RANA HASAN[*]
Asian Development Bank

YI JIANG[†]
Asian Development Bank

DEBOLINA KUNDU[‡]
National Institute of Urban Affairs

Growth of Indian Cities and "Good" Jobs: Evidence from the 2000s[§]

ABSTRACT This paper seeks to understand the drivers of population growth in Indian cities and to explore how these influence the quality of jobs in cities, as proxied by employment in enterprises with 10 or more workers. It finds that various dimensions of economic activity, including the share of young, formal firms, and better connectivity to other locations, are positively associated with the population growth of cities. These factors are also strongly associated with better quality jobs. Together with a preliminary exercise that suggests that more flexible labor and urban land markets are associated with faster employment growth across cities, the authors take the results to emphasize the importance of cities as centers of production. Policies aiming to make the most out of the Indian urbanization process must be designed accordingly.

Keywords: *City Growth, Migration, Human Capital, Market Access, Economic Activity, Economic Census*

JEL Classification: *R11, O14, O18*

[*] *rhasan@adb.org*
[†] *yijiang@adb.org*
[‡] *dkundu@niua.org*

[§] This paper has benefited from comments and suggestions provided by the discussants and participants at the India Policy Forum in New Delhi in July 2017 and by Giles Duranton. The authors thank Vishal More for his extensive support in organizing the Economic Census (EC) data by disaggregated geographic locations across EC rounds and Arvind Pandey, Pris Villanueva, and Radine Michelle Rafols for their excellent assistance in the analysis of population and EC data. Thanks are also due to Rhea Molato and Janine Lazatin. The views expressed in this publication are those of the authors and do not necessarily reflect the views and policies of ADB or its Board of Governors or the governments they represent. ADB does not guarantee the accuracy of the data included in this publication and accepts no responsibility for any consequence of their use.

1. Introduction

Cities are widely believed to be engines of economic growth and good jobs. In this context, the urbanization process under way in India is good news. According to data from the Census of India, the share of India's population residing in urban areas increased from 20 percent in 1971 to 31 percent in 2011, with the urban population of 377 million representing the second largest urban community in the world. All expectations are that the process of urbanization in India will continue, if not accelerate, with various models suggesting that the country might add another 400 million people to its cities by 2050 (United Nations 2014). However, the link between urbanization and economic dynamism is not assured, and a number of urban experts have raised concerns about the nature of urbanization underway in the developing world. For example, Gollin, Jedwab, and Vollrath (2016) bring up the case of two cities, Shanghai and Lagos. Both are large cities in countries with similar urbanization rates. However, it is highly unlikely that their potential to deliver on better economic outcomes for their residents is the same. Similar concerns are raised by Henderson (2014).

In India's case, several commentators have noted that the potential of its cities and towns to spur economic activity may not be met fully due to several factors, including limited investment in infrastructure, the prevalence of unsynchronized spatial and economic planning systems, and suboptimal land-use management (Ahluwalia, Kanbur, and Mohanty 2014; Mathur 2016; McKinsey Global Institute 2010). Thus, while India's cities may continue to grow in terms of population, whether they will fulfill their potential as engines of growth and generators of productive and well-paying jobs—good jobs, for short—remains to be seen.

In this paper, we carry out several exercises to shed light on India's urbanization process by combining information on city (and town) characteristics from three main sources of data: the 2001 and 2011 population censuses, which provide demographic, amenities, and infrastructure-related information; the 1998 and 2013 economic censuses (ECs), which provide information on the structure of economic activity; and the geographic information system (GIS)-enabled road network data, which is used to construct transport connectivity and market access (MA) measures. Our main objective is to understand the determinants of urban success, understood as not simply reflecting city population growth, but also cities' ability to provide better jobs to urban residents. This enables us to say something about the conditions under which India's urbanization will be growth-promoting and will generate more productive and better-paying jobs.

The first exercise we carry out examines the determinants of population growth in Indian cities between 2001 and 2011. Compared to the previous literature, our main contribution is to focus on cities as the unit of analysis and consider a relatively wide range of determinants of city growth, including cities' economic structure and infrastructure, connectivity to other cities and MA, and human capital. Second, we examine the association between these determinants and city-level employment in firms with 10 or more workers, which we take to be a proxy for good jobs. Finally, we carry out a preliminary assessment of whether cities in Indian states with more flexible labor and urban land regulations see faster growth of employment.

We find a positive association between various dimensions of economic activity and population growth of cities, emphasizing the importance of cities as centers of production. We also find connectivity between cities to influence city growth. We further find that the factors that contribute to population growth are also associated with our proxy for good jobs, especially in cities with a population of less than 100,000 (also referred to as towns in official census documents). Finally, employment growth, especially that driven by firms with 10 or more workers, tends to be larger in cities belonging to states with more flexible labor and urban land regulations. While the causal interpretation of our results is tentative at this point of time, we believe that the analysis contained here is useful in building a set of stylized facts about Indian urbanization and in pointing to the need for Indian policymakers to look at the urbanization process underway in India as an unprecedented opportunity for generating higher economic growth and creating good jobs.

The remainder of this paper is organized as follows. Section 2 motivates the need for understanding the drivers of city growth and jobs, and discusses the related literature. Section 3 provides a discussion of our data and the empirical framework we use to analyze various aspects of urban growth. Section 4 describes the results of our analysis while Section 5 concludes the paper.

2. Urban Growth

2.1. Urbanization and Development

Historically, urbanization has played a crucial role in generating economic growth and has been central to the development process. Industrialization has been a crucial link between urbanization and growth, with cities being the locations where factories clustered and prospered. The urbanization

underway in the developing world appears to be different. In particular, while developing a country, cities are expanding and getting denser, deficiencies in infrastructure investments—including in transport connectivity both within cities as well as across other cities and locations—sub-optimal land-management systems, and uncoordinated economic and spatial planning seem to be making it difficult for cities to thrive economically and to address the negative externalities that can arise with growing population and density (Glaeser and Henderson 2017).

Similarly, though many developing countries are becoming highly urbanized, they lack large industrial sectors. Gollin, Jedwab, and Vollrath (2016) find that the historically tight relationship between urbanization and industrialization breaks down for much of the developing world. They draw a distinction between "consumption cities," where a larger fraction of workers are employed in nontradable services, and "production cities," where a larger fraction of workers are engaged in manufacturing or in tradable services. They note that the nature of urbanization may well have implications for the subsequent role of urbanization in driving growth and job opportunities. For example, Shanghai and Lagos are cities in countries with similar urbanization rates, but with probably very different potentials for driving growth.[1]

In a nutshell, while developing country cities may be drawing more and more residents and getting larger, whether they are doing so in a manner that enables cities to play their role as engines of growth, and generators of productive and well-paying jobs is unclear.

Similar concerns have been raised about the urbanization process in India. According to the McKinsey Global Institute (2010), Indian cities have been spending only around $17 per capita on urban infrastructure versus their benchmarked needs, estimated at $100 per capita. Similarly, restrictive regulations on land use and building regulations, including very low floor-area ratios, seem to have made urban land artificially scarce, especially in the most economically dynamic cities (Brueckner and Sridhar 2012; Clarke Annez et al. 2014; Sridhar 2010a). And, notwithstanding the 74th Constitutional Amendment of 1992, aimed at revitalizing and strengthening urban governments, India's urban governance framework continues to be characterized by a lack of empowered city officials who are incentivized to promote rapid local economic development.

1. Gollin, Jedwab, and Vollrath (2016) note that cities can switch from being of one type to another. For example, cities such as San Francisco, Denver, and Houston could be considered "consumption cities" in the past. But over time, they have developed into "production cities".

Finally, the broader policy environment may not be supportive of the development of economically dynamic cities. To the extent that urbanization–industrialization linkages matter for urban success, the policy obstacles to industrial development may lead Indian cities to lack an economic base that fosters dynamism. In the Indian context, rigidities in labor regulations are widely believed to have constrained the growth of Indian manufacturing (Panagariya 2008). Further, it is believed that India has some of the most onerous land-use restrictions in the world, including restrictive floor-area ratios, and complex and time-consuming processes for conversion of land to different uses (Brueckner and Sridhar 2012; Clark Annez et al. 2014). Together, restrictive labor and land-use regulations are particularly likely to be binding constraints on industrial investment in cities as opposed to rural areas, abetted by poor city-level infrastructure and connectivity.

In summary, Indian cities, or at least some subset, may well be consumption cities and not production cities, using the language of Gollin, Jedwab, and Vollrath (2016). They may be absorbing more and more residents, but are not necessarily providing the conditions for the emergence of more dynamic firms and jobs.

2.2. Understanding City Growth: Recent Literature

While there is a large literature that looks at the determinants of urban growth (see Duranton and Puga 2014), the literature on the determinants of city growth in a developing country context is more limited. Understanding these determinants is, however, important for various reasons (Duranton 2016a). First, the infrastructure needs of cities are massive. An understanding of which types of cities are likely to grow most rapidly can inform policymakers as they decide on how to allocate scarce funds for investments in urban infrastructure across urban centers and types of cities. Second, understanding what types of factors are more or less important in driving city growth—along with a sense of how these affect development outcomes at the city level (e.g., wages and job opportunities)—can inform policymakers about what types of investments and interventions are needed. As noted above, the cross-country evidence clearly indicates that there is considerable diversity in the urbanization process underway across the developing world. Cities that are growing due to their role as centers of consumption, with the main engines of economic activity located elsewhere, may need different types of investments and interventions to improve living standards and serve as future engines of growth, as compared to cities that are primarily centers of production.

Perhaps the most comprehensive study on the determinants of city growth in a developing country is that of Duranton (2016b), who assembles a large and comprehensive dataset on Colombian cities. Focusing on city growth from 1993 and 2010, he argues that Colombian cities are best viewed as local labor markets. Cities that offer higher wages grow more, with high wages themselves driven by conditions in local labor markets and, to some extent, a more educated workforce. City roads also matter, most likely because better city-level transport infrastructure benefits commuters going to work.

Another notable study is that by da Mata et al. (2007), who construct a dataset of 123 Brazilian agglomerations to examine the determinants of Brazilian city growth between 1970 and 2000. They find that city growth is positively influenced by reductions in intercity-transport costs, increases in market potential for goods, and better labor force quality. Decreases in rural income opportunities also foster city growth while local crime and violence detract from city growth.

The literature on the determinants of growth of Indian cities is quite sparse, though it is growing. Sridhar (2010b) examines the determinants of urban population growth as well as urban economic output mainly at the district level, which as administrative subdivisions of Indian states usually cover around 5,500 km² of rural and urban area, and, based on the 2011 census, on average have around eleven cities and towns.[2] Sridhar finds that at the district level, a higher proportion of manufacturing to service employment, proximity to large cities, and primary school coverage per population are associated with larger urban populations. As for the growth of district-level urban population, a key finding is that higher temperature differences (a proxy for unfavorable climatic conditions) do not discourage growth. (As for urban output, a higher literacy rate and a greater share of manufacturing employment, relative to services across cities and towns that constitute the urban district, are associated with higher non-agricultural output per capita. This suggests that the nature of economic activity that residents are engaged

2. While urban output is captured by non-agricultural net district domestic product for districts belonging to eight of India's states in 2003–04, the analysis of urban population growth is based on aggregating city- and town-level information from the census town directories for 2001 and 1991 published by the Census of India. The determinants include district-level averages of city/town-level indicators of human capital (e.g., literacy rates), climatic conditions (temperature differences), distance to the nearest large city (with population of 100,000 or more), road length per 1,000 people, and the ratio of manufacturing to services employment. Some of Sridhar's analysis is conducted at the city level, for which a measure of land use controls as captured by the urban land ceiling legislation is also introduced in accounting for the size of cities in 2001 and population growth in cities from 1991 to 2001.

in, and not just population growth, is important to consider when thinking about what constitutes urban success.)

At the city level, Sridhar (2010b) finds that population growth is faster in cities with manufacturing rather than agriculture as their economic base, and where there have been reforms relaxing the strong land use controls associated with the Urban Land Ceiling Regulation Act (ULCRA), consistent with the view that these have detracted from city growth by artificially creating a scarcity of urban land.

Further city-level analysis of how the city shape and the spatial dimensions of land use and land-management-related policies affect city population growth is provided in Sridhar (2010b) and Harari (2016). Sridhar (2010b) finds that ULCRA and building height restrictions imposed through floor-area ratio restrictions limit the population that can be accommodated by a city and encourage sub-urbanization. Harari investigates how city shape affects the location choices of consumers and firms. Consistent with it being a consumption amenity, cities with more compact shapes are found to be characterized by larger populations (as well as lower wages and higher housing rents). Building height restrictions exacerbate the modeled welfare costs of poor city shape, while road infrastructure mitigates it.

2.3. Related Literature: Cities and Economic Dynamism

There are a large number of studies that consider whether and how cities might contribute to economic dynamism, though most focus on developed country experiences (see Duranton 2014 for a detailed review of the literature). One way to study the link between urbanization and growth is to examine the effects of cities on wages and/or productivity. The evidence, mostly from developed countries and based on estimating the elasticity of wages or firm productivity, with respect to city employment or urban density, suggests that larger cities offer workers jobs that allow them to be more productive (Duranton 2014).

In the case of India, Chauvin et al. (2017) and Hasan, Jiang, and Rafols' (2017) examination of the size of agglomeration economies associated with urban locations yields a mixed picture. While the urban district-level analysis of Chauvin et al. yields elasticities of nominal wages to urban density of around 7–8 percent, Hasan et al.'s analysis at the city level, but using a proxy measure of average city-level wages, suggests lower elasticities. Using data from Indian manufacturing, Lall, Shalizi, and Deichmann (2004) find a weak relationship between urban density and firm performance using the district as their locational unit of analysis.

Another rationale for the existence of urban agglomeration (UA), that is, more intense input-output linkages, thicker local labor markets, and deeper and wider knowledge spillovers, is that larger cities may encourage innovative activities and entrepreneurship. Cities' role in fostering entrepreneurship, usually measured by average establishment size and age, may be particularly important in both developed countries and developing ones.[3]

Glaeser et al. (1992) find a strong correlation between growth in a US city and sector over a period of time, and the presence of small establishments in that city and sector at the beginning of the period. Extending Glaeser et al. (1992), Faberman (2011) shows through correlations and ordinary least squares (OLS) regressions that, on average, high-growth cities have relatively younger establishments (less than five years old). Furthermore, Glaeser, Kerr, and Kerr (2015) address causality concerns in estimation and show that entrepreneurship (as defined by start-up employment shares) positively affects city growth. The literature is thinner for developing countries. A study that examines the relationship between entrepreneurship from a spatial dimension for India—albeit at the district level and not at the city level—is that of Ghani, Kerr, and O'Connell (2014), who find entrepreneurship in the services sector (as captured by the level of employment in rural or urban firms less than three years old) to respond more than proportionately to district population. In the case of manufacturing, supportive incumbent industrial structures for input and output markets are strongly linked to higher establishment entry rates.

Infrastructure, especially transport infrastructure, is likely to be an important influence on how incomes evolve within and across cities. Indeed, transport infrastructure is especially important for the evolution of urban systems—that is, groups of cities interconnected economically. Lall, Wang, and Deichmann (2010) examine how both local infrastructure (availability of municipal roads, street lighting, water supply, and drainage) as well as national infrastructure (proximity to international ports and highways) improve city competitiveness, measured as the city's share of national private investment. They find the latter to have a larger effect on a city's attractiveness for private investment. Alder, Roberts, and Tewari (2017) exploit changes in transport infrastructure that have led to reductions in travel time from 1996 to 2011 to examine how changes in MA (i.e., how well

3. The literature for developed countries provides evidence that cities affect the propensity for innovation—research and development activities tend to be more concentrated locationally than production in most industries, and often occur in large metropolitan areas (Audretsch and Feldman 1996). This may also carry over to some developing country cities. See Sridhar (2010b) for a discussion of the role of knowledge-based industries such as IT in driving the growth of Indian cities like Bangalore and Hyderabad.

a location is connected to all other domestic locations) influence urban and rural incomes as captured by nightlights data for 5,900 sub-districts. Their results suggest that improvements in both urban and rural MA should raise incomes. Relatedly, Ghani, Goswami, and Kerr (2016) show that a major upgrade of India's main highways raised productivity and encouraged entry of plants in districts close to the improved highways.

Finally, turning to how urban form might affect economic growth, a key contribution is that of Tewari et al. (2017), though Harari 2016 and Sridhar 2010b are also related. Tewari et al. (2017) examine the relationship between various dimensions of a city's urban form and its subsequent growth for a sample of 479 cities. They find measures of urban sprawl to detract from city economic growth as proxied by nightlights data over 2002–11.

With some notable exceptions, such as the recent work of Harari (2016) and Tewari et al. (2017), most of the studies treat districts as the urban unit of analysis (see, e.g., Chauvin et al. 2017; Ghani, Kanbur, and O'Connell 2013). This is because most of the available data from labor and enterprise surveys contain geo-information down to the district level. However, as noted earlier, Indian districts often cover large rural areas, and many contain multiple, geographically independent urban areas, that is, cities and towns. A more appropriate unit of analysis is thus likely to be the city or town, which is the administrative division below the district level and whose rural counterparts are villages.

In addition, in several of the studies noted above, in exploring the relationship between urbanization and economic activity, the city-level measures of economic activity are captured in a highly aggregated manner (e.g., through the intensity of nightlights as a proxy measure for aggregate economic activity, as in Tewari et al. 2017). Nuances that arise from different structures of economic activity—for example, differences in the relative importance of manufacturing and the role, if any, of formal and large firms—get missed in this approach.

3. Analyzing Urban Growth: Data and Empirical Framework

As noted in the introduction, our interest is to understand the factors that drive not only the growth of cities, but also the cities' ability to provide better jobs. Our empirical analysis consists of three inter-related exercises. First, we examine the determinants of population growth in Indian cities between 2001 and 2011. Compared to the previous literature, our main contribution is to focus on cities as our unit of analysis rather than urban districts as some

previous literature on urbanization has done, and to consider a wider range of determinants of city growth, especially those pertaining to cities' economic structure. Second, we examine the association between these determinants of city population growth and the proportion of "good" jobs as proxied by the share of city-level employment in firms with 10 or more workers (and which Mincerian wage regressions suggest is a good proxy for better wages). Finally, we carry out a preliminary assessment of whether cities in states with more flexible labor and land regulations see faster growth of employment.

3.1. Data

Our main sources of data include the Census of India for 2001 and 2011; establishment-level information from the fourth and sixth ECs carried out in 1998 and 2013, respectively, and GIS-enabled road network data to measure a city's connectivity and MA. We also use data from various rounds of the National Sample Survey Office's Employment–Unemployment Surveys (NSS-EUS), namely the 55th (1999–2000), 64th (2007–08) and 68th (2011–12) rounds. These provide us information on educational attainment, the extent of migration, and correlates of wages, respectively, for the urban parts of districts.

Before getting into details on the data, a few remarks on the population census are in order. First, at least since the 1961 census, the criteria for identifying urban areas have been broadly similar. In addition to statutory towns (i.e., locations with a municipal corporation, municipality, cantonment board, notified town area committee, or town council), urban areas include census towns, that is, places that meet three criteria: (a) a population of at least 5,000; (b) 75 percent or more of the male population (working population in 1961 and 1971) engaged in non-agricultural pursuits;[4] and (c) a population density of at least 400 persons per square kilometer. Census towns are administrative units formally classified as villages but which satisfy the three criteria above simultaneously.

In addition to statutory and census towns, the census also tracks UAs, continuous urban areas consisting of a town and its adjoining "outgrowths" or two or more physically contiguous towns (with or without the outgrowths of such towns).[5] UAs must consist of at least a statutory town, and its total

4. The definition for the 2011 census entailed a slight change. Now, only "main" male workers are included in the criteria; marginal workers are excluded.

5. An outgrowth is an area lying outside the boundary of the town and within the "revenue limit" of a village or group of villages. It should possess various urban features in terms of infrastructure and amenities. Examples of outgrowths include railway colonies, university campuses, and port areas.

population (all constituents combined) should not be less than 20,000. In our analysis, we treat UAs the same as towns.

Second, the terms town and city are often used interchangeably by researchers. We also adopt this practice for ease of exposition. However, in official terminology, cities refer to towns with a population of 100,000 or more. We rely on this distinction for our econometric analysis, and estimate various regression models by both pooling across cities and towns, as well as considering them separately.

Finally, while the concepts of census town, outgrowths, and UAs try to capture the urbanization process taking place outside the administrative boundaries that mark statutory towns, it is possible that official statistics understate the extent of urban growth. One reason is the complex and time-consuming process of redrawing municipal boundaries as cities and towns expand (Colmer 2015). Thus, reliance on census data may understate the true extent and growth of urban areas and, therefore, urban population.[6]

3.1.1. Census of India Data

The Census of India provides "town directories" that include detailed tabulated information not only on demographic variables, but also on geographic features (including land area), climatic amenities, infrastructure provision, social and educational services, and government revenues and expenditures at the city/town (and village) level.

We limited our attention to towns with a population of at least 10,000 in 2001 and used town IDs to merge the town directories of the 2001 and 2011 censuses. For those with inconsistent IDs, we made use of town names, district names and state names to verify and match. We treated each UA as an integrated city, aggregating variables of the constituent towns and areas to reach the final values for the UA. By matching cities and towns across the 2001 and 2011 censuses, we are able to generate both our main dependent variable (growth in a city/town's population) and various town-level characteristics as explanatory variables, such as city amenities, infrastructure, and access to state headquarters, among others.

6. An increasingly popular approach around this problem is to use data on nighttime lights and use these to delineate urban areas. The underlying logic is that urban areas are characterized by a concentration of people, and incomes such that a threshold value of the luminosity of nighttime lights can be used as a proxy for whether a given area of land is urbanized or not (see, e.g., the application of this approach by Tewari et al. 2017). However, the use of nighttime lights in the analysis of urbanization issues is not without its own drawbacks. For example, from the perspective of defining urban areas, scattering of nightlights over a wide geographic area could well lead some rural areas adjacent to cities to be defined as urban. More importantly, it is very difficult to match population and other variables used in urban analysis to such areas.

The population and area of each city or town were carefully examined. The population and data of each sample town and city from the town directory were cross-validated with the so-called A-4 tables of Census of India, 2001 and 2011. These provide the reasons for variations in the population and area of each town/city. The towns in which anomalies have been found in area and population were excluded from the sample. Both declassification and addition of towns were examined to finalize the sample. We also encountered towns/cities with extreme population and/or area growth between 2001 and 2011 in the data. After careful validation of the data using other sources to verify population figures, we restricted our analysis to those towns/cities with a population growth greater than –50 percent and less than 500 percent, and an area growth greater than –10 percent and less than 500 percent.

Our final dataset allows us to work with a maximum of 2,426 Indian cities and towns with a population of at least 10,000 in 2001. These belong to 502 districts across 21 states and 4 union territories that are well integrated to India's major markets from the perspective of transport connectivity.[7] It is important to note that our sample does not capture the urban growth that results from the transformation of a rural area in 2001 into a standalone urban center, such as a census town, in 2011. We also do not capture towns in 2001 that get classified as a new UA in 2011. This is because we have no way to split a new UA into its 2001 constituents.

Nevertheless, our sample should capture key features of urbanization and its growth between 2001 and 2011. The Census of India reports that a majority (86.1 percent) of the cities in India had a population of less than 100,000 in 2001. There were only 73 cities in India, or 2.6 percent of the total, at the other end of the spectrum with a population of 500,000 or more. In our final sample, 85.5 percent of the cities had a population of less than 100,000 in 2001 and 2.8 percent had a population of 500,000 or more. Table 1 presents the distribution of our sample cities by population and compares it with the overall, all-India numbers reported in the census. As may be seen, our sample mimics key features of the urban population and its distribution well.

Table 2 reports some summary statistics for our sample and by different city size groups. The top panel of the table is based on information provided in the population census tables, and in addition to population, also covers

7. The included states and Union Territories are: Andhra Pradesh, Assam, Bihar, Chandigarh, Chhattisgarh, Dadra & Nagar Haveli, Daman & Diu, Delhi, Goa, Gujarat, Haryana, Himachal Pradesh, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Odisha, Puducherry, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, Uttarakhand, and West Bengal.

TABLE 1. Urban Population Growth, 2001–11

Official Numbers					*Our Sample*				
City Type by Population Size	*2001*	*2011*	*Population Growth (%)*	*Contribution to Growth (%)*	*Population Size*	*2001*	*2011*	*Population Growth (%)*	*Contribution to Growth (%)*
Class I: 100,000 and above	200,098,105	264,745,519	32.3	71.1	100,000 and above n = 351	182,313,024	230,215,938	26.3	82.0
Class II: 50,000 to 99,999	27,192,982	32,179,677	18.3	5.5	50,000–99,999 n = 340	23,581,458	27,353,627	16.0	6.5
Class III and IV: 10,000 to 49,999	51,988,323	65,846,155	26.7	15.2	10,000–49,999 n = 1,735	40,455,141	47,168,309	16.6	11.5
Total	286,119,689	377,106,125	31.8	100.0	Total	246,349,623	304,737,874	23.3	100.0

Source: Census of India 2001 and 2011, and authors' estimates.

TABLE 2. Summary Statistics (Means) by Cities and Towns

Variables	*Full Sample*	*Towns*	*Cities*
Number of Cities	*2,426*	*2,075*	*351*
Population, 2001	101,546	30,861	519,410
Population growth 2001–11	0.152	0.141	0.219
Literacy rate, 2001	0.744	0.741	0.778
Electricity connections per 10,000 population, 2001	1,704	1,690	1,785
Pucca road density (kms of road per sq. km of area), 2001	4.006	3.653	6.093
D: Age 0 to 1 (ratio to total population), urban 2001	0.031		
D: Age 0 to 5 (ratio to total population), urban 2001	0.117		
D: Age 0 to 17 ratio to total population), urban 2001	0.390		
D: Internal migration (ratio to total population), urban 2007–08	0.342		
D: Intra-district migration, urban 2007–08	0.169		
D: Inter-district migration, urban 2007–08	0.122		
D: Inter-state migration, urban 2007–08	0.051		
D: Years of schooling, urban 1999–2000	5.925		
Diversity index, mfg, 1998	0.799	0.772	0.956
Diversity index, mfg and srvcs, 1998	1.649	1.587	2.012
Manufacturing employment share, 1998	0.266	0.264	0.277
Share of young firms in mfg and srvcs, 1998	0.396	0.394	0.405
Share of young firms with 10+ workers in mfg, 1998	0.009	0.009	0.014
Share of young firms with 10+ workers in mfg and srvcs, 1998	0.005	0.004	0.007
Employment share of 10+ worker firms, mfg, 1998	0.223	0.205	0.328
Employment share of 100+ worker firms, mfg, 1998	0.076	0.065	0.145
Employment share of 10+ worker firms, mfg and srvcs, 1998	0.224	0.213	0.286
Distance to state highway, 2001 ('000 km)	0.006	0.006	0.004
Distance to expressway or national highway, 2001 ('000 km)	0.013	0.013	0.007
Market access from G-roads data, 2001 (100k persons per kilometer)	0.195	0.194	0.199
S: Land regulation score	61.07		
S: Land regulation dummy (1 = Good Land Management, 0 = Weak)	0.665		
S: Flexible labor regulation (1 = Flexible, 0 = Inflexible)	0.455		

Source: Authors' estimates.

Note: "D" denotes district-level statistics. "S" denotes state-level statistics.

literacy rates and measures of city-level infrastructure provision. Although not shown in the table, the diversity of these characteristics across Indian cities is noteworthy[8] (For a detailed discussion on the distribution of urban infrastructure across Indian cities and towns, see Bhagat 2011a; Kundu, Bagchi, and Kundu 1999; NIUA 2017.)

8. For instance, the 5th percentile literacy rate across cities is around 54 percent while the 95th percentile literacy rate is almost 89 percent.

The second panel of Table 2 provides district-level information on proxies for birth and fertility rates based on age-group data (0–1, 0–5, and 0–17 years) from the 2001 population census, various measures of migration from the NSS-EUS carried out in 2007–08, and average years of schooling from the NSS-EUS carried out in 1999–2000.[9] The data pertain to only the urban parts of districts.[10] The age-group data from the population census shows that on average, the 0–17 years age-group comprises a little more than one-third of urban district population to which a city or town belongs, suggesting that natural growth is an important component of urban growth in India (Bhagat 2011b; Kundu 2007). However, migration also contributes to urban growth. According to Bhagat (2014) and Kundu and Mohanan (forthcoming), the contribution of rural–urban migration in urban growth during 2001–11 was in the range of 15–20 percent. Turning to the migration data from the NSS-EUS of 2007–08, on average one-third of the population in an urban district to which our cities and towns belong is comprised of migrants.[11] Much of the migration was driven by within-district migration and migration across districts within the same state (i.e., intra-district or inter-district migration). Migration across districts (inter-state migration) contributes much less.

3.1.2. Measures of Economic Activity Using the Economic Census

To capture the nature of economic activity at the city level, we use the EC, which is conducted by the Central Statistics Office of the Ministry of Statistics and Programme Implementation. The EC is a countrywide census of establishments engaged in all economic activities except crop production and plantations.[12] The key purpose of the EC is to provide a sampling frame for follow-up sample surveys that collect more detailed sector-specific information on the non-agricultural economy. In this study, we use public-use micro records from the fourth and sixth editions of the EC carried out in 1998 and 2013, respectively (henceforth EC 1998 and EC 2013). The data allows for the geographic location of establishments to be identified at the

9. The average years of schooling is computed by assuming the following correspondence between the level of education survey respondents report and years of education: 0 years for those who are illiterate; 5 years for those with primary education; 8 years for middle school; 10 years for secondary education; 12 years for senior secondary education; and 17 years for college graduates.

10. None of these variables was available at the city level.

11. The NSS-EUS of 2007–08 defines a household member as a migrant if their last "usual place of residence" at any point of time in the past was different from the present place (i.e., at the time of the survey). The last usual place of residence is a place (village/town) where the person had stayed continuously for a period of six months or more.

12. We use the terms establishment, enterprise, and firms interchangeably.

town and city levels.[13] Both rounds of the EC provide information on an establishment's number of employees and major economic activity, while only the micro-records of EC 1998 provide information on the age of the enterprise (as captured by information on the number of years of operation of the enterprise). The EC 1998 dataset covers 12.6 million establishments for urban India, and 17.7 million establishments for rural India, while the EC 2013 dataset covers 24 million and 35 million establishments for urban and rural India, respectively.

In order to include the EC variables into our analysis, we had to match administrative geographic boundaries across census years. Since the EC of 1998 follows spatial boundaries as of August 1997, significant manual effort was made to match city-level units to their 2001 population census equivalent. Secondary sources such as the administrative atlas of India, pin code database, digital maps, and other online resources were utilized to track geographical variabilities between both years.[14] This effort resulted in a one-to-one mapping of urban areas in the EC of 1998 with those of the 2001 population census. Similar efforts were made for the EC of 2013 whereby spatial boundaries following the 2011 population census had to be reclassified back to their 2001 status.

In constructing the various measures of economic activity across cities, we limit our attention to 22 manufacturing and 22 service industries at the two-digit level.[15] This means that we omit one manufacturing industry (recycling) and several service industries (public administration and defense, other services, activities of domestic households as employers of domestic staff, and extra-territorial organizations and bodies).

Given that EC data have rarely been used to examine the structure of economic activity (at the city or town level), we compared sectoral totals and composition of the EC 1998 with Round 55 of the NSS-EUS carried out in 1999–2000. Although level differences in employment across the two sources of data are apparent, correlations of employment distribution across the 44 two-digit industries considered here are very strong (correlation of 0.94) across the NSS-EUS and EC 1998.[16] We also find the manufacturing

13. A village is the rural counterpart to a town.

14. We are grateful to Vishal More for his support in carrying out this exercise.

15. Industrial classifications used in this paper are based on National Industrial Classification (NIC) 2004. The three-digit NIC 2008 and four-digit NIC 1987 codes were harmonized to their two-digit NIC 2004 equivalent.

16. Since the EC 1998 uses the 1987 four-digit NIC (1987) to record an establishment's economic activity, we first transformed the four-digit 1987 industry codes to their two-digit NIC 1998 equivalent.

to services sector employment ratios from EC and EUS to closely mirror each other. The main difference between the EC and NSS-EUS can be found in estimates of the share of total employment accounted by formal firms (which we take here to be firms with more than 10 workers, an assumption consistent with Indian regulations governing the manufacturing sector). This is likely to be due to under-coverage of small firms in the EC. With this caveat, the EC appears to be a good source for the sectoral distribution of firms and employment at the city level. It is certainly the only one available at this point of time.

The middle panel of Table 2 presents the variables of interest from the EC, including total employment in the 44 manufacturing and service industries we consider, the sectoral employment share from formal (10+) and large (100+) firms, and the share of young firms (i.e., enterprises reporting 5 years or less of operations).[17] We also constructed indices of diversity to capture the role of industrial variety in driving city growth and influencing employment-related outcomes. Following the previous literature (Duranton and Puga 2000; Ghani, Kerr, and Tewari 2014), our indices are defined as $Div_c = 1/\Sigma_j|s_{jc} - s_j|$, wherein s_{jc} refers to the share of industry j in city c's total employment, and s_j is industry j's share in national total employment.[18] We construct these indices separately for manufacturing as well as for manufacturing and services together. Note that the absolute difference between s_{jc} and s_j in any given city captures the degree to which a given industry j is over- or under-represented in the city relative to the country as a whole (in terms of shares). Summing across industries and taking its inverse yields the index. If a city has all of its manufacturing (or manufacturing and services) employment in one industry that is also very small in size nationally, the denominator of the index becomes large, driving the diversity index to take a very small value that approaches zero. Alternatively, if the city's employment structure is identical to that of the country as a whole, then the denominator of the index takes on small values, driving the diversity index to be large.

As with the other variables, the values taken by these variables can vary widely across cities (not shown). Moreover, they also vary by city size. For example, larger cities are more diversified in terms of industrial activity; they tend to have a larger share of manufacturing relative to employment

17. Note that for a household establishment carrying out a business for generations, the years of operation refers to the number of years since the current owner or operator took over. This is not ideal for our purposes but unavoidable.

18. We also constructed indices of specialization. Due to the weak results obtained using these, we omit these from our discussion to conserve space.

in manufacturing and the restricted set of services we consider here, and a larger share of younger firms.

3.1.3. Road Connectivity and Market Access

We obtain road network vector data from ML Infomap, which contains geographically referenced state highways, national highways, and expressways of India since the late 1990s. With the aid of GIS software, we compute the straight-line distance from a city's centroid to the nearest state highway and to the nearest national highway or expressway by 2001 as two measures of a city's road connectivity.

For MA, we consider each city's access to the largest 74 cities with population above 500,000 in 2001. Specially, MA is calculated with the formula:

$$MA_c = \sum_{i=1}^{74} \frac{1}{d_{ci}} POP_i,$$

where d_{ci} is the distance from city c to one of the 74 large cities travelled through the available road network, and set to 1 if $c = i$.[19] The MA variable is expected to capture the demand for a city's output from major markets within the country.

The penultimate panel of Table 2 shows that the average distance to a state highway and national highway or an expressway is 6 km and 13 km, respectively. In general, cities are more closely located to state and national highways than towns. As far as MA is concerned, the sample average is about 20,000 people per kilometer for a city/town. That means that different city size groups deviate from the sample average only moderately, and though not shown here, it may be noted that larger cities do not necessarily have better access to the country's major markets.

3.1.4. Measures of Labor Regulation and Land Management

Some of our analysis explores whether employment across sectors and in firms of different sizes varies by the nature of labor- and land-management-related regulations across states/cities. For a measure of labor regulation across states, we rely on those constructed by Gupta, Hasan, and Kumar

19. The road network data from ML Infomap have many disconnected or isolated roads, possibly because it does not include small local roads. This causes the GIS software—which uses an algorithm that accounts for junctions to generate the shortest path—to fail to calculate road network distance for many cities. As an alternative, we use the GRoads road data to calculate the MA for each city. The concern is that GRoads only contains 2010 information without data on historical network. We thus compared the distance computed with ML Infomap and GRoads data, respectively, for cities with information from both sources. Fortunately, the two sets of distance are highly correlated.

(2009). Flexible states have regulations that are pro-employer. On the other hand, states with inflexible labor regulations are those whose regulations are coded as being either neutral or pro-worker by Gupta, Hasan, and Kumar.

For land management-related regulations, we utilize a 2014 reform appraisal report from India's Ministry of Urban Development that had evaluated a sample of 66 cities with regards to their compliance on 23 items of planning and land-use reforms suggested by the central government. Of the 23 reforms, we consider the following six to be relevant to the nexus between urban land management and urban growth: (a) reforms in rent control; (b) the repeal of ULCRA; (c) achieving threshold measures of coverage and collection efficiency of property taxes; (d) reduction of stamp duty (to 5 percent or less); (e) simplification of legal and procedural framework for conversion of agricultural land for non-agricultural purposes; and (f) introduction of computerized processes in the registration of land and property (Jawaharlal Nehru National Urban Renewal Mission 2014).

The 2014 reform appraisal assessed the 66 cities on a 10-point scale on each reform element, with complete implementation of the reform assigned a score of 10. Based on the six reform elements we consider, the maximum possible score received by each city is thus 60. State scores, S_s, are then calculated as follows,

$$S_s = \sum_{i=1}^{n} \frac{pop_i}{pop_s} c_i$$

where pop_i refers to the city-level population, pop_s is the state-level population of the state that the city belongs to, and c_i corresponds to the individual city scores. We treat states that had scored above the median score (63.9) as those whose cities have "good" land-management policies and practices. In our regressions, we also used the raw aggregate scores as our measure of land-management practices. The results were qualitatively similar.

3.2. Empirical Framework

There are several factors that influence how much growth a city experiences (through migration and/or expansion of a city's de facto boundaries).[20] First, there is the structure and nature of economic activity undertaken

20. In an accounting sense, an increase in the growth of the urban population is the result of an increase in the population of existing cities—that is, the sum of births minus deaths plus net migration as rural residents move to urban areas—and the reclassification of rural areas to urban areas. The latter process can involve either the creation of entirely new cities and towns, or the expansion of existing cities and towns as rural hinterlands get absorbed into adjoining urban areas.

in a city. The relative importance of manufacturing jobs compared to services employment, the share of employment in formal firms (proxied here by employment in firms with 10 or more employees), the presence of large firms and young firms, and the diversity of economic activity, all are potential drivers of productivity and labor demand, and thus the wages a city offers. The latter, in turn, will drive how much migration a city experiences; it may also influence urban growth through an expansion of the physical size of a city.[21]

Second, there are factors such as the skills and education available in a city, transportation and connectivity (to other cities and locations, but also within the city), amenities, and the cost of living. A third factor is the spatial layout of cities, which, as the recent work of Harari (2016) has shown, interacts with geography and regulatory factors (such as land use regulations), and can influence the location choices of consumers and firms.

In our analysis of city growth, we focus on the first two sets of factors, examining their empirical importance by estimating regressions of the following type:

$$\ln pop_{ct} - \ln pop_{ct-1} = a^{*}\ln pop_{ct-1} + X_{ct-1} + e_{ct}, \quad (1)$$

where pop represents population, X is a vector of city characteristics, and *c* and *t* denote cities and year, respectively. As discussed in Duranton (2016b), this formulation of city growth being regressed on the initial levels of the determinants of city growth (rather than growth of the determinants) is justified when labor mobility is imperfect, so that city populations adjust slowly.

It must be noted that while the explanatory variables of Equation (1) pertain to the initial period, the estimated coefficients on the vector X cannot be interpreted as causal. For example, decisions by employers and households may be made based on future city growth. Unfortunately, natural experiments are quite difficult to find in the context of city growth, as are plausible instruments for many potential drivers of urban growth. Nevertheless, as a robustness check we introduce regressors that may act as controls for endogeneity. To control for unobservable city-level characteristics, we include historical population growth from 1991 to 2001 into our analysis as in Duranton (2016a). We additionally introduce initial land area as a control since the initial physical size of cities may offer advantages in terms of a city's potential to grow. That being said, the lack of good instruments is a limitation of this paper and its usefulness has to be seen more on the lines

21. This expansion may be captured by an expanding administrative boundary of the city, or by urban outgrowth (i.e., the urbanization of areas outside the administrative boundaries of the city).

of uncovering important patterns in the data that can inform future research priorities.

Finally, as noted earlier, the mere fact that cities are registering robust population growth is but one measure of urban success. Whether the factors that drive population growth also contribute to productive and well-paying jobs is a question of considerable importance. In principle, information on city-level productivity or wages should allow us to estimate variants of Equation (1) to address this question. However, data availability presents a challenge. In particular, the most disaggregated geographic level for which employment-related information is available from India's labor force surveys is the district.

Fortunately, a good proxy seems to be available at the city level: employment in firms with 10 or more workers. In the case of India's manufacturing sector, firms with 10 or more workers (20 if the production process does not use electricity—a rarity these days) come under the purview of the Factories Act and are deemed to be registered (or formal). There is no such hard and fast cut-off in the cases of services, but, in general, a threshold of 10 employees seems reasonable in distinguishing between firms that are more or less likely to use modern production and managerial technologies, and thus have both higher productivity and higher wages.

We use data on individual workers and their wages from the 2011–12 NSS-EUS to confirm that employees in firms with 10 or more workers are paid more even when controlling for observable individual characteristics (age, gender, and educational attainment), industry and district of employment, and whether a respondent is employed in a city with a population of one million or more as per the 2001 population census. The results, reported in Section 4, give us confidence that the share of city employment-generated by firms with 10 or more workers is a good proxy for productive and well-paying jobs.

4. Results

4.1. City Size across India: Some Patterns

Given the novelty of our dataset, it is interesting to examine whether and how city size varies with economically relevant variables. To conserve space, we do not report the results of regressions of city size but only note our key findings. First, larger cities tend to have more literate and educated populations. Second, they are also more likely to have a greater number of colleges—though not necessarily as a proportion of city population.

Third, and surprisingly, larger cities are not systematically associated with either higher road density or electricity connections. Larger cities, are, however, better connected to state and national highways (or expressways), though MA is not necessarily better for larger cities. Fourth, larger cities tend to have higher diversity indexes. This result tends to be driven by the largest Class 1 cities, that is, those with a population of 100,000 or more. Larger cities also tend to have a higher share of employment coming from manufacturing, formal sources of employment (as proxied by firms with 10 or more employees), and large firms (as proxied by firms with hundred or more employees). As in the case of the diversity index, these relationships are mainly driven by the so-called Class 1 cities with populations of a 100,000 or more.

4.2. City Growth

We now turn to estimating various versions of Equation (1) above. This formulation, whereby city growth is regressed on the initial levels of its potential determinants (rather than growth of the determinants), is justified when labor mobility is imperfect, so that city populations adjust slowly.

4.2.1. Natural Growth and Migration

Table 3, which relies on our district-level proxies for fertility and migration, indicates that migration does contribute to city growth. In all columns presented, initial city size as captured by the 2001 city population is included as a regressor. The first column does not include any of the demographic- and migration-related measures and the positive and statistically significant coefficient on initial city population suggests a failure of Gibrat's law, which holds that the population growth of cities is orthogonal to their initial level.[22] More generally, the results indicate that both natural growth (as captured by our proxy for birthrates) and migration contribute to city growth. At the same time, it must be noted that the R-squares of all the regressions are relatively low. This is likely a reflection of the fact that our explanatory variables are at the district level, while our dependent variable is at the city level.

4.2.2. Human Capital

Research on the determinants of city growth in developed countries often finds an important role for human capital. Corroborative evidence from developing countries is scarce, but studies based on Colombian and Brazilian data indicate its relevance to the developing country context (see Duranton 2016a and de Mata et al. 2007, respectively). Table 4 reports our findings on

22. Further exploration revealed that this result tends to be driven by the relationship between city population and initial population levels in cities with less than 100,000 persons in the initial year.

TABLE 3. City Growth and Demography and Migration

	I	*II*	*III*	*IV*
	Dependent Variable: Population Growth Defined As Log(2011 Population/2001 Population)			
Log 2001 population	0.027***	0.029***	0.029***	0.028***
	(0.005)	(0.005)	(0.005)	(0.005)
D: Ratio of young to total population		1.959***	2.036***	2.104***
		(0.519)	(0.511)	(0.525)
D: Internal migration (ratio to population), urban			0.049*	
			(0.025)	
D: Intra-district migration (ratio to population), urban				–0.002
				(0.025)
D: Inter-district migration (ratio to population), urban				0.051
				(0.050)
D: Inter-state migration (ratio to population), urban				0.184***
				(0.057)
Constant	–0.129**	–0.327***	–0.360***	–0.367***
	(0.049)	(0.073)	(0.068)	(0.066)
Observations	2,426	2,119	2,110	2,110
R-squared	0.103	0.130	0.128	0.133
State dummies	YES	YES	YES	YES

Source: Authors' estimates.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes:

1. "D" denotes district-level data.
2. For the variable "ratio of young to total population," "young" refers to the portion of population from ages 0 to 5.

the importance of human capital on city growth in India. As noted earlier, most of our measures for human capital are less than ideal, given that they pertain to the district level (albeit its urban component) rather than the city level. Nevertheless, the overall picture we get is one of weak association between human capital and city growth. In fact, the literacy rate appears with a negative and statistically significant sign for our full sample. Further exploration suggests that this relationship is driven by towns, that is, cities with less than 100,000 people in 2001.

4.2.3. Infrastructure and Connectivity

Table 5 describes the relationship between city growth, and our various measures of infrastructure and connectivity. Regardless of which sample we consider, our measure of MA enters the various city growth regressions with a positive and statistically significant coefficient. Its estimated coefficient implies that a doubling of MA from its sample mean (0.195) would increase city growth by eight percentage points (0.195×0.426) in our full sample. Interestingly, the coefficient of MA is larger for cities than towns,

TABLE 4. City Growth and Human Capital

	Full Sample					*Towns*					*Cities*				
	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*	*X*	*XI*	*XII*	*XIII*	*XIV*	*XV*
	Dependent Variable: Population Growth Defined as Log(2011 Population/2001 Population)														
Log 2001 population	0.016**	0.027***	0.019***	0.027***	0.027***	0.013*	0.015**	0.016*	0.015**	0.016*	0.037	0.022***	0.039	0.024***	0.022**
	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)	(0.029)	(0.008)	(0.029)	(0.007)	(0.008)
PCA: Human capital 1	–0.004					–0.004					–0.000				
	(0.003)					(0.003)					(0.007)				
PCA: Human capital 2		0.001					0.002					0.002			
		(0.003)					(0.002)					(0.020)			
Literacy rate			–0.181***					–0.189***					–0.010		
			(0.034)					(0.035)					(0.140)		
D: Years of schooling				–0.001					–0.000					–0.004	
				(0.002)					(0.002)					(0.004)	
PCA: College density					0.001					0.002					0.010
					(0.003)					(0.002)					(0.023)
Constant	–0.016	–0.136**	0.118	–0.131**	–0.128**	0.016	–0.014	0.150	–0.013	–0.012	–0.312	–0.101	–0.337	–0.070	–0.097
	(0.059)	(0.051)	(0.071)	(0.051)	(0.050)	(0.072)	(0.068)	(0.088)	(0.072)	(0.075)	(0.340)	(0.109)	(0.423)	(0.098)	(0.109)
Observations	1,838	2,119	2,112	2,119	2,426	1,704	1,799	1,967	1,799	2,075	134	320	145	320	351
R-squared	0.110	0.109	0.113	0.109	0.103	0.105	0.095	0.109	0.094	0.086	0.200	0.298	0.207	0.299	0.295
State dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Source: Authors' estimates.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes:

1. "D" denotes district-level statistics while "PCA" refers to principal component analysis.
2. Human capital1 was calculated using 2001 city-level components on years of schooling, literacy rate, and college school density (refer to Note 4).
3. Human capital2 was calculated using 2001 city-level components on years of schooling, literacy rate, and college school density (refer to Note 4).
4. College density was calculated using 2001 city-level components on the number of colleges, polytechnics, and professional schools.

TABLE 5. City Growth and Infrastructure

	Full Sample *I*	*Towns* *II*	*Cities* *III*
	Dependent Variable: Population Growth Defined as Log(2011 Population/2001 Population)		
Log 2001 population	0.026***	0.016**	0.022**
	(0.005)	(0.007)	(0.008)
PCA: Infrastructure	–0.007	–0.007	0.002
	(0.005)	(0.005)	(0.009)
Distance to state highway (in'000 km)	–0.020	–0.347	0.581
	(0.420)	(0.372)	(0.790)
Distance to expressway or national highway	–0.410***	–0.620***	0.655
(in '000 km)	(0.146)	(0.160)	(0.528)
Market access (100,000 persons per km)	0.426***	0.363***	0.616***
	(0.080)	(0.121)	(0.202)
Constant	–0.174***	–0.052	–0.204*
	(0.052)	(0.067)	(0.106)
Observations	2,403	2,054	349
R-squared	0.146	0.123	0.389
State dummies	YES	YES	YES

Source: Authors' estimates.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Note: "PCA" refers to principal component analysis with the measure on infrastructure calculated using 2001 city-level components on electricity density and *pucca* road density.

implying that growth of cities has been influenced to a greater extent by their accessibility to major markets than the growth of towns. On the other hand, access to a national highway or expressway has a significant impact on the growth of towns but not that of cities. A reduction of 10 km in the distance to a national highway or expressway results in a 0.6 percentage point increase in the growth rate of a town. In contrast, a higher percentage of households with electricity connections and greater density of paved roads has no significant relationship with subsequent city growth. This result holds even when we substitute the principal component analysis (PCA)-infra measure with its two constituent components and introduce these separately.

4.2.4. Characteristics of Economic Activity

We finally turn to the relationship between city growth and the nature of economic activities in cities as described in Tables 6a–6c. Table 6a, which pertains to the full sample of cities and towns, reveals that a higher share of manufacturing employment and a higher share of young, formal firms (i.e., firms aged 5 years or less and having 10 or more workers) in manufacturing or manufacturing and services (the 22 services we consider) are

TABLE 6a. City Growth and Economic Activities, Full Sample

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*	*X*
	Dependent Variable: Population Growth Defined as log(2011 Population/2001 Population)									
Log 2001 population	0.027***	0.029***	0.027***	0.026***	0.028***	0.025***	0.026***	0.026***	0.022***	0.021***
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)
Diversity index, mfg	0.002								0.027	0.027
	(0.019)								(0.020)	(0.025)
Diversity index, mfg and srvcs		–0.008								
		(0.008)								
Manufacturing employment share			0.066***						0.067**	0.052**
			(0.024)						(0.026)	(0.025)
Emp. share of 10+ worker firms, mfg				0.020					–0.033*	–0.036**
				(0.016)					(0.018)	(0.017)
Emp. share of 10+ worker firms, mfg and srvcs					–0.033					
					(0.031)					
Dummy: Mfg firms with 100+ workers						0.031*			0.016	0.019
						(0.015)			(0.018)	(0.019)
Dummy: Srvcs firms with 100+ workers						–0.009				
						(0.012)				
Share of young firms with 10+ workers in mfg							0.752***			
							(0.162)			
Share of young firms with 10+ workers in mfg and srvcs								1.796***	1.786***	1.471***
								(0.395)	(0.326)	(0.303)
D: Years of schooling										–0.003
										(0.002)
PCA: Infrastructure										–0.007
										(0.004)

Distance to expressway or national highway (in '000 km)										–0.443*** (0.121)
Market access (100k persons per km)										0.397*** (0.092)
Constant	–0.129** (0.050)	–0.132** (0.051)	–0.136** (0.051)	–0.123** (0.051)	–0.128** (0.049)	–0.103** (0.040)	–0.117** (0.052)	–0.119** (0.052)	–0.114** (0.044)	–0.132** (0.049)
Observations	2,424	2,424	2,424	2,424	2,424	2,426	2,424	2,424	2,424	2,096
R-squared	0.103	0.103	0.108	0.104	0.104	0.106	0.120	0.118	0.123	0.171
State dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Source: Authors' estimates.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes:

1. "D" denotes district-level statistics.
2. "PCA" refers to principal component analysis, with the measure on infrastructure calculated using 2001 city-level components on electricity density and *pucca* road density.
3. "Dummy: Mfg firms with 100+ workers" and" Dummy: Srvcs firms with 100+ workers" are assigned a value of 1 if a city has at least three firms with more than 100 employees in their respective industries.
4. Firms are considered "young" if they have been in operation for at most five years as of 1998.

TABLE 6b. City Growth and Economic Activities, Towns

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*	*X*
	Dependent Variable: Population Growth Defined as log(2011 Population/2001 Population)									
Log 2001 population	0.013*	0.015*	0.016**	0.015*	0.017**	0.015*	0.016**	0.015**	0.010	0.008
	(0.007)	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Diversity index, mfg	0.033**								0.069***	0.074***
	(0.012)								(0.015)	(0.019)
Diversity index, mfg and srvcs		0.005								
		(0.006)								
Manufacturing employment share			0.060*						0.077**	0.077**
			(0.029)						(0.034)	(0.033)
Emp. share of 10+ worker firms, mfg				0.016					–0.027	–0.030**
				(0.016)					(0.017)	(0.014)
Emp. share of 10+ worker firms, mfg and srvcs					–0.043					
					(0.031)					
Dummy: Mfg firms with 100+ workers						0.050*			0.039	0.043*
						(0.028)			(0.024)	(0.024)
Dummy: Srvcs firms with 100+ workers						–0.024				
						(0.015)				
Share of young firms with 10+ workers in mfg							0.663**			
							(0.248)			
Share of young firms with 10+ workers in mfg and srvcs								1.451**	1.308**	1.042**
								(0.603)	(0.538)	(0.455)
D: Years of schooling										–0.003
										(0.002)
PCA: Infrastructure										–0.007
										(0.004)

Distance to expressway or national highway (in '000 km)										–0.671***
										(0.136)
Market access (100k persons per kilometer)										0.321**
										(0.129)
Constant	–0.013	–0.009	–0.022	–0.005	–0.005	0.001	–0.012	–0.013	–0.027	–0.039
	(0.071)	(0.072)	(0.070)	(0.073)	(0.071)	(0.072)	(0.070)	(0.070)	(0.065)	(0.060)
Observations	2,072	2,072	2,072	2,072	2,072	2,074	2,072	2,072	2,072	1,777
R-squared	0.087	0.085	0.091	0.086	0.087	0.091	0.097	0.094	0.107	0.155
State dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Source: Authors' estimates.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes:

1. "D" denotes district-level statistics.
2. "PCA" refers to principal component analysis, with the measure on infrastructure calculated using 2001 city-level components on electricity density and *pucca* road density.
3. "Dummy: Mfg firms with 100+ workers" and" Dummy: Srvcs firms with 100+ workers" are assigned a value of 1 if a city has at least three firms with more than 100 employees in their respective industries.
4. Firms are considered "young" if they have been in operation for at most five years as of 1998.

TABLE 6c. City Growth and Economic Activities, Cities

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*	*X*
	Dependent Variable: Population Growth Defined as log(2011 Population/2001 Population)									
Log 2001 population	0.024**	0.024**	0.018**	0.015*	0.016**	0.027**	0.015**	0.014*	0.019**	0.030***
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.011)	(0.007)	(0.007)	(0.008)	(0.008)
Diversity index, mfg	–0.033								0.007	–0.014
	(0.039)								(0.034)	(0.032)
Diversity index, mfg and srvcs		–0.022								
		(0.020)								
Manufacturing employment share			0.104						0.066	–0.060
			(0.091)						(0.089)	(0.085)
Emp. share of 10+ worker firms, mfg				0.064					0.010	–0.010
				(0.051)					(0.055)	(0.075)
Emp. share of 10+ worker firms, mfg and srvcs					0.081					
					(0.094)					
Dummy: Mfg firms with 100+ workers						–0.001			–0.028	–0.018
						(0.018)			(0.029)	(0.028)
Dummy: Srvcs firms with 100+ workers						–0.031				
						(0.025)				
Share of young firms with 10+ workers in mfg							0.955***			
							(0.277)			
Share of young firms with 10+ workers in mfg and srvcs								2.569***	2.483***	1.945***
								(0.772)	(0.851)	(0.566)
D: Years of schooling										–0.004
										(0.004)
PCA: Infrastructure										0.004
										(0.015)

Distance to expressway or national highway (in '000 km)										0.641
										(0.627)
Market access (100k persons per km)										0.638**
										(0.263)
Constant	−0.094	−0.090	−0.056	−0.042	−0.062	−0.126	−0.016	−0.006	−0.050	−0.227*
	(0.094)	(0.093)	(0.091)	(0.090)	(0.087)	(0.104)	(0.087)	(0.083)	(0.086)	(0.116)
Observations	352	352	352	352	352	352	352	352	352	319
R-squared	0.295	0.298	0.300	0.298	0.296	0.298	0.329	0.332	0.337	0.431
State dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Source: Authors' estimates.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes:

1. "D" denotes district-level statistics.
2. "PCA" refers to principal component analysis, with the measure on infrastructure calculated using 2001 city-level components on electricity density and *pucca* road density.
3. "Dummy: Mfg firms with 100+ workers" and" Dummy: Srvcs firms with 100+ workers" are assigned a value of 1 if a city has at least three firms with more than 100 employees in their respective industries.
4. Firms are considered "young" if they have been in operation for at most five years as of 1998.

all associated with faster city growth.[23] To some extent, the presence of at least a few large manufacturing employers is also associated with faster city growth. Higher educational attainment and city-level infrastructure (as captured by PCA-infra) do not seem to matter. On the other hand, better connectivity, as captured by a smaller distance to the nearest expressway or national highway, matters since it drives city growth (as seen by the negatively signed coefficient on the distance variable), as does greater MA.

Taken at face value, these results suggest an important role for manufacturing, entrepreneurship, and connectivity to markets in driving city growth. Nevertheless, there are a couple of puzzling features of the results, namely, the lack of any role for educational attainment of the urban population and city-level infrastructure. Also, a higher share of employment accounted for by firms with 10 or more workers is negatively associated with city growth when we consider the various determinants together (columns 9 and 10 of Table 6a).

Considering towns and cities separately (Tables 6b and 6c, respectively), we see that the previous results pertaining to the structure of economic activity seem to be driven by the growth experience of towns. Indeed, the diversity index for manufacturing is now positive and significant, indicating that towns that had a greater variety of employment across manufacturing industries in 1998 experienced faster population growth between 2001 and 2011. Moreover, as may be expected, the impact of being closer to an expressway or national highway matters more for towns (a larger negatively signed coefficient on the distance variable). Turning to the results for cities alone, the only variables that matter for city growth are greater MA and our measure of entrepreneurship, that is, a greater share of young, formal firms.

As noted earlier, we are not fortunate enough to have a natural experiment that we could exploit to help us establish causality. Additionally, finding plausible instruments for our variables is not easy. Nevertheless, it is important to check the robustness of our results to the inclusion of some controls that might alleviate potential biases on account of omitted factors that might be driving city growth. We introduce two sets of controls, one for the size of cities in land area in 2001 and the second for city growth over the previous decade. Our results are reported in Appendix Tables A1.1–A1.3. Essentially, we find that controlling for (log) area does not affect the estimated coefficients on the share of young, formal firms in any of the full, town, or city samples.

23. If we consider all young firms, and not just those with 10 or more workers, we get estimated coefficients of a smaller magnitude, some of which are also statistically insignificant. This is not surprising since, often, entrepreneurs that operate very small sized firms may have neither the managerial or technical abilities nor the finances required to start and operate economically dynamic firms with high growth potential.

Similarly, controlling for city growth over the past decade does not affect the estimates in full and city samples, but the estimates in town samples lose their significance when introduced along with other determinants of city growth. Overall, it appears that entrepreneurship as proxied by a higher share of young, formal firms has an important role to play in promoting city growth.

Before we turn to whether these determinants of city growth also promote better-paying jobs, it is useful to consider the economic significance of our results, especially that on entrepreneurship. Specifically, under the assumption that our regression estimates can be given a causal interpretation, we can explore the effects on city growth of a hypothetical policy that fosters entrepreneurship and formalization. Table 7 illustrates an example involving the share of young, formal firms in manufacturing and services. Suppose the policy to promote entrepreneurship leads to a 100 percent increase in the share of young, formal firms in towns (left column). We can calculate how many new young, formal firms this increase translates into (327) and how many additional urban residents it would lead to (571,308) across towns. A question of interest is what additional urbanization would be generated if we were to "relocate" the same number of young, formal firms to the big cities.

The simulation presented in the right column of Table 7 suggests that the average share of young, formal firms would increase from 0.72 percent

TABLE 7. Policy Simulation: Share of Young Firms with 10+ Workers, Manufacturing and Services

	Towns	*Cities*
Share of young firms with 10+ workers (average per city)	0.0046	0.0072
Total number of firms with 10+ workers (average per city)	34.124	843.230
Total number of young firms with 10+ workers (average per city)	0.156	6.101
2001 population (average per city)	35,852	654,487
Estimated coefficients of share of young firms with 10+ workers	1.451**	2.569***
Original predicted population growth (average per city)	0.1407	0.2197
Original predicted 2011 population (average per city)	41,267	815,326
Simulation		
Share of young firms with 10+ workers (average): 100% increase	0.0092	–
Resulting number of additional young firms with 10+ workers (average per city)	0.1578	–
Resulting sum of additional young firms with 10+ workers	327	327
Resulting share of young firms with 10+ workers	–	0.0083
Resulting predicted population growth (average per city)	0.1473	0.2226
Resulting predicted 2011 population (average per city)	41,542	817,646
Resulting additional population in 2011 (average per city)	275	2,319
Resulting sum of additional population in 2011	571,308	814,131
Number of Towns/Cities	2,075	351

Source: Authors' estimates.

to 0.83 percent in cities in 2011. The predicted urban population by 2011 would increase by 2,319 per city on average and 814,131 in total for the large cities. The results are remarkable in that the same number of young, formal firms could lead to an urban population that is 43 percent higher if the increase in such firms were concentrated in large cities rather than in towns. The gap is mainly driven by the different elasticities of urban population to the share of young, formal firms across small and large cities. For policymakers interested in promoting urbanization by fostering entrepreneurship in cities, the simulation exercise suggests that focusing on large cities may be more effective.

4.3. Are the Drivers of City Growth Promoting Better Jobs?

The lack of data on wages at the city level requires us to shed light on this question in an indirect way, using a proxy for productive and well-paying jobs available at the city level. Table 8 shows the results of simple Mincerian wage regressions using the urban sample of the 2011–12 NSS-EUS data. The key result is that, controlling for age, gender, and educational attainment, workers employed in firms with 10 or more employees receive a considerable wage premium. This premium ranges from approximately 35 percent in column 2 to 25 percent when controls for industry and district of employment are added in column 5. Admittedly, there are potentially important omitted variables here—such as the quality of education and individual abilities—but for our limited purpose, these results lend credibility to the idea that employment in formal/larger firms is more productive and better-paying for roughly equivalent workers, thereby suggesting that the share of employment accounted for by firms with 10 or more employees is a reasonable proxy for the quality of jobs in a city.

With this, we examine the relationship between the share of employment in firms with 10 or more employees in 2013 (the EC year closest to the population census year of 2011) and the determinants of city growth from 2001 to 2011. Table 9 describes the results for two sets of regressions. In each set, the results are presented separately for the full sample, the town sample, and the city sample.[24] Overall, the results are broadly consistent with the city growth regressions. Thus, cities that in the initial years have

24. We also ran these regressions including the 1998 value of employment share of firms with 10 or more workers as a control. Its own estimated coefficient is positive and significant, and most of the results reported in Table 9 are preserved. The only exception is that the diversity index for manufacturing is positive and significant for the full sample and cities; however, the dummy for large manufacturing firms is negative and significant in cities.

TABLE 8. Wage Regressions, Urban India 2011–12

	NSS Round 68				
	I	*II*	*III*	*IV*	*V*
	Dependent Variable: Log Wage (Real)				
Age	0.045***	0.043***	0.043***	0.040***	0.043***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age^2	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sex (male = 1)	0.088***	0.128***	0.116***	0.146***	0.150***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
D: Years of schooling	0.083***	0.077***	0.077***	0.062***	0.060***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Dummy: Employed in firm with 10+ workers		0.345***	0.334***	0.280***	0.251***
		(0.000)	(0.000)	(0.000)	(0.000)
Dummy: Big town	0.144***		0.111***	0.097***	0.031***
	(0.000)		(0.000)	(0.000)	(0.001)
Constant	3.518***	3.476***	3.433***	3.534***	3.640***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.019)
Observations (population weighted)	29,035,826	29,035,826	29,035,826	29,035,826	29,035,826
R-squared	0.428	0.459	0.463	0.523	0.579
Industry dummies	NO	NO	NO	YES	YES
District dummies	NO	NO	NO	NO	YES

Source: Author's estimates from the National Sample Survey, Employment-Unemployment 2011–12. Estimates are calculated using urban data from the 25 states/UTs.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes:

1. "D" denotes district-level statistics.
2. Coefficients on Age^2 are reported as null due to rounding.
3. Dummy: Big town is assigned a value of 1 if respondent resides in a city with a population of one million or more based on the 2001 population census.

greater entrepreneurship, are closer to expressways or national highways, and have large manufacturing firms present, tend to have a larger share of good jobs in 2013. Moreover, some of the patterns observed earlier for towns and cities are repeated. Thus, the presence of large manufacturers and distance to national highways or expressways matter more for towns than cities. Entrepreneurship as captured by a larger share of young, formal firms, whether restricted to just manufacturing (columns 1–3) or not (columns 4–6), matters in both towns and cities, though the size of the estimated coefficient is larger for towns rather than cities. This is in contrast to the case of the city growth regressions. For the first time, our human capital measure—the average years of schooling in the district's

TABLE 9. Employment Share in Firms with 10 or More Workers in 2013

	Full Sample	*Town*	*City*	*Full Sample*	*Town*	*City*
	I	*II*	*III*	*IV*	*V*	*VI*
	Dependent variable: Employment share in 10+ Firms in 2013					
Diversity index, mfg	0.003	−0.016	0.013	0.003	−0.015	0.013
	(0.010)	(0.019)	(0.022)	(0.010)	(0.019)	(0.022)
Manufacturing employment share	0.003	−0.003	0.011	0.007	0.000	0.015
	(0.036)	(0.038)	(0.054)	(0.036)	(0.038)	(0.052)
Dummy: Mfg firms with 100+ workers	0.066***	0.071*	0.030*	0.066***	0.071*	0.030*
	(0.014)	(0.037)	(0.015)	(0.015)	(0.039)	(0.015)
Share of young firms with 10+ workers in mfg	1.038***	1.093***	0.844*			
	(0.282)	(0.257)	(0.457)			
Share of young firms with 10+ workers in mfg and srvcs				2.768***	3.009***	2.100*
				(0.845)	(0.873)	(1.128)
D: Years of schooling	0.005**	0.004**	0.010**	0.005**	0.004**	0.009*
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.005)
Market access (100k persons per kilometer)	0.015	−0.032	0.138	0.005	−0.047	0.141
	(0.073)	(0.049)	(0.164)	(0.076)	(0.052)	(0.163)
Distance to expressway or national highway ('000 km)	−0.506***	−0.489***	−0.238	−0.502***	−0.489***	−0.241
	(0.125)	(0.124)	(0.383)	(0.127)	(0.124)	(0.389)
PCA: Infrastructure	−0.001	−0.002	0.009	−0.002	−0.003	0.007
	(0.001)	(0.002)	(0.009)	(0.001)	(0.002)	(0.009)
Constant	0.235***	0.248***	0.304***	0.230***	0.242***	0.305***
	(0.020)	(0.019)	(0.053)	(0.020)	(0.019)	(0.053)
Observations	2,087	1,769	318	2,087	1,769	318
R-squared	0.295	0.267	0.425	0.301	0.274	0.425
State dummies	YES	YES	YES	YES	YES	YES

Source: Authors' estimates.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes:

1. "D" denotes district-level statistics.
2. "PCA" refers to principal component analysis, with the measure on infrastructure calculated using 2001 city-level components on electricity density and *pucca* road density.
3. "Dummy: Mfg firms with 100+ workers" is assigned a value of 1 if a city has at least three manufacturing firms with more than 100 employees.
4. Firms are considered "young" if they have been in operation for at most five years as of 1998.

urban population—turns out to be statistically significant and signed in the expected direction. Finally, it may be noted that the R-squares are considerably higher than those of the city growth regressions, suggesting that while the drivers of good jobs do overlap with the drivers of city growth, there are clearly some important factors that drive city population growth that do not necessarily lead these cities to also provide their residents good jobs.

4.4. City Employment Growth and Labor and Land Regulations

Our results so far point to the importance of cities as centers of production. We now carry out a simple check of whether employment growth in cities is influenced by state- and city-level policies that are likely to affect business decisions, namely labor- and land-management-related regulations. Table 10 describes the results based on regressing city-level employment growth over 1998 and 2013 across three industrial groups and based on aggregating employment across all firms, firms with 10 or more workers, and firms with 100 or more workers. All regressions control for initial employment and city population growth over 2001 and 2011. The upper panel A includes a dummy for whether or not the city is based in a state with flexible labor regulations while the lower panel B includes a dummy for states where cities have more flexible land-related regulations.

The coefficients on the labor regulation dummy are all positive and significant, indicating that city employment growth has been faster in states with flexible labor regulations. Moreover, the effect of labor regulations is stronger for the manufacturing sector and when restricting employment growth to just that, based on firms with 10 or more and 100 or more employees. This is as expected given that India's labor regulations are more binding on the manufacturing sector as compared to the services sector and on larger firms.

The coefficients on the land regulation dummy are also positive and statistically significant. We also find city-level employment growth to be faster in firms with 10 or more and 100 or more employees in states with effective land-management policies, especially more so amongst the labor-intensive manufacturing sector. Overall, the results are consistent with the idea that better functioning labor and land markets encourage city-level economic activity and lead to faster employment growth.[25]

25. We conducted robustness checks by using the raw score of our land management reform measures instead of dummies. The results are very similar (Appendix Table A2).

TABLE 10. Sectoral Employment Growth and Labor and Land Regulations

	Employment Growth in Manufacturing and Services			*Employment Growth in Manufacturing*			*Employment Growth in Labor-intensive Manufacturing*		
	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*
	...all firms	*...from firms with 10+ workers*	*...from firms with 100+ workers*	*...all firms*	*...from firms with 10+ workers*	*...from firms with 100+ workers*	*...all firms*	*...from firms with 10+ workers*	*...from firms with 100+ workers*
A. With Flexible Labor Regulation Dummy									
1 = Flexible labor regulation	0.064***	0.097**	0.178**	0.115***	0.283***	0.229*	0.115***	0.268***	0.373**
	(0.020)	(0.041)	(0.084)	(0.030)	(0.072)	(0.136)	(0.030)	(0.074)	(0.171)
City population growth (1998–2013)	1.036***	1.227***	1.891***	1.403***	1.838***	2.720***	1.244***	1.654***	2.042***
	(0.076)	(0.153)	(0.273)	(0.115)	(0.273)	(0.357)	(0.114)	(0.276)	(0.418)
Initial 1998 log employment	−0.088***	−0.168***	−0.340***	−0.193***	−0.325***	−0.409***	−0.186***	−0.325***	−0.395***
	(0.009)	(0.013)	(0.027)	(0.011)	(0.019)	(0.041)	(0.012)	(0.022)	(0.059)
Constant	0.851***	0.841***	1.971***	1.234***	1.007***	2.095***	1.208***	1.026***	2.145***
	(0.071)	(0.094)	(0.185)	(0.076)	(0.121)	(0.313)	(0.079)	(0.129)	(0.428)
Observations	2,403	2,332	733	2,403	1,540	320	2,403	1,333	200
R-squared	0.098	0.083	0.199	0.140	0.177	0.313	0.118	0.159	0.271
B. With Land Management Dummy									
1 = Good land management	0.089***	0.134***	0.140*	0.170***	0.432***	0.383***	0.191***	0.466***	0.536***
	(0.020)	(0.040)	(0.081)	(0.030)	(0.071)	(0.134)	(0.030)	(0.073)	(0.166)
City population growth (1998–2013)	1.022***	1.210***	1.831***	1.390***	1.834***	2.763***	1.234***	1.630***	2.158***
	(0.076)	(0.153)	(0.272)	(0.114)	(0.270)	(0.358)	(0.114)	(0.273)	(0.415)
Initial 1998 log employment	−0.089***	−0.170***	−0.341***	−0.198***	−0.336***	−0.413***	−0.190***	−0.335***	−0.401***
	(0.009)	(0.013)	(0.027)	(0.011)	(0.019)	(0.041)	(0.012)	(0.022)	(0.060)
Constant	0.842***	0.829***	1.982***	1.230***	0.952***	1.981***	1.184***	0.938***	1.995***
	(0.071)	(0.093)	(0.186)	(0.076)	(0.120)	(0.318)	(0.078)	(0.128)	(0.443)
Observations	2,412	2,341	740	2,412	1,548	321	2,412	1,338	200
R-squared	0.099	0.084	0.192	0.145	0.188	0.319	0.125	0.174	0.288

Source: Authors' estimates.

Robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5. Conclusion

The key takeaways from our analysis are as follows. First, we find a positive association between various dimensions of economic activity and population growth of cities. Significantly, some of these dimensions are also associated with a higher prevalence of better quality jobs (as proxied by the share of employment in firms with 10 or more workers). Second, in terms of the dimensions of economic activity that matter, Indian towns with a more vibrant manufacturing sector (more diversity in manufacturing, a few large manufacturers, and larger shares of manufacturing employment), with a larger share of younger, formal firms, and with greater proximity to national highways, tend to grow faster. For cities with a population of a million or more, manufacturing does not seem to play a special role in city growth. Only those with more young, formal firms and greater MA grow faster.

To what extent are these variables also associated with our proxy for good jobs? We find that the presence of large manufacturers and a higher share of young, formal firms are both associated with a higher share of better quality jobs in both cities and towns. In the case of towns, being closer to a national highway also helps.

These results are consistent with the notion that modern, large-scale manufacturing and better transport connectivity to other locations can be an important factor in spurring urban success in smaller cities. At the same time, our proxy for entrepreneurship is equally important for urban success in both small and large cities, a finding very much in line with Duranton's (2014) observation that the constant entry of new entrepreneurs appears to be a key mechanism that allows cities to sustain their greater productivity.

Third, we find cities in states with more flexible labor and urban land regulations exhibiting faster employment growth in aggregate, as well as those regulations based on employment in larger firms where productivity and wages are likely to be higher. While the findings for labor regulation are not surprising, given the literature on the effects of India's labor regulations on employment and related outcomes, those on land regulations are novel. Overall, we take our findings to suggest that urban success requires cities to function well as centers of production. This is important from a policy perspective since programs for spurring entrepreneurship and economic activity are usually thought to be distinct from policies for promoting urban development, and vice versa.

References

Ahluwalia, I.J., R. Kanbur, and P.K. Mohanty. 2014. "Challenges of Urbanisation in India: An Overview," in Isher Judge Ahluwalia, Ravi Kanbur, and Prasanna Kumar Mohanty (eds.), *Urbanisation in India: Challenges, Opportunities, and the Way Forward*, pp. 1–28, New Delhi and Los Angeles, CA: SAGE.

Alder, S., M. Roberts, and M. Tewari. 2017. "The Effect of Transport Infrastructure on India's Urban and Rural Development," Unpublished *Policy Research Working Paper*.

Audretsch, D.B. and Maryann P. Feldman. 1996. "R&D Spillovers and the Geography of Innovation and Production," *The American Economic Review*, 86(3): 630–40.

Bhagat, R.B. 2011a. "Urbanisation and Access to Basic Amenities in India," *Urban India*, 31(1): 1–14.

———. 2011b. "Emerging Pattern of Urbanisation in India," *Economic & Political Weekly*, 46(34): 10–12.

———. 2014. "Urban Migration Trends, Challenges and Opportunities in India," *Background Paper, World Migration Report*. Geneva: International Organization for Migration, https://www.iom.int/sites/default/files/our_work/ICP/MPR/WMR-2015-Background-Paper-RBhagat.pdf (accessed June 23, 2018).

Brueckner, J.K. and K.S. Sridhar. 2012. "Measuring Welfare Gains from Relaxation of Land-use Restrictions: The Case of India's Building-height Limits," *Regional Science and Urban Economics*, 42(6): 1061–67.

Central Statistical Office. 1998. *Fourth Economic Census 1998–1999*. New Delhi: Ministry of Statistics and Programme Implementation, Government of India.

Chauvin, J.P., E. Glaeser, Y. Ma, and K. Tobio. 2017. "What is Different About Urbanization in Rich and Poor Countries? Cities in Brazil, China, India, and the United States," *Journal of Urban Economics*, 98: 17–49.

Clarke Annez, Patricia, Alain Bertaud, Bimal Patel, and V.K. Phatak. 2014. "Working with the Market Approach to Reducing Urban Slums in India," in Isher Judge Ahluwalia, Ravi Kanbur, and Prasanna Kumar Mohanty (eds.), *Urbanisation in India: Challenges, Opportunities, and the Way Forward*. New Delhi and Los Angeles, CA: SAGE.

Colmer, J. 2017. "Urbanisation, Growth, and Development: Evidence from India," *Working Paper*. Washington, D.C.: World Bank, https://www.gov.uk/dfid-research-outputs/urbanisation-growth-and-development-evidence-from-india (accessed June 23, 2018).

Da Mata D., U. Deichmann, J.V. Henderson, S.V. Lall, and H.G. Wang. 2007. "Determinants of City Growth in Brazil," *Journal of Urban Economics*, 62(2): 252–72.

Duranton, G. 2014. "Growing Through Cities in Developing Countries," *Policy Research Working Paper No. WPS 6818*. Washington, D.C.: World Bank.

———. 2016a. "Agglomeration Effects in Colombia," *Journal of Regional Science*, 56(2): 210–38.

Duranton, G. 2016b. "Determinants of City Growth in Colombia," *Papers in Regional Science*, 95(1): 101–32.

Duranton, G. and D. Puga. 2000. "Diversity and Specialisation in Cities: Why, Where and When Does it Matter?" *Urban Studies*, 37(3): 533–55.

———. 2014. "The Growth of Cities," *Handbook of Economic Growth*, 2B: 781–853.

Faberman, R.J. 2011. "The Relationship Between Establishment Age Distribution and Urban Growth," *Journal of Regional Science*, 51(3): 450–70.

Ghani, E., A.G. Goswami, and W.R. Kerr. 2016. "Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing,'' *The Economic Journal*, 126(591): 317–57.

Ghani, E., R. Kanbur, and S.D. O'Connell. 2013. "Urbanization and Agglomeration Benefits: Gender Differentiated Impacts on Enterprise Creation in India's Informal Sector," *World Bank Policy Research Working Paper No. 6553*. Washington, D.C.: World Bank, http://documents.worldbank.org/curated/en/245001468040574027/Urbanization-and-agglomeration-benefits-gender-differentiated-impacts-on-enterprise-creation-in-Indias-informal-sector (accessed June 23, 2018).

Ghani, E., W.R. Kerr, and S. O'Connell. 2014. "Spatial Determinants of Entrepreneurship in India," *Regional Studies*, 48(6): 1071–89.

Ghani, S.E., W.R. Kerr, and I. Tewari. 2014. "Regional Diversity and Inclusive Growth in Indian Cities," *World Bank Policy Research Working Paper. No. 6919*, Washington, D.C.: World Bank, https://openknowledge.worldbank.org/handle/10986/18760 (accessed June 23, 2018).

Glaeser, E.L. and J.V. Henderson. 2017. "Urban Economics for the Developing World: An Introduction," *Journal of Urban Economics*, 98: 1–5.

Glaeser, E.L., H.D. Kallal, J.A. Scheinkman, and A. Shleifer. 1992. "Growth in Cities," *Journal of Political Economy*, 100(6): 1126–52.

Glaeser, E.L., S.P. Kerr, and W.R. Kerr. 2015. "Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines," *Review of Economics and Statistics*, 97(2): 498–520.

Gollin, D., R. Jedwab, and D. Vollrath. 2016. "Urbanization With and Without Industrialization," *Journal of Economic Growth*, 21(1): 35–70.

Gupta, P., R. Hasan, and U. Kumar. 2008–09. "Big Reforms but Small Payoffs: Explaining the Weak Record of Growth in Indian Manufacturing," India Policy Forum, 5: 59–108. New Delhi: National Council of Applied Economic Research.

Harari, M., 2016. "Cities in Bad Shape: Urban Geometry in India," *Working Paper*. Wharton School, University of Pennsylvania, http://marroninstitute.nyu.edu/uploads/content/CityShapeHarariMarch2016.pdf (accessed June 23, 2018).

Hasan, R., Y. Jiang, and R.M. Rafols., 2017. "Urban Agglomeration Effects in India: Evidence from Town Level Data," *Asian Development Review*, 34(2), 201–28.

Henderson, J.V. 2014. "Urbanization and the Geography of Development," *Policy Research Working Paper No. 6877*. Washington D.C.: World Bank, https://openknowledge.worldbank.org/handle/10986/18747 (accessed June 23, 2018).

Jawaharlal Nehru National Urban Renewal Mission. 2014. "Reform Appraisal Reforms." New Delhi: Ministry of Urban Development, Government of India, www.jnnurm.nic.in (accessed February 24, 2015).

Kundu, A. 2007. "Migration and Urbanisation in India in the Context of Poverty Alleviation," Paper presented at the International Conference and Workshop on Policy Perspectives on Growth, Economic Structures and Poverty Reduction. Beijing, China, http://networkideas.org/ideasact/jun07/Beijing_Conference_07/Amitabh_Kundu.pdf (accessed June 23, 2018).

Kundu A., S. Bagchi, and D. Kundu. 1999. "Regional Distribution of Infrastructure and Basic Amenities in Urban India: Issues Concerning Empowerment of Local Bodies," *Economic & Political Weekly*, 34(28): 1893–906.

Kundu, A. and P.C. Mohanan. Forthcoming. "Is There a Cohesive Narrative on Internal Migration in India?" Paper presented at Authors' Workshop on *Handbook of Internal Migration*. Mumbai: Indira Gandhi Institute of Development Research.

Lall, S.V., Z. Shalizi, and U. Deichmann. 2004. "Agglomeration Economies and Productivity in Indian Industry," *Journal of Development Economics*, 73(2): 643–73.

Lall, S.V., H.G. Wang, and U. Deichmann. 2010. "Infrastructure and City Competitiveness in India," *World Institute for Development Economics Research Working Paper No. 22*. Helsinki: UNU-WIDER.

Mathur, O.P. 2016. *Cities and the New Economic Vibrancy*. New Delhi: Institute of Social Sciences.

McKinsey Global Institute. 2010. *India's Urban Awakening: Building Inclusive Cities Sustaining Economic Growth*. New York, NY: McKinsey & Company, http://www.indiaenvironmentportal.org.in/files/MGI_india_urbanization_full-report.pdf (accessed June 23, 2018).

National Sample Survey Organization. 1999. *National Sample Survey 1999–2000, 55th Round, Schedule 10—Employment and Unemployment*. New Delhi: Ministry of Statistics and Programme Implementation, Government of India.

NIUA. 2017. *Urban India: Status of Demography, Economy, Social Infrastructure, Housing and Basic Infrastructure*. New Delhi: National Institute of Urban Affairs and Human Settlement Management Institute, Housing and Urban Development Corporation.

Panagariya, A. 2008. *India: The Emerging Giant*. New Delhi: Oxford University Press.

Sridhar, K.S. 2010a. "Impact of Land Use Regulations: Evidence from India's cities," *Urban Studies*, 47(7): 1541–69.

———. 2010b. "Determinants of City Growth and Output in India," *Review of Urban and Regional Development Studies*, 22(1): 22–38.

Tewari, M., S. Alder, M. Roberts, K. Onda, and A. Pearson. 2017. "India's Urban and Spatial Development in the Post-reform Period: An Empirical Analysis Based on Nightlight Data," Unpublished *Research Paper*.

United Nations. 2014. Revision of *World Urbanization Prospects*, https://esa.un.org/unpd/wup/publications/files/wup2014-highlights.pdf (accessed June 23, 2018).

Appendix

TABLE A1.1. City Growth and Economic Activities with Historical City Controls, Full Sample

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*
	Dependent Variable: Population Growth Defined as Log(2011 Population/2001 Population)							
Log 2001 population	0.023***	0.023***	0.019***	0.018**	0.019***	0.019***	0.013***	0.012***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.005)	(0.005)	(0.004)	(0.004)
Diversity index, mfg			0.029	0.028			0.040**	0.039*
			(0.020)	(0.025)			(0.019)	(0.022)
Manufacturing employment share			0.068**	0.053**			0.074**	0.065**
			(0.025)	(0.025)			(0.027)	(0.027)
Emp. share of 10+ worker firms, mfg			–0.034*	–0.036**			–0.030*	–0.035**
			(0.017)	(0.016)			(0.015)	(0.014)
Dummy: Mfg firms with 100+ workers			0.015	0.018			0.023	0.027
			(0.018)	(0.020)			(0.017)	(0.018)
Share of young firms with 10+ workers in mfg	0.745***				0.494***			
	(0.160)				(0.161)			
Share of young firms with 10+ workers in mfg and srvcs		1.780***	1.779***	1.463***		1.134***	1.043***	0.872**
		(0.390)	(0.325)	(0.302)		(0.383)	(0.370)	(0.368)
Land area (log)	0.004	0.004	0.005	0.004				
	(0.004)	(0.004)	(0.004)	(0.005)				
Historical city growth (2001–1991)					0.238***	0.239***	0.242***	0.234***
					(0.042)	(0.042)	(0.042)	(0.034)
D: Years of schooling				–0.003				–0.003
				(0.002)				(0.002)

(Table A1.1 Continued)

(Table A1.1 Continued)

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*
	Dependent Variable: Population Growth Defined as Log(2011 Population/2001 Population)							
PCA: Infrastructure				–0.006				–0.004
				(0.005)				(0.004)
Distance to expressway or national highway (in '000 km)				–0.434***				–0.341**
				(0.123)				(0.131)
Market access (100k persons per km)				0.400***				0.266***
				(0.092)				(0.064)
Constant	–0.099	–0.100	–0.093	–0.112*	–0.085*	–0.086*	–0.070*	–0.087**
	(0.060)	(0.060)	(0.055)	(0.058)	(0.046)	(0.046)	(0.037)	(0.038)
Observations	2,424	2,424	2,424	2,096	2,397	2,397	2,397	2,074
R-squared	0.120	0.118	0.124	0.172	0.206	0.204	0.213	0.253
State dummies	YES	YES	YES	YES	YES	YES	YES	YES

Source: Authors' estimates.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes:

1. "PCA" refers to principal component analysis, with the measure on infrastructure calculated using 2001 city-level components on electricity density and *pucca* road density.
2. "Dummy: Mfg firms with 100+ workers" and "Dummy: Srvcs firms with 100+ workers" are assigned a value of 1 if a city has at least three firms with more than a hundred employees in their respective industries.
3. Firms are considered "young" if they have been in operation for at most 5 years as of 1998; and/or considered "formal" if it employs 10 or more employees.
4. Unless otherwise indicated, total employment (firms) refers to overall employment (firms) in the manufacturing and services sectors.

TABLE A1.2. City Growth and Economic Activities with Historical City Controls, Towns

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*
	Dependent Variable: Population Growth Defined as Log(2011 Population/2001 Population)							
Log 2001 population	0.013*	0.013	0.007	0.006	0.008	0.008	0.002	0.001
	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.006)	(0.005)
Diversity index, mfg			0.071***	0.076***			0.074***	0.076***
			(0.015)	(0.020)			(0.014)	(0.019)
Manufacturing employment share			0.079**	0.078**			0.082**	0.083**
			(0.034)	(0.033)			(0.035)	(0.034)
Emp. share of 10+ worker firms, mfg			–0.029*	–0.031**			–0.023	–0.028**
			(0.016)	(0.013)			(0.013)	(0.011)
Dummy: Mfg firms with 100+ workers			0.039	0.044*			0.036	0.042*
			(0.024)	(0.024)			(0.023)	(0.023)
Share of young firms with 10+ workers in mfg	0.655**				0.413*			
	(0.250)				(0.224)			
Share of young firms with 10+ workers in mfg and srvcs		1.435**	1.295**	1.031**		0.808	0.610	0.405
		(0.603)	(0.539)	(0.456)		(0.540)	(0.503)	(0.444)
Land area (log)	0.004	0.004	0.005	0.005				
	(0.005)	(0.005)	(0.005)	(0.005)				
Historical city growth (2001–1991)					0.209***	0.210***	0.209***	0.202***
					(0.044)	(0.044)	(0.043)	(0.035)
D: Years of schooling				–0.003				–0.003
				(0.002)				(0.002)
PCA: Infrastructure				–0.006				–0.004
				(0.004)				(0.004)
Distance to expressway or national highway (in '000 km)				–0.664***				–0.589***
				(0.139)				(0.159)

(Table A1.2 Continued)

(Table A1.2 Continued)

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*
	Dependent Variable: Population Growth Defined as Log(2011 Population/2001 Population)							
Market access (100k persons per km)				0.324**				0.254**
				(0.128)				(0.108)
Constant	0.002	0.001	–0.009	–0.023	0.030	0.030	0.015	–0.001
	(0.071)	(0.072)	(0.068)	(0.064)	(0.068)	(0.068)	(0.061)	(0.052)
Observations	2,072	2,072	2,072	1,777	2,050	2,050	2,050	1,759
R-squared	0.098	0.095	0.109	0.156	0.163	0.161	0.175	0.219
State dummies	YES	YES	YES	YES	YES	YES	YES	YES

Source: Authors' estimates.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes:

1. "PCA" refers to principal component analysis, with the measure on infrastructure calculated using 2001 city-level components on electricity density and *pucca* road density.
2. "Dummy: Mfg firms with 100+ workers" and "Dummy: Srvcs firms with 100+ workers" are assigned a value of 1 if a city has at least three firms with more than 100 employees in their respective industries.
3. Firms are considered "young" if they have been in operation for at most 5 years as of 1998; and/or considered "formal" if it employs 10 or more employees.
4. Unless otherwise indicated, total employment (firms) refer to overall employment (firms) in manufacturing and services sectors.

TABLE A1.3. City Growth and Economic Activities with Historical City Controls, Cities

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*
	Dependent Variable: Population Growth Defined as Log(2011 Population/2001 Population)							
Log 2001 population	0.018	0.016	0.020	0.016	0.010	0.009	0.004	0.014
	(0.019)	(0.018)	(0.017)	(0.021)	(0.008)	(0.008)	(0.008)	(0.009)
Diversity index, mfg			0.007	−0.011			0.029	0.013
			(0.034)	(0.032)			(0.030)	(0.026)
Manufacturing employment share			0.066	−0.060			0.075	0.007
			(0.088)	(0.085)			(0.064)	(0.058)
Emp. share of 10+ worker firms, mfg			0.011	−0.019			0.001	0.005
			(0.055)	(0.072)			(0.053)	(0.075)
Dummy: Mfg firms with 100+ workers			−0.028	−0.017			−0.002	−0.002
			(0.029)	(0.028)			(0.021)	(0.024)
Share of young firms with 10+ workers in mfg	0.957***				0.654***			
	(0.290)				(0.210)			
Share of young firms with 10+ workers in mfg and srvcs		2.571***	2.481***	1.954***		1.797***	1.695**	1.596**
		(0.793)	(0.826)	(0.571)		(0.567)	(0.684)	(0.602)
Land area (log)	−0.003	−0.002	−0.001	0.015				
	(0.017)	(0.016)	(0.016)	(0.019)				
Historical city growth (2001−1991)					0.301***	0.300***	0.304***	0.295***
					(0.069)	(0.069)	(0.069)	(0.078)
Years of schooling				−0.004				−0.002
				(0.005)				(0.004)
PCA: Infrastructure				0.009				0.013
				(0.017)				(0.013)
Distance to expressway or national highway (in '000 km)				0.696				0.983
				(0.670)				(0.640)

(Table A1.3 Continued)

(Table A1.3 Continued)

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*
	Dependent Variable: Population Growth Defined as Log(2011 Population/2001 Population)							
Market access (100k persons per km)				0.649**				0.240
				(0.259)				(0.157)
Constant	−0.038	−0.020	−0.060	−0.122	−0.040	−0.032	−0.007	−0.143
	(0.176)	(0.170)	(0.158)	(0.191)	(0.098)	(0.094)	(0.083)	(0.106)
Observations	352	352	352	319	347	347	347	315
R-squared	0.329	0.332	0.337	0.432	0.482	0.485	0.488	0.540
State dummies	YES	YES	YES	YES	YES	YES	YES	YES

Source: Authors' estimates.

Robust standard errors in parentheses, cluster at state level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Notes:

1. "PCA" refers to principal component analysis, with the measure on infrastructure calculated using 2001 city-level components on electricity density and *pucca* road density.
2. "Dummy: Mfg firms with 100+ workers" and" Dummy: Srvcs firms with 100+ workers" are assigned a value of 1 if a city has at least 3 firms with more than 100 employees in their respective industries.
3. Firms are considered "young" if they have been in operation for at most 5 years as of 1998; and/or considered "formal" if it employs 10 or more employees.
4. Unless otherwise indicated, total employment (firms) refer to overall employment (firms) in manufacturing and services sectors.

TABLE A2. Sectoral Employment Growth and Land Regulations

	Employment Growth in Manufacturing and Services			*Employment Growth in Manufacturing*			*Employment Growth in Labor-intensive Manufacturing*		
	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*
	...all firms	*...from firms with 10+ workers*	*...from firms with 100+ workers*	*...all firms*	*...from firms with 10+ workers*	*...from firms with 100+ workers*	*...all firms*	*...from firms with 10+ workers*	*...from firms with 100+ workers*
Land management score	0.007***	0.017***	0.021***	0.013***	0.037***	0.019	0.015***	0.037***	0.037**
	(0.002)	(0.003)	(0.007)	(0.002)	(0.006)	(0.011)	(0.002)	(0.006)	(0.014)
City population growth (1998–2013)	1.029***	1.227***	1.849***	1.397***	1.876***	2.710***	1.243***	1.668***	2.096***
	(0.076)	(0.152)	(0.271)	(0.114)	(0.269)	(0.360)	(0.114)	(0.273)	(0.417)
Initial 1998 log employment	–0.090***	–0.169***	–0.342***	–0.196***	–0.332***	–0.414***	–0.187***	–0.330***	–0.402***
	(0.009)	(0.013)	(0.027)	(0.011)	(0.019)	(0.042)	(0.012)	(0.022)	(0.061)
Constant	0.496***	0.008	0.964**	0.631***	–0.822**	1.223*	0.472***	–0.806**	0.371
	(0.109)	(0.198)	(0.402)	(0.145)	(0.332)	(0.711)	(0.149)	(0.345)	(0.927)
Observations	2,412	2,341	740	2,412	1,548	321	2,412	1,338	200
R-squared	0.099	0.090	0.198	0.143	0.191	0.307	0.124	0.173	0.275

Source: Authors' estimates.

Robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Comments and Discussion*

Rakesh Mohan
Yale University

Let me congratulate the authors for doing a mountain of work. It is not easy to put together these large sets of data and then, of course, to match the economic census and the population census. It has probably not been done before.

Introduction

This paper is the result of very careful work indeed, but as a discussant I will concentrate on what was not done. The authors basically want to understand the determinants of city growth, what kind of cities and towns grow in India, and how. Their approach is essentially that of data mining, which is the main question I am going to raise throughout.

It might have been more useful to first refer to the generally accepted theories of urban growth and urbanization and then to see how the data can be best utilized to test those theories. Such an approach would first include consideration of what the appropriate unit is for the analysis of urban growth. Is it cities and towns? Or should it be an economy-wide approach? After all, it is overall economic policies, economic patterns, and economic growth that determine urban growth. So it might have been useful to start at the macro level and then successively embed the economy-wide findings into a more disaggregated framework. Once one has understood the macroeconomic determinants of urban growth, one can shift one's attention to the within-the-economy distribution of that growth. It would then be useful to give some consideration to the spatial disaggregation—the key theme of the paper—relevant at the state, district or the city/town level.

The key issue is to really think of the independent or exogenous variables and the policy variables that determine the spatial distribution at each level.

* To preserve the sense of the discussions at the IPF, these discussants' comments reflect the views expressed at the IPF and do not necessarily take into account revisions to the conference version of the paper in response to these and other comments in preparing the final, revised version published in this volume. In this particular case, the paper was revised very substantially due to the availability of new economic census data. The original conference version of the paper is available on www.ncaer.org.

It is not clear which of the variables used in the paper are endogenous and which are exogenous. There is extensive literature at the global level on the exogenous and endogenous determinants of the urban growth process across countries, which could be utilized in the paper.

What Would be Useful?

The questions that the authors could address are as follows: What variables have explained the differences between economies that have exhibited higher rates of urbanization relative to others? Was it just income growth, or was it change in the structural features of the economy, such as higher relative growth in manufacturing, or in services? One could then use those insights for understanding the decadal changes in urban growth in India.

One of the interesting things that have happened in India, which the paper refers to as one of its stylized facts, is that the highest urban growth rate took place between 1971 and 1981, but, remarkably, it has been slowing down since. The chart in the paper showing Asian countries did show this slow growth of Indian urbanization. This has been a continuing surprise over the last 30 years, because it goes against our conventional understanding of urbanization. These are the 30 years when economic growth accelerated. We would have expected urbanization also to accelerate. If one understands the determinants of urban growth globally, then those insights could be used to understand the inter-state variation and the dispersion of growth across states in India. Then it would be easier to understand the city-level work that the authors have done.

As far as I understand, net rural–urban migration has also slowed down. This also goes against conventional wisdom. During the last 30 years or so, the difference between agricultural and non-agricultural wages or incomes has actually been widening. This should have led to the acceleration of rural–urban migration. But this has not happened. Once again, this is very different from most other countries. So a better understanding of such economy-wide determinants of urban and city growth would be useful in looking at the city-level findings of the paper.

In understanding the slowdown of urban population growth, one determinant is the slowdown in overall national population growth. But that does not imply that urbanization or rural–urban migration growth slows down. In fact, it could be the opposite. As urban natural population growth comes down, per capita income growth in cities and in some rural areas could accelerate. That gap has been widening, including due to rural–urban differences in family size, but urbanization has not accelerated.

One clue to this conundrum is provided by the authors from their finding on the importance of manufacturing in urban growth. One distinguishing feature of Indian development over the last three decades is the slow growth in manufacturing employment despite acceleration, except to a certain extent in recent years, of value added and output in manufacturing. This is the key issue that has come out in the paper. Have urban policies had an effect on manufacturing growth in cities? In fact, all kinds of land and labor policies have been against labor-using manufacturing. Is this issue more related to economy-wide policies or to town- and city-level variables? Which issues are more important for explaining town and city growth—those that have to do with the efficient operation of factor markets, or those related to local infrastructure and local human capital endowments? If macro, economy-wide policies are common across the country, what are the town- and city-level variables that we then want to look at?

Some Specific Issues

I will end with three specific issues. *First*, a common error, which the authors have also committed in Tables 1a and 1b, is caused by the way census data are presented. When looking at the growth in urban population by the size class of towns, it is important to recognize that towns/cities graduate to upper size class between censuses. So as you go up the size ladder, there is an increasing number of towns/cities in the larger size classes. In every size class, some cities come in and some exit to the upper level. At the top level, there is no exit. So one would always find that even if larger cities do not actually grow faster than smaller ones, they appear to grow faster as shown in Table 1.

Second, more thought needs to be given to what is exogenous and what is endogenous. For example, are new entrepreneurs and firms causing higher city growth or is there greater opportunity for new entrepreneurs in the cities that are growing faster? Which way is the causation? The authors say that if you encourage more entrepreneurs to come in, that will cause higher city growth, but I am not so sure, because if there is an "oversupply" of entrepreneurs, they could lose their shirts and growth would fall. Does higher education follow higher economic activity, or vice versa? These are just a couple of examples. The authors really have to think what the exogenous drivers of city growth are. One surprising thing is that in the service sector, public administration and defense have been excluded. One of the key determinants of city growth has always been administration and defense, which is an exogenous variable. The new capital in Andhra Pradesh, for example, will cause higher city growth.

Third, there is a surprising result: larger cities have a smaller number of higher education institutions per unit of population. It may be the case that higher education institutions are larger in larger cities. The correct measure should have been the number of students in higher education as a proportion of the city population rather than the number of colleges.

Pronab Sen

International Growth Centre, India Central Programme

Let me compliment the authors on a paper I enjoyed thoroughly. The amount of care that has gone into the data work is really impressive. I have a bit of a disadvantage. Rakesh Mohan is an urban economist, I am not. But I was glad to find that the paper actually quoted two of the things I happened to have written in two of our five-year plans, both of which actually express the sense of bewilderment about urbanization, which I think is what the authors have tried to explain. So I come from the perspective of not knowing any answers but having a lot of questions that have remained unanswered.

First, we observe that migration does not explain very much of Indian urbanization. Does it require a complete change in our theory of urbanization? Otherwise, we have taken a theory based on the migration experiences of other countries and applied it to a country where migration simply isn't happening. A simple way of getting at this, and in a sense I thought the paper would get to it, is to look at two different variables, growth in population and growth in area. The migration paradigm tells us that we start with a particular geographical area, people migrate in, we get spillovers, and neighboring areas then become urbanized. Now think of urbanization without significant migration. We have an area that has a certain number of people. The city is absorbing such areas along with their people, which is the agglomeration that is causing city growth. Migration and agglomeration are different, and we have to think about them differently. I am not sure of the appropriate theory to handle this, but I am almost certain that theories that relate wage differences and productivity differences to city growth are probably not relevant. So I would suggest that in the next iteration, the authors should try and look at the change in population density, as cities are growing as an indicator of urbanization type. In a migration-driven model, this should rise; but in an agglomeration model, this would fall.

The second point is that the authors need to be extremely careful in terms of what they say about the migrating population. What we do know from the data is that practically the entire in-migration for economic reasons is

male. Men are moving from villages into towns. However, if we look at total migration, the majority is women moving from villages to towns for marriage. The economics and the demographics are being driven by two different factors. So in looking at the composition of cities and towns, we need to keep gender in mind. We cannot and should not ignore it, because the effects are different.

The third consideration is a political one. At what stage does an area actually get counted as an urban area in terms of then changing its governance structure? What drives the period for which a census town remains a census town, instead of becoming part of a statutory town? Until we have answers to these questions, I don't think we will have a good picture of urban India, because India's urbanization, from the data that I have seen, is different from experiences elsewhere. This is something I have been puzzling about for 20 years, and I am convinced that "we are really like that only." India's urbanization is different, but we really haven't come up with a theory that is sui generis. We have taken theories developed elsewhere, we have tried to shoehorn the data to fit those theories, and the net result is that we keep getting more puzzles rather than enlightenment.

General Discussion

Karthik Muralidharan observed that the paper focused on urbanization from the perspective of its contribution to the growth in jobs and industries, but it is also the case that service delivery is much cheaper in more densely populated areas. He wondered if in a future extension of the paper, selecting a few services and holding their quality constant, if the authors could measure and benchmark how much cheaper it is to deliver services as a function of spatial density. There is also a lot of value from the diversity in the consumption basket that comes about with greater spatial density.

T.N. Srinivasan agreed with Rakesh Mohan in expressing concern about the need to distinguish more clearly between endogenous and exogenous determinants of urban growth. Second, he thought there should be a clearer distinction between birthplace migration and other forms of migration in discussing urbanization.

Devesh Kapur noted that every micro village study reports high rates of migration, whereas every macro study reports little or no migration. How do we understand this severe mismatch? He suggested a greater focus on mobility, as opposed to migration. For example, improved roads have made it possible to travel over much greater distances on a daily basis. Migration

is a process where people were born in place X, and now live in place Y. His research suggested that people often hide their birthplace and report that they have always lived in their current place. He suggested greater caution in using census data as the primary measure of migration.

Rohini Pande applauded the creation of the dataset, which is a lot of work, and looked forward to its future use. It has long been thought that government policies are pro-rural and urbanization has not had a level playing field. For example, the introduction of MGNREGA between the two censuses changed the incentives to move to urban areas. Some of the work around the way MGNREGA was introduced and the identification strategies that this made possible could help to get a handle on the migration data in the paper: work by Clément Imbert may be relevant here. Second, she pointed to a major difference between India and most other countries—the urban female labor force participation rate is lower in urban areas as compared to rural areas—that would have an impact on urbanization.

Suman Bery suggested that the pattern of energy consumption was an important dimension of the urbanization process and could be included in the study, especially using geospatial data. Second, he suggested that while this study relied almost exclusively on census data, NCAER might consider collecting sample survey data that might usefully complement this work and could provide additional insights.

Anant Sudarshan pointed to the analysis of night light data as indicating a more continuous spatial distribution of the population, as opposed to clustering in clearly defined urban areas or cities. He cited the night light data from Delhi to Lahore as an example. In evaluating the process of urban agglomeration, it may be important for the paper to take account of growing road networks and the size of the city units that are being defined as urban.

Dilip Mookherjee liked the emphasis on new firms and entrepreneurship. He thought they were critical to understanding the growth of cities. But he also thought that it would be difficult to infer very much about entrepreneurship from the information in the census. It might be possible to resolve some of the issues around endogeneity perhaps using a vector-auto regression approach.

Anupam Khanna thought it would be useful to compare this work with prior studies by Chinmaya Tumbe at IIM Ahmedabad. This work also focused on migration and urbanization processes, and argued that there were important regional differences between the North and South. In addition, some prior work with which he was involved in NASSCOM suggested that the IT industry played a major role in the growth of the largest metropolitan areas after about 2004–05.

Rajnish Mehra expressed some confusion about whether the authors were referring to the growth in the number of small cities or the growth rate of small cities. The literature suggests that the growth rate of cities is independent of size, in line with Gibrat's law. Second, it suggests that the growth rate of cities is not continuous in size and that when a city comes up to a certain size—say, when it has an airport—it will move up to the next size. Hence, there is a need to control for the phase of the transition that a city's growth is taking place in.

Shekhar Shah suggested that the challenges on the data side, particularly on the census, that are obvious from this paper should lead to a discussion of how the next census could be supplemented. It would be very useful for the paper to highlight suggestions in this regard.

B.N. Goldar noted that the Economic Census data has a lot of problems, and compared to that the Annual Survey of Industries (ASI) data is much more reliable. ASI data by cities may not be easily available to researchers, but the sample frame contains the entire universe of firms with 10+ workers, and if there is a focus on manufacturing, then it might be worth looking into whether the frame could be used.

K.P. Krishnan (chair) cautioned the authors about the data on urban infrastructures. Urban infrastructure data is often distorted by the funding objectives of urban local bodies that report the information, and the data are often adjusted to suit the funding purpose. Second, city growth is measured in the paper by population growth and not GDP growth. But city growth may not come from changes in economic structure leading to in-migration but from changes in the activities of existing workers, so economic structure may only be a partial explanation for city growth. Third, he thought it is always difficult to identify the precise sources of urban agglomeration, since the impetus can come from a number of factors. His sense was that politics always has a major role in urban agglomeration. He cited the example of the dramatic growth of Bengaluru during 1998–99 to 2004–05. The growth took place in eight peripheral areas outside the main city, and this was entirely driven by the politics of urban governance.