THE SPATIAL DEVELOPMENT OF INDIA[*]

Klaus Desmet   Ejaz Ghani   Stephen O'Connell   Esteban Rossi-Hansberg

September 2013

Abstract

This paper studies the recent spatial development of India. Services, and to a lesser extent manufacturing, are increasingly concentrating in high-density clusters. This stands in contrast with the United States, where in the last decades services have tended to grow fastest in medium-density locations, such as Silicon Valley. India's experience is not common to all fast-growing developing economies. The spatial growth pattern of China looks more similar to that in the U.S. than to that of India. Our findings suggest that certain frictions are keeping medium-density places in India from growing faster.

JEL Classification: O1, O14, O18, O53, R11, R12.

Keywords: Spatial Growth; Economic Development; Manufacturing; Services; Economic Geography; India.

1 INTRODUCTION

In the last two decades the Indian economy has been growing unabatedly, with memories of the *Hindu rate of growth* rapidly fading. But that development has led to widening spatial disparities. While cities such as Hyderabad have emerged as major clusters of high development, certain rural areas have been left behind. India's mega-cities have continued to grow, fed by a steady stream of migrants from the countryside. This situation raises a number of important policy questions. Should India aim to spread development more equally across space? Are India's cities becoming too large? Should the government invest in infrastructure in the large cities to reduce congestion or in medium-sized locations to facilitate the emergence of new economic clusters?

Though such spatial inequalities are not unfamiliar from other countries, there is a relevant difference: India's growth has mainly stemmed from a rapidly expanding service sector. This

is important in the light of work by Desmet and Rossi-Hansberg (2009) who have shown that manufacturing and services exhibit very different spatial growth patterns in the U.S. and Europe. In recent decades U.S. and European manufacturing has been dispersing from high-density clusters to less dense areas, whereas services have been experiencing increasing concentration, except for the densest locations where congestion is the dominating force.

Desmet and Rossi-Hansberg (2009) relate these opposing patterns to the differential impact of ICT (Information and Communication Technology) on both sectors. They argue that the diffusion of general purpose technologies, such as ICT, leads to knowledge spillovers that are enhanced by spatial concentration and the emergence of high-density clusters of economic activity. In recent decades we have seen this phenomenon unfold mainly in services, as ICT is disproportionately benefitting that sector. More generally, they suggest that an industry's spatial growth pattern is determined by its "age", defined as the time since the industry was last impacted by a general purpose technology. By that token, in the U.S. today, services can be viewed as "young" and manufacturing as "old", since the last GPT to have benefitted manufacturing was electrification at the beginning of the 20th century. Consistent with this view, in the U.S. the spatial growth pattern in manufacturing in the first decades of the 20th century, when that sector was young, looked very similar to the spatial growth pattern in services at the end of the 20th century. Given the success of this theory in explaining the different spatial growth patterns of manufacturing and services in the U.S. and Europe, we use it as a lens through which to interpret our findings for India.

A first question, then, is whether India exhibits the same distinction between manufacturing and services as the U.S. or Europe. In India we find evidence of increasing spatial concentration in the service sector, and to a lesser extent, in the manufacturing sector. In light of our theory, this suggests services are "young", whereas manufacturing is not as "mature" as in the U.S. or Europe. The service sector in India is clearly benefitting from ICT, so that its increasing spatial concentration is what we would expect. The finding that manufacturing is younger than in the U.S. should not come as a surprise either. Recall that in the U.S. manufacturing only started dispersing in the post-World War II period. In addition, the technology shocks that determine the age of an industry should not be interpreted in a narrow sense. For example, the delicensing of the Indian manufacturing industry, which was completed at the beginning of the 1990s, has been leading to a geographic deconcentration of manufacturing activity (Fernandes and Sharma, 2012).

A second question is whether the tradeoff between agglomeration economies and congestion costs in India is similar to the one in the U.S. or Europe. Casual observation suggests that the costs of congestion in India's mega-cities are huge, implying that there should be decreasing returns to further expansion. However, these mega-cities may also benefit from relatively large agglomeration economies, compared to medium-sized cities that might suffer from market access problems, lack of intermediate goods and local infrastructure, and other impediments to growth. In the developed world this problem may be less severe, thus providing growth opportunities to medium-sized locations that are not present in India.

Comparing the service sectors of the U.S. and India, we indeed find that intermediate-density locations experience faster growth in the U.S., whereas the large mega-cities are the winners in India. In particular, in the U.S. the service clusters that experience fastest growth have a service employment density in the range of 50 to 150 employees per square kilometer. This is the case of three of the most well known high-tech clusters: California's Silicon Valley, Boston's Route 128 and the North Carolina Research Triangle. In contrast, the same medium-density locations in India fare poorly in terms of attracting employment growth, and agglomeration economies seem strongest in service clusters with employment density reaching into the thousands.

This finding is not common to all emerging economies. Although for want of high quality sectoral employment data at the local level we refrain from an in-depth study of China, our exploratory analysis suggests that China looks more similar to the U.S. in that decreasing returns dominate in high-density cities. In both China and the U.S., agglomeration economies are strongest in medium-density clusters, quite different from India. If we view the U.S. as the efficient benchmark, this suggests that India's medium-density locations are facing certain barriers and frictions that keep them from growing faster. When exploring what those barriers might be, we find that many of the possible culprits, such as the distance to large cities or the access to certain basic utilities, can be ruled out. The two that we cannot discard as being important in explaining the relative advantage of high-density clusters are the percentage of highly educated and the household access to telecommunication services. Although both variables are at least partly endogenous, they give some indication of where future research should focus if it wants to convincingly identify the barriers to growth that afflict medium-density locations.

The rest of the paper is organized as follows. Section 2 briefly summarizes the spatial growth

model of Desmet and Rossi-Hansberg (2009) to provide a framework to interpret our results. Section 3 describes the data. Section 4 analyzes spatial development in India, and compares it to the U.S. and China. Section 5 concludes.

2   A SIMPLE CONCEPTUAL FRAMEWORK

Before we present our empirical findings, it is important to have a theoretical lens through which we can interpret the results. Desmet and Rossi-Hansberg (2009) provide a theory of the spatial evolution of economic activity in which the relationship between local employment growth in an industry and the density of employment in that location is the result of three main forces. First, technological diffusion leads to geographic dispersion of economic activity. Low productivity, and thus less dense, locations benefit more from this effect. Second, knowledge spillovers, pecuniary externalities, labor market pooling, etc., all facilitated by high density, constitute an agglomeration force that leads to geographic concentration of employment. Third, congestion costs in large locations, due to high transport costs, pollution and local fixed factors, constitute an additional dispersion force.

Looking at the relationship between employment growth and employment density can then reveal which of these forces dominates for particular levels of density. For example, a declining relationship between employment growth rates and employment densities would imply that the two dispersion forces dominate the agglomeration force. As argued before, this will tend to be the case in "old" industries. In contrast, if the observed relationship is positive, it implies that the agglomeration force dominates the dispersion forces. This will tend to be the case in "young" industries. We would expect the relationship between employment density and employment growth to vary across locations of different densities, across sectors, and across time. This, then, describes an economy's spatial evolution.

We now briefly sketch the main elements of the Desmet and Rossi-Hansberg (2009) framework. We limit ourselves to highlighting those elements that will help us interpret our empirical results. For more details we refer the interested reader to the original paper. The economy consist of a continuum of locations in a closed interval $[0, 1]$ with one unit of land at each location ℓ. There are two sectors, manufacturing and services, indexed by a subscript i. Workers are freely mobile

and work where they live. Agents have CES preferences over manufactured goods and services. Firms are perfectly competitive, and use land and labor to produce. The production function is Cobb-Douglas with the share of labor μ_i less than one, so that land acts as a congestion force. TFP is sector- and location-specific and changes over time.

The economy starts off with an initial spatial distribution of TFP in each industry. This determines the initial distribution of sectoral employment across space. Productivity then evolves across time and space because of both technology diffusion and knowledge spillovers. On the one hand, the best practice technology from period $t-1$, defined as $\bar{Z}_i^{\max}(t-1) = \max_\ell Z_i(\ell, t-1)$, diffuses (imperfectly) across space by period t. On the other hand, a location ℓ benefits from knowledge spillovers, $(\int_0^1 e^{-\delta_i|\ell-r|} L_i(r,t)\, dr)^\gamma$, a function of the weighted average of employment L_i at all locations, where the weights decline with distance from ℓ. We set $\gamma + \mu_i < 1$ so that there is local congestion, thus ensuring that not all employment concentrates in one location.

Combining technology diffusion and knowledge spillovers, TFP in sector i and time t is assumed to take the form of

$$
(1) \qquad Z_i(\ell, t) = \max \left[\begin{array}{c} \rho \bar{Z}_i^{\max}(t-1) + (1-\rho)\, (\int_0^1 e^{-\delta_i|\ell-r|} L_i(r,t)\, \theta_i(r,t)\, dr)^\gamma, \\ (\int_0^1 e^{-\delta_i|\ell-r|} L_i(r,t)\, \theta_i(r,t)\, dr)^\gamma \end{array} \right],
$$

where $\rho \in [0,1]$ determines the ease of diffusion. To see the role of ρ, consider the two extremes. If $\rho = 0$, there is no diffusion whatsoever, whereas if $\rho = 1$, the technology of the most productive place today becomes freely available in all locations tomorrow.

The above expression implies that the relative importance of diffusion and knowledge spillovers will differ across different types of locations. In low-density areas, knowledge spillovers are weak, and TFP will be dominated by diffusion. In medium- and high-density areas, knowledge spillovers play an increasingly important role. In fact, when knowledge spillovers become strong enough to imply a higher TFP than the (imperfect) access to last period's best practice technology, diffusion ceases to play a role. Lastly, in the highest-density places local congestion starts to dominate, negatively impacting TFP.

The relative importance of diffusion and knowledge also differs with an industry's age. To describe the type of spatial dynamics the model implies, an example may be useful. Suppose a

new general purpose technology, such as ICT, is introduced in the service sector. Following our definition, this makes services "young". Being a new technology, it has not yet had the time to geographically diffuse, and its impact is therefore initially limited to a few areas. Because of their high TFP, those select locations will have a high concentration of services employment. The low- and medium-density areas will benefit from the diffusion of ICT. Because land is costly, the lowest-density areas will grow fastest. As a result, for those locations where diffusion is the dominating force, we will witness spatial dispersion. In high-density areas, in contrast, TFP will mostly depend on local knowledge spillovers. This leads to spatial clustering, as high-density locations attract employment in nearby places. As these neighbors mutually benefit from each other's knowledge spillovers, spatial concentration is strengthened, with high density locations growing faster. Of course, because of local congestion, the locations with highest density may grow slightly slower.

Therefore, when a new general purpose technology is introduced, the model implies an S-shaped relation between employment density and employment growth. Because of technology diffusion, low-density locations grow faster than medium-density locations (downward-sloping first part of S-shape); because of knowledge spillovers, high-density locations grow faster than medium-density locations (upward-sloping second part of S-shape); because of local congestion, the highest density locations grow somewhat slower (downward-sloping third part of S-shape).

Of course, how these different forces play out, and therefore how important the different parts of the S-shaped relation are, depends on an economy's specific conditions. A couple of examples may help to highlight the possible variations in the relation between density and growth. As a first illustration, suppose technological differences across space are very large, as is the case in many developing countries. In that case the scope for diffusion is large, so that we would expect spatial dispersion in low- and medium-density locations to be stronger. The downward-sloping part of the S-shaped curve would be steeper and extend to the high-medium-density locations. As a second illustration, suppose medium-density places have not made the necessary productive investments to be able to take advantage of knowledge spillovers. Those frictions would put the highest-density places at a relative advantage, implying that people would concentrate too much in large, dense cities. As a result, the S-shaped curve would not turn downward for those highest density places, which would continue to enjoy relatively high growth rates. Another reason for why we may witness an absence of congestion is because we are in the very early stages of clustering.

6

The highest-density locations are still attracting economic activity to nearby areas. The increasing density of those close-by areas benefits the highest-density locations, which continue to grow fast.

Over time, as ICT matures, and services become "old", two things happen. First, knowledge spillovers weaken in the high-density clusters, making congestion increasingly important. The S-shape relation between density and growth weakens and becomes downward-sloping. We get geographic dispersion as economic activity becomes more equally spread across all locations. Second, the relation between density and growth flattens. This reflects the fact that technological differences across space drop, thus reducing the scope for technological diffusion. Taken together, the theory therefore predicts that in the decades after the introduction of ICT — when services were "young" according to our definition — we should find evidence of geographic concentration over at least part of the distribution. In the subsequent decades, as ICT matures — and services become "old" according to our definition — the incentive for concentration should disappear, and we should see dispersion, which gradually weakens, as activity becomes more equally spread over space.

The explanatory power of the theory to account for the different spatial growth patterns of manufacturing and services throughout the 20th century in both the U.S. and Europe forms the motivation to use this framework to interpret our results for India. Before doing so, a couple of remarks are in place. First, we do not have a precise way of determining the age of an industry in India. Whereas in the service sector it is clear that India has been benefitting from ICT, so that services can be thought of as being young, it is less obvious how to determine the age of India's manufacturing sector. India's structural transformation out of agriculture is still under way, so that its manufacturing sector is probably younger than the one of the U.S. Moreover, certain policy shocks, such as the end of the "License Raj" at the beginning of the 1990s, have had a profound impact on productivity in manufacturing, and could further contribute to making manufacturing younger. Given the difficulty of relying on some exogenous measure to determine the age of manufacturing in India, we will use the theory as a way of identifying the age, rather than using the age to test the theory. In that sense our analysis does not constitute a test of the theory. Rather, we use our theory of spatial growth as a lens through which to interpret the results. Second, independently of the question of age, our framework provides a useful tool to help us think about the different agglomeration and congestion forces that affect the scale dependence of

7

growth. For example, if we observe high-density places growing faster than medium-density places, it must be because knowledge spillovers are stronger than congestion costs, or if medium-density places grow slower than low-density places, it must be because technology diffusion is greater than agglomeration forces.

3 DATA

To study employment dynamics across space in India, a first issue is to decide on the level of spatial disaggregation at which we have reliable data. India is divided into 35 states (or union territories) and 640 districts. While certainly the quality of the data is more reliable at the state than at the district level, having a high degree of spatial disaggregation is important. Indeed, agglomeration economies and congestion effects may get lost at higher levels of aggregation, so that focusing on districts is better. In addition, having a broad distribution of places (going from small to intermediate to large) is also important, since previous work for the U.S. has shown that the scale-dependence of growth may be non-linear. There is of course a tradeoff to be faced. By going to the district level, we need to keep the sectoral information at a high level of aggregation to keep the data from becoming less precise. Findings for the U.S. and Europe suggest that going to a finer spatial level is more relevant than going to a finer sectoral level. We follow this evidence and therefore focus on two broad sectors, manufacturing and services, at the district level.

India does not directly provide comprehensive manufacturing and services employment data at the district level. We therefore rely on micro-data from surveys. India runs two firm-level surveys, the Annual Survey of Industries (ASI) and the one conducted by the National Sample Survey Organisation (NSSO). The ASI survey has information on the so-called organized manufacturing sector (essentially comprising of firms with more than 10 workers), whereas the NSSO covers the unorganized manufacturing sector and the services sector. Both surveys, the ASI and the NSSO, overlap for the fiscal years 1989-90, 1994-95, 2000-01 and 2005-06. However, the service sector has only been surveyed more recently, in fiscal years 2001-02 and 2006-07. Given that part of our focus will be on the difference between manufacturing and services, we will use 2000-2005 for manufacturing and 2001-2006 for services.

For the case of manufacturing, the ASI covers all registered factories, and uses a sampling

frame that is stratified at the state and the four-digit National Industry Classification (NIC) level. We complement these data by the NSSO which covers all unorganized manufacturing enterprises. In the case of the NSSO the sample stratification is more sophisticated and includes both the district level and the two-digit NIC sectors. For the case of services, the NSSO follows a similar stratification, inluding the district and the two-digit NIC sectors. Note, however, that some service subsectors, such as retail, wholesale and financial services, are excluded in at least one of the two available years.[1] Furthermore, as the NIC definitions have changed over time, we make them consistent using concordances that come with the data.[2]

Sampling weights provided by the separate survey datasets are then applied to create estimates of total employment by district and sector. One obvious issue regards the reliability of this procedure and possible measurement error. To address this issue, we do a number of robustness checks. In particular, we complement our estimation of district-level sectoral employment from firm surveys by an alternative measure using the Employment-Unemployment Survey, frequently referred to as the Labor Force Survey (LFS), carried out by the NSSO in fiscal years 1999-2000 and 2004-05. This survey collects individual-level information on location, occupational status and industry of occupation, detailed enough to allow an estimation of employment by NIC industry and district. The sample stratification is similar to that of the NSSO. We run robustness checks using both major sources of data, the one based on firm-level surveys and the other based on individual-level surveys.

A last concern is that sometimes districts have been redefined, combined, or split. For these types of changes we follow a simple strategy for assuring consistent district definitions over time. In the case of a single district being divided into two or more new districts, we recreate the original district by combining the new districts (backward-compatibility). When two or more previous districts are combined, we recreate the new combined districts in the earlier years (forward-compatibility). In the case of transfers of land between districts we combine the districts involved in all periods.

4 THE SPATIAL DEVELOPMENT OF INDIA

This section analyzes the spatial evolution of employment in India. Although most research on India has focused on the manufacturing sector, we will distinguish between manufacturing and services for two reasons. First, given the emergence of India as a service-based economy, it is key to understand which types of locations are benefiting from the country's structural transformation (Ghani, 2010). Second, as already pointed out, the work by Desmet and Rossi-Hansberg (2009) has documented important differences between the spatial dynamics of manufacturing and services in the U.S. We want to see whether the same patterns show up in India.

Given that the theory emphasizes the importance of possible nonlinearities in the scale-dependence of growth, we run nonlinear kernel regressions of the form

$$(2) \qquad\qquad L_i(\ell, t+s) = \phi(L_i(\ell, t)) + \varepsilon_i(\ell, t),$$

where $L_i(\ell, t)$ is the log of sectoral employment density in year t, district ℓ and sector i. The estimation uses an Epanechnikov kernel with bandwidth 0.8.[3] Because the distribution of employment density levels is approximately log-normal, we focus on the log of employment density. To facilitate interpretation, in the figures we will plot annual employment growth as a function of initial log employment density in the same industry. In this case, a negative slope indicates geographic dispersion (convergence) and a positive slope indicates geographic concentration (divergence).

Scale Dependence in India

Figure 1 shows annual manufacturing employment growth as a function of initial manufacturing employment density (in logs). In this benchmark exercise the employment data at the district level have been constructed from firm-level surveys (NSSO and ASI). The picture suggests that manufacturing is dispersing through space. Low-density manufacturing districts are growing faster than high-density manufacturing districts. Note, however, that the 95 percent confidence intervals are extremely large in the upper tail, suggesting a rather weak relation between scale and growth for high-density locations. Indeed, as can be seen from the bottom panel of Figure 1, some of the large cities, such as Kolkata and Mumbai, are experiencing higher growth than that predicted by the kernel regression.

Services show a distinctly different pattern. As can be seen from Figure 2, although low and medium-density service locations exhibit spatial dispersion, for the high-density service locations we observe increasing concentration. That is, the high-density service clusters are gaining relative to those locations with slightly lower employment density. The bottom panel of Figure 2 shows that many of the well-known IT clusters are in the upward-sloping part of the estimated relation, suggesting that they continue to benefit from agglomeration economies. For example, service employment in Hyderabad and Chennai is growing at an annual rate of, respectively, 11 percent and 4 percent. If we were to run a simple regression, the predicted growth rate of these two cities would be, respectively, -7.1 percent and -8.2 percent. This underscores the importance of taking into account non-linearities in the scale-dependence of growth. Note that the upward-sloping part is also driven by some of the country's largest cities, such as Mumbai. But not all large cities exhibit high growth in services, as illustrated by Delhi.

One concern with the service data from the NSSO is that large service firms are underrepresented in 2001 whereas they are not in 2006 (Dehejia and Panagariya, 2010). One way to check whether this introduces a bias in our results is to make both years more comparable by leaving out all service firms above a certain employment threshold. Using two alternative thresholds of 500 and 2500 employees, the results are unchanged: services continue to become more concentrated in high-density clusters.

Another way of dealing with this and other potential drawbacks of the firm-level data is to construct employment figures from the Labor Force Survey. Figures 3 and 4 show the results of re-running the same kernel regressions, using sectoral employment at the district level from the LFS. In the case of services, we confirm our previous findings: there is clear evidence of increasing concentration in the upper tail. However, for manufacturing the results look somewhat different. While in Figure 1 we observed spatial dispersion throughout the distribution (though insignificant in the upper-tail), we now find, as in services, evidence of spatial concentration for high-density manufacturing clusters. This is consistent with our observation that some of the large cities continue to experience relatively strong manufacturing employment growth. According to the LFS, Kolkata, for example, is growing at an annual rate of 4.8 percent.

To further compare the results from the NSSO and the LFS, we run a number of additional checks by, for example, taking the average of district employment coming from the NSSO and the

LFS, or by dropping all observations for which the difference in growth rates in the NSSO and the LFS is above a certain threshold. Doing so confirms the strong evidence of agglomeration economies for high-density service clusters, and the weaker evidence of the same phenomenon in the manufacturing sector.

So far our analysis has focused on manufacturing and services as a whole. A reasonable question may be whether our results are driven by particular subgroups of manufacturing or services firms. To answer this question, we split up our sample in two different ways. In a first exercise we distinguish between the formal and the informal sector, referred to in India as the organized and the unorganized sector. This distinction may be relevant, since firms in the informal sector are less subject to laws and regulations, and thus perhaps more free to operate and choose their location.[4] Differentiating between the unorganized and the organized sectors does not change our main finding. The service sector is becoming increasingly concentrated in high-density clusters, whereas in manufacturing the picture is more mixed, with the unorganized sector becoming more concentrated and the organized sector becoming more dispersed.[5]

In a second exercise we redo our kernel regression for all 2-digit sectors for which we have sufficient data (22 manufacturing subsectors and 12 services subsectors). As mentioned in Section 3, given that we rely on survey data, sectorally disaggregating at the district level risks making our measures of employment density less precise, so that these results are statistically less robust. Still, we find that our main result — services, and to a lesser extent manufacturing, are becoming increasingly concentrated in high-density clusters — is not driven by a few subsectors. In the case of services, around 90 percent of employment is in subsectors that exhibit increasing concentration in high-density clusters, whereas the corresponding figure in manufacturing is around 60 percent.[6]

In our theory employment changes are closely related to productivity changes. We therefore would like to see whether similar growth patterns show up when analyzing productivity instead of employment. Figures 5 and 6 show the relation between labor productivity growth and employment density. As can be seen, in the service sector productivity growth increases with density, thus reinforcing the finding that high-density areas is where growth is happening. In manufacturing, however, productivity growth peaks in medium-density locations, consistent with a weaker link between density and growth in that sector.

The strong evidence of agglomeration economies in the service sector is consistent with

findings for the U.S. and Europe. Given the impact of ICT in India's rapidly growing service sector, this is what we would have expected. Being a "young" industry, services benefit from knowledge spillovers, leading to the emergence of high-density service clusters. In contrast, the evidence for such agglomeration economies in manufacturing, though weaker than in services, differs from the tendency towards dispersion across the entire distribution in the case of the U.S. and Europe. In light of our theory, this suggests that manufacturing in India is not as mature as in the U.S. or Europe. Recall that any shock that has an important impact on an industry's productivity — whether a GPT or a change in policy — may make an industry "young". One such important shock to India's manufacturing industry was the end of the so-called "License Raj," a system that required manufacturing firms to apply for a license to operate. Delicensing started in the 1980s and was largely completed by the beginning of the 1990s (Aghion et al., 2008). The elimination of these distortions unleashed productivity growth in India's manufacturing sector. According to Fernandes and Sharma (2012), delicensing has led to manufacturing becoming spatially more disperse. One reason they offer is that the "License Raj" artificially created inefficient clusters which are now breaking up.[7]

Comparing Services in India, the U.S. and China

Although the service sector in India shows some similarities with the service sector in the U.S. — both exhibit agglomeration economies — there are also some relevant differences. Focusing on U.S. counties, Figure 7 shows annual employment growth in services between 1980 and 2000 as a function of initial employment density in services in 1980.[8] Comparing Figure 7 with Figure 2, it becomes apparent that in the U.S. agglomeration economies in services dominate for medium-density locations, whereas in India agglomeration economies dominate for high-density locations.[9] In particular, Figure 7 shows that agglomeration economies in the U.S. service sector peak at a density of between 50 and 150 employees per square kilometer. Three of the main high-tech counties in the U.S. fall within that range: Santa Clara, Calif. (Silicon Valley), Middlesex, Mass. (Route 128) and Durham, NC (Research Triangle). In contrast, in India, Figure 2 shows that agglomeration economies increase in the upper tail of the distribution, in places such as Hyderabad and Chennai, with service employment densities reaching into the thousands. For those levels of density, U.S. locations exhibit substantial congestion.

When looking at these findings through the lens of the theory of Desmet and Rossi-Hansberg (2009), there are several possible interpretations. First, high-density locations suffer from local congestion, but benefit from the knowledge spillovers from nearby locations. As long as those neighboring areas gain in employment, spillovers continue to strengthen, thus allowing the high-density locations to grow at a fast pace. Therefore, in the early stages of spatial clustering, knowledge spillovers are likely to dominate congestion, even in the highest-density districts. However, given that the highest-density districts in India are in general denser than the highest-density counties in the U.S., this is an unlikely explanation. Second, it might be the case that the high-density clusters in India are more successful, not because its mega-cities are not congested, but because of the absence of agglomeration economies in medium-sized locations, implying higher-than-normal congestion in those places. Certain policies or frictions, such as a lack of general infrastructure, may prevent these medium-sized cities from growing faster.

Precisely identifying what those frictions or barriers might be goes beyond the scope of this paper. However, we can obtain some suggestive evidence by controlling for certain district-level characteristics. We run separate conditional kernel regressions,[10] controlling for (i) the percentage of the population with a high school degree or more and the percentage of the population with post-secondary education; (ii) household access to infrastructure (percentage of households with electricity, percentage of households with toilet, percentage of households with telecommunication services, percentage of households with tap water); (iii) travel time to a top-10 city; and (iv) distance to a top-7 or a top-3 city.[11]

When exploring which of these controls can explain the advantage of high-density clusters, we can rule out most. For example, being close to a major city or having access to some of the basic utilities, such as tap water or toilets, do not seem to matter. Only two variables, the percentage of the population with post-secondary education and the percentage of households with access to telecommunication services, have the potential of accounting for the relative advantage of high-density clusters. As can be seen in Figure 8, once we control for either of these two variables, there is no longer evidence of high-density service clusters growing particularly fast. In other words, if all locations had the same percentage of their population with post-secondary education or if in all locations the households' access to telecommunication services were the same, then high-density service clusters would lose their attractiveness.

The controls we are using, with maybe the exception of the distance to a major city, can in general not be viewed as exogenous characteristics of a location. For those variables that are unable to explain the advantage of high-density clusters, such as the access to tap water or toilets, their potential endogeneity is not a problem: if those controls do not matter in the presence of endogeneity, they are even less likely to matter if their effect is purged from endogeneity. For the only two controls that do seem to explain the advantage of high-density clusters, endogeneity is more of a concern. It is only to the extent that the choice of residence and the quality of telecommunication infrastructure are exogenous to a location's density that we are identifying a causal channel. Given that there are good reasons to believe that this need not always be the case, more work is needed to convincingly identify the exact growth barriers afflicting medium-density locations. However, by being able to discard many of the standard suspects, such as proximity to large urban centers, our analysis provides a strong indication of what those possible barriers are likely to be.

If part of the worse performance of India's medium-density locations is their deficient local infrastructure, it may be useful to compare India's experience, not just to that of the U.S., but also to that of the other large emerging economy, China. Figure 9 compares India and the U.S., whereas Figure 10 compares India and China. Before discussing the results, a word of caution about the data we use for China: the employment figures measure the number of "staff and workers", also referred to as "formal employment", rather than total employment. This leads to underreporting of employment, especially in rural areas, as it excludes, amongst others, workers employed in township and village enterprises. In as far as the share of staff and workers in total employment is not orthogonal to size, this will introduce a bias in our results.[12] Subject to this caveat, Figures 9 and 10 show that China looks very different from India. Once a threshold of around 150 employees per square kilometer is reached,[13] agglomeration economies start dominating in India, whereas the opposite happens in China. For Chinese locations with a density above 150 employees per square kilometer, service employment growth becomes strongly decreasing with size, indicating important congestion costs.[14] Along that dimension, China looks more like the U.S., where congestion costs also dominate for locations above the 150 employees per square kilometer threshold. Given that the overall level of local infrastructure is better in China than in India, this finding is consistent with the interpretation of frictions holding back the growth of medium-density locations in India,

but not in China.

Although in terms of the tradeoff between agglomeration economies and congestion costs in high-density places China and the U.S. look similar (and different from India), there is another dimension along which the U.S. looks different from both China and India. As can be seen from Figure 9 and Figure 10, the difference in growth rates between fast-growing places and slow-growing places in India and China is much larger than in the U.S. This is not because the differences in growth rates across space are irrelevant in the U.S., but because they are enormous in India and China. In those countries the difference in annual employment growth rates between fast-growing locations and slow-growing locations is greater than 20 percentage points. Compared to the U.S., the spatial distribution of economic activity in both India and China is changing at a break-neck pace. This underscores the importance of this type of spatial analysis in developing economies. As discussed in our theory, this difference might well be related to the greater differences in productivity across space in China and India, compared to the U.S. This gives more scope to technology diffusion, allowing low-density places to grow much faster than medium- or even high-density places.

If we view the U.S. as the efficient benchmark, it is interesting to see how spatial growth would look like in India if it had the same relationship between density and growth as the U.S.. This counterfactual exercise is represented in Figure 11, where the left-hand panel shows the predicted growth rates of Indian districts, based on the estimates for India, and the right-hand panel shows the counterfactual growth rates of Indian districts, based on the estimates for the U.S.[15] When comparing the maps, two features stand out. First, many of the relatively slow-growing Indian districts would grow much faster. These correspond to medium-density places, similar in density to places such as Silicon Valley. As mentioned before, with few exceptions, these districts in India do not seem to be able to take advantage of the service revolution. Second, if India had the same scale dependence in growth rates as the U.S., different areas of the country would benefit from growth in the service sector. Growth would be more concentrated in the coastal regions, especially in Southern states such as Tamil Nadu and Kerala, as well as in Northern states such as West Bengal, Bihar and Uttar Pradesh. Of the well known IT clusters in India, the medium-density places such as Ahmedabad and Pune, and especially Bangalore, have high growth rates in the counterfactual, whereas the high-density places, such as Chennai and Mumbai, do not.

5 CONCLUSIONS

This paper has analyzed the recent spatial evolution of manufacturing and services in India. To guide the analysis and to interpret the results, we have relied on the spatial theory of Desmet and Rossi-Hansberg (2009), which argues that spatial growth patterns depend on an industry's age. In particular, the theory predicts that "young" industries tend to become spatially more concentrated, whereas "old" industries have a tendency towards greater dispersion. Evidence for the U.S. and Europe has suggested that an industry becomes "young" when it is impacted by a GPT. Given the important role played by ICT in India's service revolution, this theory provides a good lens through which we can interpret spatial growth patterns. Since the theory says that the relation between density and growth may be non-linear, we have followed the methodology previously employed to analyze the U.S. and Europe, and used kernel regressions to evaluate the spatial growth patterns in different industries. An additional advantage of using the same methodology is that it has allowed us to compare India's spatial development to that of the U.S.

The evidence we have provided and the accompanying theory that helps us interpret it suggest that the spatial evolution of India continues to favor districts with high levels of employment density. This is especially the case in services. Given the role played by ICT, this is consistent with services being a "young" sector. The evidence in manufacturing is more mixed, and depends on the particular dataset we use. Overall, our findings demonstrate that in service sectors agglomeration forces still dominate dispersion forces in high density areas. In other words, these high density clusters of economic activity continue to be India's engines of growth.

The above conclusion confronts us with a policy dilemma. Should India focus the development of urban infrastructure, and in general facilitate the location of employment, in its large cities in order to exploit the still important agglomeration effects? Or should India develop infrastructure in medium-density locations in order to remove some of the impediments of growth present in these areas?

To shed light on these questions we have compared the experience of the U.S. with that of India. The results are striking in that the evidence of agglomeration in the U.S. service sector is all concentrated in locations with densities of employment *below* 150 employees per square kilometer, while in India the evidence of agglomeration is found in locations with densities *above* this

17

threshold.[16] In other words, if we take the U.S. as the efficient benchmark, then 150 employees per square kilometer is the *ideal* density to take advantage of agglomeration economies. In India, however, these medium-density location are the worst places. This suggests that the costs of congestion in India are either much smaller than in the U.S., the agglomeration forces are much larger than in the U.S., or that there are some frictions, policies, and a general lack of infrastructure in medium-density cities that prevents them from growing faster, therefore favoring concentration in high-density areas. It is not obvious to us why Indian individuals should dislike congestion less than Americans or should benefit more than Americans from agglomeration economies. These forces seem to be more technological and universal. Therefore, the likely culprits are restrictions to economic growth in intermediate-density cities or districts. In this context, we have found some evidence suggesting that two such barriers may be the low level of highly educated and the deficient local infrastructure, in particular poor access to telecommunication services. Our findings for China, an emerging economy that has suffered less from a lack of infrastructure, support this interpretation. Similar to the U.S., congestion in the Chinese service sector is strong in locations with high employment density.

What is therefore preventing medium-density locations in India from growing and taking full advantage of agglomeration forces? Why is their evolution, relative to low- and high-density areas, so different from that in advanced economies? This paper identifies this specific issue as a major question in India's spatial development. Although conditioning on local characteristics points to some possible explanations and discards others, having confident answers to what the sources of these distortions are will lead to better informed and more effective urban and regional policy. Future research should therefore focus on identifying what these barriers to growth in medium-density locations are.

As an endnote, the success story of Bangalore — the Silicon Valley of India — is one of the notable exceptions to our general findings: that district has a density level of similar magnitude as the high-tech clusters in the U.S. Interestingly, it traces its history back to the so-called Electronics City, set up in the 1970s as an industrial park 18 kilometers south of the city. Perhaps this particular example points to a promising way to eliminate the growth restrictions that we have uncovered in many other intermediate-density districts.

Notes

[1]When comparing the results with the U.S., we will make the definition of services in both countries comparable.

[2]Nataraj (2011), Kathuria et al. (2010), Hasan and Jandoc (2010) and Dehejia and Panagariya (2010) provide detailed overviews of similarly constructed databases. See also Fernandes and Pakes (2010) for a description of the Indian manufacturing sector.

[3]We also experimented with using an optimal bandwidth. This does not change the qualitative results, but makes the comparison between graphs more difficult. Further details of this methodology can be found in Desmet and Fafchamps (2006).

[4]In manufacturing the unorganized sector consists of establishments with either less than ten workers (if they use electricity) or less than 20 workers (if they do not). These firms are not required to register and do not pay taxes. In services no legal distinction between the organized and the unorganized sectors exists, so we use a ten worker threshold, and define all firms below that threshold as being part of the unorganized sector. This is similar to the approach in Ghani, Kerr and O'Connell (2011) who describe the difference between organized and unorganized in more detail.

[5]This latter finding is consistent with results in Ghani, Goswami and Kerr (2012).

[6]These percentages are obtained by classifying each 2-digit sector into one of three categories: those that exhibit increasing concentration in high-density clusters, those that exhibit dispersion in high-density clusters, and those that exhibit no clear pattern. Depending on how we deal with those sectors that show no clear pattern, the percentages differ slightly.

[7]Of course there may be other reasons for why manufacturing has not shown the same tendency towards further concentration. One such reason are increasing measures to move polluting industries out of urban areas.

[8]To make all the figures comparable, the scale of the horizontal axis is always the same (i.e. observations below 0 are not shown). If we were to show smaller places, we would find evidence of convergence in low-density counties.

[9]Our regressions for the U.S. take counties as the unit of observation. To make the definition of services as similar as possible to the one in the U.S. we are using the sum of transport & utilities and other services from the BEA. Using broader definitions of services by including, say, retail and wholesale, do not change the findings.

[10]See Desmet and Fafchamps (2006) for further details on conditional kernel regressions in a similar context.

[11]All data come from the 2001 Population Census for India, with the exception of travel time to a top-10 city and the distance measures, which comes from Lall et al. (2010). For more details, see Ghani, Kerr and O'Connell (2011).

[12]Data for China come from the China City Statistical Yearbooks with prefecture-level cities as the unit of observation. A second caveat, in addition to the one already mentioned, is that services refer to the "tertiary sector" implying a broader definition than the one used for India and the U.S. We use this broader category because of changes in the definitions of different service subsectors over the time period under consideration. Using alternative definitions of services in China does not change the qualitative results though.

[13]In the figures this corresponds to a log employment density of 5. Because the Chinese service data are not exactly comparable to those of India (on the one hand, they are more inclusive by considering the tertiary sector, and on the other hand, they are less inclusive because they only measure "formal" employment), not too much should be read into the exact level of this threshold.

[14]Note that in our data aggregate tertiary employment went down in China between 2000 and 2007. Indeed, one of the effects of liberalization was a reduction in the share of formal employment (i.e., a reduction in "staff and workers").

[15]The counterfactual growth rate has been multiplied by the mean growth rate of Indian districts relative to the mean growth rate of U.S. counties. Given that differences in growth rates across locations of different densities in the U.S. are much smaller than in India, differences in the counterfactual growth rates are also much smaller than actual differences. When interpreting the results, what matters are the relative differences across locations.

[16]We also experimented with European regions, and found similar results to those in the U.S.

References

Aghion, Philippe, Robin Burgess, Stephen J. Redding, and Fabrizio Zilibotti. 2008. "The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India," *American Economic Review*, 2008, 98, 1397-1412.

Rajeev Dehejia and Arvind Panagariya. 2010. "Services Growth in India: A Look Inside the Black Box," Working Paper 2010-4, School of International and Public Affairs, Columbia University.

Desmet, Klaus and Marcel Fafchamps. 2006. "Employment Concentration across U.S. Counties," *Regional Science and Urban Economics*, 36, 482-509.

Desmet, Klaus and Esteban Rossi-Hansberg. 2009. "Spatial Growth and Industry Age," *Journal of Economic Theory*, 144, 2477-2502.

Fernandes, Ana M. and Gunjan Sharma. 2012. "Together We Stand? Agglomeration in Indian Manufacturing," Policy Research Working Paper #6062, The World Bank.

Ghani, Ejaz, 2010. "The Service Revolution in India," VoxEU.

Ghani, Ejaz, William R. Kerr, and Stephen D. O'Connell. 2011. "Spatial Determinants of Entrepreneurship in India," NBER Working Paper #17514.

Ghani, Ejaz, Arti Grover Goswami, and William R. Kerr. 2012. "Is India's Manufacturing Sector Moving Away from Cities ?," NBER Working Paper #17992.

Hasan, Rana and Karl Robert L. Jandoc. 2010. "The Distribution of Firm Size in India: What Can Survey Data Tell Us?" ADB Economics Working Paper #213.

Kathuria, Vinish, Rajesh Raj Seethamma Natarajan, and Kunal Sen. 2010. "Organized versus Unorganized Manufacturing Performance in India in the Post-Reform Period," MPRA Working Paper No. 20317.

Lall, Somik V., Hyoung Gun Wang, and Uwe Deichmann. 2010. "Infrastructure and City Competitiveness in India," UNU-WIDER Working Paper.

Nataraj, Shanthi, 2011. "The Impact of Trade Liberalization on Productivity and Firm Size: Evidence from Indias Formal and Informal Manufacturing Sectors," *Journal of International Economics*, 85, 292-301.
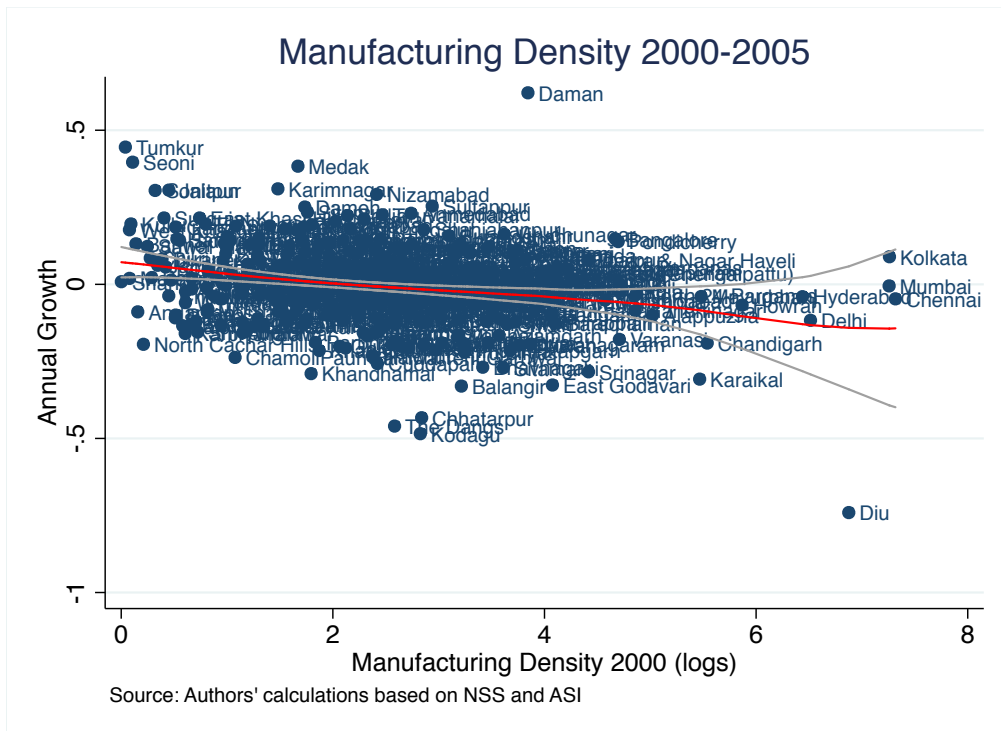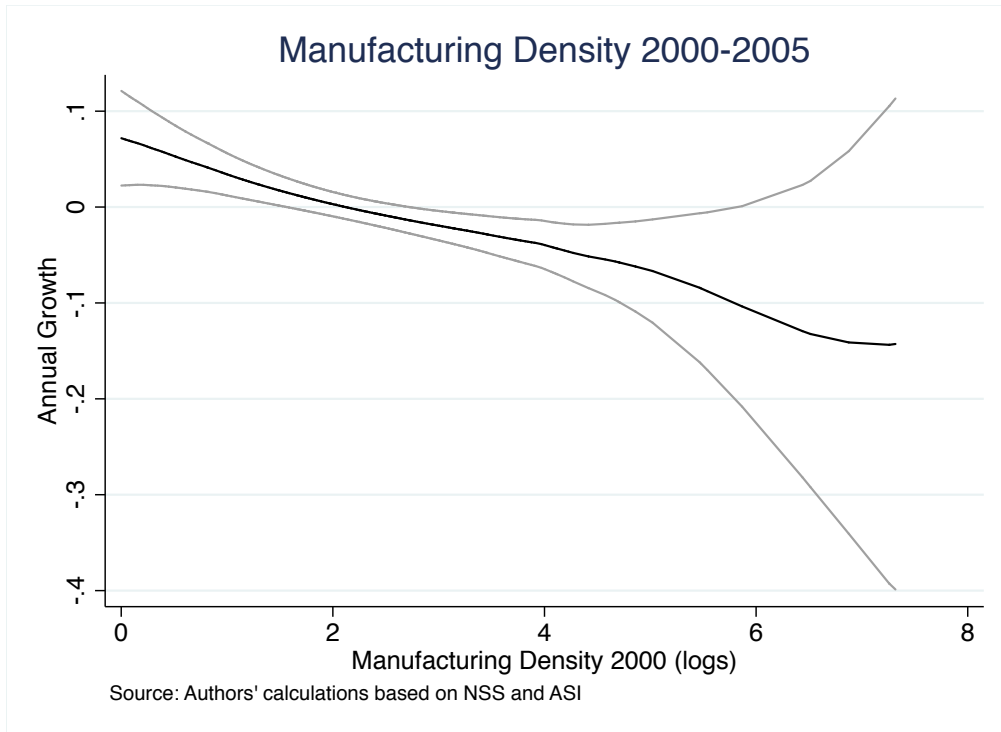
Figure 1: Annual manufacturing employment growth as a function of initial manufacturing employment density (logs), based on NSS and ASI, 2000-2005.
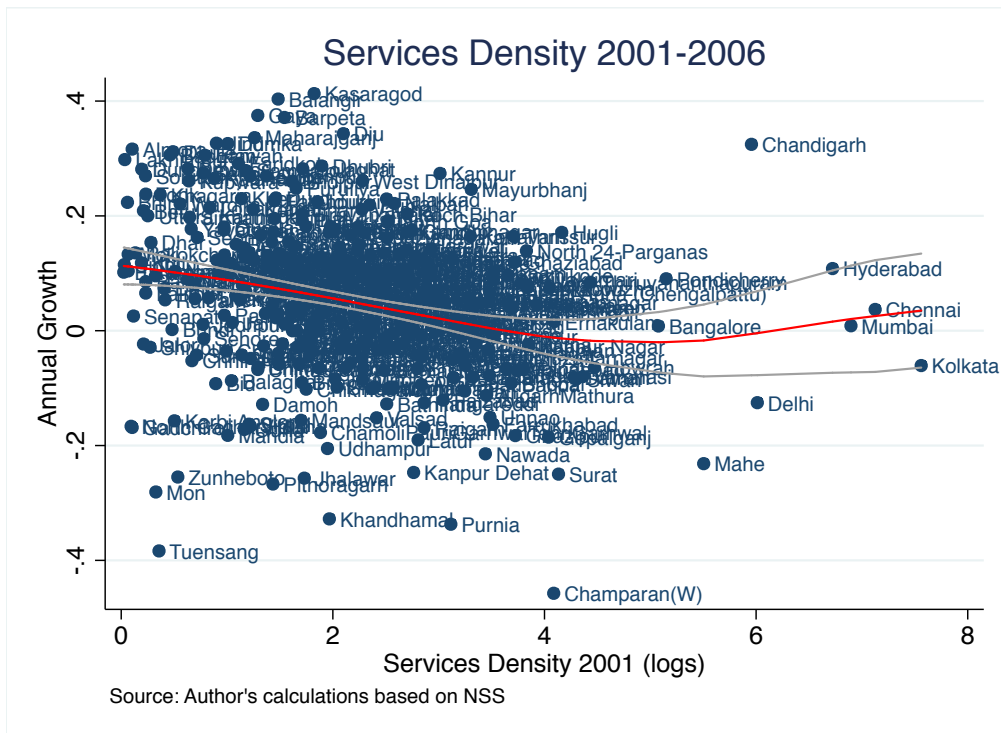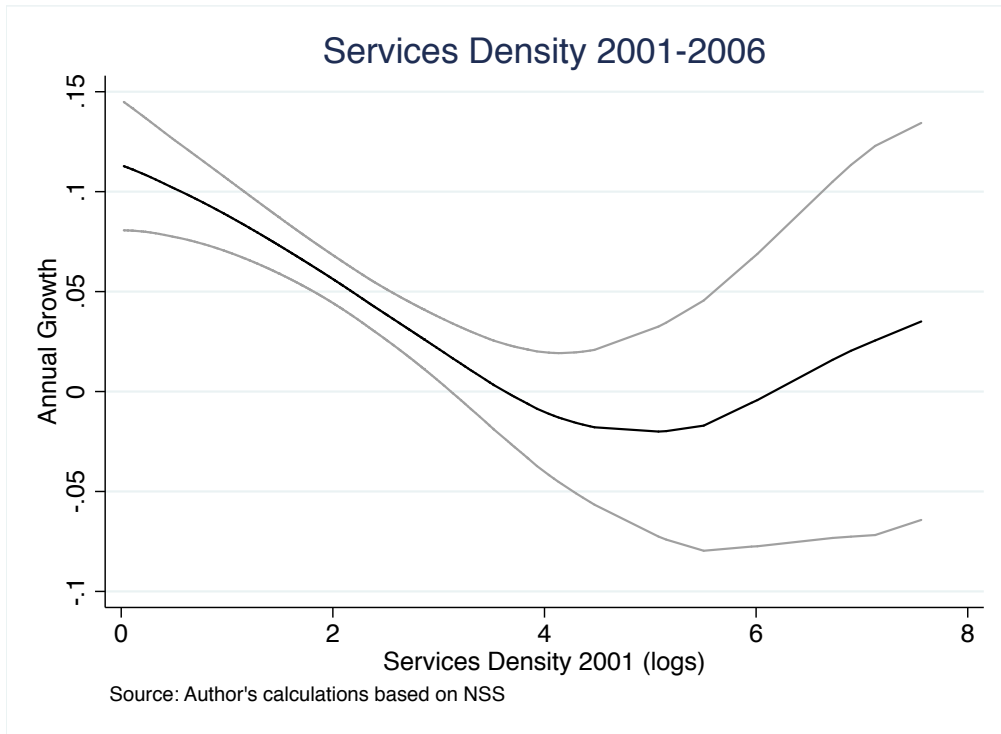
Figure 2: Annual services employment growth as a function of initial services employment (logs), based on NSS, 2001-2006.
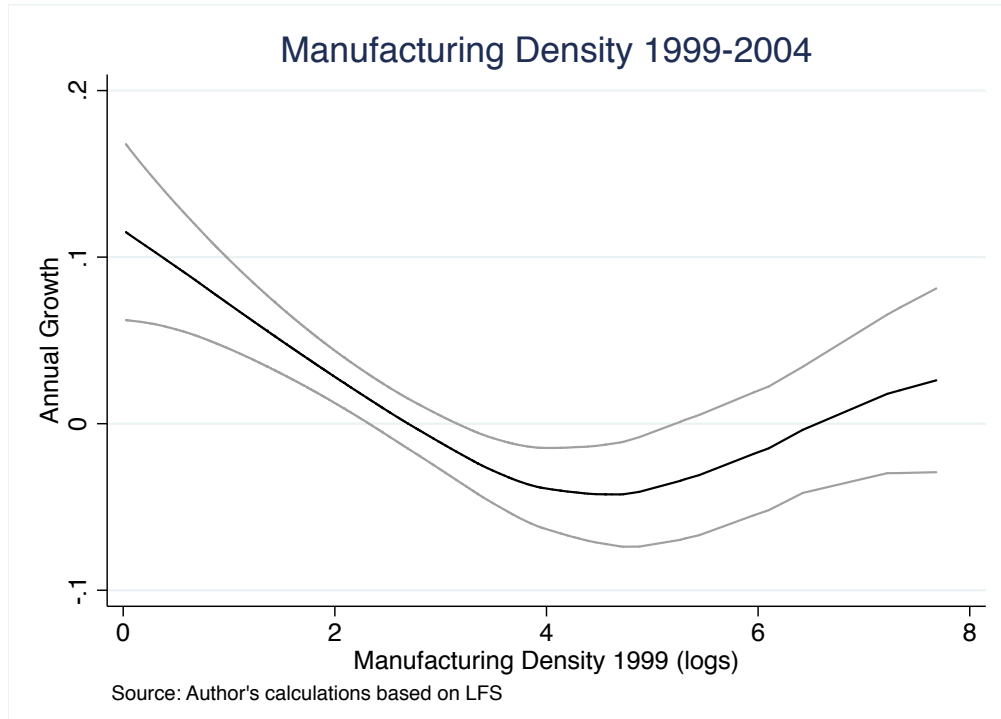
Figure 3: Annual manufacturing employment growth as a function of initial manufacturing employment density (logs), based on LFS, 1999-2004.
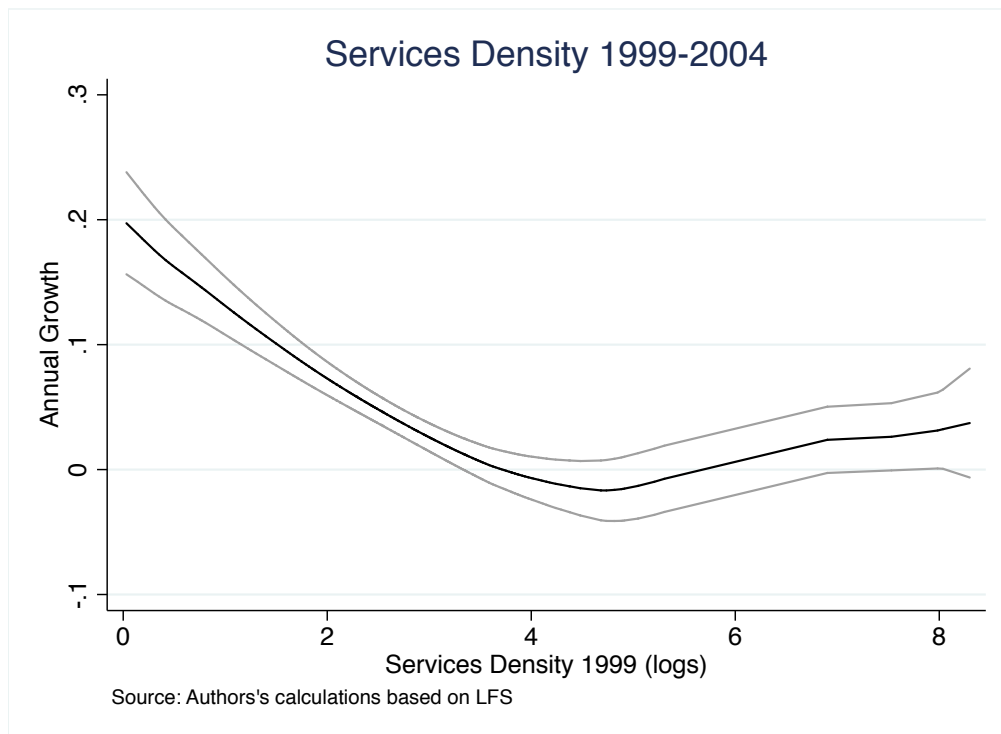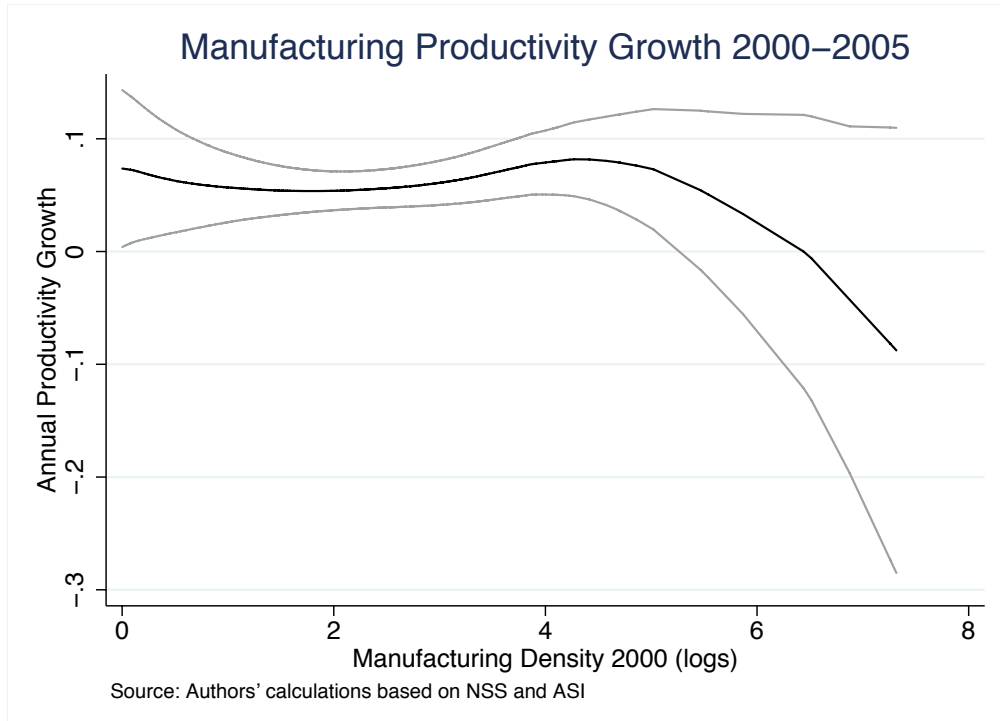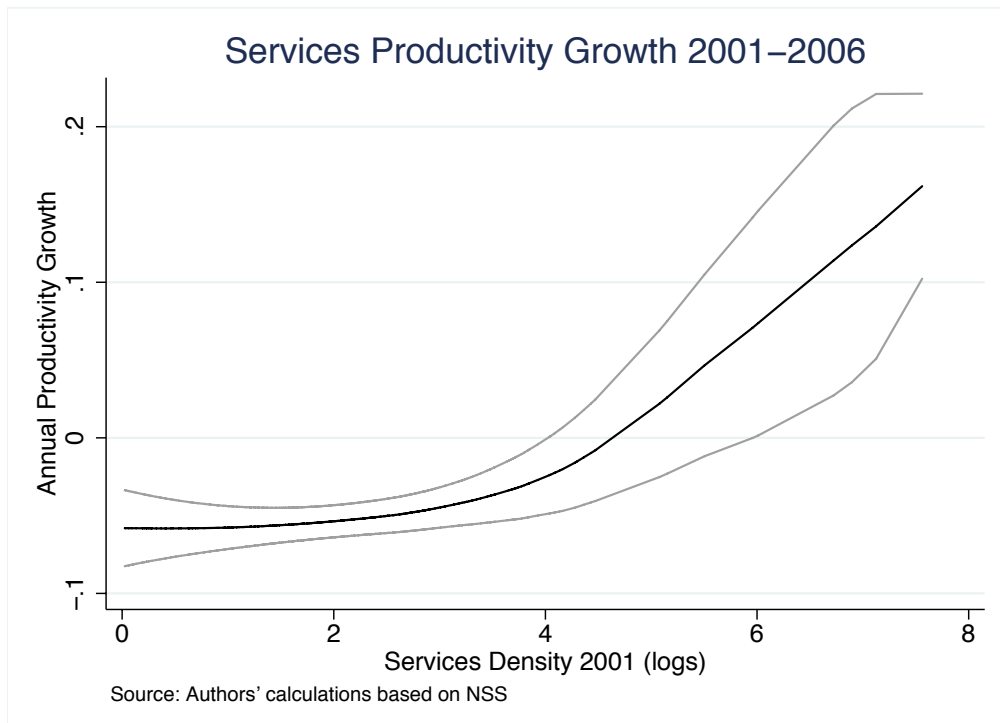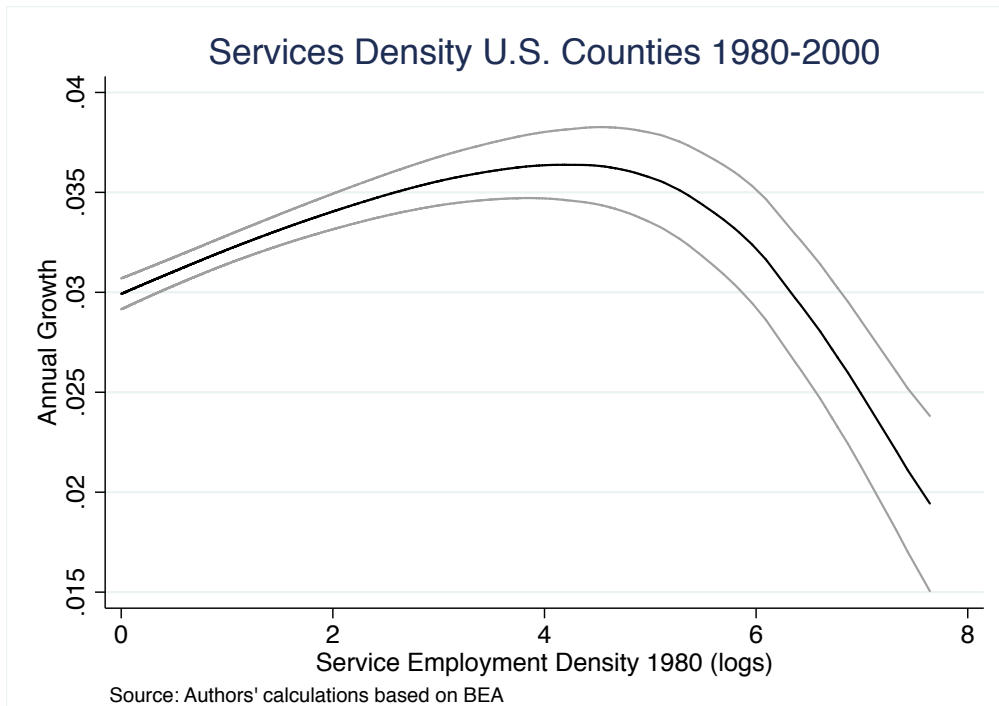

Figure 4: Annual services employment growth as a function of initial services employment (logs), based on LFS, 1999-2004.

Figure 5: Annual manufacturing labor productivity growth as a function of initial manufacturing employment density (logs), based on NSS and ASI, 2000-2005.

Figure 6: Annual services labor productivity growth as a function of initial services employment density (logs), based on NSS, 2001-2006.

Figure 7: Annual service employment growth as a function of initial service employment density (logs), U.S. counties, 1980-2000.
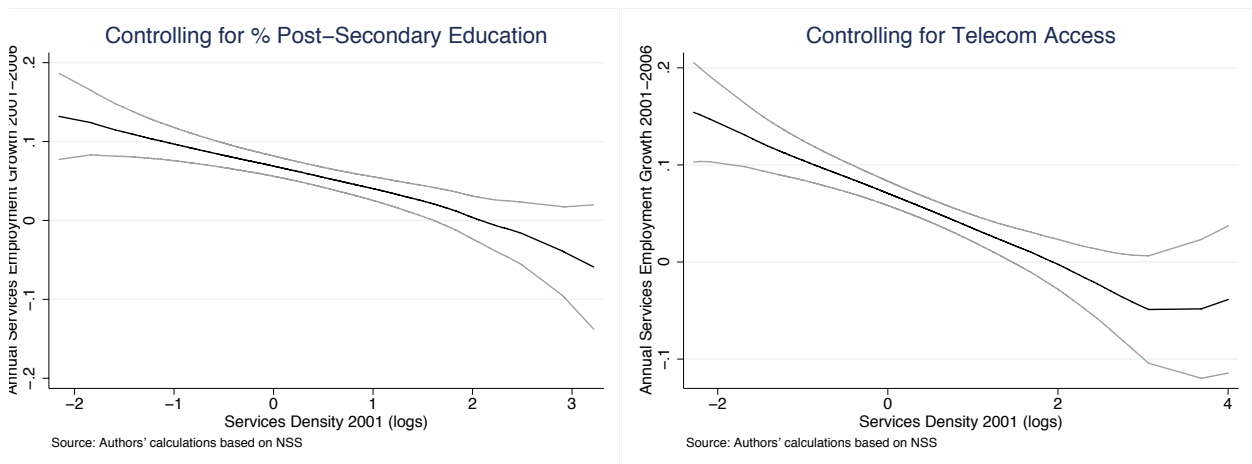

Figure 8: Annual service employment growth as a function of initial service employment density (logs), controlling for percent of population with more than secondary education (left panel) and percent of households with access to telecommunication services (right panel), based on NSS, 2001-2006.
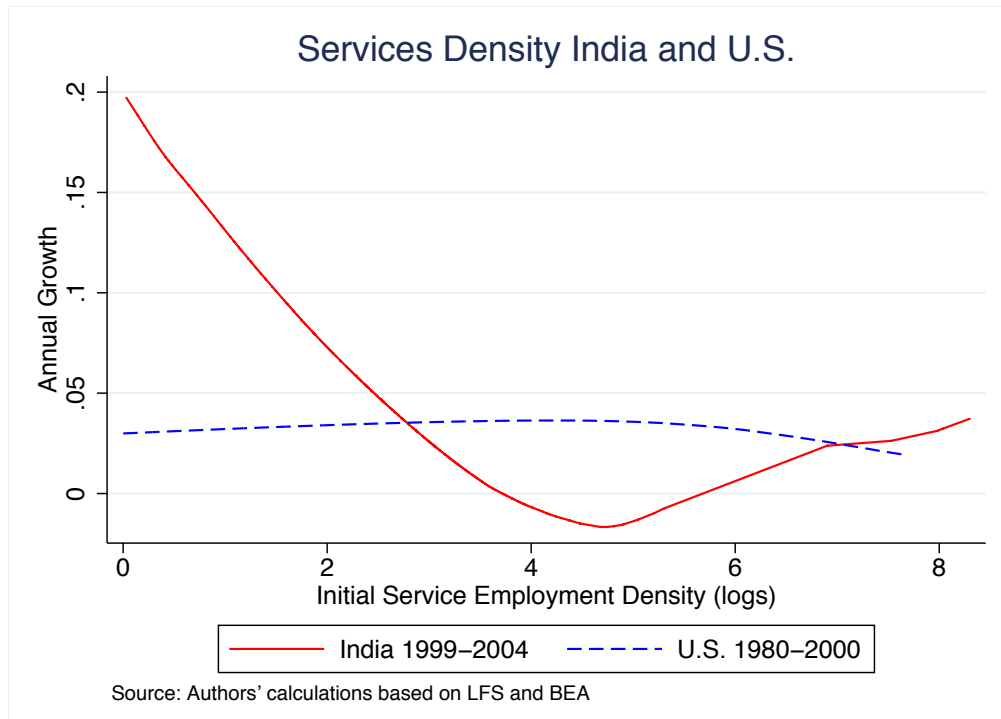
Figure 9: Annual service employment growth as a function of initial service employment density (logs), U.S. counties, 1980-2000, and Indian districts, 1999-2004.
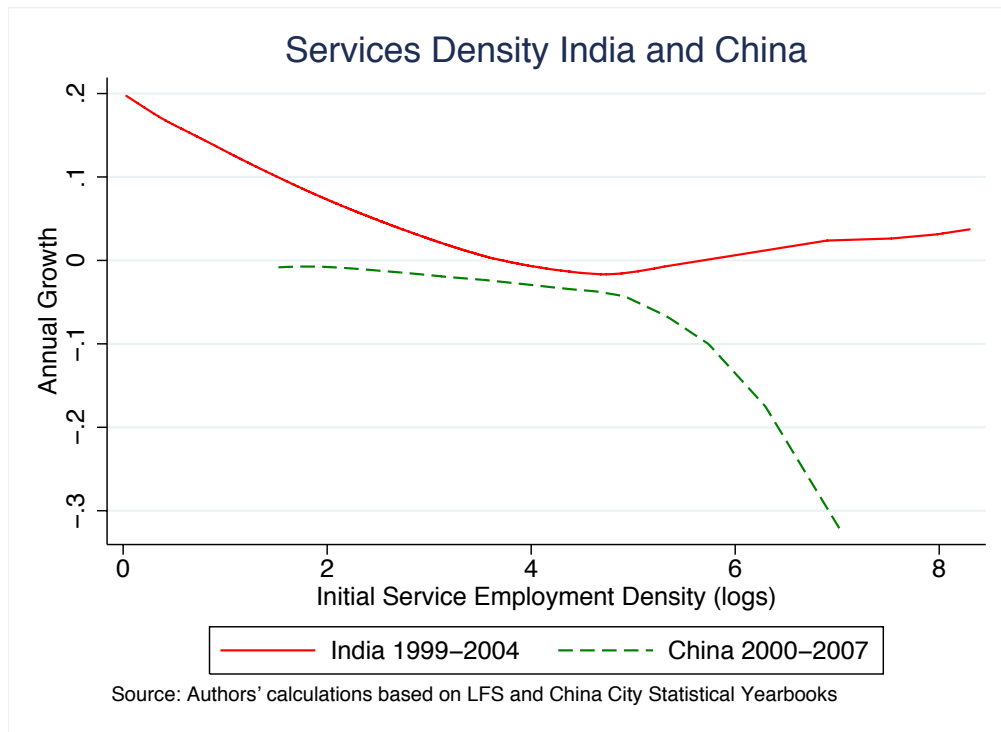

Figure 10: Annual service employment growth as a function of initial service employment density (logs), Indian districts, 1999-2004, and Chinese prefecture-level cities, 2000-2007.
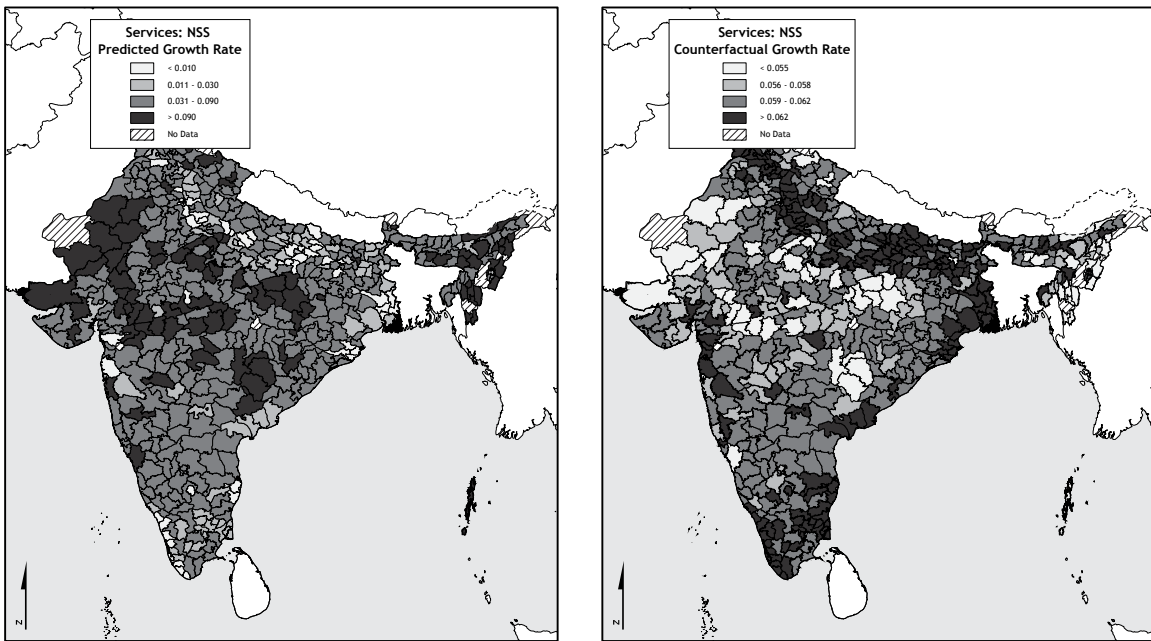
Figure 11: Growth in services employment, predicted based on NSS data (left panel) and counter-factual based on U.S. counties (right panel).