delta L_1} \qquad (27)$$

according to normal marginal productivity conditions. The competition for land is between the CBD industry and suburban uses; thus land use is determined by the rent at the edge of the CBD k_o from the center.

Land being the only transportation input,

$$P_2(u) = a_1r(u) \qquad (28)$$

i.e., the cost per passenger mile depends only on the rent r(u) at that u, a_1 being a constant.

Housing is produced with competitive input as well as output markets. We have

$$w = \beta_3 \frac{p_3(u)x_3(u)}{N_3(n)}, p = \gamma_3 \frac{p_3(u)x_3(u)}{K_3(u)},$$

$$r(u) = \alpha_3 \frac{p_3(u)x_3(u)}{L_3(u)} \tag{29}$$

and $p_3(u) = \overline{A}_3 r(u)^{\alpha 3}$

where $\overline{A} = \left\{A_3 \alpha_3^{\alpha_3} \beta_3^{\beta_3} \gamma_3^{\gamma_3}\right\}^{-1} w^3 p \gamma_3$ (30)

Other Conditions

The main equilibrium condition is that a worker at u cannot decrease his location costs by moving toward the city center. The decrease in transportation costs would be exactly balanced by an increase in housing costs:

$$P_2\,(u) + p3'\,(u)X_3 = 0 \tag{31}$$

where

$$p_3'(u) = \frac{dp_3(u)}{du}$$

This is a crucial condition and should really be derived from some maximization conditions. It has embedded in it notions concerning the disutility of transportation and relative prices of housing and transportation.

The rent at the edge of the city is given exogenously and we can assume it to be agricultural rent, i.e.,

$$r(k_1) = r_A \tag{32}$$

In equilibrium all land must be used up:

$$L_1(u) + L_2(u) = 2\pi u \qquad 0 \le u \le k_o \tag{33}$$

in the CBD; and

$$L_2(u) + L_3(u) = 2\pi u \qquad k_0 \le u \le k_1 \tag{34}$$

in the suburbs.

Finally,

$$N_1 \equiv \int_0^{k_0} N_1(u)du$$

$$= (1-\delta)\int_{k_0}^{k_1} N(u)du \tag{35}$$

which merely makes sure that all workers live somewhere.

Solution

Even though this model is based on highly simplified notions of the structure of the city it does not have a straightforward solution.

The solution should provide us with:

a. All output quantities and prices.
b. All the input quantities and prices.
c. Parameters k_o and k_1 for the size of the city.

It is useful to note that we have a large amount of information given exogenously:

a. All the parameters in production
b. functions.
c. Demand function parameters for

d. goods.
e. Rental rates of labor and capital.
f. Fraction of labor force employed in the suburbs.
   Demand per worker for housing.
   Value of agricultural land.

Inspection of the model shows that the rent-distance function r(u) for land is the critical function to be determined; from which many of the other functions can then be derived. Mills provides some interesting insights from the model but does not solve the whole model.

Within the CBD,

$$L_1(u) = \frac{2\pi}{\lambda r(k_0)}\left(1 - e^{-\lambda r(k_0)u}\right) \tag{36}$$

i.e., the amount of land used in production increases at a decreasing rate as one goes out from the city center even though the amount of total land grows with u.

As a consequence,

$$L_2(u) = 2\,\pi\, u - L_1(u) \tag{37}$$

hence the land required for transportation increases at an increasing rate up to the edge of the CBD. Mills notes that for a very large city $r(k_o)$ and k_o are both large and in the limit

$$L_1(u) \to 0$$
$$\text{and} \quad L_2(u) \to 2\pi u$$

i.e., all land at the edge of the C B D is required for transportation. We can visualize this result as the requirement for a ring road around the CBD.

The implications of the growth of the city can be found by varying k_o and analyzing the results. Optimal reallocation of land will be provided - given the assumptions of the model.

For the suburbs, i.e., the city outside the CBD Mills derives

$$r(u) = (A_0 + B_0 u)^{-\frac{1}{1-\alpha_3}} \tag{38}$$

where $A_o = A_o(r_A, \alpha_3, X_3, k_1)$
$B_o = B_o(\alpha_3, X_3)$

which shows that rent declines as one moves towards the edge of the city, but not exponentially. An exponential decline results only when $\alpha 3 = 1$, i.e., that land is the only input in housing. The implication is that factor substitution makes the land rent profile flatter. We can observe here that this is more likely in earlier stages of development when land **is** the major input in housing construction.

Finally, Mills derives the population density function:

$$\frac{N(u)}{L_3(u)} = (C+Du)^{-1} \qquad (39)$$

where C and D are functions of r_A, α_3 and k_1, i.e., rent of agricultural land, size of the city and techniques of house construction. The density is thus declining with distance but **not** exponentially as is often argued, e.g., Clark (1951).

It is of interest to compare this model with some of Mills' own later work. The later modifications are simple in some ways but more refined in others. This model is curious in a number of ways.

a. The transportation production function has only land as an input.

b. The city is artificially divided into the CBD and suburbs. Workers in the suburbs are imagined to work on housing and transport even though transport has only land in the production function.

c. Housing has a constant per worker demand. This robs the model of part of the flexibility given by the possibility of utility optimization between consumption of space and transportation. In a later model (Mills, 1969) he simplifies it into 2 sectors -- goods and transportation - where goods are interpreted to include housing. The production function is made a constant returns to scale function. All the employment is not in the CBD and the center is only seen as a major transportation node through which all exports pass - this is seen as the justification for the existence of the city. Transportation is now produced by a Cobb-Douglas production function with constant returns to scale. The demand for transportation is now linked to the production of $X_1(u)$ with the assumption that each unit of X_1 generates a fixed demand for transportation. As a result the rent of land, wages and the rental rate of capital are all linked with the cost of transportation - which itself is a function of land rental. w and p are exogenously given. This model is more internally consistent than the other one and the economy is more integrated. The model is mainly used to derive the rent-distance function. It is of a form similar to the earlier model. Land rents decline exponentially if the two production functions (for goods and transportation) have equal exponents, i.e., use equal shares of land. We then have

$$r(u) = r_0 e^{-Au} \qquad (40)$$

where A is a function of all the other parameters of the two production functions. Everything else, e.g., land use intensity (capital/acre), land used for transportation and output/acre can be derived as a function of land rent. If, then, land rent is a negative exponential function so are all the land use functions.

Finally, Mills in **Studies of the Structure of the Urban Economy (1972d)** suggests a complex model including congestion. Here he ignores the CBD and concentrates on the suburbs. All employment is now in the CBD; consequently, both the housing and transportation functions do not use any labor. Transportation is again produced with land only. Housing is produced with a Cobb-Douglas production function using land and capital. The demand side is richer: housing demand per worker is made price and income elastic. This way one can investigate the effects of general income changes on the structure of a city. The cost of transportation is now affected by congestion. The congestion function is taken by Mills from earlier work of others, (e.g., Walters, 1961) and is really not very different from that of Wingo. It is

$$P_2(u) = \bar{P}_2 + C\left(\frac{X_{2D}(u)}{X_{2S}(u)}\right)^D \qquad (41)$$

where P_2 (u) is cost of transportation at u per mile (as before)
$\bar{P}_2$ is some constant (free of congestion) cost; and
C and D are parameters determined technologically by the transportation system.

Much of Wingo's work was interpreted to be the specification of these parameters. D for automobiles is believed to be about 2. The point to note is that congestion cost is seen to be a power function of excess demand. The equilibrium conditions of this model are essentially the same as in the earlier model (Mills, 1967).

Mills finds that the introduction of congestion and of elastic housing demand makes the model impossible to solve analytically. The rest of the book is devoted to a numerical solution and a demonstration of how sensitivity analysis can be performed on such a model. Some of the interesting results are:

a. Technical progress in transportation is seen to induce workers to use more transportation by moving farther out and expanding the city. Land rents fall in the city center.

b. An increase in income elasticity of housing demand increases size of total area and reduces population density - not surprisingly. However, the magnitude of the effect is surprising: a 9 to 10 percent increase in elasticity causes a 90 percent increase in city area. The disutility of travel is perhaps not taken into full account as suggested earlier for the 1967 model. An increase in income has somewhat similar effects.

c. An increase in D in equation 41 (the elasticity of congestion cost with respect to amount of congestion) has a somewhat paradoxical result. Travel cost increases near the CBD but decreases farther out. This is because this increase amounts to "decongestion" farther out and thus people move farther away. CBD rents rise while they decrease farther out, i.e., the rent distance function increases in curvature.

Much of this model is geared to the explanation of decentralization of U.S. urban areas over time as observed empirically. Increases in incomes and population and technical change in transportation are seen as the main causes.

In summary, Mills' work is interesting because it attempts to see the urban area as a whole in general equilibrium models. It is a further demonstration that even extremely simple models tend to be mathematically cumbersome. Each addition of complexity results in simplification somewhere else in the model. The more realistic the assumption, the more unlikely it is that an analytical solution is possible. However, an equally important demonstration is that even simple models provide us with insight into prevailing urban structures. He has also incorporated into his general equilibrium model some of the notions of the housing market from Muth, the land market from Alonso and transportation characteristics from Wingo.

These are the 'classical' urban economic models. They are 'classical' because they are pioneering attempts at modelling cities from the economist's vantage point. They are still the most influential in urban economic model building. Even many of the policy-oriented models derive much of their methodological base from these models, as will be evident in later sections. These models bear the same relationship to large economically oriented policy models as do the basic Keynesian macro-economic models to the large macro-econometric models.

The more recent and more theoretical models which have not been quite so influential are reviewed in the following section.

2.2 The "New" Urban Economics

This section reviews a spate of urban modelling work which has appeared in recent years and one which has already been dubbed as "the new urban economics" (Mills and Mackinnon, 1973). It is distinguished from earlier work by being more rigorously theoretical, with a higher disregard for reality. The attempt is to explore the extent of possible conclusions from simple formulations. While the work of Wingo, Muth, Alonso and Mills is also theoretical, it is rooted in extensive empirical work carried out by them. While little operational relevance can be derived from these new urban economic models it is of interest to review them because:

a. Some of them are the work of distinguished economists shifting from other fields.

b. They do offer some counter-intuitive results, e.g., that it is optimal to have an unequal distribution of utility even where all households have the same tastes and income.

c. They demonstrate again the intrinsic complexity of urban areas. Even highly simplified assumptions often lead to models that do not have analytical solutions.

d. The analytical innovation is in the use of control theory or the calculus of variations which is in many ways similar to the growth theory literature. Here, space is the crucial variable over which optimization of one kind or another is taking place - like time in growth theory.

Almost all the models are monocentric. The city is in a flat plain, travel is equally costly in all directions, and all travel is from home to work. These assumptions make it possible to use one dimensional analysis with distance from the CBD acting as the main spatial variable. The issues explored are the distribution of rent, residential density and space, consequences of travel congestion on city structure and rents, implications of individuals having different incomes usually on their location pattern, the ability of competitive markets to sustain optimum city structure. All these models are static and thus have no implications for urban growth. The reason this is probably technical. It is difficult enough to optimize over space; adding time would make the exercise impossible.

These models will be categorized and reviewed under three headings:

a. Distribution of Land Rent, Population Density and Income

Beckmann's (1969) article "On the Distribution of Urban Rent and Residential Density" can probably be regarded as the first in the "New Urban Economics." Beckmann attempted to derive rent, population density and distribution of income groups as functions of distance. Delson (1970) and Montesano (1972) have pointed out various errors in Beckmann's analysis. Here we review the altered results. The model assumes a Pareto income distribution.

$$N(m) = Am^{-a} \tag{42}$$

where m is household income;
N is number of households with incomes greater than m
A, a are positive constants (empirically a has been found to be in the region of 2).

The utility function is

$$U = C_0 \log q + C_1 \log u + \sum_{i=2}^{n} C_i \log Z_i \tag{43}$$

where q is amount of land occupied by a household;
u is distance from CBD; and
Z_2 to Z_n are all other goods.

The problem is to maximize (43) subject to the budget constraint

$$r(u)q + tu + \sum_{i=2}^{n} P_i Z_i = m \tag{44}$$

where t is daily transportation cost per unit distance and P_i is the price of the i^{th} consumption good.

The general problem is similar to that investigated by Alonso but with the crucial difference that everyone does not have the same income now although they do have the same utility function.

In this solution, Beckmann assumes the CBD to have a radius of 1. Delson's correction (1970) asserted that this was an unnecessary assumption. It really does not make much difference if regarded as a normalization procedure making the radius of the CBD the measure of distance. Montesano (1972) provides the correct and complete solution to the problem. Firstly, Beckmann asserts that all households with the same income will locate at the same distance. However, the first order conditions can be used to show only that all households at the same distance have the same income. We have

$$q = -\frac{tu}{\frac{r(u)}{C_0} + u\frac{dr}{du}} \tag{45}$$

r(u) is only a function of u and so q is also only a function of u. Hence a household located at u pays the same rent, occupies the same space and has the same transportation costs and must therefore have the same income. The converse, however, does not follow.

Secondly, Montesano asserts that we need the explicit assumption $\frac{dm}{du} > 0$, i.e., that income increases monotonically with distance, to solve the model. Montesano calls this an assumption and then proceeds to show at the end of the paper that second order conditions for maxmization of utility require that this be so. Our interpretation here is that this is indeed a result of the model **not** an assumption. The second order conditions depend on the form of the utility function; the result therefore probably depends on the particular form used.

Montesano shows that the assumption of t=0, i.e., money costs of transport equal to zero leads to multiple solutions. We can see this intuitively as Montesano does at the end of his article after having derived multiple solutions. If t=0 we can derive an expression for the utility function which is only dependent on income; households are then indifferent to location and it is then not surprising that this assumption leads to multiple solutions. The lesson here is that it is sometimes useful to look at second order conditions before tediously solving a model. To the extent that the model tells us anything about urban patterns we can observe that in LDCs where the money costs of transport, (i.e., walking to work) are indeed zero for the poor we may have greater difficulty in "optimum" urban design.

With the assumption of $t > 0$ Montesano does obtain a unique solution which shows:

a. That r(u) is convex but decreasing less than proportionally.
b. That q(u) is increasing, i.e., residential density declines with distance.
c. That y(u) is increasing.

All these are expressed in rather complicated functions with no simple interpretations. None of these results are, of course, surprising; but what is surprising is the complexity of the analysis given a straight - forward utility function, budget constraint and description of the income distribution. This does not augur well for the inclusion of income distribution in models of LDC cities.

b. Congestion and Transportation

The next group of models are concerned with the optimal allocation of the urban area to transportation; the costs of congestion are usually given particular attention. Solow (1972, 1973) uses the standard model, i.e., maximizes utility with consumption and housing space as arguments in a logarithmically additive utility function subject to a budget constraint. He derives the declining rent-distance function and concludes that "the rent profile must fall fast enough that those living farther from the center and spending on space a fixed fraction of income after transportation cost occupy more space than those living closer to the center." (Solow, 1972) Solow then introduces congestion in the following way. The aggregate width of the road network at distance u is

$$2\pi u(1-b(u))$$

where b(u) is space devoted to housing. Then annual cost of round-trip travel per person-mile at distance u is

$$C\left(\frac{N(u)}{2\pi u(1-b(u))}\right)^{D}$$

where N(u) is the number of people living beyond u. Total cost per person is then

$$t(u) = C(2\pi)^{-m}\int_{1}^{u}\left(\frac{N(u)}{u(1-b(u))}\right)^{D}du \qquad (46)$$

where radius of the CBD is normalized as 1. This merely states that congestion costs are proportional to traffic density. This formulation is no different from Mills (1972d) (Equation 41) which itself was taken from Walters (1961). The point to be made here is that this shows explicitly some of the interdependence characteristics of urban areas. N(u) - the number of people living beyond u are dependent on t(u) which itself depends on N(u). Solow solves for the unknown functions t(u), N(u) but has to make the following assumptions:

a. The same fraction of land area is devoted to housing at every distance; i.e., b(u) = b.
b. The typical person spends half of his total income on housing. Both are assumptions that deprive the model of its interesting components - in particular assumption a. Solow himself addressed this problem in a more simplified context of the long narrow city (Solow and Vickrey, 1971). The conclusion reached was that a higher proportion of land would be needed for transportation near the center of the city in the case where city size was limited.

Solow solves his model numerically and finds one interesting result. The introduction of congestion makes the rent profile more convex: the rent falls more sharply as one leaves the CBD and then less sharply near the city limits. It may be recalled that Mills' model, discused on page 48, found a similar result. Solow investigates this result more thoroughly in the later article (1973). Assumptions (a) and (b) are both relaxed and numerical solutions obtained for different parameter values. The main result is that adding roads near the CBD flattens the rent gradient most strikingly. Since congestion is greatest near the CBD adding roads there reduces congestion costs, hence transport costs and therefore the flattening of the rent gradient, the rent differences being transport cost differences. Solow finally does some cost-benefit calculations on the allocation of land to roads. In the absence of congestion tolls, market land values reflect differences in private transport costs, not total social costs. Land values do not fall as fast as they should and the market rent function lies every-where below the "correct" rent function. Land is therefore undervalued and if these values were used in benefit-cost calculations too much land would end up being used for roads.

While this second article does provide some interesting numerical results these models do not provide any new insights into urban structure. Their use, perhaps, lies in Solow's pedagogical style which is a good example of a rather gradual, step-by-step approach to model building. Solow himself suggests that the model could be improved by explicit inclusion of housing in addition to land as a residential cost; the addition of time costs in transportation costs; and the existence of two or more income classes. To this we can add the inclusion of production functions for transportation and housing, i.e., a richer specification of the supply side.

More recent analyses of the congestion cost problem have emanated from the Berkeley group which is heavily influenced by the control engineering approach. Optimizing is now done from the social point of view: total costs of some kind are minimized. Since these are in the form of a city-wide integral, the problem is fairly straightforward one in the calculus of variations. The solutions are not straightforward and economic implications are often difficult.

Mills and de Ferranti (1971) can really be said to have posed this problem first. Their concern was to find the optimum allocation of land to transportation in the suburbs in the presence of congestion. We use the more general formulation of Livesey (1973) to illustrate this class of models. The usual circular city assumptions are made with N people being given as working in the CBD. The model itself is rather simple:

$$L_1(u) + L_2(u) = \theta u \qquad (47)$$

where $L_2(u)$ denotes land used for transportation at radius u;

$L_1(u)$ is land used for residence in the suburbs and for business in the CBD; and

θu is land available.

The density of workers in business is constant at a_c and residential density in the suburbs, a_s is also constant. Thus the number of people working at radius u

$$N_c(u) = a_cL_1(u) = a_c(\theta u - L_2(u)) \qquad (48)$$

and the number residing at u are

$$N_s(u) = a_sL_1(u) = a_s(\theta u - L_2(u)) \qquad (49)$$

congestion costs are

$$t(u) = \bar{t} + C\left(\frac{T(u)}{L_2(u)}\right)^D \qquad (50)$$

where $\bar{t}$ is some constant cost here assumed zero and T(u) is number of travellers at u. This is now a familiar formulation.

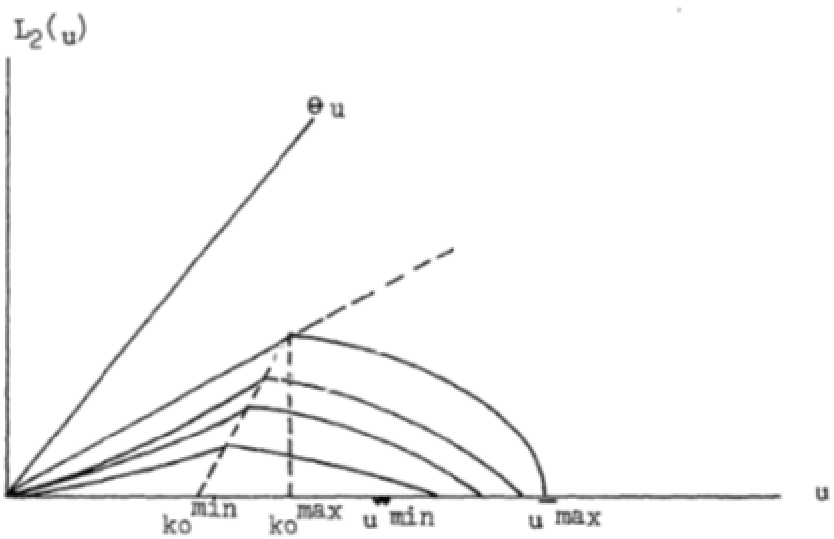


Figure 2: OPTIMAL ALLOCATION OF LAND FOR TRANSPORTATION IN BOTH THE CBD AND THE SUBURBS FOR A FIXED WORKING POPULATION
(Livesey Model)
(ko -radius of CBD)

Land value is taken as a constant R_a which is opportunity cost of agriculture use regarded as the relevant alternative use.

Hence total social cost is

$$\int \{C\ T(u)C(u)^D + R_a \theta u\}\, du \tag{51}$$

when congestion cost $C(u) = \frac{T(u)}{L_2(u)}$

This is the integral to be minimized subject to the given constraints. The problem is first tackled separately for the CBD and suburbs and then in a unified way for both parts of the city. The form of the solution is best seen in a diagram (Figure 2) for the analytical expressions are quite cumbersome and uninformative. Figure 2 shows that the optimal allocation of land for transportation as we move out from the center to the edge of the CBD is a monotonically increasing concave function and then a monotonically decreasing convex function until the boundary of the city is reached. The maximum is at the edge of the CBD.

This is really quite an uninteresting model for it has very little economic and behavioural content. The opportunity cost of land being taken as constant deprives the model of any pretense as a serious (though abstract) model of an urban area. It would be consistent if all business activity were equally spread over the city and all employment were local. But there would then be no transportation either. The assumptions of constant business and residential densities are equally restrictive. The model would clearly be made too complex, perhaps unmanageable if more reasonable assumptions were used. As it is it should only be called a problem in the calculus of variations and not suggested as shedding any light on the structure of a city.

A subsequent model from the Berkeley group by Legey, Ripper and Varaiya (1973) extends the analysis by allowing for substitution between land and capital. Housing and transportation are produced by Cobb-Douglas production functions using land and capital. Demand for housing and transportation is perfectly inelastic, i.e., everyone demands the same amount of each commodity. The total social cost now includes capital costs. The interest rate on capital is taken as given. Land value is again taken to be constant reflecting the alternative agricultural value. The sum of transport costs (which have the usual formulation (Equation 50), capital costs and land costs is to be minimized for optimality. The solution gives the magnitudes of land and capital devoted to housing and transportation and the optimum size of city given the population. Two solutions are obtained: the optimal solution corresponding to what a central planning board would do and a market solution. The central result is that a market city would be more spread out than an optimal city. If, however, congestion tolls are charged, the market city could be the same as the optimal one.

This model is of somewhat greater interest even though the demand side is primitive. The criticism of using constant agricultural value holds again.

Another model inspired by Solow-Vickrey's "Long Narrow City" is Marvin Kraus' "Circular City" (1973). It is concerned with questions of optimal land use in an urban environment with particular emphasis on transportation rights of way and the pricing procedures necessary to utilize them efficiently. This model has no residential sector and business valuation of sites reflects only travel on the city's roads. Demand for trips between any two units of business area is inelastic. The only costs of roads are the value of the land they cover and tolls can be levied on all roads in a costless way. All intersections are signalized and these are costless too. The object is to minimize total transport costs in the city incurred per unit of time. The total area to be allocated to business is given but the city's radius is to be determined as is the distribution of business area with the circumference. Each trip uses a route which minimizes the price to the trip taken which is taken to include money as well as time costs. Optimization of the signalization and toll system minimizes the total cost since trip demands are inelastic.

Radial as well as circumferential traffic is allowed in this model. Each unit area generates a demand of g trips per hour which are uniformly distributed over all units of business area. All land within a central disc, a circle of radius u_o concentric with the city's circumference (radius u1) is allocated to a circumferential inner road.

Let y(u) be area of land devoted to radial roads and s(u) be the area devoted to business within the ring bounded by circles of radius u_o and u.

Now

$$f_y(u) = \frac{y'(u)}{2\pi u}, y'(u) \geq 0 \tag{52}$$

and

$$f_s(u) = \frac{s'(u)}{2\pi u}, s'(u) \geq 0 \tag{53}$$

These functions characterize the intensity with which land is allocated to alternative uses. We also have

$$y(u) + z'(u) \leq 0 \tag{54}$$

Traffic: Let $V_1(u)$ be volume of radial traffic through an arc of radius r and length dr; and $C_1(u)du$ be the capacity of such an arc.

Similarly V_2 (u) is the volume of circumferential traffic through a radial segment of infinitesimal length dr at radius r; and C_2 (u)du the capacity of such a segment.

The cost per trip mile in direction i, u miles from the center

$$AC_i(u) = f\left(\frac{V_i(u)}{C_i(u)}\right) \quad i = 1,2 \tag{55}$$

which is the familiar volume divided by capacity type function. Here,

$$\begin{aligned} f'(.) &> 0 \\ f''(.) &> 0 \\ \text{and } f(0) &> 0 \end{aligned} \tag{56}$$

i.e., f(.) is an increasing strictly convex function and is non-zero at zero density; V_0 is the hourly flow of circumferential traffing crossing any radial line segment of the inner ring road and ring road hourly travel costs are given by a function

$$2\pi G(V_0, U_0)$$

Thus total travel costs for the city per unit time are

$$2\pi \int_{u_0}^{u_1} \sum_{i=1}^{2} uv_i(u) AC_i(u) dr + 2\pi b(V_0, U_0) \tag{57}$$

This is the expression to be minimized.

Further specification of road capacity relates it to land and green time (when signals are green) allocation patterns. Further

$$P_i(u) = AC_i(u) + T_i(u) \qquad i = 1,2 \tag{58}$$

where $P_i(u)$ is the price of a trip mile and $T_i(u)$ is the toll per trip mile.

Before the problem can be solved rules for trip patterns are provided.

In the solution, analytical expressions are first obtained for

$$V_1(u),\ V_2(u) \text{ and } V_0$$

i.e., the traffic follows in each of the directions.

We can summarize the final results for optimality:

a. In the absence of an inner ring road, **any** configuration of trip prices inducing travel through the city center leads to explosive travel costs.

b. Not surprisingly, toll charges should be the difference between average and marginal costs.

c. On every circle, the marginal rates of substitution of land for variable trip costs be equal in the production of radial and circumferential travel. This ensures the optimal allocation of land between radial and circumferential roads.

d. Lastly, a similar condition holds for the allocation of green time to radial and circumferential roads. The marginal rates of substitution of the value of green time for variable trip costs should be equal for travel in both directions.

This model is not noted for its realism either. It is however, a noteworthy attempt to relax the general assumption of all travel being radial. The treatment of the two directions in travel can be extended to the modelling of different modes of travel. This would be particularly important for LDC cities where modes are, indeed, heterogeneous.

C. 'Optimal Towns'

The last group of models considered in this section are those concerned with deriving the conditions and consequences of "optimal towns". They are somewhat difficult to compare since optimality clearly depends on the welfare function used and the type and extent of constraints in the model. The welfare function reflects the moral or other preferences of the modeller while the constraints constitute his conception of the city.

Mirrlees (1972) is the originator of this group of models and his approach is somewhat different from the others. He poses the problem in almost the simplest form possible. The welfare function is the sum of all individual utilities. Individual utilities depend on consumption of goods, space of residence and location:

$$U = U(x,q,u) \tag{59}$$

where x is consumption of goods other than housing;
q is amount of space used in housing; and
u the distance from the center is the location variable.

We may recall that this is the same formulation as Alonso (1964). Further,

$$q = q(u)$$
$$\text{and} \quad x = x(u) \tag{60}$$

are assumed but are also required for optimality. The main conclusion of the paper is that optimal allocation requires that utility is a function of distance. We have the seemingly surprising result that with identical individuals welfare maximization requires unequal treatment except in some special cases. In fact, this turns out not to be so surprising, because u is included explicitly in the utility function. As Mirrlees shows, when u is not explicitly included, it is possible to achieve equal utility. The problem arises because identical individuals have to be placed in different locations. They can all have the same utility if some trade-offs are possible. If, for example, more space is traded off against higher transport costs equal utility becomes possible; but explicit preferences for location have to be dismissed. A richer specification of locational preferences which can be traded would also allow equal utilities. When location is dependent on u alone only a special case where the rent gradient and the utility function are such that the changing consumption of goods and housing are always peculiarly balanced against distance changes allows equal utilities.

This is an interesting theoretical exercise even though the nature of an urban area is in a very rudimentary form: distance from the center is the all important variable. Otherwise the problem is straightforward utility maximization (summed to form the social welfare function) with budget constraints. It is of interest because it shows forcefully that:

a. Cities imply inequality even if all individuals are identical **and** if the city is uniform.

b. With identical individuals this is possible to mitigate only if locational preferences are more complicated than distance from the city center. This implies that the city would then be less homogeneous.

c. As a corollary, we can also have equality of sorts if people have different utility functions; equality is then difficult to define operationally except by income.

Riley (1973) has a similar model but includes the number of leisure hours explicitly in the individual's utility function. Since these vary with transportation time, distance is now a "pure" location preference variable in the utility function. It does not include the disutility of commuting. Riley makes the social welfare function rather more egalitarian by making it logarithmically additive, i.e.,

$$W = \pi_{i=1}^{N} U^i \tag{61}$$

where W is total social welfare;
U^i is the utility of individual i; and
N is the population of the city.

Riley finds that utility increases exponentially with distance, i.e., optimality requires unequal treatment of identical individuals even with an egalitarian social welfare function. The reason here is not quite the same as in Mirrlees' model since distance is now a "pure" location variable in the utility function, depending on the parameters

$$\frac{\delta U}{\delta u} > 0$$

$$\text{or} \frac{\delta U}{\delta u} < 0$$

in this problem. Riley's explanation of his result is that the fact that an individual can live at only one location (and not at two) causes a non-convexity which is not present in the usual case of utility and welfare maximization. Here everyone does have the same marginal utility of income but since everyone cannot have an identical consumption bundle (nor can he have an identical total utility level) the degree of inequality is primarily governed by the elasticity of utility with respect to distance. Riley uses

$$U = x^{\alpha} q^{\beta} h_3^{\lambda} u^{\delta} \qquad \alpha, \beta, \lambda > 0 \qquad (62)$$

where h_3 is number of leisure hours and the other symbols are the usual ones. Inequality is an increasing function of δ. If this is positive, the degree of inequality decreases with β - the elasticity of utility with respect to residence area. If δ is negative the degree of utility increases with β. Congestion costs are not considered and transport costs are linearly proportional to u in this model.

Dixit (1973) and Oron, et.al., (1973) have the same concern as Mirrlees and Riley but have more developed models of the urban area. Both use substantively similar models, Dixit following Oron, et. al.

Dixit's main theme is that optimum city size is determined by the balance between economies of scale in production and diseconomies in transport, congestion being an important part of the latter. We can summarize the two models, as follows:

Goods are produced with increasing returns to scale

$$X = AN^{\alpha}L^{\beta} \qquad (63)$$

where $\alpha + \beta > 1, 0 \le \alpha, \beta \le 1$ and N is number of man-hours worked. (Oron, et. al., have only labor as input since they fix CBD size).

Housing and transportation are produced with land only: L_1 (u) and L_2 (u), respectively at distance u.

Dixit's major additions are:

a. He includes congestion in the conventional way, i.e., a power of traffic density plus some constant as in Equation (41) (Mills). Oron, et. al., assumed that these costs were directly proportional to traffic density.

b. The individual's utility function is

$$U(u) = x^{\sigma/1+\sigma} q^{1/1+\sigma} \tag{64}$$

where x=x(u) and q=q(u) follow our usual symbols. Oron, et. al., assume σ=1.

This form of the utility function means that expenditure on goods and housing occurs in the proportion σ:1 with a given income aud prices. Dixit's point, which is well taken, is that σ is at least 3.

Workers supply a fixed number of hours devoted to work and commuting. It would be more realistic to assume that work hours are fixed.

Dixit's social welfare function is

$$\int_{k_0}^{k_1} -U(u)^{-m} \{-n'(u)\} \tag{65}$$

where $m > 0$,
n(u)du is the number of people between u and u + du;
k_0 is the radius of the CBD, and
k_1 is the radius of the city.

m is a parameter which controls the level of inequality in the optimum allocation. A higher m means less inequality and in the limit, $m \to \infty$ means full equality. Oron, et. al., constrain their model to this case.

Dixit derives analytical functions for traffic density, residentital density and utility level, all in terms of consumption x(u). x(u) is then expressed in terms of u for which analytical expressions are obtained for the case of pure congestion costs in travel.

Residential density:

$$\text{const}(C_1 + C_2 u)^{-\left\{1 + \frac{a}{D(a-1)}\right\}} \tag{66}$$

where C_1, C_2 are constants;
D is the exponent in the congestion function and

$$a = \frac{D(1+m)\,(1+\sigma)}{(1+D)m}$$

This (66) is a negative exponential form if a=1 but this case has no straightforward interpretation. Solow regards (66) as the more general form. Dixit's innovation over Mills (1967) is in making the land proportion used in housing and transportation endogeneous and in using a social welfare function. According to Dixit, a more developed housing production function yields similar results.

The rent function

$$r(u) = \text{const}(C_1 + C_2u)^{\frac{-(1+D)a}{D(a-1)}} \qquad (67)$$

and

traffic density =

$$\text{const}(C_1 + C_2u)^{-\frac{a}{D(a-1)}} \qquad (68)$$

are again negative exponential if a=1.

Finally,

$$U(u) = \text{const}(C_1 + C_2u)^{-\frac{1}{m(a-1)}} \qquad (69)$$

which says unambigously that

$$\frac{dU}{du} > 0$$

i.e., households located farther away have higher utility. The degree of inequality, of course, depends on m. Only $m \to \infty$ implies equality; hence more usual values like m = 1 imply considerable inequality.

Dixit's model is important in several respects:

a. It is the most developed model of its kind combining the approaches of Mills and Mirrlees, i.e., general equilibrium and optimality.

b. More variables are made endogenous than most models.

c. Analytical results are obtained even with congestion.

d. The allocation of income between housing and other goods is realistic.

It would be useful to bring capital into a similar model to get a "more" complete general equilibrium analysis. The curious assumption of fixed leisure hours should be dropped for fixed work hours. It is difficult to see how the assumptions affect the results.

Dixit does some numerical calculations on his model. He shows how transport costs are crucial in determining optimum city size. Lower transport costs are instrumental in making possible larger cities which allow greater advantage from economies of scale.

2.3 Summary

This section has reviewed the main strands of the development of urban economic models over the past decade and a half. There was almost no work of this kind before this period.

It is worth noting that there is a surprising unity of concerns among the different models that have been reviewed. Almost all these models investigate optimal residential location - either from the point of view of the household itself or for maximization of social welfare as expressed in a welfare function. Particular attention is paid to the operation of the land market and the effects of congestion in transportation. Here, the salient features of these models are brought together to give a better idea of what has been gained from them.

Muth illustrates how relatively simple economics can be utilized to understand the structure of urban areas - in particular the housing market. He maximizes utility which has only housing and other things as arguments: location is not an explicit argument in the function. He finds that a consequence of market equilibrium is declining rent gradient with distance from city center. He also shows how capital/land substitution operates to make the housing price gradient much less (by an order of magnitude) steep than the land price gradient. Wingo concentrates on transportation and illustrates the complexities of urban modelling by deriving a plausible specification of a transportation cost function. He also obtains a decline in rent gradient as a consequence of transportation costs. Alonso turns the problem around and emphasizes the uniqueness of each urban location. This argument makes it invalid to derive aggregated demand functions for urban land. The consequence of this argument is that a 'pure' location variable must be included in the utility function. The implications of doing this are brought out in later work by Mirrlees, Dixit and others who show that this characteristic of urban land makes inequality inevitable if everyone has the same utility function. A corollary is that different utility functions would make equality possible.

Mills' contribution was to bring together many strands of work in a general equilibrium model of a city. His conclusion also is that the land rent profile is crucial to the allocation of activities within a city, that rent declines with distance from the city center and that factor substitution makes the land rent profile flatter. He also finds that congestion makes the land-rent distance function more convex - a finding corroborated in later work by Solow. Mills also demonstrates that even a highly simplified urban general equilibrium model becomes mathematically cumbersome very quickly. Beckmann extends earlier work by positing an income distribution function for households and concludes that income increases with distance as Muth had by comparative static analysis along with additional assumptions. Beckmann's work also illustrates how a relatively simple income distribution function makes the mathematics cumbersome. Solow rediscovers the rent-distance relation but his contribution is in a lucid exposition of how a step by step approach to urban model building can be followed. In a benefit cost framework he finds that if congestion is neglected land would be undervalued. This conclusion is taken further by Legey, Ripper and Varaiya, who conclude that a market city is more spread out, if congestion costs are not somehow internalized by the actors in the urban market. Kraus' circular city makes a significant attempt to allow other than radial travel and the resulting model illustrates the costs of adding such simple attributes of reality into a model.

Though these models are at a sufficiently theoretical plane they should be regarded as conceptual building blocks towards more operational models. Each of the concerns exhibited in these models: the land/transport trade-off; the uniqueness of location; the effect of transport congestion on city form; the consequences of egalitarian welfare functions, etc., is a real problem which has first to be dealt with at a general level before operationalization into policy models.

In concluding this discussion of theoretical urban economic models we can remark that higher levels of generality and general equilibrium type models do yield some insights into urban form, as distinguished from particularistic and/or partial equilibrium models. This is what we would expect from highly complex and inter-related phenomena.

III. OPERATIONAL OR POLICY-ORIENTED MODELS

This section reviews, cursorily, some of the work done on policy-oriented models over the past fifteen years. These models are rather more difficult to review than the explanatory models because of their sheer size. In a paper such as this it is not practicable, nor of benefit, to present all the technical details. Indeed, it is difficult to present them technically at all because of the complexity of their notation. Here our objective is to appreciate the essence of the methodology used in each of these models rather than achieving a detailed understanding of each. Thus technical details are provided wherever it is considered necessary for this objective and symbolic notation is used only when it facilitates exposition.

The models reviewed in this section are mostly large in the sense that the only practical way to operate them is on a computer. They are spatially disaggregated to a greater or lesser extent and allocate activities to geographic zones. They pertain to metropolitan areas and their concerns are intra-metropolitan. Regional models are not considered here.

Such models have had a checkered history. They came to the fore in the early 1960s in the U.S. as concern over the declining central cities mounted and various planning solutions were sought. It was thought that these models would help planners in their professional roles as advisors to public decision-makers, with emphasis on objective plan evaluation. It was also expected that they would have an educational role in developing better theory of urban spatial structure as well as in giving planners as well as decision-makers more systematic ideas of urban areas. These models have failed with respect to the first objective and been only partially successful with respect to the second. Great disillusionment had set in in the U.S. by about 1968, but the challenge was enthusiastically taken up across the Atlantic in Britain at about the same time. Meanwhile, the new urban modelling in the U.S. has been done more by economists than others.

Most urban models are focused on land utilization, the types of structures erected on the land, the prices of the land and structures, the types of households which occupy these structures, and the impact of changes in the transportation network on this system. Part of production and employment often called ’basic’ or ’export-based’ is regarded as exogenous. The number and types of households in the city and their living place locations are derived from the number and types of workers and their workplace locations. This is usually the most important part of large policy-oriented urban models.

While by no means exhaustive this section reviews some of the major strands of policy-oriented modelling. The Lowry model is presented first since it was the first of its kind and is still regarded as the high point of modelling experience in 15 years. Many of the distributions generated by it were based on the gravity concept of interaction. Later developments in Britain have used the entropy maximization technique to give better theoretical basis to these techniques. The concept of entropy maximization is therefore introduced in an elementary way to facilitate understanding of the basis of the models that follow. These are models developed by Marcial Echenique and his associates.

These models probably represent the highest development of the Lowry framework and are representative of many such efforts in Britain as well as the U.S. Moreover, they are of special relevance to us because they are among the very few comprehensive modelling efforts so far attempted in the LDCs. The NBER model reviewed next is the most ambitious effort based on the use of behavioural relationships-mostly economics- that has yet been developed. This model derives much of its rationale from the 'classical' economic models discussed earlier. Finally Edwin Mills' policy-oriented planning model is reviewed. This is an optimizing model using a mixed integer programming framework.

3.1 The Lowry Model

The model was developed as part of a study of the Pittsburgh region with the purpose of aiding the regional planning effort. The objectives of this model are best stated in Lowry's words:

> "The object of this research has been the development of an analytical model capable of assigning urban activities to sub-areas of a bounded region in accordance with those principles of locational interdependence that could be reduced to quantitative form. The model is not designed to project regional aggregates such as total employment or population, but rather to allocate such aggregates to locations within the region. Properly adapted, it should be useful for the projection of future patterns of land development and for the testing of public policies in the fields of transportation planning, land use controls, taxation, and urban renewal."
>
> (Lowry, 1964, p. 2)

The model has 3 sectors: a basic sector, a retail sector and a household sector. The basic sector is the export sector whose employment and location is not affected by local events. These are activities whose location and employment levels are assumed to be given. The retail sector has local clients whose employment levels and location are closely tied in with access to local residents. The location of house-holds is powerfully influenced by the residents' place of work. In addition, the location and number of households also depend on the location of retail establishments and vice versa, i.e., they are interdependent. The structure of the model is therefore quite simple. It follows the methods of social physicists more than those of economic theorists. In other words, it seeks to replicate the urban environment by observing statistical regularities rather than explaining them. The main principle used in location of retail enterprises and distribution of households is an analogue to Newton's law of gravity.

The level of interaction is directly proportional to the mass of interacting bodies and inversely proportional to the distance between them -- usually the square of distance.

The city was divided into a grid composed of one mile squares and these were the smallest areas that the model handled. The model distinguished four types of land use:

$$A_j = A_j^U + A_j^B + A_j^R + A_j^H \tag{70}$$

where A_j is area of tract j;
U refers to unusable land;
B to area used by the basic sector;
R to retail sector; and
H to household sector, i.e., residential

A_j and A_j^B are given as is the employment provided by A_j^B, i.e., E_j^B

The retail sector is divided into types of establishment each of which has an employment function of its own:

$$E^k = a^k N \tag{71}$$

i.e., population N of the city generates employment E^k for type k retail establishments. For each tract j

$$E_j^k = b^k \left\{ \sum_{i=1}^{n} \left(\frac{c^k N_i}{T_{ij}^k} \right) + d^k E_i \right\} \tag{72}$$

This is in many ways the central part of the model. The size of establishments of type k in tract j is determined by accessibility of house-holds over the whole city but only local employment in the tract E_j. T_{ij}^k is a measure of distance between tracts i and j and N_i is population of tract i. a, b, c, and d are constants. This says that the likelihood of household shopping trips declines with distance and market potential for a tract is a weighted index of the number of households in surrounding tracts. Locally employed individuals make only short range retail trips.

$$E^k = \sum_{j=1}^{n} E_j^k \quad (73)$$

$$E_j = E_j^B + \sum_{k=1}^{m} E_j^k \quad (74)$$

i.e., total employment in tract j is a sum of basic and total retail employment.

The land A_j^R occupied by the retail sector in each tract is then determined through an exogenously specified employment density co-efficient (e^k) for each type of establishment. Thus

$$A_j^R = \sum_{k=1}^{m} e^k E_j^k \quad (75)$$

Total population is simply a function of total employment in the city.

$$N = f \sum_{j=1}^{n} E_j \quad (76)$$

The number of households in each tract is a function of that tract's accessibility to employment opportunities.

$$N_j = g \sum_{i=1}^{n} \frac{E_i}{T_{ij}} \quad (77)$$

Total population, is, of course, the sum of tract populations

$$N = \sum_{j=1}^{n} N_j \quad (78)$$

The model then has some constraints to control establishment size and densities.

$$E_j^k \geq Z^k \qquad \text{or} \qquad E_j^k = 0 \quad (79)$$

The size of type k establishment must be greater than some number Z^k.

$$N_j \leq Z_j^H A_j^H \quad (80)$$

places a constraint on maximum residential density for each tract (which may vary from tract to tract).

$$A_j^R \leq A_j - A_j^U - A_j^B \quad (81)$$

restrains the amount of land used by retail establishments to that available.

The model is shown to satisfy the necessary conditions for solution, i.e., the number of unknowns is equal to the number of equations, and a solution method is suggested using the constraint inequalities.

We note that even such a simple model with few behavioural relationships is quite demanding in terms of data and computer capacity. The city was divided into 456 tracts, it had 1.5 million people divided into 448,000 households and 550,000 jobs. It distinguished between 5 land uses: basic, residential, retail unusable and agricultural or vacant. It was found that retail trade had to be clustered into only 3 types: neighborhood facilities like food stores and gasoline services; local facilities like eating and drinking places, medical and health services, etc.; and metropolitan facilities with larger versions of local facilities like department stores, financial services, etc. Almost all manufacturing was regarded as basic. A great amount of data were needed to generate trip distribution functions. Space use standards had to be derived to generate estimates of area demanded by retail employment. It is clear that such data are difficult to find in developing countries. Furthermore, gross coefficients, (e.g., space standards) would be difficult to observe because of a far greater heterogeneity in types of retail establishments.

The model was successful in generating plausible co-distributions of employment and residential population given its very simple structure and methodological underpinnings. Lowry himself is very cautious in claiming usefulness of the model and really regards it as a first generation effort leading to better work. Its map of the city is filled partly by hand and only partly by its own structure. It is not easy to transplant from one environment to another since the structure of the model is very sensitive to the data base on which it is built.

While the Lowry model was seen to have great promise it has not been possible to use it operationally in many places. Goldner (1971) reviews the aftermath of the model in an appreciative vein but really ends up hoping that future models could be more useful. The descendants of Lowry's model observed the basic/retail dichotomy; the causal chain from basic employment, to residential population to retail employment; and the multiplier relationships of all other employment to basic employment. Despite a great amount of modelling effort in the U.S. it is remarkable that only one reached anywhere near operational use. Innovations suggested have ranged from higher disaggregation of tracts and model parameters to finer specifications of household and employment types.

Host of the operational models have emerged in Britain where it is worth noting that the number of tracts used in most models is in the region of 100 as compared with Lowry's 456. Wilson

(1974) and Batty (1972) provide a good review of the theoretical and practical developments in this field of modelling in Britain: though there was an explosion of model building based on the Lowry framework in the U.S. too it was largely unsuccessful and few models, if any, reached the operational stage. Useful reviews of these developments are found in Kendrick (1972) and Brewer (1973) in addition to Goldner (1971) mentioned above. It would appear that American disillusionment, in large measure, resulted from unrealistic expectations about what could be **quickly** learned from urban simulation models, serious underestimates of the difficulties of constructing truly useful models and the lack of an appropriate and long term financial commitment to their development. Being tied to policy and planning requirements of particular communities and studies, virtually all efforts to date have had to deal with unrealistic deadlines and other limitations. Judging from the British experience one could also say the opposite: there has been too much money available for the development of these models in the U.S. The result has been unwieldy models which never succeeded in being operational for policy use. There were therefore many disasters and the model 'movement' died in 1968. Many obituaries and post mortems have been written[1] but D.B. Lee's (1973) is perhaps, the most insightful.

In Britain, on the other hand, descendents of the Lowry model only started to appear around 1968 and were developed by Michael Batty and A.G. Wilson under the auspices of the Centre for Environmental Studies in London. Practical work on the models has been proceeding in concert under the direction of Lionel March and Marcial Echenique and their associates at the Centre for Land Use and Built Form Studies in Cambridge. They have been made operational by the Cambridge group in 5 towns in England - all of which have been of less than 500,000 population. In Britain it is conjectured that constraints of time, money and computer use had led to models which are smaller and therefore easier to put into effect. Even so it is not clear from any of the documentation if they were actually used for policy purposes. Batty (1975) has estimated that total resources expended in urban modelling in the last 8 years in Britain have not amounted to more than $600,000. About 20-30 models have been developed during this time. If this estimate is correct, the achievements

1. See Brewer (1973), Department of Transportation (1973) among others.

have been truly remarkable. A review of British modelling is found in Batty (1972).

The Echenique group has gone on to develop models in Latin America, starting with Santiago and Caracas, and is now developing one in Sao Paulo. These models are extensions of their work in the 5 towns in Britain and are essentially Lowry derivatives with Wilson's entropy maximization formalism. We review the Santiago and Caracas models in some detail but present a simple introduction to entropy maximization techniques before proceeding further.

3.2 Entropy Maximization

A.G. Wilson introduced the concept of entropy maximization to urban modelling in the context of Lowry type models. He wanted to find a better theoretical basis for the use of gravity models in distribution and allocation. He explains and develops this in his 1970 publication. The gravity model is based on a Newtonian analogy. A characteristic function of urban models is to find levels of interaction between spatially separated zones, e.g., the pattern of movements of people or goods between zones. The gravity model posits the interaction between zones i and j as proportional to each of a mass at i, a mass at j, and inversely proportional to some function of the distance (or travel cost) between them -- such as equation (72) -- in the Lowry model. The gravity analogy deals in aggregates and the formulation is deterministic. Entropy maximization deals directly with the components of the system of interest and obtains interactions as statistical averages. If we are interested in the journey-to-work, the gravity model takes a residential population in zone i and jobs in zone j as, its masses; the entropy maximizing method deals with individuals, assesses their probability of making a journey and then obtains statistical averages. This formulation is therefore probabilistic.

The aim of this section is to give an intuitive understanding of the entropy maximizing method and of its role in urban modelling. Before this can be done, entropy itself needs to be defined and purged of its thermodynamic connotations.

Entropy is a precise mathematical concept measuring the "amount of uncertainty" represented by a probability distribution. Nothing more nor less can be read into the concept. Given a probability distribution:

$P_1, P_2 ....P_n$
associated with a random variable
$$X_1, X_2, ... X_n \tag{82}$$

$$S = -\sum_{i=1}^{n} P_i \quad \ln \quad P_i$$

is the entropy of the system. It is a unique measure of the amount of uncertainty in the given distribution. It measures how uniform a distribution is. Our intuition or common sense tells us that a uniform distribution has a large amount of uncertainty. Furthermore, the higher the number of possible states, i.e., larger the n, the more uncertainty there is. The probability of any state occurring is

$$\frac{1}{n} \qquad \text{i.e.,} \quad P_i = \frac{1}{n} \tag{83}$$

Entropy should therefore be a monotonically increasing function of n. Equation (82) obeys this rule as well as some other conditions which make it a unique, unambiguous measure. Any change in the direction of equalising the different probabilities will increase the entropy.

Maximizing entropy amounts to saying that we want to make the distribution as uniform as possible subject to whatever constraints exist. The most familiar use of entropy is in thermodynamics. The characteristic problem there is that some average level of energy, say $\overline{E}_i$ in a system is known; there are many different quantum levels E_i and we want to assign probabilities to each of these quantum levels. This problem is solved by maximizing entropy subject to the condition that the expected value of E_i is $\overline{E}_i$. The solution to this constrained maximization problem assigns these probabilities to each quantum level. What the method achieves is that, given the mean value, it provides the most uniform distribution possible.

Another example presents a more intuitive understanding of the entropy maximization technique. Suppose there are 1000 cars parked bumper to bumper and they occupy the full length of, say, 3 miles. We also know the total weight of these 1000 cars. We can also find out the length and weight of each make of car that may be in this cluster of cars. The problem is: given only this information, can we make any inferences about the number of cars of each make that are in this cluster. We can convert this into an entropy maximization problem and the solution will give us the most likely distribution of makes. If n_i is the number of cars of make i, N the total number of cars, then the distribution of

$$\frac{n_i}{N}, \quad i = 1, \quad \ldots m$$

where there are m makes, is the distribution we are looking for. The total weight and length of all cars and the weight and length of each make is the information we have. This information comprises the constraints to the entropy maximization problem.[1]

The function of the technique is therefore to provide the most plausible unbiased distribution given rather sketchy information. It makes sure that no other information or bias other than that subsumed in the constraint set occurs in the predicted distribution.

In the context of urban systems, the use of entropy maximization is well illustrated in relation to a system dealing with movements of people, for example the journey to work. A **state** of the system is defined as an assignment of individual persons to the movement channels in the system such that it does not violate any of the constraints on movements. A **distribution** of the system is a macro-property of the system -- a distribution of total movements regardless of the movements of individual persons. There are three levels of resolution in this analysis. First, the micro-state is the assignment of **each** individual to particular work trip categories. Second, the meso-state is the number of individuals going from each origin i to each destination j. Third is the macro-state which describes only how many people work in each destination j and how many live in each origin i. Many combinations of micro-states can give rise to the same meso-state and many combinations of meso-states can give rise to the same macro-state. This problem is rather similar to the car problem described above. Given a macro-state we can use entropy maximization to find the most likely distribution of meso-states and given a meso-state we can predict the most likely micro-states. Since many states of the system can form one distribution, then on the assumption that all the states are equiprobable, the model is based upon the most probable distribution of person movements subject to any constraints.

Generating a journey-to-work distribution through the use of entropy maximization amounts to saying: if we repeatedly ask the population of a city to choose work and residence locations **aimlessly (or randomly),** though subject to some constraints, the distribution that we get most often is the distribution that entropy maximization provides. The constraint set can, of course, contain information which divides people according to income classes, race, employment type, etc., analogous to the information on length and weight of each make in the car example above. The journey-to-work distribution that is

1. This method of presentation is taken from: E.T. Jaynes "***Probability Theory in Science and Engineering***" Dallas, Texas: Feld Research Laboratory, 1958.

obtained is in this sense not totally random. However, beyond the constraint set, the 'randomness' prevails and it is this implication of the entropy maximization technique that is seen here as the most problematical in the context of urban modelling. Even within classifications of employment, income, race, etc., people do not locate themselves aimlessly but through the operation of some preference functions (implicit though they may be), and subject to the external market forces. If **all** this information can be included in the constraint set of the entropy maximizing procedure the same result is obtained. That this is unlikely will now be demonstrated mathematically. To this it should be added that intuitively this would appear to be an extremely difficult task, if indeed it is possible at all. In any case in practice thus far, the constraint set has seldom included any preference functions and even the market information included has been rudimentary. Thus, characteristically, entropy maximization models which utilize mean values to generate various location distributions can scarcely be regarded as having any predictive value.

Mathematically, entropy maximization does the following:

Given a random variable

x_1, x_2 .... x_n

with probability

$p_1\ p_2$ ....p_n

We maximize entropy.

$$S = -\sum x_i p_i, \quad \ln \quad p_i \qquad (85)$$

$$\text{s.t.} \sum p_i f(x_i) = E\{f(x)\} \qquad (86)$$

$$\text{and} \quad \sum p_i = 1 \qquad (87)$$

This gives

$$P_i = \exp\{-\lambda - \mu f(x_i) \qquad (88)$$

$$\text{and} \quad e^{\lambda} = \sum \exp\{-\mu f(x_i) \qquad (89)$$

(since P_i = 1)

where λ and μ are the Lagrange multipliers.

This shows that if we know averages we can generate trip distributions. This is the basic entropy maximization procedure: clearly the number of constraints can be increased to include more information.

The following now demonstrates the conditions in which a utility maximizing technique gives similar results. This treatment follows Wilson (1970).

In the standard utility maximizing problem, we maximize:

$$U = u\ (X_1, X_2, \ldots\ldots\ldots X_n, M) \qquad (90)$$

s.t.

$$x_i P_i = M \qquad (91)$$

where U is the utility of the consumer

x_i are amounts of goods consumed
p_i are prices
and M is the income of the consumer

To obtain a solution we maximize , yet first order conditions, solve the resulting equation system and then find the optimal quantities.

If we now define:

$$y_i = \frac{x_i P_i}{M} \tag{92}$$

y_i can be a probability distribution and we have

$$S = -\sum y_i \quad \text{In} \quad y_i \tag{93}$$

We can also write and maximize

$$U = U\left(\frac{y_1 M,}{P_1}, \frac{y_2 M}{P_2}, \ldots., \frac{y_n M}{P_n}, M\right) \tag{94}$$

$$\text{S.t.} \qquad \sum_i y_i = 1 \tag{95}$$

and find a solution

$$y_i = y_i\,(p_1,\ P_2 \ldots., P_n\ M) \tag{96}$$

which defines the same system as the standard utility maximization system.

Now let there be constraints

$$f_k\,(y_1, y_2 \ldots.. y_n) = g_k \tag{97}$$

$$k = 1, \ldots m$$

and maximize

$$S = -\sum y_i \ln y_i \tag{98}$$

s.t.

$$y_i = 1 \tag{99}$$

and constraint set above (97)

Wilson then suggests that maximizing S is equivalent to maximizing

$$U = S + \sum \mu_k (g_k - f_k) \tag{100}$$

$$\text{s.t.} \qquad \Sigma\, y_i = 1 \tag{101}$$

Here under certain conditions μ_k are the Lagrange multipliers of the earlier maximizing S problem.

If entropy S plays no role in the utility function this will exhibit itself in two ways:

(i) The parameters μ_k will be large compared with S, thereby reducing S to insignificance;

or (ii) There will be as many constraints (components of the utility function) as there are variables y_i, in which case the set of constraint equations can be solved directly for the y_i s without reference to entropy.

What is really being said here is that if we put all the information that we have in a utility function into constraints in the entropy maximizing problem the solution will be the same. To understand this better: if a utility function orders behavior in a rather deterministic fashion and this can be transformed into a constraint set there will be little allowance for uncertainty in the system. Hence maximizing a measure of uncertainty (entropy) would make no difference and the same answer would be obtained. The argument is therefore that if the utility function does not have this kind of information then uncertainty should be taken account of: entropy maximization does this while utility maximization does not. Therefore the former is a more general procedure.

The problem is that it may not always be possible to transform the information in a utility function into a usable constraint set for entropy maximization. If this in fact can be done there would be no difference in the solution and there would be no meaningful choice, However, maximizing entropy in a system which is completely determined by the constraint set serves no purpose. Thus the correspondence between utility maximization and entropy maximization is illusory: it happens only when entropy maximization loses its meaning.

It is hoped that the above gives an intuitive as well as partial mathematical understanding of the entropy maximization approach. It gives some idea of the role of mean values and of the various parameters that are used in distributions in the Echenique et. al models that are reviewed next. Some basis is also provided for the methodological criticism that is offered for these models.

3.3 The Echenique et al Models

Some of the main attempts to apply modelling techniques to L.D.C. cities have been made by Marcial Echenique[1] and his associates in cities in Latin America. They have built and calibrated models for Santiago and Caracas and are now doing so for Sao Paulo. These models are essentially built around a Lowry framework embellished by entropy maximization distribution techniques while each model has its special characteristics and objectives. Here we describe in some detail the Caracas and Santiago models though technical detail is kept at a minimum. The Sao Paulo model is the latest one being developed and no documentation is available as yet.

These models are extensions of the urban stock and activity model developed at Cambridge (Echenique et al., 1969) which had been applied to new towns in Britain. The Santiago model expands the original framework of the Cambridge model by modelling the inputs to the urban model; that is, the location of basic employment, the transport cost and parameters such as the labor participation

1. Of the Martin Centre for Urban and Architectural Studies, Cambridge University, England.

rate and service employment ratio. This is done by coupling the urban model with a regional model and a detailed transport model. The original model gives at every time interval the changes in employment by economic sectors and the population change. The transport model interacts with the stock and activity model determining the level of accessibility in each zone by a detailed assignment of vehicles to the networks.

The Caracas model attempts to disaggregate the urban model in order to explore policies related to income distribution and squatter housing. In order to do this an economic framework is used to explain the urban processes: given the location of basic employment, the socio-economic groups and income groups of these employees are calculated. The income of workers determines the housing type and transport that they can afford. The model attempts to simulate the operation of a simplified land market, establishing rents and land values by competition of different land uses.

The Santiago Model

The structure of the model is illustrated in Fig. 3.

The Regional Model predicts changes in population given birth and death rates, migration data and demand for labor in each region. Employment is regarded as proportional to the amount of investment which is considered exogenous. In this case only public sector investment is taken into account since it comprises 77% of total investment and is, perhaps, more predictable. The output from the regional model is fed into the intra-urban model at each time period. The results of different investment policies, e.g., different allocation mixes between regions can be tested. The model gives employment in each region by different sectors: agriculture, mining, manufacturing and administration.

The Intra-urban model is divided into various parts:

(i) Basic Employment Model: This distributes the total basic employment given by the regional model through zones and employment sectors. Agricultural employment is a function of

-agricultural land available
-fertility
-accessibility to consuming zones

and is subject to maximum density constraints.

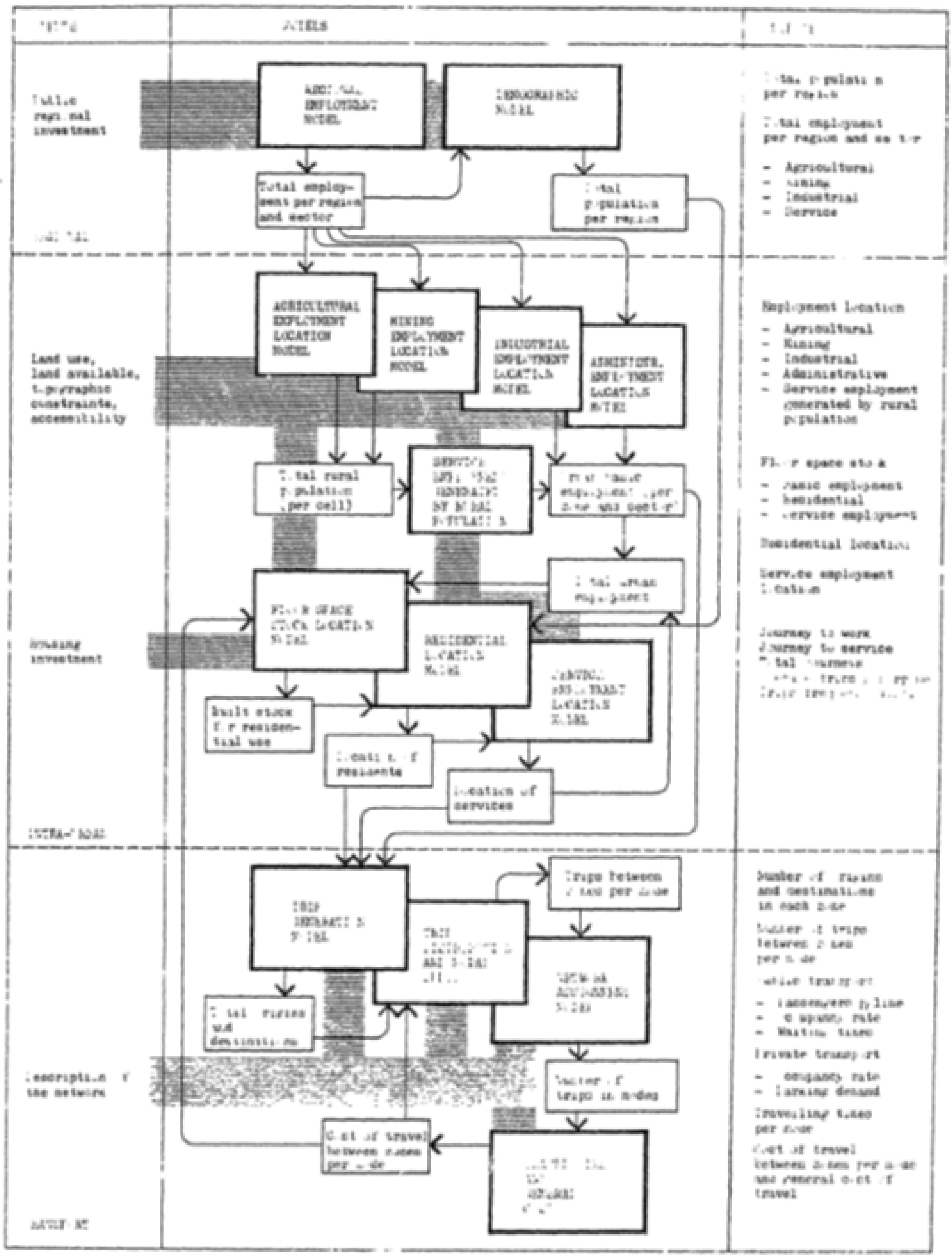


Figure 3: THE STRUCTURE OF THE SANTIAGO MODEL

(ii) Industrial employment is a function of

-industrial land available
-regional accessibility of zones

and an additional parameter is introduced to simulate the existing cluster behavior of industry.

Administration and finance employment is merely a function of urban accessibility. Finally, some service employment which is generated by the rural population (agricultural and mining) is also regarded as basic employment.

(iii) The urban stock and activity model: this takes account of the spatial characteristics of the city to distribute the population to residential areas and to generate service employment. It takes as inputs
- amount of land available for development in each zone
- transportation network
- total amount of floor space per employee
- basic employment space standards

The model first distributes the existing stock of floor space, then the location of residential activity and finally the location of service employment generated by the residential activity. This process has to be iterated until the service employment reaches the total given by the regional model. These distributions are done by the entropy maximization techniques.

(iv) The transportation model:
Given the distribution of activities and the transportation network the transportation model can simulate the travel activities of the city. Since the distribution of activities depends on the transportation network and vice versa the transportation model can be linked to the stock and activity location model and iterated to simulate the changes resulting from changes in either. The model has four stages:

(a) Trip generation: this gives the total number of trips emanating from and coming cc each zone.

(b) Distribution and Modal Split: this provides the distribution T^k_{ij}, i.e., the number of trips from zone i to zone j by mode k. This requires a cost matrix to operate. The modes considered were bus, motor car and pedestrian.

(c) Network assignment: this assigns the trips through the network defined as a set of modes and links. Each mode has a different network: buses follow pre-established routes, cars avoid congestion while pedestrians follow the shortest routes since no congestion is allowed for them.

(d) Generalized cost; this assigns costs to each of the modes and routes, These costs are then fed back into step b. and iterated until convergence.

The output of the transportation model is essentially the inter-zonal accessibility matrix which is fed back into the activity stock and location model and the whole model is then iterated.

Simulation for Santiago: Since the governmental changes in Chile this model has been abandoned. It would appear that not much work was done on it after calibration. Work on the model stated in 1970 and it was finally operational in 1973.

The simulation was done over a sixty zone irregular pattern. Data were obtained from various sources.

- population from the 1970 national survey
- employment from 1969 and 1972 origin destination surveys and made compatible with the 1970 census
- floor space from the taxation office
- roads from the Ministry of Public Works and Transport.
- maps.

All information was based on 1970. The regional model was run to simulate the period 1965-70. It was run, under 3 different investment assumptions for prediction up to the year 2000.

- continuation of part investment pattern
- no investment in Santiago
- no investment in service activity in Santiago

All the simulations were found to be quite accurate as compared with actual data for the relevant years. There is no evidence nor information on the efficiency of the model as a predictive device. Since the model is calibrated on 1970 data it is not surprising that it simulates the past well.

The Caracas Model: A Disaggregated Model of a Metropolitan Area

This model is said to combine the macro-scale or social physics approach with the micro-scale or economic approach. It attempts to simulate interaction of supply and demand in the land market. Given the location of 'basic' employment (manufacturing, government and agriculture) the model distributes employees to their residential places. As in the Santiago model and other models with the Lowry structure, the residential population generates demand for services which generates more employment and more services and iterates until equilibrium is reached.

It has 5 sub-models:

(i) Employment: This determines the socio-economic group according to employment type. Each socio-economic group has an income distribution, which determines the housing type they can afford as well as the transport they use. The model has 4 employment types

-service
-industrial
-government
-agricultural

and 3 socio-economic group types,

- managerial
- professional and technical
- clerical
- manual
- agricultural worker

We have

$E_i^{xy} = E_i^x p^{E(xy)}$

where E_i^{xy} is number of employees working in zone i of employment type x and socio-economic group y. Given E_i^x the number of employees of type x in zone i, and $p^{E(xy)}$, the proportion of employees of type x who are in socio-economic group y, E_i^{xy} is obtained.

Once this is obtained the income of each group is determined. To do this two parameters have to be known.

mean income of employees in group y,
a parameter of the distribution in group y.

The distribution is done over 5 income ranges. The next step in the procedure is to calculate E_i^{zho} the number of employees working in zone i in income group z living in house type h using transport type o. Here there are two types of housing h

- normal housing
- squatter housing.

Transport types ire also two

- car owners
- non car owners,

This calculation can be done with a knowledge of the following parameters:

- value of mean income in income group z
- mean income of those using transport type 0.

The distribution parameters generating the distributions from these mean values have to be calibrated.

The output of this sub-model therefore gives (a) the number of employees

$$E_i^{xy}$$

working in zone i of employment type x and socio-economic group y

(b) E_i^{zho}

those working in zone i income group z living in home type h and travelling by transport type o

(c) various combinations and summations of the above.

(ii) *Land:* This sub-model distributes land for different uses depending on rent paying ability and total supply of land.

Symbolically, the model calculates L_j^g the quantity of land used by activity g in each zone j. The supply of land is considered as a function of total land available in the zone, L_j, and the demand by each activity. Here the activities are of six types

- service
- industrial
- government
- agricultural
- normal housing
- squatter housing

i.e. the sum of employment and housing categories.

Once again various mean values and standards are required and distribution parameters to be calibrated:

- land standards according to each activity
- level of each activity in each zone
- mean rent paying ability or each activity in each zone
- amortization rate of expenditure in land according to activity.

The total value of land in each zone is calculated as the sum of all expenditures on land in each zone.

(iii) Residential location: This set of equations determines E_{ij}^{zhok}, i.e., the likely distribution of employees working in zone i living in zone j in income group z, housing type h, transport type o and using transport mode k. For this to be accomplished, various other calculations have to be made first.

- monthly cost of location which consists of monthly cost of transport to work according to mode, monthly average cost of transport to services according to income groups, value of time spent in travelling according to income group and mode, and the average rent for housing type h in zone j
- mean location cost for each income group

Each of the location cost components are calculated as usual with the help of various mean values and distribution parameters.

Once the employees are distributed to residential zones according to income group, housing type and mode of transport the model transforms them into households and thence the total population in each zone.

(iv) Transport: This sub-model calculates the transport costs needed for the residential sub-model.

The cost of the journey to work is merely an average cost per mile for each mode multiplied by the distance and frequency of trips between zone i and j.

The cost of transport to service locations is somewhat more complicated since it has to be calculated according to household location and income groups. Furthermore, service trips from workplaces have to be accounted for also.

(v) Service location: This sub-model calculates

$$p_{ij}^{sfhok}$$

i.e. the number of people living in zone j, travelling to zone i for services, belonging to household income group f, housing type h, having transport type o and travelling by mode k.

The calculation involves a knowledge of mean monthly transport costs to services and the calibration of a distribution parameter.

Having calculated the distribution of population making service trips the model generates

E_i^s

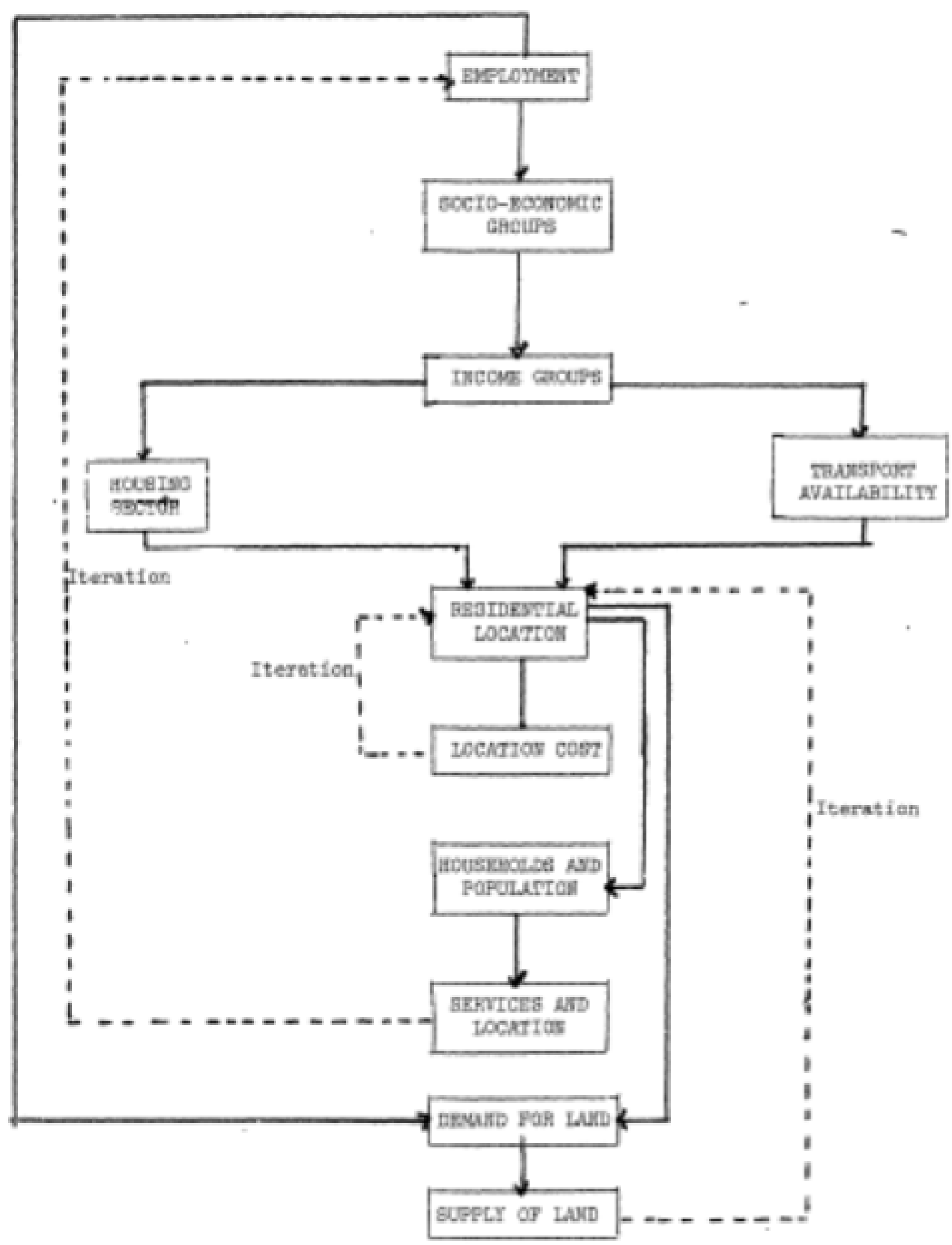
the number of employees required for services in zone i.

Solution of the Model

Figure 4 illustrates the structure of inter-relationships in the model and the iterative process used to solve it. It also well illustrates the central importance of the residential location sub-model. A solution of the model requires the following equilibria.

(a) Employment-Population: Given employment, residential population is generated resulting in a demand for services, hence service employment is added to original employment and the process iterated until equilibrium is reached.

Figure 4: CARACAS MODEL: STRUCTURE OF ITERATION PROCESS



(b) Location cost - residential location: is largely determined by location costs of which one component is rent. The rent is a mean for all residents and therefore depends on who lives in the zone. This process is iterated to reach an equilibrium.

(c) Residential Location - Land: The demand for land is generated by mean rent which determines the amount of land residents are able to buy in competition with other activities. The new land is distributed until there is no further change in the distribution of residents across zones.

The order of solution is as follows:
Given employment

(i) location cost - residential location equilibrium is achieved

(ii) service location and distribution of land is achieved

(iii) new increase in employment causes more residents and the whole process is repeated again and again until full equilibrium is achieved.

It is not clear from the summary mathematical statement of the model why the model should be expected to converge to an equilibrium.

Application to Caracas

The city was divided into 30 zones. Service and government employment was concentrated in one zone, industrial employment in three and agricultural employment only in peripheral zones.

All land with less than 60% slope was regarded as available but excluding land for such public purposes as parks, military uses, cemeteries, university and for roads. The employment structure was found to be

51% service
32% manufacturing
16% government and services
1% agriculture

Most of the data came from a 1966 5% sample survey covering some 60,000 people. The data required were

(a) Basic inputs:
- employment located in each zone according to employment type
- land available in each zone
- distance matrix for each pair of zones

(b) Coefficients:
The nature of these has been mentioned in the description of the model.

(c) Output values: essentially the outputs are

E_{ij}^{zhko} - employees working in i living in j, income group z, living in housing type h, having transport availability o and travelling by mode k

H_j^{fho} - households living in j, income group f, living in housing type h and having transport availability o.

$c_{ij}^{loc \quad zhko}$ - cost of location of an employee living in j, working in i, income group z, house h, using transport mode k and owning transport type o

Various summations of the above can also be obtained for different variables.

In addition, transport costs, housing costs, land values, land availability, trip lengths, etc., can also be obtained in disaggregated form.

Simulation of the model gives good results in terms of the closeness of output values to the actual 1966 data.

Evaluations

The development of the Caracas model was begun in 1968 and, perhaps, finished in late 1973 or early 1974. It is not clear how much it cost nor how many professional man-years were spent on its development. The total financial as well as professional costs are likely to be substantial, especially if the basic model development costs in Cambridge, England are also included.

The basic criticism of these models is of their model structure and methodology. Despite Wilson's innovative explanation and development of entropy maximization techniques, their basis for urban simulation remains suspect. What the technique essentially does is to take the mean value of some variable and then generate some most likely distribution around it subject to whatever constraints are chosen.

To illustrate, one step in the procedure in the employment sub-model (p. 94) is to calculate E_i^{zho} - the number of employees working in zone i, in income group z, living in house type h, and using transport type o. This calculation is carried out by using:

- value of mean income in income group z
- value of mean income in housing group h
- value of mean income of those in transport group o

and two distribution parameters for distributing housing group h and transport group o. These parameters are calibrated from the base data. Thus, various mean values pertaining to different groups are used along with calibrated distribution parameters to obtain the required E_i^{zho} distribution. A similar procedure is employed in all the other distributions obtained in the model.

As such, the model merely replicates observed data and i, is then not surprising that its simulation is remarkably close to actual values. At best, such a model can be viewed as a set of reduced form estimates of behavioral structures that are not specified, at worst it is a collection of spurious, accidental or temporary relations between variables. If the latter is the case, the model has no content and could be dangerously misleading in forecasting. If the former is true, then the model is useful as long as the underlying behavioral structure is unaltered. The problem is that it is impossible to judge which is the case and even if the more optimistic assumption is true, we have no basis for knowing when or for what reasons the unspecified underlying structure changes.

The usefulness of this model as a policy tool is therefore under serious question. The Caracas model was calibrated for 1966 data. It has not been updated for any later year because of lack of data. If this had been done there would have been some basis for an informed evaluation. Model simulation for a later year based on 1966 calibration could have been tested on actual values. On the basis of this test some faith in the invisible underlying behavioral structure might have been generated. As it is there is no basis for even this modest assurance and the lack of post 1966 data merely underscores the difficulty of obtaining data for such a model of an LDC city.

The Santiago model never went into real operation because of the governmental changes in Chile. It was developed under governmental sponsorship so the expectation was that it would have been used for policy purposes. While it is true that the governmental change in question was of a radical nature, this circumstance nevertheless illustrates an important problem in the use of such a large model for policy purposes. Any ‘good’ large model requires time to develop as well as for data collection. Three years is almost a minimum for the model to become operational. Yet policy makers change over such periods as a matter of routine. LDCs in particular are more prone to rapid change. Models such as the Santiago and Caracas models, which have a rather opaque methodological base, are then all the more difficult to explain and ‘sell’ to successions of competing policy makers. If their bases were less opaque, it could be easier to convince new policy makers of their continuing validity.

The Echenique group is now developing a model for Sao Paulo. It is expected that this model will take less than a year to develop, utilizing about six professional man years in the process. This model will be a more aggregated version of the Caracas effort. Documentation is still not available, so not much can be said. If it is developed within a year, that would be an encouraging sign. However, the strictures against its methodology would remain.

It must be pointed out that the Caracas model was developed as an educational device in a university environment. Strictures against its use as a policy tool may therefore be unfair to the authors, although they proffer good advice to potential users.

The problem of data has already been mentioned. It has two dimensions. First is the sheer lack of and the difficulty of collecting such detailed data. Second is the time-factor. With the rapid pace of change in LDC cities, no sooner is a detailed set of data collected, than it is obsolete. What this implies is not that data should not be collected and modelling given up but that smaller, more manageable sets should be sought for smaller, more manageable models.

The basic use of the Echenique models for policy purposes is as forecasting tools. They can give planners information on the possible consequences of their actions as well as stimulate thought about new directions, They have no normative content. They do not help in evaluating any consequences. To the extent that the planners have faith in their forecasting structure the models are clearly useful, to the extent that they do not, the models are unusable. It is then left to the modellers to find ways of defending their model structure and explaining it to planners. This process is easier for models with more comprehensible structures. Models placing greater stress on behavioral relationships can be just as misleading if these relationships are badly estimated, but they are easier to test since these relationships are easier for the non-modeller to appreciate or reject. In this respect it is encouraging to note that the Caracas model has as an objective simulation of the land market according to micro-economic concerns, though the actual market simulation is somewhat primitive.

Before concluding this section, t is worth mentioning that Nathaniel Lichfield and Partners have developed similar models somewhat further as described in Christopher Turner (1975) and Lichfield (1975). In addition to a few structural modifications, their major extension in the Urban Growth Simulation Model for North Central Texas is the application of evaluative sub-models to the output of the main model. These sub-models evaluate the consequences of alternative policies on the cost of public utilities, on air pollution, accessibility to urban resources and 'social deprivation'. This is an encouraging development in making these models more directly relevant for policy concerns.

3.4 The N.B.E.R. Simulation Model [1]

This is perhaps the most ambitious of all urban modelling attempted. It follows a somewhat different family of models pertaining to transportation and urban land use. Six of these built in the late fifties and through the sixties for the Puget Sound, Southeastern Wisconsin, Atlanta, Detroit and the San Francisco Bay area (2) are well reviewed in Brown., et. al. (1972). Here we merely describe some of their unifying characteristics as a prelude to the N.B.E.R. simulation model.

The objective of these models was to help policy makers in the planning of transportation. They are characterized by the assumption of an undirectional relationship between land use and transportation. Thus considerable effort is devoted to modelling land use in some detail to derive transportation requirements. Regional population and employment forecasts are taken as exogenous. Input-output methods are used to forecast future employment by industry. Retail employment is derived from these forecasts and households located according to family type. Different models employ various levels of disaggregation for dwelling types and household types and their assignation. These results for projected land use then provide the basis for future transportation plan design. Some use supply and demand concepts for equilibrating the housing market, taking into account volume of housing stock and pattern of filtering with age of structures. The San Francisco models are somewhat different in that they follow the Lowry model structure.

1. This section reviews the work described in Ingram et., al. (1971, 1972), Kain and Ingram (1974) and Brown et.al. (1972).

All these models have heavy data requirements, but still have few behavioral relationships embedded in them. They are in the genre of mechanistic forecasting models which find it difficult to cope with technical changes and innovations which affect the structure of cities in a crucial way. Interdependencies are usually modelled in a sequential manner: the tension between sequential and simultaneous relationships has been observed earlier. All of these models were expensive:

(i)	Atlanta $1.75 m	
	Data collection	36%
	Analysis and models	24%
(ii)	South Eastern Wisconsin $ 1.99 m	
	Data collection	62%
	Analysis and models	14%
(iii)	Bay Area Transportation Study $5.54m	
	Data collection	60%
	Analysis and models	18%
(iv)	Detroit Talus $ 4.70m	
	Data collection	46%
	Analysis and models	19%
(v)	Puget Sound Transportation Study $1.7 m	
	(further breakdown not available)	

We note that a major part of the expense was always on data collection thus underscoring the importance of looking critically at data needs of models.

The NBER simulation model was embarked on to improve on earlier work. It was deeply rooted in economic theory with utility maximizing households and profit maximizing firms. Its goals were to:

(a) enrich economic theory;
(b) advance the art of model building;
(c) evaluate problems of urban growth and decay; evaluate specific problems and policies; and consider broad strategies fur dealing with U. S. cities.

This effort has been quite successful in realising (a) and (b), but somewhat less so in realizing (c), despite the fact that it has been handsomely supported financially as well as intellectually. The model does not appear to have reached the stage of evaluating public policies.

We now describe the model itself:

The Model

The NBER model incorporates the theoretical approach of the traditional analytic models of residential location and urban spatial structure into a framework with more realistic and less restrictive assumptions. It is more realistic in the following ways:

(i) It drops the monocentric assumption of analytic models and explicitly incorporates multiple work places.

(ii) It abandons the long run equilibrium framework of analytic models which characteristically ignore the effects of durability of capital. The NBER model overcomes this by representing the standing stock of physical capital in the city and models the supply side of the housing market in detail.

(iii) Finally, it takes account of externalities such as neighborhood effects and racial discrimination.

It was designed to simulate major changes in urban spatial structure that occur over periods of from 10 to 50 years. It simulates effects on spatial structure of long term trends in the level and distribution of employment, changes in transportation, technology and increases in income. It provides a description of spatial structure at a point in time and modifies this over a period of years by simulating location and investment decisions of firms, households and home suppliers.

The model, is primarily a model of urban housing markets. It does represent other urban phenomena, such as industry location and changes in the demographic structure of the population, but the behavior of the housing sector is its central concern. This involves modelling the behavior of housing consumers, suppliers and the 'market' in some detail. In doing this it is claimed to improve on previous models (most of the social physics variety) which were elaborate statistical descriptions with little or no theoretical justification.

The model can be described in terms of a demand sector, a supply sector and a market clearing sector. The activities in each of these sectors are carried out in one or more of seven sub-models.

(a) Demand Sector

(i) *Employment Location Sub-Model:* Given an exogenous change in total employment, it revises the level and composition of employment at each work place and by each of nine industry types. It translates employment changes by industry to changes in employee characteristics.

(ii) *Movers Sub-Model:* According to the results of (i) above this model generates movers, i.e., households vacating housing units in response to employment changes, and modifies them to produce households seeking housing.

This has the effect of preserving some of the city structure and changing it only incrementally.

(iii) *Demand Allocation Sub-Model:* This is where the model allocates households to housing types. It takes account of the costs of the journey to work and expected housing prices to form gross housing prices. Thus each housing type is associated with a gross housing price in relation to every work place. Finally, household allocation is performed by the use of demand functions.

(b) The Supply Sector

(i) *Vacancy Sub-Model:* As a result of the movers (a, ii above) this sub-model generates vacancies in each zone by housing type. This also includes new construction.

(ii) *Filtering Sub-Model:* In response to expected housing prices formed in (a, iii) above quality classifications of available housing stock are changed. This also takes account of expected maintenance cost and therefore relative profitability of different maintenance strategies.

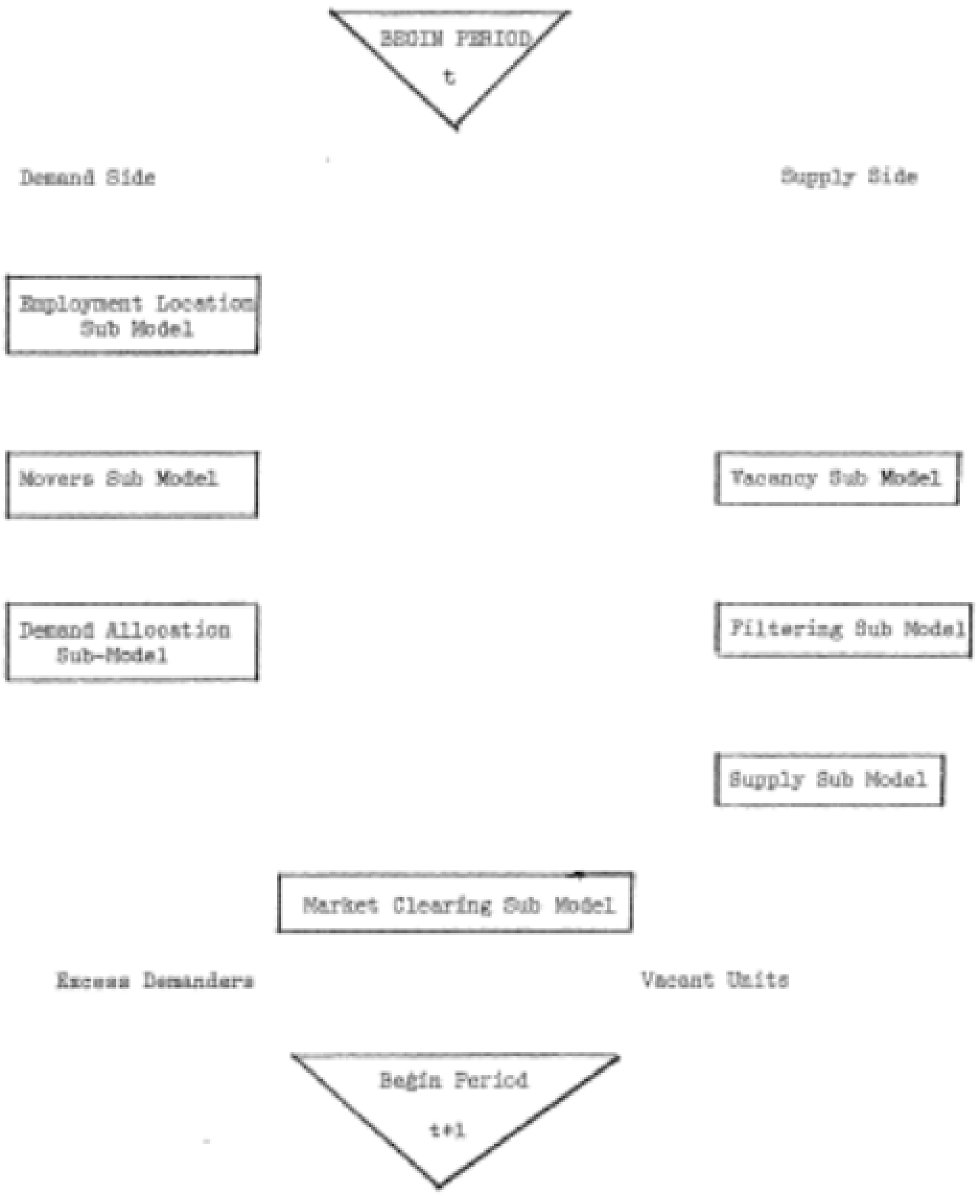
(iii) *Supply Sub Model:* This also responds to (a, iii) above and performs stock transformation according to profitability of construction and transformation activities from expected prices and exogenous building costs.

(c) Market Clearing Sector

Market Clearing Sub-Model : Moving households are matched to available units of the types chosen in the demand allocation sub-model. Each house-type is solved for as a separate sub-market. Shadow prices are used to generate prices for the nest Lime period. Work-trip patterns are also updated.

The model structure described above is illustrated in Figure 5. Embedded in the model structure described above are the following special features not found in most other models:

Figure 5: STRUCTURE OF THE N.B.E.R. MODEL



(i) Most other models locate the entire population at once to produce target year solutions. This model allows adjustments in employment and residential locations, alterations to the housing stock and changes in the distribution of work trips in an incremental fashion only. This has the effect of letting the existing city structure influence the future which is a pleasing simulation of reality.

(ii) The model does not force equilibrium of each housing sub-market in each period. Individual housing sub-markets are allowed to have excess supply of units (vacancies) or excess demand which, in turn, affect prices in the next period.

(iii) The model produces expected housing price by housing type and zone for each period. These prices affect the behavior of households seeking housing as well as firms producing them.

In the 'Detroit Prototype' version of the model the region is divided into 19 workplaces and 44 residence zones, the former being aggregates of 32 inner residence zones. It distinguishes households by family size, family income and education and age of head resulting in 72 household classes. Housing is distinguished by structural type, number of rooms, quality and lot size, resulting in 27 types. There are two modes of travel.

We now describe some of the processes determining the distributions generated by the model.

(a) Behaviour of the Consumer: The consumer is essentially seen as the classical utility maximizing household. In this model the following assumptions are made:

(i) The household has a fixed and predetermined set of demands for travel to known destinations. The journey to work predominates in travel costs and households place monetary value on travel costs.

(ii) Households have preferences for housing 'attributes' which they buy in a finite number of combinations of 'housing' bundles.

(iii) Housing bundle prices vary by location and these price surfaces are known to consumers who act as price takers.

The consumer's problem is then posed as a cost minimization problem. Since travel costs are subsumed in gross housing prices, the optimal location for the household is that location which has a 'housing bundle' whose price is the minimum. These minimum prices, furthermore, are the result of its demand based on income, household characteristics and taste in the traditional demand analysis way.

(b) Determination of Housing Prices and Quantities

As mentioned earlier, the NBER model differs radically from earlier approaches in recognizing the durable nature of urban structures and allowing city structure to change only marginally in each period. This also has the result of allowing disequilibrium in various housing sub-markets. However, it follows traditional economic theory in assuming that housing production is responsive to market demands and prices, that suppliers are profit maximizing and that households are price takers in both output and factor markets. It assumes further that

(i) Housing outputs are heterogeneous and are produced using combinations of existing durable structures, current inputs and neighborhood attributes.

(ii) Most of the supply in each period is from used structures.

(iii) Some of the housing attributes are not produced by competitive firms, e.g., they are supplied by local governments. This has the effect of placing constraints on suppliers' actions.

The production function for housing consequently reflects these assumptions in including existing structures and neighborhood characteristics as inputs in the function. This is a radical departure from standard practice which usually includes land, capital and labor in inputs in a production function.

(c) Market Clearing

The household's selection of a housing bundle is represented in the NBER model by econometrically estimated demand functions. These equations express the probability that a particular household will consume a particular type of housing as a function of the household's socio-economic-demographic characteristics and the minimum gross price of the bundle. Solution of these functions gives the number of persons employed at each workplace who will demand each type of housing bundle. The demand for each type of bundle is summed over all workplaces, and firms then attempt to satisfy this demand in each location on the criterion of profit maximization.

Spatial competition among households competing in the same housing sub-market is represented in this model by a linear programming algorithm which minimizes aggregate travel expenditure for households competing in the same sub-market. Thus the procedure also yields shadow prices for each bundle type in each residence zone in the solution.

The output of the model essentially gives the workplace and location of each household and the bundle of housing it consumes. Further it produces expected prices for each type of housing in each residence zone during each period.

Included among the inputs to the model are employment; costs of performing various supply activities, (i.e., an array of supply costs with a given technology and fixed factor costs); zoning constraints like limits in residential density; amount of available land in each zone; and forecasts of demand over current period.

Evaluation

The NBER model is being developed still further and has now reached a third version, 'Pittsburgh II', but little documentation is available after the Detroit Prototype. The major modification is an expansion of the number of neighborhoods in the model which (because of the way housing bundles are defined) reduces the number of housing bundles that have to be coped with in each zone. The model has really developed into a research process rather than a product.

It must be recognized in conclusion that the NBER model is an admirably ambitious attempt to incorporate urban realities within a basic framework of urban economic theory to produce a model of city structure. It is not, however, clear yet how successful or how expensive this effort has been. Its development was begun in 1968 and has cost, at a minimum, one million dollars. There has not yet been a policy simulation with the full model. Its data requirements are clearly high: detailed demographic data are needed to generate household types and their characteristics; housing price data and construction costs data are needed in detailed fashion for the supply sector. The model, as a whole, is quite unwieldy and does not yield many fruitful results in comparison with the effort involved in building it.

It is difficult to enumerate reasons why an enterprise such as the NBER model effort has not become an "operational" model. It was started with clear-cut objectives, had ample and long term financial support and has had some of the best urban practitioners involved in its conception. One reason is merely that it was over-ambitious. Another reason is, perhaps, that its objectives have changed over the course of its development. While it was originally envisaged that it would be a useful policy-making or policy-helping device, it has slowly become an almost wholly research device. The specification of each sub-model has been a project in itself and has illustrated the use of economic theory in modelling. In this respect it has been very instructive. It has spawned a host of side studies which have enriched the state of the art of urban studies.

It is probably fair to conclude that the NBER model has been more successful as an analytical or explanatory model rather than as a policy-oriented model. This is as should have been expected since its size is too large and data requirements too intensive and detailed for the time and resource constraints of policy making. Its usefulness for IDC cities lies in some of the ways it has adapted existing theory to reality. In particular, its enumeration of housing-bundle types may be a good approach to the extremely varied bundles found in LDC cities. Partial research efforts in these cities could be modelled after parts of the NBER model - applying it fully would be foolhardy.

3.5 Mills Optimizing Programming Model (Mills, 1975).

Finally, in this section we present yet another kind of model: Mills' recent optimizing programming model. Although one would particularly expect policy oriented models to have some normative content, such models have been more the exception than the rule.

The paradigm for normative models in economics is that first some market resource allocation is derived, then tested against some welfare criterion and then government policy prescribed if this allocation turns out to be sub-optimal. The assumption is always that the government operates in the public interest while everyone else furthers his own interests. In this model Mills explores the issue of the desirable extent of governmental regulation and interference in the urban system. It introduces assumptions that indicate why it is desirable for government to undertake certain activities. It then specifies the effect of government activity on private resource allocation, and the effect of private activity on the use of the public service. It permits calculation of an optimum allocation of both public and private resources. Finally it demonstrates that competitive markets sustain an optimum allocation of resources if the public sector provides its service in optimum fashion.

The model concentrates on the provision and pricing of transportation. It is formulated in a non-linear programming framework and is a development of earlier models by Mills (Mills, 1972b, 1974).

The Model

Unlike other models presented in this section this one makes a larger number of simplifying unrealistic assumptions but it still is rather complex. The city is seen as a homogeneous plain stretching in all directions from a central export point. The model determines the amount and production technique of each of an arbitrary number of goods and services to be produced at each location in the urban area. Each good that is produced must be exported or transported to the point where it is consumed. This model unlike others therefore also takes account of traffic in goods.

Space in the area is represented by a square grid centered on the central export point. All squares at a given distance from the center have identical patterns of production, consumption and transportation. Transportation is only between squares and only in the north-south and east-west directions.

$\bar{r}$ goods are produced in the city. Of these, two are high income housing and low income housing. All other goods are exported and the export amounts are given exogenously. Apart from the center goods can also be exported from suburban nodes a miles from the center. The first set of equations in the model are identities to ensure that

(a) total exports of each good are at least as much as exogenously prescribed and

(b) shipments into each square plus production in the square equal outward shipments plus use as input or final consumption in the square. It is important to note that this implies a set of $\hat{u}$ x r equations in each square.

The next set of inequalities builds the transportation system. They determine $t_{rk}(u)$ which is the number of unit miles of good r shipped at congestion level k per square at u. The congestion level helps in determining the width of roadway required - higher congestion means a narrower roadway. Once again, this is a set of r x $\hat{u}$ inequalities. In addition there is one to ensure-that land demanded by transportation is less than total land for transportation. A further set of inequalities and equations ensures that there is only one level of congestion in each square. Finally, there is an inequality to ensure that the land used for all purposes does not exceed the total available.

A feature of the model is that it permits an arbitrarily large number of production techniques used to manufacture goods. This is done by concentrating on the capital-land ratio which is represented by the height of buildings. A large number of storeys means high capital-land ratio. Labor input-output coefficients are independent of building heights.

Such a view of the production process implies that a city's labor force and total output of all goods are determined by the export requirements via the input-output matrix. Transportation costs, land and capital requirements, however, are endogenously determined. The objective function is then to minimize the sum of land, capital, transportation and exports costs needed to produce the required export goods. It is assumed that R_A rent for land, at the city periphery is uniform, and the city can therefore be expanded at this uniform unit cost. Similarly, capital can be acquired without limit at a fixed rental rate R.

The model is solved using mixed integer programming techniques. The only non-linearity in the model is the integers used to represent congestion levels.

Solution of the Model. The model is solved for a hypothetical U.S. city of about 1 million population.

It has 5 sectors, i.e., $\bar{r} = 5$

- office activities
- retail firms
- manufacturing firms
- low income housing
- high income housing

$\bar{u} = 11$ implying an urban area of about 250 sq. miles.

$\bar{s} = 20$ implying that the highest building has 20 floors.

Income groups are the lower half of the population and the richer half

R_A = \$4000/acre and
R = \$1m rental rate per \$10 m worth of capital value.

An entire input-output matrix was constructed for the 5 sectors. Approximate values of automobile travel and goods movement were used. Time costs varied with income.

With such a simplified model there are 788 variables excluding slack variables. Of these, 30 are integer valued. There are 219 constraints in the example. The model was solved on an IBM 360/91 computer and took 6 minutes (\$70) of computer time and used about 400 K of the core of the computer. (The capacity of the core is about 1100K).

The solution results in a city rather typical of a U. S. city as intended. The central square is devoted to office activities in 17 story buildings and transportation is at the highest congestion level. The next square has retail firms, some low income housing, in 5 story buildings. Square two has high and low income housing as well as some office activities. The city center is ringed by low income housing as is typical of U.S. cities. All other squares have some of all activities except that the suburban export nodes (û = 7) has no manufacturing firms.

The congestion pattern is unsatisfactory: high in square zero, low in square one and than high again after square five. Land rent falls rapidly near the center and then slowly as is typical of most cities.

Evaluation

In many ways this model can interestingly be compared with the Echenique et. al., approach. Conceptually, they do rather similar things. Both specify some basic sector exogenously around which a city is generated. Echenique et. al., specify this sector as one that is unrelated to local or residential location while Mills specifies exports. Both models then distribute activities and transportation by maximizing some function subject to some constraints. Both involve the guessing of a large number of mean values and parameters specifically. This model must be fed with

- total export of each good
- input-output matrix coefficients, i.e., input of good q per unit output of good r using production technique/s.
- land standard for transportation at each congestion level.
- capital-land ratio for transportation
- unit cost of exporting each good from squares at each distance
- total cost of shipping goods per unit at each congestion level
- rental rate of peripheral lands
- rental rate of capital.

Mills' model is more simplified by regarding the city as symmetric, thereby reducing the spatial problem to a unidimensional one. His model permits generalizing by allowing the identification of each square uniquely, but computation will then become cumbersome.

The key conceptual difference between the two models is in their objective functions. Entropy maximization merely does some 'most probable' distribution within given constraints, while here costs are minimized. The latter is more appealing because it is easier to understand clearly what is being minimized. One therefore appreciates the reason behind whatever distribution of activities, transportation, etc., is generated. Furthermore, in the policy making context the policy maker understands such a procedure readily. With entropy maximization, at best, the modeller himself has some idea of the generating force in his model. It is easier for a modeller to interact with a policy maker when he also understands what is going on. Parameters in this model also have easy meanings while they are merely some 'distribution parameters' that are calibrated in the entropy maximization approach. As a result, interaction with a policy maker could yield changes in parameters and in the objective function that could easily be incorporated into Mills' model while it would be difficult to do so in the Echenique et al efforts.

A caveat here is necessary. Not intuitively understanding the procedure in a model does not necessarily mean that its results are less reliable than one that is comprehensible. It does, however, mean that it is more difficult to evaluate and that its results demand a greater degree of faith.

The main drawbacks in Mills' model are its monocentric assumption and its unimodal transportation. As has been argued earlier in this paper, such assumptions may have some validity in the U.S. context but almost none in a L.D.C. situation. Further, its production functions are of the fixed coefficient type which do great violence to urban reality. Mills recognizes these defects and suggests the following by way of extensions of the model:

(i) Transportation: introducing a second mode which has a different right of way like a subway is easy since we merely have another set of I.O. coefficients, etc., which can be used to represent the second mode. Buses, rickshaws, pedestrians, etc., would be more complex since the same roadway would be used.

(ii) Scale Economies: Mills suggests that these are possible to include by the use of more integer variables and constraints specifying threshold levels of production. However, he conjectures that such modifications would require identification of each square and the computations required could well exhaust even a large computer.

In its present form the model can be used easily to test policies such as land pricing, transportation pricing, land taxation and land use controls with manageable modifications. In each case, the primary output of the model is the amount and production technique of each good to be produced in each square, the origin and destination of each shipment and the congestion level in each square. This suffices as the profile of a city.

Before concluding it is worth noting that the number of variables and constraints in this model is large despite its conceptual simplicity. It needs a late generation large computer like the IBM 360/91 which is often still not available in LDCs.

In conclusion, we regard such a model as an interesting departure from other standard approaches with some advantages over them. Its drawbacks, however, do not make it a fruitful approach to LDC cities at present.

3.6 Summary

This section has reviewed policy-oriented models which are regarded as representing the main strands of urban modelling. Each of the models reviewed is quite different in approach from the others and none are regarded as successful policy models - particularly for possible uses in LDC cities.

The objective of each of the models is to distribute activities spatially in a city. A typical output of a model would be the allocation of residential and employment location by zones in the city. This would be disaggregated by socio-economic types of households, types of residential structures, types of employment, etc. The main features of each of the models are now summarized.

The Lowry model is the parent of many models that have followed it since its first publication. Its conceptual contribution was to posit some basic sector of activity which is exogenously given and is the driving force behind a city. This activity requires households to supply labor which, in turn, require retail services. The provision of retail services requires more households and the process is iterated until equilibrium is reached. The principle behind the generation of location distributions is the gravity model of interaction. The original Lowry model was partly filled in by hand to produce plausible employment-residential co-distributions. Lowry himself was not too sanguine about its potential usefulness but his contribution must be regarded as the single most important and influential one in non-economic urban modelling. The Lowry model descendants in the U.S. have largely been failures and have been subject to heavy criticism from within the profession, consequently leading to disillusionment.

More innovative developments have been carried out in Britain where A.G. Wilson has given a more respectable theoretical base to the gravity model through the method of entropy maximization. Echenique and his associates have developed models based on Lowry and entropy maximization and attempted to apply them to South American cities. While these models have reproduced city structures fairly faithfully after careful calibration their value as planning or predictive tools remains suspect. They have scant behavioral underpinnings which makes one sceptical about their use as predictive devices. Moreover, they are heavily data intensive which makes their applications in LDCs limited.

The NBER model can be regarded as the economists' answer to urban planners. This model has a behavioral structure based on micro-economic theory and borrows much from the analytical models of Muth and Alonso discussed in the last section. It makes them more realistic by dropping the mono-centric assumption; by taking account of the existence of disequilibrium in markets, the effects of the durability of capital structures and of externalities like neighborhood effects. Despite these pleasing features it can scarcely be called an operational model. It is unwieldy to operate, it is not evaluative and there are some indications that it has developed into a research process rather than a product. Its data requirements are even more intensive than the Echenique et. al., models. However, much can be learned from the formulation of its sub-models.

The last model discussed - Mills' programming model - differed from the others in that it is an optimizing model and has much more simplifying assumptions. It regards the city as mono-centric and allows a unimodal form of transport. Despite its simple and unrealistic assumptions it becomes mathematically complex and large. It has 788 non-slack variables of which 30 are integer valued and has 219 constraints. Its solution requires non-linear programming techniques and a late generation computer.

Large models of urban areas are expensive in terms of data as well as technical expertise, particularly for LDCs, while their use is of dubious value. The essential conclusion of this section is a somewhat paradoxical one. While large scale urban models may be of some use as research methods exploring urban form and structure and their underlying rationale they are of little use as practical policy devices. This is more so in the context of LDCs where change is so fast that models are soon rendered obsolete for policy purposes. The next section suggests some fruitful approaches which are more modest in scope though not without their own limitations.

IV. SOME FRUITFUL APPROACHES TO URBAN MODELLING

Finally, in this section we review some work whose approaches though quite different in each case seem to give some hope for useful urban modelling. It should be regarded as an illustrative section merely suggesting **some** fruitful approaches.

It must be emphasized here that these models are not considered directly transferable to LDC cities. They are too rooted in their particular institutional settings for such transferability, as indeed they should be. We start with a description of some of the work of the Transportation and Location Analysis Group in the Master Planning Commission of Stockholm as reported in Lundquist (1973a and 1973b) and Andersson (n.d.). Their model is an evaluative one and is particularly suited to the Swedish environment where the government has considerable control over urban form. The second model discussed is the Urban Institute Housing model as described in various publications of that Institute. This model reflects the U.S. institutional system focusing on the urban housing market. It was originally designed to help the Department of Housing and Urban Development in testing policies such as the housing subsidy scheme. It also is therefore highly policy oriented and is quite different in structure from other models discussed.

Lastly, we discuss an explanatory model which seeks to find behavioral relationships which explain the residential pattern in a city in the British context. It is presented here to illustrate how the development of policy-oriented and explanatory models should go hand in hand to increase our understanding of cities. The relationships and parameters found in an Apps type model can then be fed into policy-oriented models as inputs. Such a model is also an example of how a data gathering effort could usefully be organized systematically.

All these models are described in semi-technical detail for ease of understanding.

4.1 The Andersson-Lundquist Stockholm Model

Their approach is a realistic one concerned with evaluating policy alternatives for the city of Stockholm. The main policy variables are the provision of transportation and residential housing. The urban problem is therefore viewed as one that can be analyzed in terms of dynamic interdependent investment planning. Scarce resources are to be allocated to physical investments in order to provide maximum social and economic welfare. The high degree of complexity in urban structure is essentially caused by strong interdependencies over space and time. Thus the problem can be made manageable only if it is decomposed hierarchically. In other words, we should first divide a city into a fairly small number of zones, study the interactions between them, derive approximate plans at this level of aggregation and only then disaggregate further. The reason for this approach lies in technical realism. Inter-dependencies that are believed to exist in the urban system can only be modelled with non-linear mathematics and these are manageable only if the problem size is kept small. We can thus have a mathematically complex model at an aggregate level and then have sequential transition to disaggregated linear models which can be of larger size.

The authors of this model see the crucial concerns of urban modelling to be:

a. Planning under uncertainty.
b. Normative welfare criteria.
c. Individual behaviour.
d. Spatial and sector disaggregation.
e. Explicit treatment of interdependencies.

Uncertainty about the future is a characteristic feature in the planning of LDC cities. This is currently caused by their explosive growth rates as well as political instability in many places. The latter is particularly important in the formulation of normative welfare criteria that are to be used in an optimizing model. If one is interested in the implementation of planning ideas, the welfare criteria must reflect those of the political policy makers. At the same time, since these can be expected to be myopic, (e.g., because of election cycles), if the modellers' interest is in longer term welfare, the results obtained from the use of different welfare criteria should be robust. In this way, uncertainty about the future is also better taken care of.

In specifying the model the city is seen to be composed of:

a. Building stock;
b. Transportation systems;
c. Recreation land.

The interplay between transportation network structure and urban location pattern is seen as fundamental. Activities are to be allocated within a slowly changing building stock. The modelling is done at 3 levels:

a. First, interdependencies between building stock capacity and communication network design are explored. This is Model 1.

b. Second, interdependencies between activity location and use of transportation network are specified. This is model 2.
c. Finally, the results of a and b are further disaggregated.

In model 1, the welfare function is defined in terms of indices of:

a. Accessibility: good opportunity for interaction that is promoted by a compact urban structure.
b. Environment: space standards measured by level of pollution, degree of segregation, population density, etc.
c. Costs of investment resources, operating activities, etc.

The Stockholm model has N disjoint zones and if we have

B_i - amount of building stock in Zone i

t_k - (0,1) variable denoting absence or existence of transportation network link we can formulate costs of interaction as a measure of accessibility:

$$I = \sum_{i,j}^{N} (B_i d_{ij}(\bar{t}) B_j)/\bar{B}^2 \qquad (102)$$

where d_{ij} is distance between zone i and j and $\bar{B} = \sum_i B_i$, total building stock being given.

Similarly, we can formulate costs of congestion:

$$C = \sum_{i}^{N} \left\{ \left(\frac{B_i}{A_i} \right) \cdot B_1 \right\} / \bar{B} \qquad (103)$$

where A_i is the area of Zone i; and $\frac{B_i}{A_i}$ is thus a measure of congestion.

The problem is then to minimize some weighted sum of these two costs. We have a trade-off between the ease of interaction and the unease of congestion.

We minimize:

$$W = \alpha I + \beta C \qquad (104)$$

subject to

$$\alpha + \beta = 1 \qquad (104.1)$$

$$\sum B_i = \bar{B} \qquad (104.2)$$

$$\text{Capital required for } \bar{B} \leq \bar{K} \qquad (104.3)$$

$$\text{Labor required for } \bar{B} \leq \bar{L} \qquad (104.4)$$

$$B_i = B_i^0 + \Delta B_i \qquad (104.5)$$

and

$$\Delta B \geq 0 \qquad (104.6)$$

Here α and β are the weights assigned to each of the costs. These weights can be varied to test the effects of different degrees of importance attached to environmental concerns and accessibility concerns. The first constraint merely ensures that the weights sum to unity. The second constraint (104.2) ensures that the amount of building stock in all the zones sums to the given total building stock.

Constraints (3) and (4) (104.3 and 104.4) are capital and labor availability constraints to reflect the local capital and labor markets. The left hand sides really specify production functions or parameters for labor and capital derived from production functions. They specify the technology being used. In this model Cobb-Douglas production functions are used to derive capital (K) and labor (L) coefficients. (104.5) shows that new building stock in each zone is existing building stock B_i^0 plus the additional building ΔB_i planned for the current period. (104.6) ensures that **some** building does take place.

The aggregate building stock demand levels, capital, labor availability magnitudes, etc., are assumed to be obtained from national or regional projections or planning models.

The objective function in this problem is a highly non-linear combinatorial one which cannot be dealt with by conventional optimization techniques. Thus a heuristic tree searching procedure is suggested for finding a solution. The interdependencies between building stock location and transportation network design are considered by a nested procedure where the layout of building stock is guided by its implications for network performance. In simpler terms, a systematic trial-and-error procedure can be used - specially since the transportation network is described by a simple (0,1) system and thus has a finite number of possibilities.

Stockholm was divided into 12 zones for the purposes of this model and 25 transportation projects were considered. Model solutions give level of building stock in each zone and the existence or non-existence of transportation links between zones. The model is flexible in the following ways:

a. Model structure: We can vary the specification of the objective function as well as the constraints to test the robustness of results. In particular, values of α and β can be systematically varied to observe the effects of differing priorities.
b. Exogenous data: Capital, labor coefficients can be varied to represent different technologies. Given desired stocks can be varied as well as growth rate projections.
c. Heuristic procedures: The solution procedure can be varied to test for sensitivity of results.

These are precisely the kinds of decision variables that one is interested in exploring in the context of LDC cities. The objective function can easily be modified to reflect concerns of such cities: investment and maintenance costs, for example, would probably be more important than certain kinds of environmental problems. The most appealing part of this (general) model is its low requirement of data. We need to divide a city into only a few zones. As is evident from (104) the kind of parameters and data needed are at such aggregated levels that it should be easy to estimate or guess for LDC cities. Furthermore, we can test the robustness of results by varying the data.

Lest this sound too optimistic we have to recall that the model is highly complex mathematically and that the solution is not straightforward. High levels of skill are therefore needed which are not often available in developed countries, let alone the less developed ones. In addition, the input requirements, and the objective function formulation is such that the modeller be intimately acquainted with the city being modelled since many judgments are involved.

We describe model 2 with even less rigour. The information of interest is posited as a vector:

$$X^T = (L^T, R^T, P^T) \qquad (105)$$

where L^T is zonal supply of land;
R^T is residential activity; and
P^T denotes production activities.

The objective function is of the form

$$\alpha X^T DX + (1-\alpha) X^T YX$$

where D is an interaction matrix;
Y is a congestion matrix; and
α is a trade-off parameter.

The constraints include:

a. Exogenously determined growth rates of production levels.
b. Limited availability of labor and capital as before.
c. Balancing of supply and demand of land.
d. Depreciation rates.
e. Activity relocation.
f. Balanced growth of residential and productive activity. This is achieved by a set of equations relating the labor market to the housing market.

The model has a quadratic programming structure with a quadratic objective function and linear constraints. Global optimum cannot be guaranteed because of the existence of non-convexities and various initial solutions have to be tried.

The model was tried for the Stockholm region using 7 zones. All economic activity was grouped into 3 production activities: manufacturing industry, retail service, public services. The data and parameters used were obtained in various ad hoc ways; e.g., from

a. Regional projections: production levels, total labor force, spatial requirements.
b. National data: technological coefficients
c. Fictive values: friction of distance, interaction intensities and regional variation of technology.

The model tells us the location of manufacturing and residential activity given the overall constraints and objective function.

Once again, as with model 1,we observe that the model appears to be of manageable size without a high degree of information requirements. We do not describe its structure as a model to be followed but as a realistic approach worth learning from. The delineation of activities and the specification of the objective function can easily be modified to reflect LDC cities.

Model 3 does not seem to have been developed and tested yet. Since it is at a highly disaggregated level it will probably suffer from the normal information and other difficulties discussed in Section 3 in the context of other policy-oriented models.

In a planning context, models such as model 1 and 2 would appear to be particularly useful since they evaluate possible courses of action in a way that can easily be discussed with a policy maker. Moreover, the policy-maker's objective function can be probed by varying the coefficients in the objective function. It could also be useful in informing policy makers what their objective functions are when these are unarticulated but implicit from appreciation of particular sets of results. These models are also at a level of aggregation at which policy decisions are often made, e.g., should a bridge be constructed or shouldn't it. Once such courses of action in terms of transportation links and levels of desired building stock by zones have been suggested, further disaggregation can either be left to normal market processes or to a next disaggregated level model. The latter requires a deeper knowledge of people's preferences with regard to housing and location: one method of investigating these is given in the Apps model in Part 3 of this section.

4.2 The Urban Institute Housing Model[1]

1. This section reviews the work described in De Leeuw (1972), De Leeuw and Struyk (1975) and De Leeuw et. al. (1974).

The Urban Institute Housing Model is presented here to illustrate yet another approach to modelling and one which augurs well for the future and for applicability to LDC cities. The model represents the population of an urban area by a relatively small number (30 to 50) of 'model' households, the housing stock by a relatively small number of 'model' dwellings, and market behavior by the interaction of these model households and dwellings over a ten year interval. Households choose among all available dwellings, taking into account their income and family type, dwelling quality and price, and neighborhood characteristics. Owners of existing dwellings decide how much to upgrade or depreciate their dwelling so as to maximize expected profits. Builders are prepared to construct new dwellings demanded at a price which covers costs and a normal profit rate. Governments can affect market outcomes in a wide variety of ways ranging from complex subsidy formulae to rent controls to outright prohibitions of certain dwellings in certain locations.

The model has been applied to six different metropolitan areas to obtain key parameters of household and landlord behavior. It is recognized that to convey a credible description of urban housing markets the model must account for a host of details and characteristics of these markets which cannot be captured in a simple way. For example, it must be recognized that there exist distinctions between new stock and existing stock as do the distinctions among housing submarkets of different qualities and locations. Particular attention is devoted to two characteristics of housing which make it different from other goods: durability and neighborhood effects.

The 'model' dwellings resemble the actual housing stock at the start of the decade by location and level of housing services. 'Model' households resemble the actual population of a metropolitan area in income distribution and in demographic and racial composition. The model, in effect, is a description of the process by which households and dwellings get matched within an urban area. The other agents in the model are the building industry which supplies the new dwellings, and the government which can regulate the housing market in a variety of ways. Thus there are four economic agents participating in the model:

- households deciding which dwellings to select
- the owners of existing dwellings deciding what quantities to supply at what cost
- the builders constructing new dwellings

- government constraining or facilitating outcomes in various ways.

The time-frame of the model is a decade i.e. the model predicts the situation of end-of-decade households and dwellings **given** the housing stock at the start of the decade and **given** travel times, construction costs and certain other information during the decade.

A solution to the model is a set of locations, quantities and prices in which no one has any incentive to change.

The model is firmly based on past theoretical work which it extends by bringing various strands together into a comprehensive housing model. The authors of the model identify three strands of past work as of particular relevance:

(a) **Alonso (1964) and Muth (1968)** and their followers who stress the interrelation between transportation costs and housing demand and the implications for rent gradients and location patterns.

(b) ‘**Filtering’:** They have drawn on a second strand of the literature which deals with quality differences among existing dwellings, their changes, and the occupancy patterns accompanying them.

(c) **Neighborhood Effects:** The development of the literature on the effects of neighborhood racial composition and wealth on household choices and household prices has been influential in the inclusion of zonal wealth and racial composition in household utility functions in this model.

Thus the construction of this model admirably shows the link between explanatory models of the simple type and policy-oriented models which typically tend to be rather more complex. It is distinguished from other such models by the following characteristics:

(a) The scope of the ideas incorporated into the model, ranging from accessibility-housing cost trade-offs to constraints on market outcomes imposed by the existing stock to the influence of neighborhood, race, and wealth on household choice;
(b) The relatively small size of the entire model compared to other simulation models;
(c) The empirical application of exactly the same theoretical model to six areas; and
(d) The range of housing, income, and land use policies which can easily be analyzed within the framework of the model.

The Model

We will now present a semi-technical description of the model structure. Space does not permit a full description while a non-technical description would not be useful here. We pay particular attention to the meanings of particular parameters used.

The model is driven by two kinds of behavior: utility maximization by households and profit maximization by owners of existing dwellings:

(i) *Household Utility Functions:* Each household maximizes utility:

$$U_{ij} = H_{ij} \quad X_{ij} \; Z_1 \quad Z_2 \quad Z_3 \text{ ....} \qquad (106)$$

where U_{ij} is the utility of the j^{th} dwelling to the i^{th} household
H refers to housing services
X to other goods and services
Z_1, Z_2, Z_3 are neighborhood or zone characteristics.

Now

$$H_{ij} = [Q_j - \gamma_1\alpha_i M_i/P_n]^{\lambda_i} \qquad (107)$$

$$X_{ij} = [(M_i - Q_jP_j) - \gamma_1(1 - \lambda_i)M_i]^{1-\lambda_i} \qquad (108)$$

where

- Q_j is quantity of housing services offered by dwelling j
- γ_1 is a parameter affecting the degree to which households will alter their housing choice in response to a price discount,
- α_i is a parameter expressing the strength of housing preferences versus preference for other goods for households of type i,
- M_i is household i's 'model' income (after adjustments for taxes and transfers)
- P_n is the price per unit of service of newly constructed dwellings and
- P_j is the price per unit of dwelling j.

'Model' income

$$M_i = (M_i^a)^{\cdot 6} (\overline{M})^{\cdot 4} \qquad (109)$$

where M_i^a is the actual annual income of household i and $\overline{M}$ is the median income of all households of type i.

Note that if $\gamma_1 = 0$ the utility function reduces to a Cobb-Douglas type utility function. Here other goods are merely represented as income minus housing with the implication that the price of other goods is normalized at 1. $\gamma_1 > 0$ has the effect of posing a minimum level of housing services only an excess over which gives positive utility. These minimum levels vary with income unlike a constant term which would be found in a Stone-Geary utility function. The value of γ_1 is the critical determinant of the degree of substitutability in demand among housing sub-markets. A high γ_1 raises the minimum level of housing and other goods that is acceptable and therefore narrows the options available to a household.

Neighborhood Characteristics

$$Z_1 = (200 - T_j)^{.5 + \alpha_i - \alpha_1} \tag{110}$$

where Z_1 represents accessibility

T_j average travel time in a month

α_i is once again the strength of housing preferences for household type i

and α_1 is the value of α_i for households of type 1, i.e., white non elderly families

200 - T_j is an approximation of the monthly leisure time available to a worker in Zone j. The exponent .5 is based on studies suggesting that people value leisure time at about half their wage rate.

Z_2 represents the average net rent (gross rent Less operating costs) of a zone relative to the average net rent in an SMSA (Standard Metropolitan Statistical Area). More precisely

$$Z_2 = \left[\frac{\{(P_j - P_0)Q_j\}\,\text{zone}}{\{P_j - P_0)Q_j\}\,\text{SMSA}}\right]^{.01\gamma_2} \tag{111}$$

where P_j is the price per unit of housing services

P_0 is minimum operating costs per unit

Q_j is quantity of housing services

The numerator is the zonal average and the denominator is the SMSA average.

λ_2 is a parameter expressing willingness to pay in exchange for living in a wealthy zone.

The higher λ_2 the higher the utility people get by living in a wealthy neighborhood and consequently more they are willing to pay for this privilege.

Finally,

Z_3 is a zonal characteristic capturing the effect of racial composition of neighborhoods on utility.

$$Z_3 = R_{ij} + [1000/(100\gamma_3 + 1)] \tag{112}$$

R_{ij} is the proportion of households in the zone of dwelling j belonging to the same racial group as household i

λ_3 is a parameter expressing the strength of preferences for racial homogeneity.

The larger λ_3 the more sensitive utility is to changes in R_{ij}. For example if $\lambda_3 = 0$

Z_3 can vary from only 1000 to 1001 while if $\lambda_3 = 1$ Z_3 can vary from 100 to 101 - a much larger percentage variation.

Thus the utility function takes account of utility gained by households due to

- quantity of housing services
- quantity of other goods
- accessibility of neighborhood they live in
- average wealth or income of neighborhood and
- racial composition of neighbor-hood

each expressed in appropriate functional form.

(ii) *Existing Dwelling Supply Functions*

The supply curve for existing dwelling j is specified as follows:

$$Q_j = [\beta_1 + \frac{2}{3}\beta_2\left(\frac{P_j - P_0}{P_c}\right)] Q_0 \qquad (112)$$

where

Q_j is the level of housing services currently orovided by dwelling j
Q_0 is the level provided ten years ago
P_j is the price per unit of housing services offered by dwelling j
P_0 is operating costs per unit

and

P_c is capital costs per unit of service for a new dwelling.
All prices are in flow terms, i.e., on a monthly basis.
β and β_2 are parameters to be determined empirically.

This expression is derived from profit maximization subject to a function. The profit maximization procedure is actually production maximization of expected profits. This embodies within it the behavior of the landlord over a long time horizon.

β_1 can be interpreted as some measure of depreciation over a decade.

β_2 is a parameter in the production function and determines the slope of the supply function.

(iii) *Model solution*

In essence, once we are given plausible parameter values
$\gamma_1, \gamma_2, \gamma_3$;

β_1 and β_2

and the α_i s

the model can be used to simulate the housing market.

The key exogenous variables in the model are

P_n the price of new housing. This includes P_c and P_0 the capital cost and operating cost components.

P_0 the operating cost component is very important as it affects strongly the number of removals from the housing stock.

Q_m the minimum quantity standards for new housing.

Y_i the household incomes.

Q_0 the quantity of initial housing services.

T_j the travel time associated with each zone.

The endogenous output of the model comprises: prices and quantities for each existing and new dwellings, numbers of new dwellings and removals, assignments of households to dwellings, zonal averages of prices, quantities, incomes and racial proportions.

Zonal averages of net rents and of racial proportions are fed back into the model and affect household behavior as explained earlier. Ultimately all the endogenous output is determined by exogenous variables, by government regulations and by the behavioral parameters for households and owners.

While the structure of the model is relatively simple and comprehensible its solution is a complex operation and will not be described here.

(iv) *Data Requirements and Inputs:*

One of the features of this model is its great flexibility. It can be run with a range of data quality. On the one hand the estimation of model parameters could be done after the availability of extensive and detailed data. The estimation of each one of the parameters: α_i s, γ_1, γ_2, γ_3, β_1, β_2 needs detailed econometric work and each can comprise a research project by itself. Similarly, the construction of ‘model’ households, household incomes, ‘model’ dwellings, travel times, etc., all need data that is readily available in the U.S. census and Federal Housing Administration (FHA) surveys. Such data are not, in general, so readily available in L.D.C. cities. On the other hand, all the parameters can be guesstimates as well as all the exogenous inputs and the model could still be used for policy purposes. The guesstimates could then be successively improved. For example, rough notions of income distribution could be utilized to construct ‘model’ households; similarly ‘model’ dwellings can be constructed from impressions of the variety of existing dwellings. This can be done primarily because the model is small in size: it is difficult to ‘construct’ plausible data when the model needs thousands of households but it can be done by hand when it only needs 30 - 50. Indeed, the entropy maximization models use similar averages and then generate the distributions in an attempt to be more realistic. Here no attempt is made to squeeze more information from the data than exists and averages remain averages.

(v) *Outputs*

The model is quite clear in identifying which are the parameters, the exogenous variables and the endogenous variables. The parameters have ideally to be estimated and reflect the behavior of households. The exogenous variables are the ones that we play with to simulate government policies. For example, a housing allowance program would be simulated by changing the Y_is, while a construction subsidy would be simulated by changing the P_n - the price of new households. Quantity constraints are subsumed by Q_m. The model then gives us the effect of different variables by forecasting changes in endogenous variables.

(vi) *Evaluation*,

The Urban Institute housing model has been calibrated for six metropolitan areas of the U.S. Durham (N.C.), Washington D.C., Chicago, Portland (Oregan), Pittsburgh and Austin. Except for some instances parameter values for different cities are not very different giving some faith in the belief that they do reflect behavioral relationships. The forecasting abilities of the model have been tested by using parameter values of the other 4 cities on Washington D.C. and Chicago with encouraging results: about 18% error in prediction. The composite results have then been used to simulate urban areas for policy testing. The simulations allow for variations in policies, types of cities and in behavioral parameters.

The suggestion here is not that the Urban Institute Housing Model is ripe for application to an LDC city. Indeed, it is too based on the character of the U.S. housing market for such a transplantation. Its institutional structure is heavily that of the U.S. - particularly that of a relatively free housing market albeit a segmented one. It is, however, an excellent example of an approach which is at once sophisticated and relatively simple as well as fruitful for policy concerns. This model does **not** give any welfare conclusions though its outputs could easily be fed into evaluative sub-models. Lest any wrong impression has been created this is not a simple model to solve. Indeed, a typical run of the model takes about $120 to run with about forty 'model' households. Increasing the number of these households to about 75 would probably quadruple costs. Thus expanding the model is expensive. It has taken about 10 professional man-years over a period of about 3 years to make the model operational at a total cost exceeding $600,000. These man-years have, moreover, been those of some of the best people in the field. Each city application costs about $20 - 40,000 and a period of about 4 months to put into effect. This presumes the existence of a U.S. type data base which can readily be adapted for the model.

The recommendation here is that the approach embodied in this model - the logical structure and its policy applicability - is a good one to follow. It was built up from a base of clear objectives and explicit recognition was made of the theoretical bases used. Such clarity is rarely found in urban models. Its decadal span is also a sensible one in view of the durable aspects of city structure. One must, however, beware of the temptation to transfer the model as it is to an L.D.C. city since it is firmly rooted in the U.S. experience and institutional system. Its methodology is what should be transplanted.

4.3 A Model of Housing Demand at a Disaggregated Level

Paradoxically enough, fruitful policy-oriented models are likely to be at higher degrees of aggregation while explanatory models may have to be disaggregated. Policy-making is time-bound and policies are characteristically made at an aggregated level. People's behavior varies in all kinds of ways with different socio-economic characteristics, family cycle concerns, etc. It is then not really surprising that models which study this behavior are likely to be more detailed than those which prescribe or test policies. Moreover, the results of detailed explanatory models make it easier to feed policy-oriented models with more realistic data, parameters and coefficients which are often otherwise mere guesstimates.

Patricia Apps developed a theoretical framework for a model of housing demand at a disaggregated level and tested it for the city of Reading, England. The structure of the model is based on the economic theory of market demand which states that the individual's consumption patterns are the result of his preferences, income constraints and housing prices. It is also based on the assumption that differences in preferences are caused by social status, household size and stage in a family cycle of the household. The model thus predicts changes in housing demand resulting from changes in factors such as:

a. Number of households.
b. Income.
c. Social status distribution.
d. Size of household or composition.

The analysis is conducted at two levels. At the first level, housing is seen as a single commodity and a set of income -- consumption on Engel curves are derived for each household type. As the next level there is a two dimensional set of Engel Functions; one dimension for different household types and the other for different housing commodities. Working through conventional consumer theory of utility maximizing house-holds with budget constraints the Engel curve is obtained

$$C_h = \pi_h X_h = C_1(M, \pi_1, \ldots \pi_n) \qquad (113)$$

where C_h, the allocation of expenditure on housing (commodity X_h) with price π_h, is specified for household i of type k.

M is household income; and

π_i, i=l,...., n are the relative prices of other commodities.

The form of the function for household type k for different incomes reveals preferences for that household type.

At the next level, utility is a function of housing characteristics considered as a set of commodities. We then have, for household j of type k

$$\begin{aligned} U = {} & U_1(X_{11}, X_{12} \ldots X1m) + U_2(X_{21}, X_{22}, \ldots X_{2m}) \\ & + \ldots + U_h(X_{h1}, X_{h2}, \ldots X_{hm}) + \ldots \\ & U_m(X_{n1}, X_{n2}, \ldots X_{hm}) \end{aligned} \qquad (114)$$

where Ui, i=l to n is the utility provided by the i^{th} group of commodities; U_h is the utility provided by housing with X_{hl} ... X_{hm} being the m housing characteristics.

We then obtain the second level of housing consumption income Engel curves

$$C_h = \sum_{f=1}^{m} \pi_{hf} X_{hf} = \sum_{f} C_{hf}$$

$$C_{hf} = C_{hf}(C_h, \pi_{h1}, \pi_{h2}, \ldots \pi_{hm}) \qquad (115)$$

where π_{hf} is the price of housing commodity f;

X_{hf} is the amount of housing commodity f consumed by household j of type k; and

C_{hf} is the expenditure on housing commodity f now expressed as a function of total housing expenditure and the relative prices of different housing commodities.

Given data on:

a. Total housing expenditures

b. Characteristics of housing

e.g. i) location variables including employment accessibility; shopping accessibility; quality of neighborhood schools;

ii) dwelling characteristics such as age of structure, type of structure and condition of structure; and

iii) space of dwelling.

Equation (115) can be estimated by multiple regression methods to obtain π_{hi} s, i.e., the implicit prices of different housing services.

Finally, having obtained these prices, demand functions for each of these characteristics can be estimated using household types as the independent variables.

This procedure therefore reveals household consumption patterns for different housing characteristics subject to income and price constraints. A by-product of this model is that the system is easier to handle since the number of housing characteristics is much smaller compared to housing types. This is particularly true of LDCs where the variety of housing types is even greater -- from straw shacks to tall apartment buildings.

Apps was quite successful in obtaining the housing characteristic prices from careful regression analysis but was not able to obtain good demand functions in the next stage. We have the normal problem of linear assumptions; housing characteristics are assumed to be non-interacting. Apps speculates that deficiencies in the data may be causing her problems.

As should be evident from the description of this model, it is extremely data intensive. Apps used a good 10 percent sample survey of Reading and a host of other data sources to build all the indices of housing characteristics. The appeal of this procedure is that it reveals the preferences of consumers. It is this kind of knowledge that is necessary for the disaggregation of results obtained from Lundquist-Andersson type models. In particular, some appreciation of the differences in behaviour of different income groups is obtained. This information is necessary for robust planning so that plans do not get subverted by disgruntled groups whose preferences have been grossly violated -- as has happened often in relocation of low income groups in various cities. However, the question of data availability remains, though consumer household surveys of a detailed nature are being made increasingly now.

Such work is particularly important for LDC cities since, as is suggested earlier, preference of different income groups, ethnic groups, etc., are likely to be more disparate in LDC cities than, say in the

U.S. The results of such models are particularly important since planners and policy makers do not generally have intuitive notions of preferences of people outside their class and income experience. With huge slum and squatter populations, such organized information would be invaluable in LDC cities. Good estimates of such magnitudes as income and price elasticities of demand for housing and transportation are the kind of inputs that both Urban Institute and Stockholm models need for useful operation.

One sector of urban activity that has been ignored here and throughout this paper is the location pattern of industry in cities. This is not accidental. It is because there is little work on the subject even in Western economics. We mention in passing here the work of Bergsman, Greenston and Healy (1972, 1975) which could usefully be applied to LDC cities. They have attempted to redefine the Standard Industrial Classification (S.I.C.) system according to location clusters of different industries. They have used data from the Census of Manufacturing and used techniques of factor analysis to obtain their clusters. Their classification turns out to be quite different from the S.I.C. which is used internationally. It is this kind of work that is needed in LDC cities to test the speculations offered in Section 1. One obtains insights into the nature of scale economies and agglomeration from such work and consequently a better understanding of the structure of cities. Their work does not shed any light on intra-urban industrial location but does offer insights into industrial cluster patterns.

V. MODELLING POOR CITIES: WHAT SHOULD BE DONE

How LDC cities should be modelled is very much dependent on what the objectives of the exercise are. The lessons from the last fifteen years experience of urban modelling in developed countries is more explicit in telling us what not to do than in what should be done. One of the main lessons is that at this stage we should be modest in our expectations of the capabilities of modelling. Consequently, limited objective models are more likely to be operationally useful than comprehensive models. Based on the following perception of the Bank's institutional role, a fruitful strategy towards the modelling of poor cities is now suggested.

The World Bank's main activity consists of lending funds for projects and programs at concessional as well as market interest rates. Thus models of urban structure should help in identification as well as evaluation of urban projects. Identification of projects requires models which predict and/or prescribe the future structure of a city. Evaluation requires cost-benefit analyses and urban models can help in tracing some inter-relations as well as in optimization. To feed such models explanatory or analytical models also are needed to promote basic understanding of the urban process. Behavioral relationships and parameters can then be used in the policy oriented models ultimately needed.

It is the major conclusion of this paper that policy-oriented models should be of the 'sketch-planning' type, i.e., small and manageable while the analytical or explanatory models can be larger and disaggregated. The reasons for this have been documented throughout the paper. Sane concerns are, however, worth emphasizing. A useful policy model needs to have fast turn-around time to make it more responsive to policy needs and issues. It is only then that modellers and decision-makers can interact and test for various alternatives. Policy makers have had no involvement in models developed in the U.S. It is then not surprising that these models have not been used for policy purposes. Thus if we are now serious about developing policy-oriented models this major shortcoming must be remedied.

The Stockholm model and the Urban Institute Housing Model are two examples of approaches that meet some of these requirements. They are both small and manageable though mathematically complex. The Stockholm model attempts to have a structure conducive to planner/modeller interaction. It is therefore a step in the right direction. It was also developed with fairly specific objectives. Its policy applications were in helping gross decisions like building or not building a bridge. That is a realistic approach in the sense that policy decisions are usually about some large investments rather than about detailed residential locations. Policy makers are interested in gross impacts rather than in disaggregated allocation. This view of the policy making process is more relevant in countries not having command economies as is the case for almost all LDCs.

The Urban Institute Housing Model is more based on behavioral relationships but it utilizes the idea of telescoping a large system into a small, more manageable and comprehensible one. Its data requirements are flexible in that it can be operated on guesstimates as well as on Thus, in LDCs, where speed in decision making needs to be accurate data. encouraged and where data is scarce, this type of a model can be usefully utilized in an incremental fashion. Systematic policy planning can be begun with rough data and the model successively refined as better and better data is available. One other aspect of the Urban Institute model deserves comment. It seeks to be a general behavioral model so that the same structure can be calibrated for different cities. It is not self-evident that such an approach can work across continents and cultures. This needs more research on the explanatory model side and should be conducted concurrently. One's presumption here is that within the U.S. where the institutional structure is essentially common across cities this is not an unjustified approach. If a model is to be useful it is important that attention be given to the particular institutional structure of the country concerned. Thus adequate account must be taken of the larger public sectors in LDCs and of the consequent constraints on the private market. Similarly, the higher segmentation of LDC markets has to be recognized. Local nationals can be expected to be more conversant with such institutional differences than expatriates. Thus it is suggested that local participation in modelling is of the utmost importance. This should aid in interaction with policy makers as well.

Urban models should be seen as a process rather than as products. The use, for example, of an Urban Institute type model which can involve successive improvements in data as well as model is clearly a process. This process is educational for modellers as well as policy-makers. If seen in this way modelling will suffer less from short time horizons. This makes it even more important that local nationals be involved in modelling efforts rather than quick kibitzing expatriates.

It is mentioned above that policy-models require inputs from explanatory or analytic models. These inputs are both data oriented as well as of behavioral relations. Continuing research aimed at achieving greater understanding of existing patterns as well as of future changes is necessary for the development of good policy models. Such research might help in appreciating what cities are for and why they exist. The Apps model which is very disaggregated has been suggested as one fruitful approach to understand the behavioral relations implicit in housing. In such efforts it is almost inevitable that extremely disaggregated data is needed since one is interested in the whys and wherefores of behavior. We can afford to do this in this context because the usefulness of such modelling is not time-bound. Specifically, the kind of information severely lacking in LDCs is of the following nature.

A. **Transportation**

Some ideas on transportation in LDC cities were discussed in the introduction. Continuing that discussion one would like detailed information on the following:

(i) Analysis of Movements:

The hustle-bustle characteristic of an LDC city has to be systematized. Information is needed on the origin destination pattern, traffic variability at different times and, as a result, the importance of the journey to work.

(ii) Costs of Different Modes:

We need detailed information on operating (and fixed) costs of different modes - walking, bicycles, automobiles, buses, etc. In addition, the income classes which use these modes have to be determined in order to calculate time costs. Using some notion of efficiency, is the prevailing pattern efficient?

(iii) Congestion:

Movement in central areas of LDC cities is painfully slow. To what extent is this relevant as a transport cost?

Such information is crucial to the understanding of location patterns in LDC cities. Various conjectures concerning the movement of people were offered in the introduction: they can be tested with the help of such information. Mean values, cost parameters and even preference patterns can be derived to be then used as inputs in operational models.

B. Housing

(i) Production:

Considerable information 4-already available on this score. Further quantitative information is needed on self-1elp housing to be able to construct production functions for the housing sector,

(ii) Preference Structures:

This is a rather muddy area of research which needs further theoretical as well as empirical clarification. One has to disentangle the effects of incomes from those of tastes in peoples choice of housing. The Apps kind of work is one approach to doing this. For a more intuitive grasp of housing preferences and patterns as they change with income and time, Homer Hoyt's studies (1959, 1966) of residential neighbourhoods in the U.S. provide good examples.

(iii) Housing-Transport Trade-Offs:

Studies of preferences and income variation with regard to housing will also provide information on the effect of transport costs on location choices. It is clear from existing models that we get different urban location patterns depending on the elasticity of demand for space and the imputed cost of time. More measurement in this area would provide valuable insights in the design of urban economic models.

Housing for the poor has become a much discussed problem. Most LDC cities have a majority of the people living in slums, squatter settlements and shanty towns. Such housing is often regarded by planners as deficient and as a situation to be remedied. Since such housing comprises a significant magnitude and hard information on it is not yet available, such information is crucial to the good design of the housing sector in urban models.

C. Industry and Employment

(i) The Informal Sector:

While more and more information is slowly being made available on the content of the informal sector not much progress is being made on its spatial characteristics. One would, for example, like to know the kind of movements of goods and of people it generates. What degree of flexibility does it have to the urban structure? To what extent is it innovative in overcoming traditional spatial problems of transport costs, etc.?

(ii) Small-Scale Industry:

Bergsman et. al., (1972, 1975) have performed a valuable service in the U. S. by attempting to find clusters of activity which always seem to go together. A similar analysis of patterns in LDCs would provide a great amount of interesting information on the nature of economic activity in these cities. The interconnection between activities within these clusters would provide clues on the reasons for agglomeration.

(iii) Links with the Hinterland:

It has often been suggested that the degree of primacy of large cities in LDCs is higher than elsewhere. This coupled with the vast dual economy literature would give some impression of self-sufficiency in these cities. Since the Lowry type models in particular and others as well often take some basic "export" sector as given and as the real driving force in the model structure, it is important to identify the nature aid extent of linkages that LDC cities have with their immediate hinterland as well as more far off ones.

The informal sector has been mentioned in the introduction. Its existence, it has been suggested, makes many market activities in LDCs different in nature from those in Western cities. Its employment patterns have a large effect on location patterns - both residential as well as employment. Thus information on its spatial characteristics is essential to the design of useful urban models.

Each of these is a substantive research project in itself. Moreover, as outlined in the introduction account has to be taken of the differences between cities within the less developed world. Information gathering must be done within some kind of framework to be most useful. That is merely another word for modelling.

The 'modelling agenda' suggested above is eminently practical. It involves concurrent investment in the accretion of knowledge as well as in policy aids. Some of this research is messy, but necessary if we are to expand our understanding of poor cities in order to articulate sensible policies.

BIBLIOGRAPHY

Adams, F. G., et al., 1968; "Underdeveloped Land Prices during Urbanization: A Micro-empirical Study Over Time." *Review of Economics and Statistics*, 50, May, 248-258.

Alonso, W., 1964a; "The Historic and the Structural Theories of Urban Form: Their Implications for Urban Renewal", *Land Economics,* 40, May, 227-231.

Alonso, W., 1964b; Location and Land Use: Toward a General Theory of Land Rent. Cambridge: Harvard University Press.

Andersson, Åke E., "Integrated Stockholm, Transportation and Location Theory and Policy", Stockholm, Master Planning Commission (Mimeographed).

Apps, Patricia, 1973a; "A Residential Model: 1. Theory", *Environment and Planning,* 5, No. 5, September-October.

Apps, Patricia, 1973b; "A Residential Model: 2. Implicit Prices for Housing Services", *Environment and Planning,* 5, No. 6, November-December.

Apps, Patricia, 1974; "A Residential Model: 3, Demand Equations for Housing Services." *Environment and Planning,* 6, No. 1, January-February.

Arrow, K. J., 1951; "Mathematical Models in the Social Sciences", In *The Policy Sciences,* Daniel Lerner and H. T. Lanvell, eds., Palo Alto, Stanford Press.

Artle, Roland, 1972; "Urbanization and Economic Growth in Venezuela", *Papers and Proceedings of the Regional Science Association* 27.

Bailey, M., Muth, R. F. and Nourse, H. D., 1963; "A Regression Method for Real Estate Price Index Construction", *Journal of the American Statistical Society,* December.

Batty, Michael, 1972; "Recent Developments in Land Use Modelling: A Review of British Research", *Urban Studies,* 9, No. 2, June 1972.

Batty, Michael, 1975a; "In Defense of Urban Modelling", *Journal of Royal Town Planning Institute,* 61 May.

Batty, Michael, 1965b; "Recent Developments in Activity Allocation Models in a British Context," International Conference on Urban Development Models, Cambridge.

Baxter, Richard, "Urband Development Models: An Overview", International Conference on Urban Development Models, Cambridge.

Baumol, W. J., 1967; "The Macroeconomics of Unbalanced Growth: the Anatomy of Urban Crisis", *American Economic Review*, 57, June.

Beckmann, M. J., 1958; "City Hierarchies and the Distribution of City Size", *Economic Development and Cultural Change*, 6.

Beckmann, M. J., 1968; Location Theory, New York, Random House.

Beckmann, M. J., 1969; "On the Distribution of Urban Rent and Residential Density", *Journal of Economic Theory*, 1, June.

Beckmann, M. J., 1971; "Von Tühnen Revisited: A Neoclassical Land Use Model", *Swedish Journal of Economics,* 74, March, 1-8.

Beckmann, M. J., and John C. McPherson, 1970; "City Size Distribution in a Central Place Hierarchy: An Alternative Approach", *Journal of Regional Science,* 10, No. 1.

Bergsman, J., P. Greenston and R. Healey, 1972; "The Agglomeration Process in Urban Growth", *Urban Studies,* 8, No. 3.

Bergsman, J., P. Greenston and R. Healey, 1975; "A Classification of Economic Activities Based on Location Patterns", *Journal of Urban Economics,* 2, No. 1, January.

Berry, B. J. L., 1964; "Cities as Systems within Systems of Cities", *Papers and Proceedings of the Regional Science Association* 13.

Berry, B. J. L., and A. R. Pred, 1961; *Central Place Studies*, Edited by H. G. Barnum, R. Kasparson, and J. Kiuchi. Philadelphia, Bibliography Series 1, Regional Science Institute.

Berry, B. J. L., J. W. Simmons, and R. J. Tennant, 1963; "Urban Population Densities: Structure and Change", *Geographical Review*, 53, July.

Böventer, von E., 1969; "Walter Christaller's Central Places and Peripheral Areas: The Central Place Theory in Retrospect", *Journal of Regional Science*, 9, April.

Bradford, D. F. and H. H. Kelejian, 1973; "An Econometric Model of the Flight to the Suburbs", *Journal of Political Economy,* 81, May-June.

Brewer, Garry, 1973; *Politicians, Bureaucrats and the Consultant,* New York: Basic Books.

Brigham, E. F., 1965; "The Determinants of Residential Land Value", *Land Economics,* 41, No. 4, November.

Brodsky, H., 1970; "Residential Land and Improvement Values in a Central City", *Land Economics,* 46, August.

Brown, H. J., *et al.*, 1972; *Empirical Models of Urban Land Use,* New York: NBER, Columbia University Press.

Bussiere, R. and Snickars, 1970; "Derivation of the Negative Exponential Model by Entropy Maximizing Method", *Environment and Planning,* 2 No. 3.

Chinitz, Benjamin, 1960; *Freight and the Metropolis,* Cambridge: Harvard University Press.

Christaller, W., 1968; *Central Places in Southern Germany,* Translated by C. Baskin, Englewood Cliffs, New Jersey, Prentice Hall.

Clark, Colin, 1951; "Urban Population Densities", *Journal of the Royal Statistical Society,* Series A, 114.

Clark, Colin, 1967; *Population Growth and Land Use,* New York: St. Martins Press.

Clonts, H. A., Sr. 1970; "Influence of Urbanization on Land Values at the Urban Periphery", *Land Economics,* 46, November.

Curran, F. B., and Stegmeier, J. T., 1958; "Traffic Patterns in 50 Cities", *Public Roads - A Journal of Highway Research,* 30, December.

Czamanski, S. "A Model of Urban Growth", *Papers and Proceedings of the Regional Science Association*, 13.

DeLeeuw, F., 1971; "The Demand for Housing: A Review of Cross Secti-Evidence", *Review of Economics and Statistics,* 53, February.

DeLeeuw, F., 1972; *The Distribution of Housing Services,* Washington D.C.: The Urban Institute Working Paper.

DeLeeuw, F., and Ekanem, N. F., 1971; "The Supply of Rental Housing", *American Economic Review,* 61, December.

DeLeeuw, F., and Struyk, Raymond J., 1975; *The Urban Institute Housing Model: Application to Chicago, Illinois,* Washington, D. C. The Urban Institute Working Paper.

DeLeeuw, F., et al., 1974; *The Market Effects of Housing Policies,* Washington, D.C.: The Urban Institute Working Paper.

Department of Transportation, 1973; *A Review of Operational Urban Transportation Models,* Washington, D.C., N.T.I.S. (D.O.T.-T.S.C.-496).

Delson, J. K., 1970; "Correction on the Boundary Conditions in Beckmann's Model on Urban Rent and Residential Density", *Journal of Economic Theory,* 2.

Dixit, A., 1973; "The Optimum Factory Town, "*Bell Journal of Economics and Management Science,* 4 Autumn.

Echenique, Marcial, 1975a; "Urban Development Models: Fifteen Years of Experience", International Conference on Urban Development Models, Cambridge.

Echenique, Marcial, 1975b; "The Application of Urban Models in South America: The Case of Santiago and Caracas", Waterloo (Ontario): Socio-Economic Systems Design in Urban and Regional Planning-A Colloquium. July.

Echenique, Marcial, et. al., 1974; "A Disaggregated Model of Urban Spatial Structure: Theoretical Framework", *Environment and Planning 6*.

Echenique, Marcial, et. al., 1975; "A Disaggregated Model of a Metropolitan Area: Caracas", Waterloo (Ontario): Socio Economic Systems Design in Urban and Regional Planning-- A Colloquium, July.

Fales, R. L., 1971; *Location Theory and the Spatial Structure of the Nineteenth Centur City,* Unpublished Ph.D. dissertation, Northwestern University.

Fales, R. L. and Moses, L. N., 1973; "Thünen, Weber, and the Spatial Structure of Nineteenth Century City", In M. Perlman and B. Chinitz (Editors).

Forbes, J. D., 1972; "Central Place Theory-An Analytical Framework for Retail Structure", *Land Economics,* 48, February.

Gilmore, H., *Transportation and the Growth of Cities,* Glencoe, Illinois: Free Press.

Goldberg, M. A., 1970a; "An Economic Model of Intrametropolitan Industrial Location", *Journal of Regional Science,* 10, No. 1.

Goldberg, M. A., 1970a; "Transportation, Urban Land Values, and Rents: A Synthesis", *Land Economics,* 46, May.

Goldner, W., 1971; "The Lowry Model Heritage", *Journal of the American Institute of Planners,* 37, No. 2, March.

Goldstein, G. S., and Moses, L. N. 1973; "A Survey of Urban Economics", *Journal of Economics Literature,* 11, No. 2, June.

Goldstein, G. S., and Moses, L. N. 1975; "Interdependence and the Location of Economic Activities", *Journal of Urban Economics,* 2, No. 1, January.

Grethers, D. M. and Mieszkowsk, P. 1974; "Determinants of Real Estate Values", *Journal of Urban Economics,* 1, No. 2, April.

Harris, B. (ed)., 1965; "Urban Development Models: New Tools for Planning", *Journal of the American Institute of Planers,* Special Issue, 31, May.

Harris, B. (ed)., 1968; "Quantitative Models of Urban Development: Their Role in Metropolitan Policy Making," in Perloff and Wingo.

Harris, Britton, 1975; "Model Building and Rationality", International Conference on Urban Development Models, Cambridge.

Harris, C.D. and Ullman E. L., 1945; "The Nature of Cities", *The Annals of the American Academy of Political and Social Science,* 242, November.

Henderson, J. V., 1972; Hierarchy Models of City Size: "An Economic Evaluation", *Journal of Regional Science,* 12, No. 3

Herbert, J. P. and Stevens, B. H., 1960; "A Model for the Distribution of Residential Activity in Urban Areas", *Journal of Regional Science,* 2, Fall.

Hoch, I., 1972; "Income and City Size", *Urban Studies* 9 No. 3, October.

Hoover, E. M., 1948; *The Location of Economic Activity,* New York: McGraw Hill.

Hoover, E. M., 1968; "The Evolving Form and Organization of the Metropolis", in *Perloff and Wingo.*

Hoyt, H., 1939; *The Structure and Growth of Residential Neighborhoods in American Cities,* Washington, D.C.: U.S.F.H.A. Government Printing Office.

Hoyt, H., 1966; *Where the Rich and the Poor Live: The Location of Residential Areas Occupied by the Highest and Lowest Income Families in American Cities,* Washington D.C.: Urban Land Institute.

Hutchinson, B.G., 1975; "Applications of Lowry Type Models in the Toronto Region", International Conference on Urban Development Models, Cambridge.

Ingram, G.K., et al., 1972; *The Detroit Prototype of the NBER Urban Simulation Model,* New York, N.Y.: N.B.E.R.

Ingram, G.K., et al., 1971; *The N.B.E.R. Urban Simulation Model,* National Technical Information Service, Washington D.C., Report No. PB-198-554-555.

I.B.R.D., 1975; "The Urban Poverty Task Force", *Urban Poverty in Developing Countries: The Challenge of the Future,* Washington D.C., mimeo.

Isard, W., 1956; *Location and Space Economy: A General Theory Relating to Industrial Location, Market Areas, Land Use Trade and Urban Structure,* New York: The Technology Press of M.I.T.

Jacobs, J., 1970; *The Economy of Cities,* London, Jonathan Cape.

Kain, J.F., 1964; "A Contribution to the Urban Transportation Debate: An Econometric Model of Urban Residential and Travel Behavior", *Review of Economic and Statistics,* 46, February.

Kain, J.F., 1967; "Postwar Metropolitan Development: Housing Preferences and Auto Ownership", *American Economic Review,* 57, May.

Kain, J.F., and Quigley, J.M., 1970; "Measuring the Value of Housing Quality", *Journal of The American Statistical Society,* 65, June.

Kain, J.F., and Ingram, Gregory K., 1974; "The N.B.E.R. Urban Simulation Model as a Theory of Urban Spatial Structure", in A.A. Brown et. al., *Urban and Social Economics in Market Planned Economics,* New York: Praeger.

Kaiser, E.J., 1968; "Locational Decision Factors in a Producer Model of Residential Development", *Land Economics,* 44, August.

Kendrick, David, 1972; "Numerical Models for Urban Planning," *Swedish Journal of Economics,* 74, March.

Keiper, J., et. al., 1961; *Theory and Measurement of Rent.* Philadelphia: Chilton Co.

Kirwan, R.M., and Martin, D.B., 1971; "Some Notes on Housing Market Models for Urban Planning", Environment and Planning, 3, No. 3.

Koopmans, T.C., and Beckmann, M.J. "Assignment Problems and the Location of Economic Activities", *Econometrica,* 25, January.

Kraus, M., 1974; "Land Use in a Circular City", *Journal of Economic Theory,* 8, August.

Lampard, E.E., 1955; "The History of Cities in the Empirically Advanced Areas", *Economic Development and Cultural Change,* 3, January

Lave, L.B., 1970; "Congestion and Urban Location", *Papers and Proceedings of the Regional Science Association,* 25.

Lee, D.B., 1972; *Requiem for Large Scale Models,* Berkeley: Institute for Urban and Regional Studies, University of California, April.

Legey, L., Ripper. M., and Varaiya, P., 1973; "Effects of Congestion on the Shape of City", *Journal of Economic Theory,* 6, April.

Livesey, D.A., 1973; "Optimum City Size: A Minimum Congestion Cost Approach", *Journal of Economic Theory,* 6, April.

Losch, A., 1954; *The Economics of Location,* Translated by W. Waglom, New Haven: Yale University Press.

Lowry, I.S., 1964; *Model of Metropolis,* Santa Monica, Calif.: Rand Corp.

Lowry, I.S., 1965; "A Short Course in Model Design." *Journal of the American Institute of Planners,* 31, May.

Lundquist, L., 1973a; "Transportation Analysis and Activity Location in Land Use Planning --with Application to the Stockhom Region", Conference on Dynamic Allocation in Space. Stockholm.

Lundquist, L., 1973b; "Integrated Location - Transportation Analysis: A Decompositional Approach", *Regional and Urban Economics,* 3, No. 3.

Mackinnon, J. 1974; "Urban General Equilibrium Models and Simplicial Search Algorithms", *Journal of Urban Economics,* 1, No. 2, April.

Mera, K., 1967; "Tradeoff between Aggregate Efficiency and Interregional Equity: A Static Analysis", *Quarterly Journal of Economics,* 81.

Mera, K., 1973; "Regional Production F-unctions and Social Overhead Capital: An Analysis of the Japanese Case", *Regional and Urban Economics,* 3.

Meyer, J.R., Kain, J.F., and Wohl, M., 1965; *The Urban Transportation Problem,* Cambridge, Harvard University Press.

Milgram, G., 1967; *The City Expands: A Study of the Conversion of Land from Rural to Urban Use,* Philadelphia: Institute of Environment Studies, University of Pennsylvania.

Mills, E.S., 1967; "An Aggregative Model of Resource Allocation in Metropolitan Area", *American Economic Review,* 57, May.

Mills, Edwin S., 1974; "Mathematical Models for Urban Planning," in A.A. Brown et.al., *Urban and Social Economics,* New York: Praeger.

Mills, Edwin S., 1975; Planning and Market Processes in Urban Models" in Ronald Grieson *Essays in Honor of W. Vickrey,* forthcoming.

Mills, Edwin S., 1969; "The Value of Urban Land," in Perloff, H.S., *The Quality of the Urban Environment,* Washington D.C., Resources for the Future.

Mills, Edwin S., 1970a; "Urban Density Functions", *Urban Studies,* 7, February.

Mills, Edwin S., 1970b; "The Efficiency of a Spatial Competition", *Papers and Proceedings of the Regional Science Association,* 25.

Mills, Edwin S., 1972a; "Welfare Aspects of National Policy Toward City Sizes", *Urban Studies,* 9, No. 1, February.

Mills, Edwin S., 1972b; "Markets and Efficient Resource Allocation in Urban Areas", *Swedish Journal of Economics,* 74, March.

Mills, Edwin S., 1972c; *Urban Economics*, Glenview, Illinois: Scott, Foreman and Co.

Mills, Edwin S., 1972d; *Studies in the Structure of the Urban Economy*, Baltimore: Johns, Hopkins Press.

Mills, Edwin S., 1974; *Mathematical Models for Urban Planning,* Princeton, N.J.: Princeton University, Mimeographed.

Mills, Edwin S., and deFerranti, D.M., 1971; "Market Choices and Optimum City Size", *American Economic Review,* 61, May.

Mills, Edwin S., and Mackinnon J., 1973; "Notes on the New Urban Economics", *Bell Journal of Economics and Management Science,* 4, Autumn.

Mirrlees, J.A. "The Optimum Town", *Swedish Journal of Economics*, 74, March.

Mohan, Rakesh, 1975; "Urban Land Policy, Income Distribution and the Urban Poor", Princeton University: Research Program in Economic Development, Discussion Paper 51, January, (forthcoming: Brookings volume).

Montesano, A., 1972; "A Restatement of Beckmann's Model in the Distribution of Urban Rent and Residential Density", *Journal of Economic Theory,* 4, April.

Moses, L.N., and Williamson, H.F. Jr., 1963; "Value of Time, Choice of Models and the Subsidy Issue in Transportation", *Journal of Political Economy,* 71, June.

Moses, L.N., and Williamson, H.F. Jr., 1967; "The Location of Economic Activity in Cities", *American Economic Review,* 57, May.

Musgrave, J. C., 1969; "The Measurement of Price Changes in Construction", *Journal of the American Statistical Society,* 64, June.

Muth, R.F., 1961a; "Economic Change and Rural Urban Land Conversions, Econometrica, 29, January.

Muth, R.F., 1961b; "The Spatial Structure of the Housing Market", *Papers and Proceedings of the Regional Science Association,* 7.

Muth, R.F., 1969; *Cities and Housing,* Chicago: University of Chicago Press.

Muth, R.F., 1971; "The Derived Demand for Urban Residential Land", *Urban Studies,* 8, October.

Nathaniel Lichfield and Partners, 1975; Urban Growth Simulation Model, London: mimeo, July.

Niedercorn, J. H., and Bechdolt, R. V., Jr., 1969; "An Economic Derivation of the 'Gravity Law' of Spatial Interaction", *Journal of Regional Science,* 9, August.

Niedercorn, J. H., and Bechdolt, R. V., Jr., 1971; "A Negative Exponential Model of Urban Land Use Densities and its Implications for Metropolitan Development", *Journal of Regional Science,* 11, December.

Oates, W. E., 1969; "The Effect of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis", *Journal of Political Economy,* 69, November-December.

Oi, W., and Shuldiner, P., 1962; *An Analysis of Urban Travel Demands,* Evanston, Illinois: Transportation Center, Northwestern University.

Oron, Y., Pines, D., and Sheshinski, E., 1973; "Optimum U. S. Equilibrium Land Use Pattern and Congestion Toll", *Bell Journal of Economics and Management Science,* 4, Autumn.

Pack, Janet Rothenberg, 1974; *Use of Urban Model in Urban Policy Making,* Philadelphia: Government Study Center, University of Pennsylvania, January.

Papageorgiou, G. J., 1971; "The Population Density and Rent Distribution Models within a Multicentre Framework", *Environment and Planning,* 3, No. 3.

Papageorgiou, G. J. and Cassetti, E., 1971; "Spatial Equilibrium Residential Land Value in a Multi-center Setting", *Journal of Regional Science,* 11, December.

Perlman, M. and Chinitz, B. (eds.), 1973; *Spatial, Regional and Population Economics: Essays in Honour of Edgar Hoover,* London: Gordon and Breach Science Publishers.

Perloff, H. S. (ed.), 1969; *The Quality of the Urban Environment,* Washington, D. C.: Resources for the Future.

Pines, D., 1970; "The Exponential Density Function: A Comment", *Journal of Regional Science,* 10, No. 1.

Pred, A. R., 1966; *The Spatial Dynamics of U. S. Urban-Industrial Growth, 1800-1914,* Cambridge: Massachusetts Institute of Technology Press.

Rancich, M.T., 1970; "Land Value Changes in an Area Undergoing Urbanization", *Land Economics,* 46, February.

Richardson, H. W., 1969; *Regional Economics: Location Theory, Urban Structure and Regional Change,* New York: Praeger.

Richardson, H. W., 1971; *Urban Economics*, Middle-sex: Penguin Books.

Richardson, H. W., 1972; "Optimality in City Sizes and Urban Policy: A Sceptic's View", *Urban Studies,* 9, No. 1, February.

Richardson, H. W., 1973; "A Comment on Some Uses of Mathematical Models in Urban Economics", *Urban Studies,* 10, No. 2, June.

Riley, J. G., 1973; "'Gammaville': An Optimal Town", *Journal of Economic Theory,* 6.

Riley, J. G., 1973; "Optimal Residential Density and Road Transportation", *Journal of Urban Economics,* 1, No. 2, April.

Rodwin, L .1966; "The Theory of Residential Growth and Structure," in W. L. C. Wheaton, G. Milgram, and M. E. Meyerson (eds.), *Urban Housing,* New York: Free Press.

Sarna, A. C., 1975; *A Disaggregated Land Use Transport Model for Delhi, India,* Waterloo (Ontario): University of Waterloo, (unpublished Ph.D. Thesis).

Scott, A. J., 1969; "Spatial Equilibrium of the Central City," *Journal of Regional Science,* 9, No. 1.

Senior, M. L., 1974; "Approaches to Residential Location Modelling of Urban Economic Models and Some Recent Developments", *Environment and Planning,* 6, No. 4.

Seyfried, W. R., 1970; "Location and the Centrality of Urban Land Values," *Land Economics,* 46, August.

Sibley, D., 1973; "The Density Gradients of Small Shops in Cities", *Environment and Planning,* 5, No. 2, March-April.

Solow, R. M., "Congestion, Density and the Use of Land in Transportation," *Swedish Journal of Economics,* 74, March.

Solow, R. M., 1973a; "Congestion Cost and the Use of Land for Streets", *Bell Journal of Economics and Management Science,* 4, Autumn

Solow, R. M., 1973b; "On Equilibrium Models of Urban Location," in J. M. Parkin (ed.), *Essays in Modern Economics,* London: Longmans.

Solow, R. M. and Vickrey, W. S., 1971; "Land Use in a Long, Narrow City", *Journal of Economic Theory,* 3, December

Stevens, B. H., 1959; "A Note on Inter-Areal Linear Programming for a Metropolitan Area", *Journal of Regional Science,* 1, No. 2, Spring.

Tiebout, C. M., 1961; "Intra-Urban Location Problems: An Evaluation", *American Economic Review,* 51, May.

Turner, Christopher G., 1975; "The Design of Urban Growth Models for Strategic Land-Use Transportation Studies," forthcoming, *Regional Studies,* 9.

Varaiya, P., and Artle, R., 1972; "Locational Implications of Transportation Costs", *Swedish Journal of Economics,* 74, March.

Vernon, R., 1959; *The Changing Economic Function of the Central City,* New York: Committee for Economic Development.

Wendt, P. F., 1957; "Theory of Urban Land Values", *Land Economics,* 33, August, 1957.

Wheaton, W. C., 1974a; *A Bid Rent Approach to Housing Demand,* Cambridge: Working Paper 135, Department of Economics, Massachusetts Institute of Technology, July.

Wheaton, W. C., 1974b; *Income and Urban Residence: An Analysis of Consumer Demand for Location,* Cambridge: Working Paper 137, Department of Economics, Massachusetts Institute of Technology, August.

Wheaton, W. C., 1974c; "Linear Programming and Locational Equilibrium", *Journal of Urban Economics,* 1, No. 3, July.

Wheaton, W. L. C., Milgram, G., and Meyerson, M. E. (eds.) 1966; *Urban Housing,* New York: Free Press.

Wieand, K. Jr., and Muth, R., 1972; "A Note on the Variation of Land Values with Distance from the CBD in St. Louis", *Journal of Regional Science,* 12, No. 3.

Wilson, A. G., 1975; "Urban and Regional Models: Some Organizing Concepts to Describe the State of the Art," International Conference for Urban Development Models, Cambridge.

Wilson, A. G., 1970; *Entropy in Urban and Regional Modeling,* London: Pion Ltd.

Wilson, A. G., 1974; *Urban and Regional Models in Geography and Planning,* London: John Wiley and Sons.

Wingo, L. Jr., 1961a; "An Economic Model of the Utilization of Urban Land for Residential Purposes", *Papers and Proceedings of the Regional Science Association,* 7.

Wingo, L. Jr., 1061b; *Transportation and Urban Land,* Washington, D. C.: Resources for the Future, Inc.

Wright, C., 1971; "Residential Location in 3-Dimensional City", *Journal of Political Economy, 79,* November-December.

Zipf., G. K., 1949; *Human Behaviour and the Principle of Least Effort,* Cambridge: Addison Wesley.

JOURNAL OF INDIAN SCHOOL OF POLITICAL ECONOMY

Index of Vol. XXXIV (2022)

AUTHOR INDEX

SUBJECT INDEX

Banking

Editorial

Industry

Documentation

INDIAN JOURNAL OF AGRICULTURAL ECONOMICS

(Organ of the Indian Society of Agricultural Economics)

Vol. 77	JULY-SEPTEMBER 2022	No. 3

CONTENTS

SUSTAINABLE AGRICULTURAL GROWTH, DIVERSIFICATION AND FARM INCOME

GENDER EQUALITY AND RESILIENT AGRICULTURE

INNOVATIONS, ACCESS TO TECHNOLOGY AND COMPETITIVENESS OF MARKETS

ISSUES IN WETLAND ECOSYSTEMS CONSERVATION AND MANAGEMENT IN THE CONTEXT OF NORTH EAST HILL REGION

RAPPORTEURS' REPORTS

Annual Subscription Rates

Individual Membership Fee	: Rs. 1,000.00; £ 50.00; $ 100.00.
Life Membership Fee	: Rs. 10,000; £ 500.00; $ 1000.00.
Institutional Subscription	: Rs. 3,000.00; £ 75.00; $ 350.00.

Please address correspondence to the Hon. Secretary & Treasurer, The Indian Society of Agricultural Economics, C-104, First Floor, Sadguru Complex I, Near Vagheshwari, Gen. A.K. Vaidya Marg, Goregaon (East), Mumbai - 400 063 (India).

Telephone: 022-28493723; Fax: 091-022-28493724; e-mail: isaeindia1939@gmail.com website: www.isaeindia.org

INDIAN JOURNAL OF AGRICULTURAL ECONOMICS

(Organ of the Indian Society of Agricultural Economics)

Vol. 77 | OCTOBER-DECEMBER 2022 | No. 4

CONTENTS

ARTICLES

Annual Subscription Rates

Individual Membership Fee	: Rs. 1,000.00; £ 50.00; $ 100.00.
Life Membership Fee	: Rs. 10,000; £ 500.00; $ 1000.00.
Institutional Subscription	: Rs. 3,000.00; £ 75.00; $ 350.00.

Please address correspondence to the Hon. Secretary & Treasurer, The Indian Society of Agricultural Economics, C-104, First Floor, Sadguru Complex I, Near Vagheshwari, Gen. A.K. Vaidya Marg, Goregaon (East), Mumbai - 400 063 (India).

Telephone: 022-28493723; Fax: 091-022-28493724; e-mail: isaeindia1939@gmail.com website: www.isaeindia.org

INDIAN SCHOOL OF POLITICAL ECONOMY

OTHER PUBLICATIONS

Marathi Monthly:

ARTHABODH-PATRIKĀ,
(Subscription Rs. 100/- per year)

Books and Booklets:

1. *Bhāratāteel Lok-Sankhyā Wādhichā Prashna: Santati-Niyamanāchee Jaruree Wa Tyāteel Yash,*
   (The Problem of Population Growth in India: The Need for Family Planning and its Success) (in Marathi)
   by Kumudini Dandekar, Pp. 60, Rs. 30/-.

2. *Compulsory Primary Education - Opportunities and Challenges,* Second Edition,
   (in English and Marathi),
   by Jayakumar Anagol, Pp. 54 and 61, Rs. 30/-.

3. *Shodh Ghetay Tay Shikshan,*
   (To Searches is Education), (in Marathi),
   by Ramesh Panse, Pp. 67, Rs. 50/-.

4. *Kartā-Karvitā, Adhunik Mendu-Sanshodhan va Aple Jeevan,*
   by Ramesh Panse, Rajyashree Kshirsagar, Anita Belhekar-Deshmukh,
   Pp. 242, Rs. 350/-.

5. *Governance, Supervision and Market Discipline: Lessons from Enron*
   by Jayanth R. Varma, Pp. 87, Rs. 100/-.
   Reprint from (*Journal of Indian School of Political Economy,*
   Vol. XIV, No. 4, October-December 2002)

6. *Poverty in India,* (Reprint 2008), by V.M. Dandekar and N. Rath,
   Reprinted by Books for Change Bangalore, Pp. vi+140, Price Rs 200/-

INSTRUCTIONS FOR AUTHORS

Please follow the instructions meticulously. It will greatly expedite the editorial process.

SUBMISSION

All manuscripts should have been proof-read before submission. Send (1+2 copies), preferably one ribbon copy and two xeroxes, to the Editor. Mimeographed copies are acceptable if clearly legible. With the manuscript, include a cover letter identifying the author with his present or past position, address and telephone number. Mention any special circumstance concerning the paper, such as its earlier presentation at a meeting or a conference. We will assume that papers submitted to this Journal are not under consideration elsewhere.

FORMAT

All text, including block quotations, footnotes, and table headings, should be double-spaced and typed on one side. Use medium-weight, opaque, bond paper. All pages should be of the same size, preferably 8-1/2" x 11", and unbound. Leave a minimum left-hand margin of one and a half inches, and a minimum right-hand margin of one inch. Number all pages, including footnotes and/or references, consecutively.

SUMMARY

In every paper, there should be a summary strictly not exceeding 100 words.

TEXTUAL DIVISIONS

If a paper is divided into major sections, the introductory section needs no heading. Subsequent sections should be given titles typed in capital letters and placed at the centre of the page. Do not use roman numerals. If there are sub-sections, the sub-titles should be underlined and placed justified with the left margin.

QUOTATIONS

All quotations should be checked carefully for accuracy, and should be unaltered except for ellipses and bracketed insertions. Block quotations should be double-spaced and set off from the text by indentation.

FOOTNOTES AND REFERENCES

All footnotes and references should be at the end, first footnotes, then references. In the text, footnotes should be numbered consecutively by superscripts following the punctuation. Reference citations in the text should be enclosed in square brackets, as follows: [Author, 1965, p. 9]. References listed at the end should be arranged alphabetically by author and should include the following information: for articles - author, year of publication, title, name of journal (underlined), volume and issue number; and for books - author, year of publication, title (underlined), publisher and place of publication, in the following format. We convert underlining into italics.

Maital, S., 1973; 'Public Goods and Income Distribution', *Econometrica*, Vol. XLI, May, 1973.

Chakravarty, S. 1987; *Development Planning: The Indian Experience*, Clarendon Press, Oxford.

If a Reference is cited in a Note, the Note may use the shortened reference form:

4. For a critique of recent industrial policy proposals, see Marshall [Marshall, 1983, Pp. 281-98].

The full name of any organization or government agency should be spelt out first if subsequent reference is to be by acronym.

MATHEMATICAL AND TABULAR MATERIAL

All tables, graphs, figures, and illustrations should be numbered and amply spaced on separate sheets of paper indicating in a marginal note in the text where such material is to be incorporated. Mathematical equations should be typed on separate lines and numbered consecutively at left margin, using Arabic numerals in parentheses. Use Greek letters only when essential. The word *per cent*, not the symbol %, should be used in the text and the tables.

COMPUTER CD/E-MAIL

If the manuscript is on a computer, send a copy of it on a CD/E-mail covering text, tables, graphs, etc., with details of the software used.

ISSN 0971-0396

Printed and published by S. Sriraman on behalf of the Indian School of Political Economy. Printed at COMP-PRINT KALPANA PVT. LTD., 461/4, Tilak Road, Lane Opp. New English School, Pune 411 030 and published at 'Arthabodh', 968/21-22, Senapati Bapat Road, Pune 411 016. Editor - S. Sriraman.

R.N. 49745/89