Labor Productivity and Density:

Examining the Economic Geography of India

In Partial Fulfillment of the Requirements for the A.B. Degree

Submitted to Princeton University

Advisor: -

Department of Economics

April 11, 2012

Abstract

The physical proximity of output, people, and firms in densely populated areas creates an environment that facilitates economies of scale, access to resources, technological progress, efficient labor markets, and flows of ideas and information. While a number of studies have examined the productivity-enhancing effects of density in the U.S. and in Europe, few have been conducted in India, a rapidly expanding economy where the levels and variation in population density are much higher. This paper presents estimates of the effects of variation in population, labor, output, and new firm density on levels of labor productivity in Indian states, following the methodology of Antonio Ciccione and Robert E. Hall in their defining 1996 paper "Productivity and the Density of Economic Activity." The empirical analysis shows that density measures of agglomeration exhibit strong causal relationships on levels of labor productivity across states in India from 1981-2001. The results are similar across population, labor, and output density, where doubling density results in a 13 to 21 percent increase in labor productivity. I also find that doubling new firm density results in a 6 to 12 percent labor productivity increase. These results are higher and more varied than the findings for the U.S. and Europe, which estimate a six percent labor productivity increase to a doubling of density, and suggest that density may have a larger effect on labor productivity in India.

Dedication

For my parents and sister, who dragged me to India.

Acknowledgements

The author would like to acknowledge the invaluable guidance of
- as my advisor for this thesis. I am also grateful for the help and suggestions of
, , and my parents. Special thanks to
my constant
thesiser extraordinaire. Last but not least, thanks to my friends and to Small World Coffee for making this possible.

Table of Contents

INTRODUCTION **1**

Figure 1. Population density in India, 1901-2011 ..... *5*

BACKGROUND **6**

MACRO TRENDS IN ECONOMIC GROWTH AND PRODUCTIVITY 6

Figure 2. Number of firms incorporated by size, 1947-2009 ..... *9*

Figure 3. GDP and firm formation in India, 1980-2010 ..... *10*

UNLOCKING THE POTENTIAL OF FIRM FORMATION 10

LITERATURE REVIEW **13**

CONCLUDING REMARKS ON THE EXISTING LITERATURE 19

DATA **20**

Table 2. Summary of key variables *21*

METHODOLOGY **23**

IDENTIFICATION AND ESTIMATION 24

INSTRUMENTAL VARIABLE REGRESSION 25

APPLYING THE EMPIRICAL SPECIFICATION TO INDIA 28

Table 3. Instrumental variable regressions of various density functions on labor productivity as regressed for each time period *30*

RESULTS **31**

Table 4. Summary of instrumental variable regression results, ... *31*

1981, 1991, 2001, and Panel *31*

Figure 4. Scatterplots of population density, labor density, and output density versus labor productivity, 1981-2001 ..... *34*

Figure 5. Scatterplot of new firm density versus output density, ... *37*

Figure 6. Scatterplot of new firm density versus labor productivity, ... *38*

A NEW INSTRUMENT 39

Table 5. Comparing instrumental variable results using two different instruments, population density (1971) and rainfall (2009) *40*

Table 5. Comparing instrumental variable results using two different instruments, population density (1971) and rainfall (2009) continued *41*

DISCUSSION AND CONCLUSIONS **41**

Introduction

In 2010, the Government of India announced an explicit focus on cities as loci of economic activity. In January 2012, the Institute for Financial Management and Research (IMFR) in India released a statement identifying the Indian economy's dependence on large, rapidly expanding urban centers as drivers of economic growth and productivity (Institute for Financial Management and Research 2012). However, it is not clear whether agglomeration economies, the benefits arising from the geographic clustering of people, firms, or activities, outweigh the negative congestion effects that arise in highly dense areas. Agglomeration increases the proximity of people and resources, leading to decreased transportation costs for goods, people, and ideas; on the other hand, congestion drives up real estate prices, creates competition for labor and other resources, and causes pollution.

While cities are hubs of economic activity, differences in density across states in India vary widely, and it is not clear whether the benefits of agglomeration economies hold true for levels of density that are much higher than those in the U.S. Uttar Pradesh, a state in north India, has over 200 million inhabitants alone, meaning it would be the fifth-largest country ahead of Brazil according to 2011 estimates. This paper addresses the gap in the existing research on India by examining the effect of agglomeration economies as measured by density on labor productivity. In India, where levels and variation of spatial disparity are far higher than in most developed economies, are agglomeration forces more powerful or are congestion effects more overwhelming?

Economies of agglomeration are an application of the larger study of economic geography, defined by Economist Paul Krugman as "the location of factors of production in space" (Krugman 1991). Contemporary agglomeration studies emphasize the roles of positive externalities that appear as a result of increased interaction in dense areas. That is, the physical proximity of output, people, and firms in urban areas creates an environment that facilitates economies of scale, access to resources, technological advancement, efficient labor markets, and flows of ideas and information, thereby enhancing labor productivity. If the cost or number of labor hours required to transport products from one stage of production to the next increases with distance, creating the technology to conduct all stages of production in one geographical area leads to increasing returns to scale. Thus, the output per worker increases as density increases. The concentration of economic activity also increases the number and variety of complementary goods and services available. Moreover, denser areas enhance knowledge dissemination in industries where information exchange is difficult, rapidly changing, or of particular importance, for instance in high-skilled industries (Ciccione and Hall 1996, Abel et al. 2010). In agglomeration economies, there are increasing benefits to proximity of goods and services, labor, and ideas.

While a number of studies have examined the productivity-enhancing effects of density in the U.S. and in Europe (Marshall 1890; Jacobs 1969; Lucas 1988; Glaeser 1991), few have been conducted in India. The distribution of Indian people, labor, output, and firms is different from that of the U.S. In the U.S., where urban areas account for almost 90 percent of gross domestic product (GDP) and 82 percent of the population, urbanization dominates the economic geography more

than in India, where only 30 percent of the population is urban.[1] India as a whole, though, has a much higher average population density than the US: nearly four times the population of the U.S. occupies approximately one-third the land area (Abel et al. 2010, CIA World Factbook). This presents an interesting situation in which urban areas are extremely densely populated, but the majority of the country lives in sparsely populous areas. In India, rapid urban expansion is also often unaccompanied by the necessary development of physical infrastructure, causing severe strains on resources and dangerous overcrowding. Is encouraging rapid urbanization and the creation of megacities desirable in a developing economy where the levels and variation in population density are extremely high?

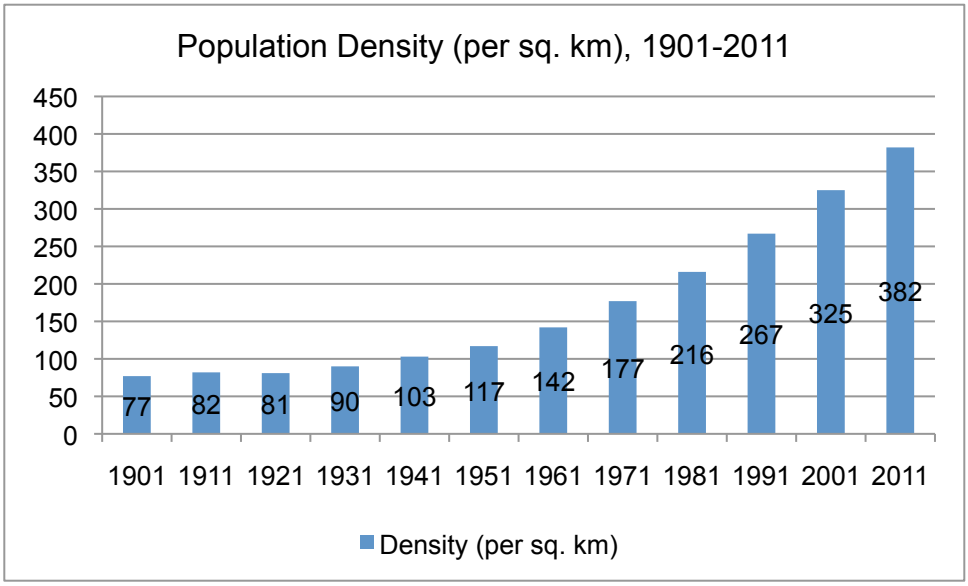
This paper finds evidence that higher spatial density of factors of production is even more effective in increasing labor productivity in India than in the U.S. and Europe. Specifically, I estimate that doubling population, labor, or output density increases labor productivity by two to three times more in India from 1981-2001 than in the U.S. in 1988. I apply the methodology of Antonio Ciccione and Robert E. Hall's defining 1996 paper "Productivity and the Density of Economic Activity" to estimate the causal effects of variation in population, labor, output, and (in an

[1] According to the Census of India Metadata, "in the Census of India 2001, the definition of urban area adopted is as follows: (a) All statutory places with a municipality, corporation, cantonment board or notified town area committee, etc. (b) A place satisfying the following three criteria simultaneously: i) a minimum population of 5,000; ii) at least 75 percent of male working population engaged in non-agricultural pursuits; and iii) a density of population of at least 400 per sq. km. (1,000 per sq. mile). For identification of places which would qualify to be classified as 'urban' all villages, which, as per the 1991 Census had a population of 4,000 and above, a population density of 400 persons per sq. km. and having at least 75 percent of male working population engaged in non-agricultural activity were considered" (Census of India 2001 Metadata).

expansion of the methodology) new firm density on levels of labor productivity in Indian states. Using a detailed dataset of Indian medium and large-sized firms, I find a similar but smaller effect of the spatial density of new firms on labor productivity. Because I am concerned about the presence of reverse causality, I use instrumental variables regression to draw causal inference from observed data. Reverse causality arises in this instance because density may affect levels of labor productivity, but increased labor productivity may also lead to increased density.

This analysis is timely and will help guide policy decisions and economic research going forward. Preliminary results from the 2011 Census of India estimate that India will be the most populous country in the world by 2020. Census reports also show that the population density has been increasing steadily since the mid-20th century at approximately 20 percent per year (Census of India 2011). The chart below exhibits the visible acceleration of growth in population density since 1901, which peaked from 1961 to 1981 with an average year-over-year population density growth rate of over 23 percent.

Figure 1. Population density in India, 1901-2011



Source: Census of India 2011.

India now has 53 cities with over one million residents, with 91 million more people living in cities now as compared to in 2001 (Nijman 2012). Furthermore, these cities produce the majority of national economic output. As of 2008, urban areas accounted for approximately 63 percent of Indian GDP, and this number is expected to increase to 75 percent by 2021 (Government of India Planning Commission 2008). These numbers are similar to those in China, where urban areas now generate over 60 percent of GDP, but less than those in the U.S., where urban areas generate 90 percent of GDP (World Bank, Abel et al. 2010).

To summarize my results briefly, the empirical analysis shows that density measures of agglomeration exhibit strong causal relationships on levels of labor productivity across states in India from 1981-2001. The results are similar across population, labor, and output density, where a doubling of density results in a 13 to 21 percent increase in labor productivity, and indicate that density has a larger

effect on labor productivity in India than in the U.S and Europe. Furthermore, I find that doubling new firm density results in a 6 to 12 percent labor productivity increase, suggesting that there are smaller but still significant returns to incorporating new firms close to each other.

This paper is structured as follows. First, I present a brief background of the rise of urbanization and entrepreneurship in the context of the Indian economic climate. Next, I review the relevant literature on density and productivity. I include studies conducted in India and the U.S. and Europe, as well as an introduction to the emerging literature on the role of new firms in spurring economic growth. In the following two sections, I describe the data and methodology applied in this paper. Finally, the last two sections conclude with the results found and a discussion of their implications.

Background

Macro trends in economic growth and productivity

In order to understand the effects of agglomeration economies in India, it is necessary to situate them within in the broader context of economic growth and productivity in the country from 1981-2001. The 1990s witnessed the effects of sweeping structural reforms in India. In 1991, the Indian government approached the International Monetary Fund (IMF) for assistance in the face of a severe balance of payments crisis that was fueled by the collapse of a major trade partner in the Former Soviet Union (FSU) and escalating oil prices following the Gulf War. The assistance was given under the condition of major structural economic reforms,

including reduction of the budget deficit, restriction of credit and devaluation, trade liberalization, and reduced government protection of national industries (Balakrishnan 2010).

Though the receipt of conditioned aid from the IMF is generally cited as the source of India's economic growth in the nineties, economic historians note that professional economists in the Indian government already acknowledged the importance of economic globalization. The 1980s saw a series of endogenous 'pro-business' policy reforms, marking the beginnings of a gradual move away from India's post-independence protectionist fiscal and monetary policies. My analysis examines the Indian economy in its transition from 'pro-business' to 'pro-market' reform from 1981 to 2001, a period of consistent growth in the Indian economy and a transition from the previous "Hindu rate of growth." The term "Hindu rate of growth" was coined by Indian economist Raj Krishna, and is controversial for its suggestion that the 'complacent' Hindu attitude was inherently linked to the slow economic growth of the 1950s-1980s. In "India since Independence: An Analytic Growth Narrative," American economist Bradford DeLong points out that in every year post-1980, the level of GDP per capita is higher than the predicted level based on the trend line from 1962-1980 (DeLong 2003, Balakrishnan 2010). Rodrik and Subramanian (2005) follow DeLong's work with a more robust methodology that demonstrates that total factor productivity (TFP) growth accelerated beginning in 1980. Rodrik and Subramanian suggest that the increasingly competitive nature of the political environment from Indira Gandhi's return to power in January 1980 forward applied pressure on political candidates to respond to citizen demand for economic reform. They loosely link the 1980 pro-business reforms by Prime

Minister Indira Gandhi, continued by her son Prime Minister Rajiv Gandhi in 1985 after her assassination, with increased TFP across both agriculture and industry in India (Balakrishnan 2010).

The 1991 reform package marked the beginning of 'pro-market' reform, resulting in the privatization of key industries previously controlled by government monopolies, with the exception of petroleum, edible oils, and fertilizer. Average tariffs fell from more than 80 percent in 1990 to 37 percent in 1996. An IMF working paper by Petia Topalova found that these reductions in trade protections led to higher levels of growth and firm productivity for private companies. There was no evidence of increased productivity in companies that continued under government ownership. This result suggests that the formation of new firms may be linked to the liberalization and economic growth of the Indian economy (Petia 2004, Balakrishnan 2010).

Figure 2 below illustrates the patterns in firm incorporation since Indian independence in 1947. According to the Prowess Indian Corporate Database,[2] a widely recognized source for information on medium and large corporations, almost no new firms with significant market capitalizations[3] were incorporated before the

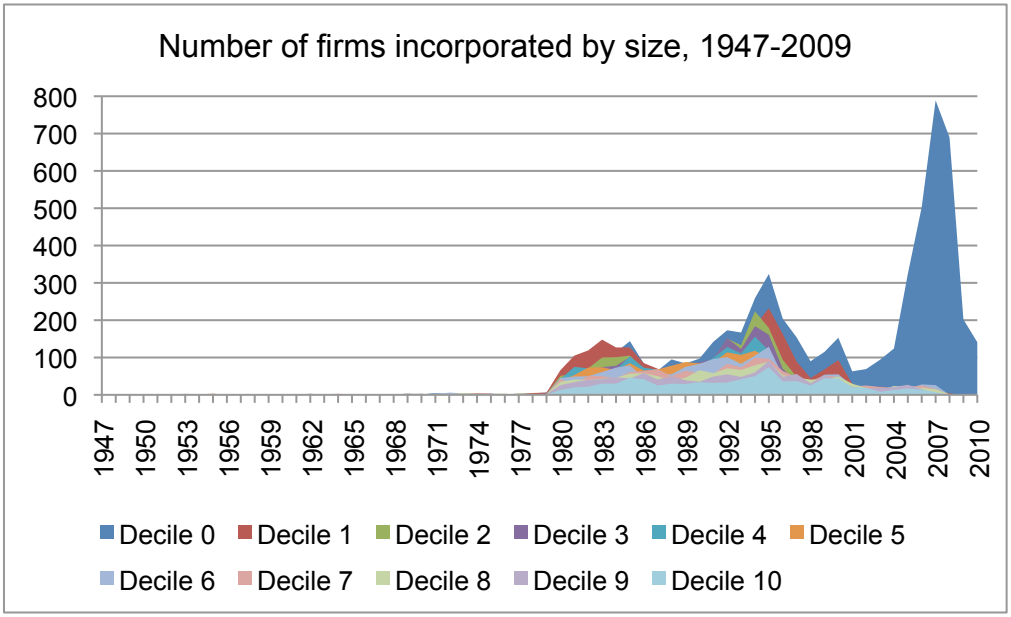
---

[2] According to its website, Prowess "contains detailed information on over 25,346 firms. The database covers most of the organized industrial activities, banking and organized financial sectors, and other service sectors in India. The companies covered in Prowess account for 75 percent of all corporate taxes and over 95 percent of excise duty collected by the Government of India." More information is available at http://www.cmie.com/database/?service=database-products/firm-level-data-services/prowess-corporate-database.htm (CMIE Prowess 2012).

[3] In accordance with the CRSP guidelines, it is assumed that the Prowess Database divides firms into size deciles calculated from quarterly portfolios of common stock (Center for Research in Security Prices). The Prowess Database lists the largest firms by portfolio capitalization in decile 1,

late 1970s, but there was a jump in firm formation from 1980-1985. Coinciding with the IMF structural reforms and the rejuvenation of the Indian economy, firm formation reached a new high from 1990-1995. We see a steep climb in formation of smaller firms as GDP entered a high-growth phase in 2004, peaking in 2007 and falling dramatically during the global recession from 2008-2009. From 2004-2010, the economy averaged a growth rate of eight percent per year until late 2011. From this plot, it appears that studying firm formation patterns may correspond with macroeconomic trends over time.

Figure 2. Number of firms incorporated by size, 1947-2009



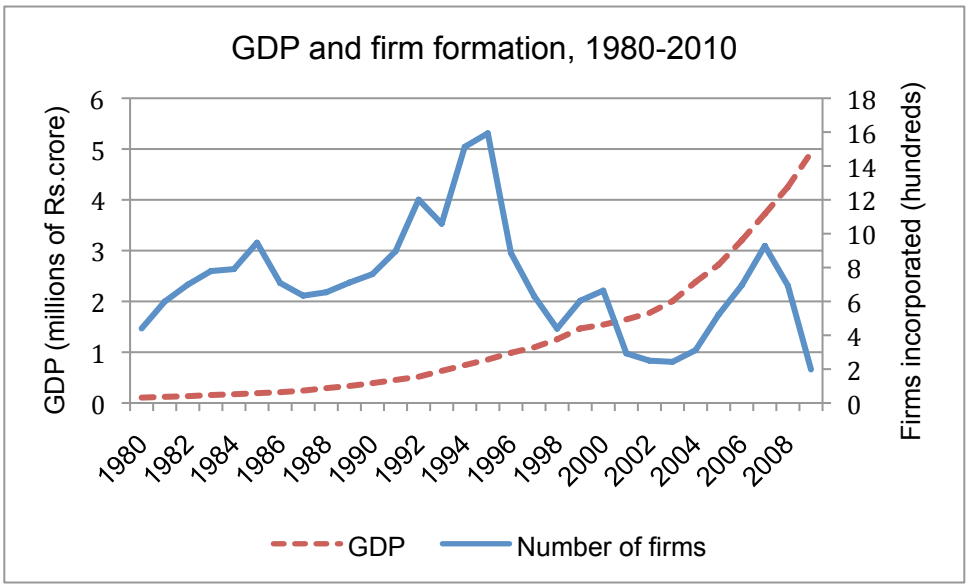
Decile 0 represents the smallest firms, and Decile 10 represents the largest firms. Source: Prowess Indian Corporate Database.

---

and the smallest in decile 10. Because some firms fall even below the threshold of the smallest decile, some liberty was exercised in redefining size "0" as size "0.001" to accommodate the logarithmic regression employed. In order for this scale produce a meaningful analysis given that the smallest firms are represented as close to zero, I have redefined the deciles from smallest (1) to largest (10). Note that this is a change of semantics rather than substance, and has no meaningful effect on results.

In the plot of national GDP versus firm formation given in Figure 3, however, the relationship is not as clear as expected. From 1980-1988, the number of new firms and GDP do not appear to be correlated. From 1988-1993, 1998-2000, and again from 2002-2007, GDP and new firm formation rise together. However, from 1994-1998, 2000-2002, and 2007-2009, new firm formation drops steeply while GDP continues to rise. It is evident from the figure that the relationship between firm formation and GDP is not obvious. In the Indian context, it is necessary to capture a more nuanced picture of newly formed firms (the "opportunity") and how they relate to output and productivity.

Figure 3. GDP and firm formation in India, 1980-2010



Sources: Reserve Bank of India, Prowess Indian Corporate Database.

Unlocking the potential of firm formation

My expansion of Ciccione and Hall's methodology examines the potential of new firms to affect labor productivity growth. It comes at a time when firm

formation as a driver of economic growth in developing economies is becoming an increasingly important topic of study. The Kauffman Foundation, the world's largest foundation devoted to entrepreneurship, uses empirical data from the United States to illustrate that firms create four times as many jobs in their first year of life than in their fourth year (The Kauffman Foundation 2010). The International Finance Corporation (IFC) and Aspen Network of Development Entrepreneurs (ANDE), among others, have researched the question: "Thirty years of research shows that a small percentage of firms create the vast majority of new jobs and GDP growth in the U.S. and Europe, but how does this translate into opportunities in emerging markets?" Specifically examining the spatial concentration of new firms and its consequences for labor productivity in a country with as much regional variation as India is especially relevant. Experts from the IFC and ANDE believe that understanding the interaction between new firm formation and level of productivity could unlock the potential to achieve development and economic growth goals in emerging markets (NextBillion.net).

However, the presence of new firms does not ensure growth or economic stability in itself. In fact, market over-fragmentation may be an indication of the businesses' inability to achieve scale or significant growth. Especially in emerging markets, a large percentage of the labor force may be employed in small firms out of lack of opportunity at larger firms. In India, though small businesses employ 90 percent of the workforce, they produce only eight percent of total GDP (Timm

2011).[4] These workers might prefer the stability and level of compensation that larger firms can offer, but few such opportunities exist.

In the last quarter of 2011, growth in the Indian economy slowed to approximately six percent. Even if the economy regains momentum soon, many worry that growth is unlikely to break seven percent. Although six percent growth may sound high, for an emerging economy with a large deficit supporting the world's second-largest population it is barely enough. India's high growth trajectory has many stakeholders. For millions of Indian entrepreneurs and business people, economic productivity is a fast-track ticket out of poverty. Multinational corporations and investors have invested billions of dollars in India as a manufacturing and service provider, and as a vast consumer market (The Economist 2012). This paper estimates the effects of density of population, labor, output, and firm formation on labor productivity to evaluate how India's increasingly dense states will influence economic growth.

[4] Data on Micro, Small, and Medium enterprises (MSMEs) is collected by the 4th All-India Census of MSMEs, 2006-2007. Footnote from source: "These figures are from http://smetimes.tradeindia.com/smetimes/news/top-stories/2009/Sep/01/sick-msmes-increase-marginally-says-4th-msme-census60059.html. While India conducts a census every four years by its Ministry for Small Businesses, the small business's share of India's GDP is likely to be higher as the MSME definition does not account for many small farmers and larger small businesses. According to some, MSMEs could account for 10% or less of all small firms (interview Fisme secretary general Anil Bhardwaj)" (Timm 2011).

Literature Review

Economist Alfred Marshall's *The Principles of Economics* is often credited with laying the foundation for agglomeration economics. In discussing localized industries, he asserts that the three main forms of transport costs—the costs of moving goods, people and ideas—are reduced by agglomeration. These benefits can be summarized as: reduced shipping costs from supplier to producer or from producer to customer; reduced labor costs due to the pooling of diverse and high-quality labor (a constant market for skill); and intellectual spillovers that accelerate learning, technological progress, and innovation (Marshall 1920, Glaeser et al. 2009).

The academic literature of agglomeration economics reveals a recent resurgence in studies on the topic, and a new focus on emerging economies. Empirical studies have generally focused on city size, industry size, or urban population as determinants of productivity, particularly in the U.S. and in Europe. The recent literature has seen a convergence on use of spatial density rather than degree of urbanization to account for geographically localized externalities (Ciccione and Hall 1996). The existing literature on India has yet to link density with labor productivity.

Ciccione and Hall's "Productivity and the Density of Economic Activity" is a key study used in this paper. Ciccione and Hall state that labor productivity varies widely by state in the U.S.: in 1988, the output per worker in the most productive state was two-thirds larger than that of the least productive state. In measuring the effects of agglomeration forces, however, they run into the classic problem of circular or simultaneous causality (simultaneity). Firms or people may

settle in an area because it is productive, or areas may become more productive and attractive because firms and people choose to locate there. Ciccione and Hall correct for this endogeneity problem by using an instrumental variables approach based on historical measures of population density, geography, and climate. They use density data from the county level and output data on a state level to find that "a doubling of employment density increases average labor productivity by approximately six percent" (Ciccione and Hall 1996). Specifically, the paper creates a function for labor density that accounts for inequality of density within a state, and regresses this on state-wide labor productivity.

Table 1 below, reproduced from the *Handbook of Regional and Urban Economics: Cities and Geography*, provides one method of organizing the literature on agglomeration economies. It splits relevant studies into four categories: Industrial scope, Temporal scope, Geographic scope, and Organizational/ competitiveness, along with their key findings. We see that Ciccione and Hall's approach can be considered an estimation of the geographic scope of agglomeration economies (Henderson and Thisse 2004, 2125). My paper regresses state-level spatial density functions on labor productivity, and can be considered an addition to the literature on the geographic scope of agglomeration economies.

Table 1. Summary of the scope of agglomeration economies

Issue	Paper	Key Finding
Industrial scope	Moomaw (1981, 1983), Kukamura (1985), Calem and Carlino (1991), Nakamura (1985), Henderson (1986), Sveikauskas (1975), Henderson, Kuncoro and Turner (1995), Rosenthal and Strange (2001), Henderson (2003a), Glaeser et al (1992)	Urbanization Localization Diversity
Temporal scope	Glaeser et al. (1992), Henderson, Kuncoro and Turner (1995), Henderson (1997), Glaeser and Maré (2001)	Past agglomeration causes growth Lags in wage effects
Geographic scope	Rosenthal and Strange (2003), Ciccione (2002) Deckle and Eaton (1999) Ciccione and Hall (1996), Ciccione (2002)	Attenuating effects National effects Employment density
Organizational/ competitiveness	Glaeser et al. (1992), Henderson, Kuncoro and Turner (1995), Combes (2000b), Rosenthal and Strange (2003) Rosenthal and Strange (2003) Henderson (2003a)	Competition Small firms generate more Small firms receive more

Source: Henderson and Thisse 2004, 2125.

Studies estimating the function of agglomeration economies in India are limited. The paper "TFPG and Urbanization Economies: A Case of Indian Industries," published in 2000, examines the industry-specific effects of urbanization and urban employment on total factor productivity growth (TFPG) from 1976-1977 to 1992-1993. In the paper, Arup Mitra of the Institute of Economic Growth, Delhi, finds that the size of urban population in a state has significant positive effects on four industries: food products, cotton textiles, rubber, and petroleum and coal products. He also finds that total urban employment in a

state has significant positive effects in eight industries, namely: beverages, cotton textiles, woolen products, non-metallic minerals, basic metal, metal products, transport, and other manufacturing industries. However, for seven of the eleven total industries the benefits of urbanization decline after urban population or employment crosses a certain threshold, suggesting an inverse U-shaped relationship between TFPG and urbanization. From Table 1, Mitra's analysis would likely be classified as addressing issues of industrial scope with findings on the importance of urbanization. My work on the geographic scope of agglomeration economies and their national effects complements Mitra's work. Though I do not engage in an industry-level analysis, I use a more robust methodology to examine the effects of *spatial density* of people, labor, output, and firm formation—as opposed to urban population or employment—on labor productivity. While Mitra finds that certain industries exhibit a U-shaped relationship between urbanization and TFPG, my work helps to situate this result within the context of statewide labor productivity returns to density (Mitra 2000).

In "Agglomeration Economies and Productivity Growth in India," Ashta Agarwalla of the Indian Institute of Management, Ahmedabad, investigates whether a study of the service sector can provide insight into role of agglomeration economies in India. Expanding upon Mitra's methodology, Agarwalla determines the effects of the proportion of urban population and diversity of urban industry on state-wise TFP in India. The paper finds that urbanization and TFP have an inverse U-shape after a certain urbanization threshold even for services, but that higher urban diversity increases TFP at all levels. Furthermore, the results suggest that regional diversity, rather than district-level diversity, may be a better unit of

analysis for future research because data on the regional level is more reliable. Finally, Agarwalla states that it is important to understand the effects of spatial density of urbanization activity on productivity (Agarwalla 2011). I build on the foundation Agarwalla lays for future research by performing a region-level analysis of spatial density on productivity. I use a density function rather than urbanization measures though, which the literature in the U.S. and Europe shows is a more sophisticated measurement of agglomeration. My analysis provides a macro-view of the labor productivity returns to density in the economy as a whole to complement Mitra and Agarwalla's micro-economic analysis.

Recent research and empirical analysis regarding entrepreneurship and firm formation has increased to complement academic studies on population density. In the 2011 paper "The Contribution of New Business to Regional Employment," Dr. Michael Fritsch and Dr. Yvonne Schindele of Friedrich Schiller University find that both quantity and quality of new business formation, as measured by size of firm and survival rates, are good indicators of economic growth, as evidenced by a study of regions in West Germany. Fritsch and Schindele also find that region-specific variables play an important role in explaining economic growth, and reinforce the importance of controlling for geographic factors (Fritsch and Shindele 2011). I apply these findings by using a dataset that includes firms over an established size threshold as a measure of firm quality and employing instrumental variable and panel data analysis to control for demographic and geographic variation. My results are consistent with those of Fritsch and Shindele in that they underscore the role of higher quantities of quality firms in increasing economic growth.

Timothy F. Slaper et al. approach the relationship between innovation and economic growth from a different angle in their 2011 paper, "The Index of Innovation: A Tool for Regional Analysis," published in *Economic Development Quarterly*. Slaper et al. compile an innovation index to assess whether and to what degree innovation explains the growth of labor productivity, as measured by GDP per worker across U.S. counties. The index is comprised of approximately 15 variables including measures of population density, education, research and development, venture capital investment, high-tech employment, and natural resources. After calculating the index across U.S. counties and regressing it on labor productivity, they find that county-specific factors such as population density have the strongest effects on labor productivity, lending further credence to the importance of this relationship. The authors also find that further studies should examine effects across regions instead of across counties, because counties are too targeted to determine useful trends of regional or geographic diversity, which is similar to Fritsch and Schindele's results (Slaper et al. 2011). The level of detail at which I have reliable data in India is the state level, but the work of Slaper et al. suggests that state-level analysis in fact enables valuable analysis of larger trends. While the innovation index takes a broad approach, examining a series of variables, I present a focused analysis using a single independent and a single dependent variable in order to quantify the specific effect of population density on labor density.

Concluding remarks on the existing literature

The existing research on density of economic activity, firm formation, and productivity suggests that density and urbanization influence labor productivity and TFP. In addition, geographic and demographic factors provide robust instruments to determine the causal relationships that exist between them. My results find that population, labor, and output density increase labor productivity by two to three times as much in India from 1981-2001 as Ciccione and Hall's findings in the U.S. in 1988. The main contribution of this paper is to apply a preferred empirical specification of agglomeration economics to India, which analyzes the effects of density rather than urbanization on labor productivity. Most of the existing literature on India focuses on industry-wide analysis of the effects of urbanization on TFP, and my analysis contributes new information on the effect of the geographic scope of agglomeration economies on labor productivity. Furthermore, the existing literature suggests that it may be important to study the role of new firm formation and to distinguish in the methodology by firm size, because metrics that include smaller firms may be as relevant for productivity as a geography of fewer, larger firms with more resources. This paper brings together Ciccione and Hall's robust methodology with the increasingly important study of whether new firm formation and labor productivity exhibit similar returns to density.

Data

This empirical analysis uses a comprehensive panel of census data for 28 states and three union territories in India from 1981-2001. I hypothesize that labor productivity in India exhibits comparable returns to density to those found in the U.S. and Europe, where doubling the density of an area results in a six percent increase in labor productivity. Because the magnitude of density in India—in terms of population, labor, and output—is dramatically higher[5] and infrastructure and city planning are lacking, I anticipate higher congestion effects of increased density of economic activity. However, the returns to living in dense areas are also amplified, because they generally feature significantly improved access to goods and services. Existing research on the economic geography of India (Mitra 2000, Agrawalla 2011) suggests that agglomeration economies do exist in India. Overall, I expect the effects of increased density on productivity to be quite similar to or slightly more than those measured in the U.S. and Europe.

To estimate the parameters of this methodology, I use state-wise data available on GDP, land area, population, number of workers, and number of new firms incorporated to construct measures of density and output per worker. The data for land area, population, and number of workers is reported in the 1981, 1991, and 2001 Censuses of India. The Reserve Bank of India reports yearly GDP data. Variables for the density of population, labor, and output are constructed by dividing state-wise population, number of workers, and GDP, respectively, by state land area in square kilometers. Labor productivity is constructed by dividing the

[5] See Appendix A for a global population density map.

state-wise gross domestic product by number of workers, yielding a measure of output per worker.

I obtain data on the number of new firms incorporated from the Prowess Indian Corporate Database, which lists medium to large public and private firms according to the guidelines of the U.S. Center for Research in Security Prices (CRSP). Noting that India only achieved independence in 1947, I assume that this data is accurate for the current definition of the subcontinent only from that year forward. For this analysis, I reduce the dataset to include only those firms that were newly incorporated in 1981, 1991, and 2001, which leaves approximately 1,800 firms. According to Prowess, no new firms have been incorporated in the Andaman and Nicobar Islands or the tribal states of Mizoram, Sikkim, and Tripura since independence. Also, no firms were incorporated in 1981, 1991, or 2001 in Jammu and Kashmir or in the tribal state of Manipur. Table 2 below contains summary statistics for the variables used in the key regressions. From the table it is evident that the logarithm of new firm density has the largest variation across states, while the logarithm of rainfall has the least. Note that rainfall is used as a check on the population density instrument and is addressed in the Results section of this paper.

Table 2. Summary of key variables[6]

Variable	Mean	Std. Dev	Min	Max
Log population density	-.6477572	.6627491	-2.122231	1.635372
Log labor density	-1.120005	.5297864	-2.40181	.4863927
Log output density	-.9825261	.8114994	-2.933054	1.57316
Log new firm density	-4.75166	1.519407	-7.646841	-1.077719
Log population density (1971)	2.100233	.5665112	.7781513	3.438067
Log rainfall (2009)	2.758101	.3472276	2.087781	3.371806
Log labor productivity	.1374791	.4541376	-.5565973	1.086767

[6] See Appendix B for table of values for all variables, states, and years included in this analysis.

Data for the year 2011 could not be incorporated into this analysis because although the 2011 Census of India results have been released in the form of a preliminary report, state-wise data has not been released for number of workers, and Prowess has only been updated to include new firms through the end of 2010. These databases are expected to be current through 2011 by 2012 year-end.

Before proceeding further, it is important to note the political-geographic breakdown of the country. India is divided into 28 states and seven union territories, where the National Capital Territory of Delhi is considered a union territory. Jammu and Kashmir, the northernmost state of India, is disputed among India, Pakistan, and China, and the Indian state of Jammu and Kashmir is referred to as Indian-administered Kashmir or Indian-occupied Kashmir, depending on the party. The "seven sister states" are the relatively undeveloped and sparsely populated northeastern tribal states: Assam, Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland and Tripura.

Three Indian states, Jharkhand, Chattisgarh, and Uttarakhand, were formed in 2000 from existing states, causing a re-allocation of national land area. Namely, Jharkhand was carved from the state of Bihar, Chattisgarh from Madhya Pradesh, and Uttarakhand from Uttar Pradesh. For the purposes of consistent analysis, I recreate the original state by combining the data from the two new states. Thus, this paper includes data from 28 states and three union territories, but they are re-grouped to form 25 consistent state regions and three union territories.

Methodology

The specific methodology followed here is that of Ciccione and Hall (1996), applied to recent Indian data on population density, output density, labor density, and output per worker. I also expand upon this methodology, applying a similar instrumental variable regression to estimate the effects of the spatial density of medium and large firms on labor productivity.

Ciccione and Hall present two theoretical models to explain how levels of density affect labor productivity as the foundation for their empirical analysis. First, the externalities model considers the positive spillover effects resulting from the density function, including those stemming from abbreviated distances, shared resources, a robust local labor market, and enhanced flows of knowledge. If the positive agglomeration effects outweigh the negative congestion effects of density, states with higher density will have higher levels of labor productivity. Second, the intermediate product theory hypothesizes increasing production returns to local intermediate goods (Abdel-Rahman 1988, Riviera-Batiz 1988). That is, labor productivity is higher where there is more variety of intermediate goods produced, and higher density areas have more variety of intermediate goods. Thus through this mechanism, denser areas will have higher levels of labor productivity. This agglomeration effect outweighs the congestion effect when differentiated intermediate goods are important and are not good substitutes for each other, and as a result the production function will have increasing returns to density. Both theories present possible theoretical foundations to support the end empirical estimation method presented here to test whether higher density leads to higher labor productivity (Ciccione and Hall 1996).

Identification and Estimation

Instrumental variable regression allows consistent estimation of causal relationships when the explanatory and dependent variables may exhibit circular causality or when we are concerned about the distortive effects of confounding variables. As mentioned earlier, the classic problem of simultaneity is relevant in measuring the effects of agglomeration forces. Firms or people may settle in an area because it is productive, or areas may become more productive and attractive because firms and people choose to locate there.

The standard OLS regression of the form $y = \beta x + u$ makes the assumption that the density function (the regressor X) is uncorrelated with the error term (u), and that density only affects output per worker through the direct effect of X on Y. In this situation, we have reason to believe that there is a problem stemming from the endogeneity of the density function. In general, this endogeneity may occur as a result of simultaneous causality, as in this case (X affects Y, and Y affects X), because relevant variables are omitted from the methodology, or due to measurement error. Use of an instrumental variable allows isolation of only the exogenous variation in the density function and thus makes causal inference possible in a non-experimental setting.

There are two necessary conditions for a valid instrument: it must be relevant (corr $(Z, X) \neq 0$) and exogenous (corr $(Z, u) = 0$). That is, instrument must be correlated with the regressor X but uncorrelated with the error term: it cannot suffer from the same problem as the original predicting variable. This says that Z (the instrumental variable) affects Y (the dependent variable) only through X (the independent variable).

Instrumental variable regression is often conceptualized in two stages. In the first stage, the instrumental variable Z is regressed on X to obtain the estimated $\hat{X}$. In the second stage, Y is regressed on the estimated $\hat{X}$ to obtain an estimate of the coefficient β_{IV}, in which the regressor X is written as a linear function of the instrumental variable Z. These equations can be written as follows:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_i + u_i \\ X_i &= \gamma_0 + \gamma_1 Z_i + \varepsilon_i \end{aligned} \qquad (1.1)$$

We can understand the logic of this method by examining the estimated coefficient β_{IV}, which represents the ratio of the causal effect of Z on Y and the causal effect of Z on X:

$$\beta_{IV} = \frac{\text{cov}(Y_i,Z_i)}{\text{cov}(X_i,Z_i)} = \frac{\text{cov}(\beta_0 + \beta_1 X_i + \varepsilon_i, Z_i)}{\text{cov}(X_i,Z_i)} = \frac{\text{cov}(\beta_0,Z_i) + \text{cov}(\beta_1 X_i,Z_i) + \text{cov}(\varepsilon_i,Z_i)}{\text{cov}(X_i,Z_i)} = \frac{\beta_1 \text{cov}(X_i,Z_i)}{\text{cov}(X_i,Z_i)} = \beta_1 \qquad (1.2)$$

Here we see that $\text{cov}(X_i,Z_i)$ must be non-zero. Essentially, instrumental variable regression is a way to obtain a consistent estimator of the unknown coefficient of the regression function.

Instrumental Variable Regression

In the instrumental variable approach, my two primary assumptions are as follows. First, I assume that density is uncorrelated with the error term; that is density does not respond to the random element of labor productivity. Second, I assume that there exists an exogenous state characteristic that can function as an instrumental variable for density, namely the time-lagged population density variable (I use population density in 1971). Because we know that there are true differences in productivity across states, it is not realistic to assume that density is uncorrelated with those differences. States that have intrinsic features that make

them more productive (e.g., climate, geography) will attract workers. The key is that the instrument should be relevant—it should affect density—but that differences in labor productivity should not affect the instrument today (Ciccione and Hall 1996).

Ciccione and Hall use the following geographic and demographic factors as instruments:

1. Presence or absence of a railroad in 1860,
2. Population of the state in 1850,
3. Population density of the state in 1880,
4. Distance from the eastern seaboard of the U.S. (1993).

These are good instruments because we can assume that they are only related to worker preferences; they are not associated directly with labor productivity. I use population density of the state in 1971 as an instrument. As specified in Ciccione and Hall's paper, the time-lagged population density variable likely does not significantly affect modern labor productivity differences except through the effect of higher future population density.

If instruments are only weakly correlated with the endogenous variables or are not exogenous, tests of significance have incorrect size and may produce misleading results. Staiger and Stock (1997) formalized the definition of "weak instruments" and most studies accept that a first-stage F-statistic greater than ten indicates adequate instrument strength. First-stage F-tests for all regressions using population density in 1971 as the instrument reveal F-statistics generally far greater than ten. In all cases we do not have reason to believe that the instrument, population density in 1971, is a weak instrument.

First, let us take a brief look at Ciccione and Hall's empirical specification. Their paper first defines output production q in an area of space a by employing labor n such that:

$$f(n,q,a) = n^{\alpha}(\frac{q}{a})^{(\lambda-1/\lambda)} \qquad (2.1),$$

where α is the elasticity of output with respect to labor and $(\lambda-1)/\lambda$ is the elasticity of output with respect to density. Thus total output on a county level is $q_c = a_c(n_c/a_c)^{\alpha}(q_c/a_c)^{(\lambda-1)/\lambda}$. If we solve this function, it can be represented as:

$$\frac{q_c}{a_c} = \left(\frac{n_c}{a_c}\right)^{\gamma} \qquad (2.2).$$

Here, γ is the product of the production elasticity α and the elasticity of output with respect to density $(\lambda-1)/\lambda$. Now, if we aggregate this equation to the state level where ϕ is a constant that depends on the interest rate, A_s^w is a Hicks-neutral technology multiplier for state s with elasticity w, and $D_s(\theta,\eta)$ represents the aggregate density function,

$$\frac{Q_s}{N_s} = \phi A_s^w D_s(\theta,\eta) \qquad (2.3).$$

The resultant equation describes the average labor productivity in a state, given by the total output Q_s divided by the total number of workers in that state N_s. If A_s is log-normally distributed and if there is a randomly distributed error term we can take the logarithm of both sides,

$$\log\frac{Q_s}{N_s} = \log\phi + \log D_s(\theta,\eta) + u_s \qquad (2.4),$$

which yields the regression off of which my paper is based.

I present a density function that builds upon the foundations of previous work. Ciccione and Hall define density as "the intensity of labor, human, and physical

capital relative to physical space. Density is high when there is a large amount of labor and capital per square foot" (Ciccione and Hall 1996, 54). I use this definition to construct measures of density of output, population, and firms, though I employ metric units. Ciccione and Hall use state-level GDP, but their density parameter is aggregated from county-level data, thus incorporating inequality of density across counties within the state as a method to weight state density (Ciccione and Hall 1996, 56-57). This is desirable because the average density of a state may not be an accurate reflection of the true distribution of concentration of factors of production. However, reliable data on India is available only to the state level, and little accurate data is available on distribution of density within the state itself, rendering a similar analysis both challenging and unreliable. India's states are large, populous, and extremely distinct from each other, which I hope makes them useful units for analysis despite the lack of accurate county-level data.

The dependent variable, expressing labor productivity, is constructed by taking the logarithm of the ratio of GDP and number of workers at the state level, shown as the left side of equation (2.4), $log\ (Q_s/N_s)$. This term represents labor productivity as average output per worker. Labor productivity is consistently used as the dependent variable for all regressions.

Applying the Empirical Specification to India

I use a series of instrumental variable two-squares least squares regressions with one independent variable (X), one instrumental variable (Z), and one dependent variable (Y) to evaluate the causal relationship between density (X) and labor productivity (Y) in India. I run this regression using a time-lagged population

density instrument for three different endogenous spatial density variables, population density, labor density, and output density, over a cross-sectional dataset for each of the years 1981, 1991, and 2001 respectively, and over a panel dataset from 1981-2001. Thus there are three main regressions for each of the four time periods, for a total of 12 main regressions. I also expand upon these regressions, introducing new data on the number of firms formed to estimate the effect of new firm density on labor productivity for each of the four time periods.

Let us take population density as the X-variable in the first regression. Without instrumental variable regression, the linear Ordinary Least Squares (OLS) regression of the form $y = \alpha + \beta x + u$ with one dependent variable (labor productivity) and one regressor representing the population density function for a single year would be specified as:

$$\log\frac{Q_i}{N_i} = \beta_0 + \beta_i(\log\frac{P_i}{A_i}) + u_i \quad (3.1),$$

where i denotes state, $\log(Q_i/N_i)$ is the labor productivity expressed as a logarithm of the ratio of state GDP divided by the total number of workers in that state in a given year (Y-var), $\log(P_i/A_i)$ is the population density expressed as the logarithm of the ratio of the state population divided by the total land area of the state (X-var), and u_i is an error term.

Now let us instead use an instrumental variable approach following the Ciccione-Hall methodology, where population density is the independent variable, labor productivity is the dependent variable, and population density in 1971 is the instrument. The instrumental variable regression is specified as:

$$\log\frac{Q_i}{N_i} = \beta_0 + \beta_{IV}\log\frac{P_i}{A_i} + u_i \quad (3.2),$$

where i denotes state, $\log(Q_i/N_i)$ is the labor productivity expressed as a logarithm of the ratio of state GDP and the total number of workers in that state in a given year (Y-var), $\log(P_i/A_i)$ represents the *estimated value* of the state population density from the regression of the instrument on population density ($\hat{X}$-var), and u_i is an error term. This is the empirical specification used for each regression in this paper. Table 3 below illustrates the instrumental variable regressions run for population, labor, output, and new firm density on labor productivity for each time period included in this analysis.

Table 3. Instrumental variable regressions of various density functions on labor productivity as regressed for each time period

Independent variable	Instrumental variable regression
Population density P_i = population	$\log\frac{Q_i}{N_i} = \beta_0 + \beta_{IV}\log\frac{P_i}{A_i} + u_i$
	Where $\log P_i/A_i = \hat{X}$
Labor density N_i = labor force	$\log\frac{Q_i}{N_i} = \beta_0 + \beta_{IV}\log\frac{N_i}{A_i} + u_i$
	Where $\log N_i/A_i = \hat{X}$
Output density Q_i = gdp	$\log\frac{Q_i}{N_i} = \beta_0 + \beta_{IV}\log\frac{Q_i}{A_i} + u_i$
	Where $\log Q_i/A_i = \hat{X}$
New firm density F_i = # new firms	$\log\frac{Q_i}{N_i} = \beta_0 + \beta_{IV}\log\frac{F_i}{A_i} + u_i$
	Where $\log F_i/A_i = \hat{X}$

In regressions where both the independent and dependent variables are logarithmic transformations, the coefficient is referred to as the elasticity of Y with respect to X. The coefficient β_{IV} is interpreted as the percent change in the dependent variable, labor productivity, when the independent density function increases by one percent.

Results

Table 4 displays the results of the instrumental variable regressions for each year across all four density measures, with population density in 1971 as the instrument. Higher levels of population, labor, output, and new firm densities result in significantly higher labor productivity over time.

Table 4. Summary of instrumental variable regression results, 1981, 1991, 2001, and Panel

X-var	Y-var 1981	Y-var 1991	Y-var 2001	Y-var Panel
Variables	Log labor productivity (1981)	Log labor productivity (1991)	Log labor productivity (2001)	Log labor productivity (1981, 1991, 2001)
log population density	0.188***	0.139***	0.176***	0.188***
	(0.0490)	(0.0474)	(0.0560)	(0.0501)
Constant	-0.208***	0.209***	0.756***	0.259***
	(0.0480)	(0.0421)	(0.0445)	(0.0410)
R-squared	0.351	0.277	0.101	
log labor density	0.208***	0.150***	0.208***	0.189***
	(0.0554)	(0.0520)	(0.0574)	(0.0524)
Constant	-0.105	0.280***	0.872***	0.349***
	(0.0733)	(0.0639)	(0.0648)	(0.0643)
R-squared	0.323	0.256	0.322	
log output density	0.172***	0.130***	0.172***	0.159***
	(0.0380)	(0.0393)	(0.0393)	(0.0371)
Constant	-0.0870	0.244***	0.721***	0.293***
	(0.0644)	(0.0457)	(0.0272)	(0.0426)
R-squared	0.536	0.438	0.536	
log new firm density	0.125***	0.0600***	0.120***	0.0959***
	(0.0375)	(0.0187)	(0.0382)	(0.0247)
Constant	0.222	0.380***	1.280***	0.593***
	(0.177)	(0.0866)	(0.199)	(0.120)
R-squared	0.145	0.402	0.105	

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Let us begin by analyzing the effects of population density on labor productivity, i.e. output per worker. Table 4, Row 3 shows that doubling the population density of a given region results in a 14 to 19 percent increase in labor productivity. Density has a significant effect on the elasticity of labor productivity

at the 99 percent confidence level in 1981, 1991, and 2001 separately and when regressed as a panel. This result is significantly higher than Ciccione and Hall's results in the U.S. in 1988, which find that a doubling of density results in a six percent labor productivity increase. The effect of population density is highest (18-19 percent) in 1981 and 2001 and lowest (14 percent) in 1991. The balance of payments crisis, uncertain economic climate, and shifting economic policy reforms in 1991 may have disrupted the relationship between density and labor productivity. The average R-squared value for population density is 0.24, which indicates that only 24% of the variance in labor productivity is explained by population density. This is a significantly smaller portion than Ciccione and Hall's R-squared values of 55%, indicating that the population density regression for India has lower goodness of fit.

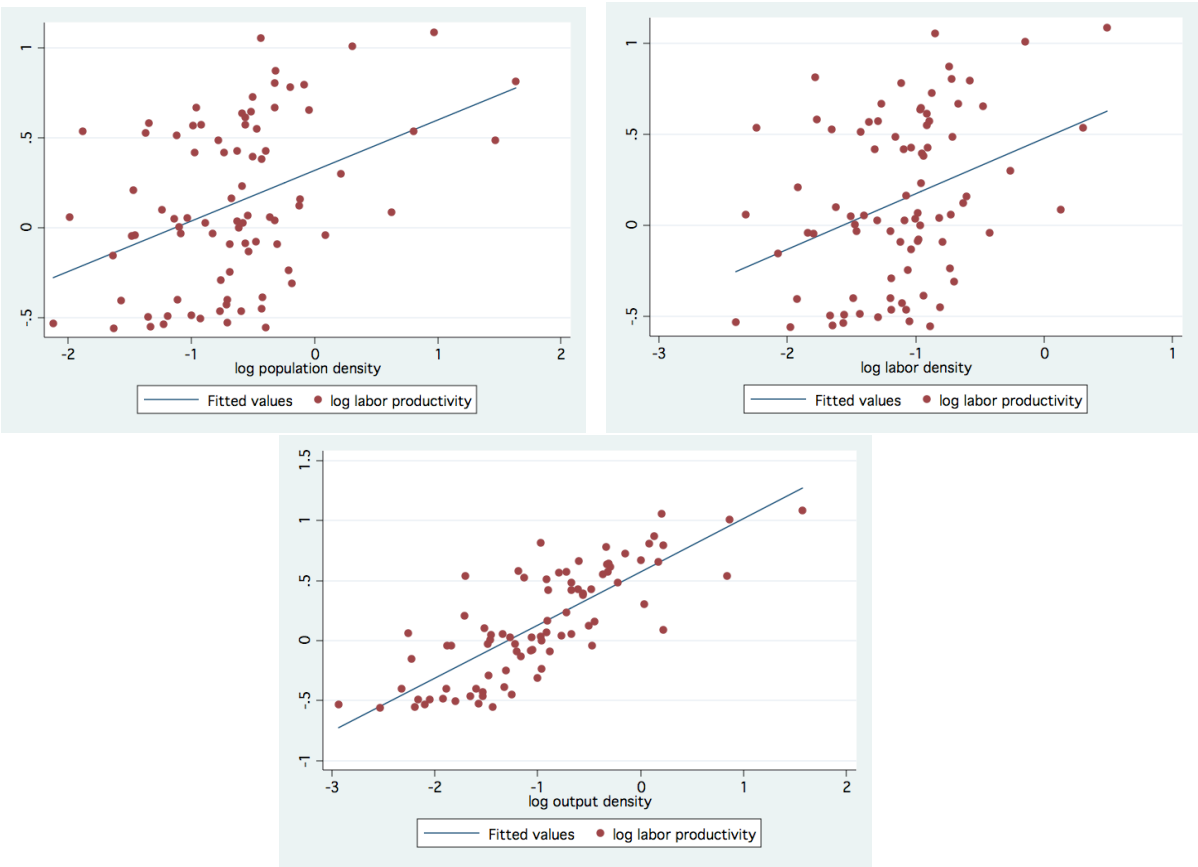
Next, labor density also exhibits a significant effect on labor productivity at the 99 percent confidence level in 1981, 1991, and 2001, both separately and when regressed as a panel. The effect is approximately 21 percent in 1981 and 2001 but falls to 15 percent in 1991. As noted above, 1991 was a year of economic turbulence in India and relatively uncertain growth, and both population density and labor density exhibit lower effect on labor productivity during this time. The R-squared values for labor density are higher than for population density, averaging at 30%, and indicate better goodness of fit.

Third, density of output significantly affects output per worker in 1981, 1991, and 2001, both separately and when regressed as a panel, at the 99 percent confidence level. The elasticity coefficient indicates that doubling output density results in a 13 to 17 percent increase in labor productivity. The effect is consistent

at 17 percent in 1981 and 2001 but dips to 13 percent in 1991, which mirrors the lower returns to labor productivity expressed by both population density and labor density in 1991. The effects of output density are the results that are most directly comparable to those of Ciccione and Hall, and illustrate that output density has a two to three-fold greater effect on labor productivity across states in India from 1981-2001 than across counties in the U.S. in 1988. The R-squared values from the output density regressions indicate the highest goodness of fit, with R-squared averaging 0.51, which is very similar to Ciccione and Hall's at 0.55.

Scatterplots of density versus labor productivity demonstrate positive correlations that reinforce my results, but indicate that instrumental variables analysis does play an important role in deciphering these relationships. A scatterplot of population density versus labor productivity reveals a clear positive relationship. An analogous plot of labor density versus labor productivity exhibits a similar relationship, though the line of best fit is slightly flatter, as shown in Figure 4 below. The strongest relationship is between output density and labor productivity. This is noteworthy in light of the results of the instrumental variable regressions, which show that labor density affects labor productivity the most (15 to 21 percent elasticity), and output density affects labor productivity the least (13 to 17 percent elasticity). From these plots we can infer that use of an instrument to capture only the effect of density on labor productivity yields a different result than would have been observed without instrumental variable regression. This reinforces the presence of simultaneous causality and substantiates the use of instrumental variables regression to draw causal inference.

Figure 4. Scatterplots of population density, labor density, and output density versus labor productivity, 1981-2001



After conducting instrumental variable regressions for each year separately and as a panel, it might also be interesting to investigate whether population, labor, and ouput density in base years 1981 and 1991 had a significant effect on growth of labor productivity over the following ten-year period. We do this cognizant of the low likelihood of conclusive results, given that our data in decade-long increments. Using the difference in logarithms to evaluate the effect of density level on labor productivity over the following decade, however, density does not appear to affect growth in output per worker during the 1981-1991 and 1991-2001

time periods. Level of population density, labor density, and output density do not significantly contribute to the change in labor productivity over the next decade.[7]

Overall, my results are consistent with previous studies (Mitra 2000, Agarwalla 2011), which suggest the presence of agglomeration economies in India. However, but the effects found are significantly higher than those found in similar studies for the U.S. and Europe, which place returns to density in the six percent range.

There are a few factors that potentially explain the difference in results between India and the U.S. First, in India there are more than ten times as many people per square kilometer than in the U.S.: in 2011, approximately 380 people lived per square kilometer in India versus 32 people per square kilometer in the U.S. (Census of India 2011, CIA World Factbook). The magnitude of Indian density is radically higher, increasing the intensity of concentration of factors of production. This implies that changes in density have a more powerful effect. Furthermore, the inequality of density across regions is greater in India. In 1981, New Delhi was more than 550 times as densely populated as the mountain state of Arunachal Pradesh. In 2001, Delhi was over 700 times as densely populated as Arunachal Pradesh. This pattern is also reflected in the inequality of labor productivity across regions. In the U.S., Ciccione and Hall found that in 1988, "output per worker in the most productive state was two thirds larger than in the least productive state" (Ciccione and Hall 1996). In India in 1981, output per worker in the most productive state, Delhi, was four times larger than in the least

[7] See Appendix C for the results of these growth regressions.

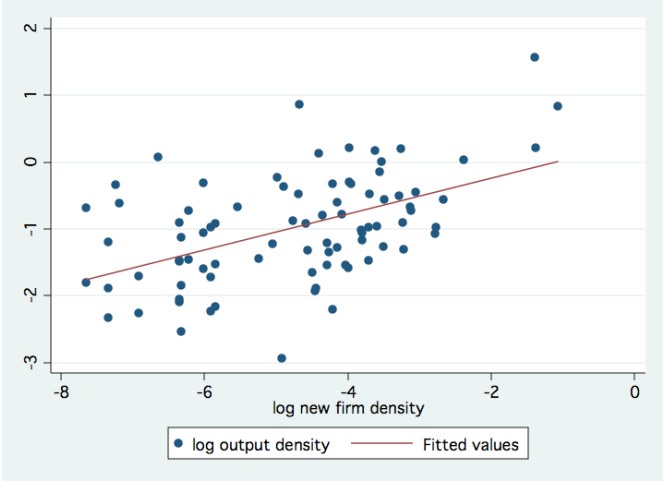
productive state, Mizoram (Reserve Bank of India, Census of India 1981). In addition, the difference between urban and rural areas is great, and sparsely populated areas are often particularly infrastructure, resource, and capital-poor. It naturally follows that the labor productivity returns to increased density are correspondingly large.

In an extension of Ciccione and Hall's methodology to include new firm density, I find that higher spatial density of new medium and large firms has significant returns to output per worker, as shown in Table 4, Row 5. We recall that the variable for spatial density of firms is constructed by dividing data on number of firms incorporated per year in Prowess by state land area data from the 1981, 1991, and 2001 Censuses of India. From this analysis, I demonstrate that the spatial density of new firms explains a significant amount of the variation in labor productivity across states in India in 1981, 1991, and 2001 separately and when regressed as a panel. The spatial density of new firms has less of an effect on output per worker than population, labor, or output densities do, but it still increases labor productivity by 6 to 12 percent. I observe the same pattern as demonstrated by the population, labor, and output density coefficients; spatial density of new firms has the strongest effect in 1981 and 2001 (approximately 12 percent), but drops in 1991 (six percent).

The new firm density regressions provide a good check on the previous regressions because we might hypothesize that density of output, or GDP per square kilometer, is highly correlated with density of firms incorporated, or number of firms per square kilometer. Intuitively, firms produce output, so new firm density should be a good predictor of output density. In fact, when we check this

hypothesis, we find that the logarithm of density of firms and the logarithm of density of output are correlated at approximately 0.5. The two variables are not correlated perfectly, but there is a positive relationship between them as illustrated in Figure 5 below.

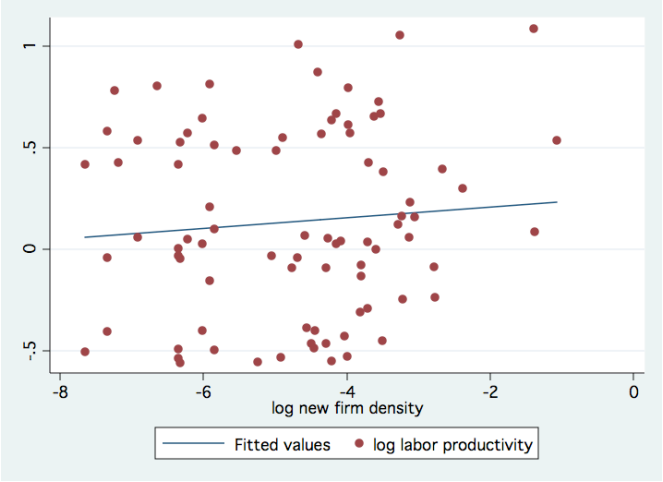
Figure 5. Scatterplot of new firm density versus output density, 1981-2001



However, the relationship between density of new firms and *labor productivity* is not necessarily intuitive. Though Figure 5 shows that new firm density might produce similar labor productivity returns as output density produces, the logarithms of new firm and labor densities are correlated with only a value of 0.2. New firms are not necessarily the most productive firms, and incorporating a company in a given state requires a high level of capital and resource investment that may not realize immediate labor productivity gains. Figure 6 below illustrates that the relationship between new firm density and labor productivity is not obvious from the observed data. The data are scattered and do not reveal a clear trend; the line of best fit is flatter than those in the Figure 4

scatterplots of population, labor, and output density versus labor productivity. The weak observed relationship between new firm density and labor productivity in Figure 6 indicates that the extended instrumental variable regressions that find significant effects of high new firm density on labor productivity represent an important contribution.

Figure 6. Scatterplot of new firm density versus labor productivity, 1981-2001



After regressing new firm density on labor productivity in each year separately, I also examine its effects on growth of labor productivity over time. The effects of new firm density in base years 1981 and 1991 on growth of labor productivity over the following ten-year period do not appear to be significant. New firm density, like population, labor, and output density, does not show significant effects in the regressions of growth of labor productivity over 1981-1991 and 1991-2001.[8]

[8] See Appendix C for the results of these growth regressions.

Before proceeding, it may be useful to place this result in context. We recall that Prowess only contains records of firms that submit tax records and report market capitalizations that qualify them as medium and large-sized firms, though they do comprise over 75 percent of corporate tax revenue. Thus, this paper's findings on new firm density are rather targeted. But the result that higher densities of medium and large-sized firms increase average state labor productivity is understandable given the lack of infrastructure, the generally poor ability of the government to foster economic growth, and the dominant role of the private sector in developing urban India. The interaction between the broader Indian economic environment, density, and labor productivity is nuanced, and I consider further implications in my discussion.

A new instrument

In considering the highly significant and large effects of density on labor productivity, a potential concern is the true exogeneity of the instrument. That is, we might wonder whether population density in 1971 is necessarily exogenous to labor productivity, or whether it might influence labor productivity though variables other than current population density such as level of development. To address this concern, I conduct the same regressions using state rainfall in 2009 as an instrument, which has the advantage of being truly exogenous and can serve as a check on the results obtained using the population density instrument.

The results from the new regressions using rainfall as an instrument are displayed in Table 5 below. First, the first-stage F-tests reveal that the rainfall instrument is weak, while the population density (1971) instrument is strong.

Table 1, Column 3 shows that the logarithm of 2009 rainfall has the lowest variation of any of the main variables in this analysis, which may partially explain the weakness of the instrument. All first-stage F-statistics of the instrumental variable regressions using rainfall (2009) are less than ten, revealing that is a statistically weak instrument; the first-stage F-statistics for population density (1971) are consistently greater than ten. The magnitudes of the coefficients are relatively similar across both instruments, but the results using rainfall as an instrument are less significant. We expect this, however, given that rainfall's low F-statistics indicate that it is a weak instrument.

Table 5. Comparing instrumental variable results using two different instruments, population density (1971) and rainfall (2009)

Year	X-Var	Instrument	Coefficient	First stage F-statistic
1981	Population density	Population density 1971	0.188***	534.42
		Rainfall 2009	0.225*	3.86051
1981	Labor density	Population density 1971	0.208***	448.12
		Rainfall 2009	0.240	0.0513
1981	Output density	Population density 1971	0.172***	265.24
		Rainfall 2009	0.193**	4.25
1981	New firm density	Population density 1971	0.125***	14.12
		Rainfall 2009	1.194**	3.98
1991	Population density	Population density 1971	0.139***	438.32
		Rainfall 2009	0.225*	4.09
1991	Labor density	Population density 1971	0.150***	382.09
		Rainfall 2009	0.230*	4.64
1991	Output density	Population density 1971	0.130***	189.26
		Rainfall 2009	0.187**	5.05
1991	New firm density	Population density 1971	0.0600***	55.36
		Rainfall 2009	0.0714**	5.79

Table 5. Comparing instrumental variable results using two different instruments, population density (1971) and rainfall (2009) continued

Year	X-Var	Instrument	Coefficient	First stage F-statistic
2001	Population density	Population density 1971	0.176***	71.85
		Rainfall 2009	0.127	2.65
2001	Labor density	Population density 1971	0.208***	199.20
		Rainfall 2009	0.117	5.77
2001	Output density	Population density 1971	0.172***	150.07
		Rainfall 2009	0.105	4.58
2001	New firm density	Population density 1971	0.120***	11.97
		Rainfall 2009	0.100	0.80
Panel	Population density	Population density 1971	0.188***	N/A
		Rainfall 2009	0.240	N/A
Panel	Labor density	Population density 1971	0.189***	N/A
		Rainfall 2009	0.192	N/A
Panel	Output density	Population density 1971	0.159***	N/A
		Rainfall 2009	0.161*	N/A
Panel	New firm density	Population density 1971	0.0959***	N/A
		Rainfall 2009	0.0833*	N/A

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Discussion and Conclusions

In this paper, I examine the role of agglomeration economies, as captured through various density functions, in increasing average labor productivity. The purpose is to determine whether agglomeration economies in India create similar labor productivity returns to those found by Ciccione and Hall in the U.S., or if congestion costs in extremely high-density areas diminish the benefits of agglomeration. My empirical results provide evidence that density, as measured by

population, labor, and output density, increases labor productivity across states in India from 1981-2001 by two to three times more than observed in the U.S. in 1988. The significant effects suggest that there are still increasing labor productivity returns to increasing the density of areas in India.

There are also significant labor productivity returns to increased density of new firms in a given state across all three years studied. As a result of using a dataset that only includes firms with medium and large-sized market capitalizations, the measure of new firm density by its nature describes new firms of a certain success level. In other words, having more new medium and large-sized firms per square kilometer in a given state increases the average labor productivity of that state. Thus, the positive agglomeration effect of increased density of new firms overpowers the potential negative congestion effect that results from many firms competing over limited resources, land, labor, capital, and customer bases. Because starting a new company often requires significant initial investment, it can be considered to some degree a measure of confidence in the future economic growth potential of a state. As more firms locate in an area, the risks, costs, and barriers to entry for other firms decline. The first medium and large firms in an area often create infrastructure, establish power plants and water pipelines, and invest in research and development, increasing the pace of technological innovation, establishing best practices and encouraging growth. The causal effect of new firm density on labor productivity is as hypothesized, and is a good check on the output density-labor productivity relationship, since output density and new firm density are positively correlated as shown in Figure 5.

We might find a different result if we were to analyze firms of all sizes, not just medium and large firms. Fritsch and Schindele illustrate that quantity of *quality* firms is an important indicator, and adhere to their emphasis on firm size since medium and large firms are most likely to drive productivity gains (Fritsch and Schindele 2011). While medium and large firms seem to flourish in densely corporate environments, small firms may struggle to succeed in the shadow of more powerful companies. If I were to include the myriad of small family-owned and other private businesses that employ the majority of the Indian labor force in my measurement of firm formation, it is possible that a higher density of new firms may not increase labor productivity. It would be useful to expand on the implications of small firm density in future work; however, it is outside the scope of this paper.

Overall, consistent patterns emerge from the twelve regressions for population, labor, output, and new firm density: across all density measures, there is a dip in the effect of density on labor productivity in 1991. This pattern may reflect the depressed and uncertain state of the economy at the time, but the magnitude of the effect (13 to 15 percent) shows that there are still significant positive returns to increased density even in a time of economic downturn. Similarly, the elasticity coefficients are extremely similar for all density measures for 1981 and 2001, as displayed in Table 4, Columns 2 and 4, respectively. In 1981 and 2001, the elasticity of population density is approximately 18 percent, labor density is 21 percent, output density is 17 percent, and new firm density is 12 percent.

The results of the panel data regressions using data for all three years and 28 areas included in this analysis are given in Table 4, Column 5. When taken over

two decades together, the elasticity coefficients of population density and labor density are the highest at 19 percent, as compared to output density at 16 percent and new firm density at 10 percent. More densely populous areas, as measured by people per square kilometer or by number of workers (labor) per square kilometer, have higher returns to labor productivity than output or new firm-dense areas do. To differentiate even further, the effects of higher new firm density on labor productivity are the smallest, with an elasticity coefficient six percent lower than that of output density. This may suggest that the goods produced matter more than the number of institutions producing them. The importance of the output, or goods produced, relates to the intermediate goods theory discussed earlier that asserts that the number and diversity of goods available increase labor productivity.

The strong effects of population and labor density on increasing labor productivity imply that the externality theory may explain more of the agglomeration effect in India than the intermediate goods theory. What we might call the 'people-effect,' or the benefit of having more people or laborers per square kilometer in an area, facilitates increased human contact and the spread of knowledge and ideas. If instead output and new firm density had higher returns to labor productivity, then it might suggest that increasing the variety of intermediate goods produced was a larger source of agglomeration benefits. But the large labor productivity returns to areas that are dense in people and workers suggest that the costs of moving people and labor are high, and that increasing their proximity reduces these costs.

The importance of the people-effect is supported by Glaeser and Kohlhase's 2003 paper "Cities, Regions, and the Decline of Transport Costs" (Glaeser and Kohlhase 2003). The authors assert that costs of moving people have not declined as quickly as costs of goods. Figure 7 depicts their main findings, that the costs of transporting goods, whether inputs or finished products, have declined by over 90% during the twentieth century. Even though technological progress and the creation of leaner, more efficient machines have enabled people to move their goods across space more quickly and cheaply, moving people is still expensive. The advantages of proximity tend to stem overwhelmingly from improving flows of knowledge, and the role of cities is shifting towards facilitating the contact of individuals, not moving goods. The higher effect of labor and population density on labor productivity in India reinforces the economic benefits of increasing the proximity of people that Glaeser and Kohlhase find (Glaeser and Kohlhase, 2003).

Figure 7. The declining share of transport costs in GDP in the US

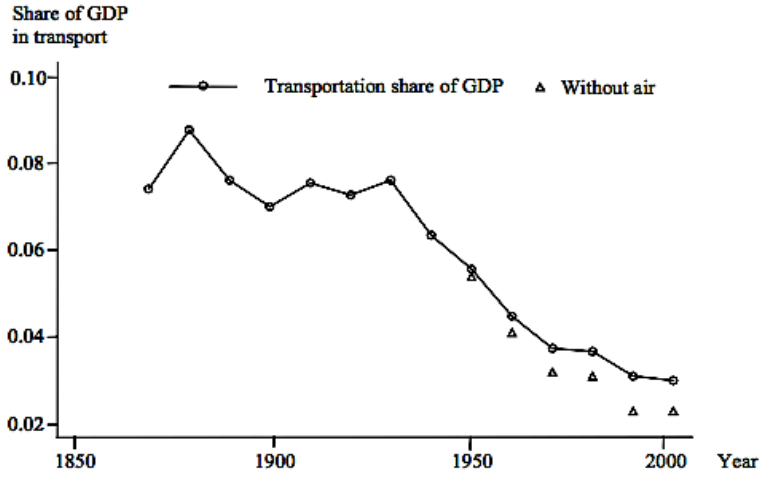


Fig. 1. The share of GDP in transportation industries
Source: Department of Commerce (since 1929), and Historical Statistics of the U.S. (Martin Series) before then.

Source: Glaeser and Kohlhase 2003.

Taken as a whole, the results presented in this paper suggest that either agglomeration economies are more influential in India than in the US, or that congestion costs are less prohibitive. There is no reason to believe that the costs of increased density, and therefore congestion, would be different in India than in the U.S. However, my findings suggest that agglomeration economies may have more impact. In India, where there is often a severe lack of infrastructure, slower technology penetration, and steep costs to getting help or information, agglomeration economies yield higher benefits. Intuitively, developing economies are farther away from their production possibility frontiers than developed countries are, and it follows that the marginal returns to density are higher than those in developed economies.

Moreover, the Indian government is largely considered ineffectual and has limited ability to provide a basic ecosystem for economic growth outside of highly dense areas. The private sector is responsible for a large share of urban and economic development. When more people, workers, gross product, or large and medium firms locate in a particular state, the private sector often has incentives to invest in large amounts of infrastructure, literally paving the way for more sources of economic growth to enter, or providing for existing people, workers, production, and firms to flourish. In Delhi, even the public, city-wide metro construction project was partially executed on a public-private partnership (PPP) basis between the government-represented Delhi Metro Rail Corporation (DMRC) and large corporate partners Reliance Energy and CAF (a Spanish company). Cyber parks, zones of concentrated, high-quality commercial infrastructure chartered by private corporations, have sprung up in cities such as Hyderabad, Bangalore, Gurgaon, and

even lesser-known cities such as Indore and Bhopal, forming population, labor, output, and firm-heavy regions within states. The Indian government is also promoting the creation of Special Economic Zones (SEZ), which are designated areas within states that have more liberal economic laws in order to attract private investment. The government is working with large local and foreign corporations including Mahindra, Biocon, Wipro, Infosys, DLF, and Nokia to construct and develop SEZ (Ministry of Commerce 2012). In SEZ and outside them, private companies frequently build roads, install water pipelines, establish power plants and grids, and invest in other physical infrastructure that increases the ability to conduct business and produce and distribute goods and services.

There are several limitations of this paper that deserve acknowledgement. First, though state-level in India is the most reliable data, a lack of county or district-level data is notable in light of Desmet and Rossi-Hansberg's work in the U.S., which shows that higher levels of spatial disaggregation are important (Desmet and Rossi-Hansberg 2009). Better district-level data will enable more precise measurement of labor productivity returns to agglomeration economies.

Second, this paper analyzes labor productivity returns to density by state and over time, but does not distinguish industry-level effects. As mentioned earlier, this paper is able to provide insight on a quality-controlled set of firms by analyzing only medium and large-sized new firms. However, this approach does not yield an industry-specific understanding of the effect of new firm formation on labor productivity. Especially in India, where the services industry has grown dramatically over the past few decades to overtake the size of the manufacturing industry, sector-wise analysis is particularly relevant. According to the Economic

Survey of India 2010-2011, the services industry is the fastest growing sector and accounts for 57.2 percent of GDP, while manufacturing and agriculture account for 28 and 14.6 percent of GDP, respectively (Economic Survey 2010, 4-5). A more comprehensive study would examine labor productivity effects across manufacturing versus services industries, mature versus early-stage industries, and formal versus informal sectors. A recent paper by Ghani, Goswami, and Kerr finds that manufacturing plants in the formal sector are moving to rural areas, while informal sector activity is migrating to urban centers (Ghani, Goswani, and Kerr 2012). Education and infrastructure are stronger in urban areas and are particularly important for promoting informal sector activity, but more mature manufacturing industries have already established best practices and thus do not benefit as much from knowledge sharing. These plants may also favor lower real estate costs in more sparsely populated areas (Ghani, Goswani, and Kerr 2012). The migration of manufacturing jobs to rural areas suggests that a policy tradeoff may be emerging between fostering employment growth in rural areas versus encouraging labor productivity growth in urban areas. Applying an industry-specific analysis to the effects of density on labor productivity would provide useful information for policymakers in determining how to best promote inter-industry economic development in both rural and urban areas.

Finally, this paper does not incorporate the possible effects of human capital on labor productivity. Some studies in the U.S. and Europe add a human capital variable to the regression of density on labor productivity, represented by a measure of educational attainment (Abel et al. 2010). Just as workers provide a different measure of agglomeration economies than population provides, the

educated population may capture yet a different effect. In "Productivity and the Density of Human Capital," Abel et. al. employ a model of urban productivity to estimate the human-capital-enhancing effect of density on TFP for U.S. data from 2001-2005. They find that the effects of density on productivity are overstated in the literature. Their results show that a doubling of density increases TFP by two to four percent as opposed to by six percent (Ciccione and Hall 1996, Abel et al. 2010). In India, where the average education level is lower and access to and quality of schooling is widely disparate, quantifying the population with higher-level degrees as a measure of human capital as Abel et. al do in the U.S. may not be useful. Future studies should carefully evaluate which education variables provide the most appropriate measures of human capital in India, as these may not be the same as in the U.S.

The results and implications of this paper present noteworthy and actionable conclusions for economists and policymakers. It is clear that large agglomeration economies are active across states in India, with multiple density measures consistently demonstrating two to three times higher effects on labor productivity over a twenty year time period than have been found in the U.S. This paper quantifies agglomeration through density measures and its effects on labor productivity for the first time in India. Using a methodology that is consistent with relevant literature in the U.S. and Europe, I provide an empirical basis upon which to guide the Government of India's stated focus on cities and mega-cities as drivers of economic growth. In India, where 70 percent of the population still resides in rural, sparsely populated areas, there is potential for the government to encourage increased density in urban centers. States that aim to increase labor productivity

might consider urbanization-promoting policies such as infrastructure or education investments that will induce the creation of more high-density areas.

However, I insert a cautionary note that my analysis does not advocate dramatic rural-urban migration to already highly-dense cities, but rather implies that increased concentration of factors of production over states as a whole exhibit increasing labor productivity returns. This paper's results urge policymakers to consider the scope of the areas over which resources are currently spread in order to identify opportunities to encourage clustering of economic activity. This may include developing more sparsely populated areas in which the number of significant firms is growing, or investing in reducing congestion costs in larger cities. Both the state and federal governments should look to the private sector as a vital partner in encouraging productivity growth.

Mindful of the delicate balance between clustering and over-urbanization, between productivity and standard of living, this paper raises the important question of how to determine the optimum urban density, or whether such an optimum exists. Questions like these deserve further study. Above all, the work presented here demonstrates that the spatial dynamics of the Indian economy is increasingly relevant and presents a vital field for future research.

References

Data Sources

Census of India
http://censusindia.gov.in/
http://censusindia.gov.in/Metadata/Metada.htm.

Reserve Bank of India
http://www.rbi.org.in/

IndiaStat
www.Indiastat.com

Prowess Indian Corporate Database
http://www.cmie.com/database/?service=database-products/firm-level-data-services/prowess-corporate-database.htm

Other

Abraham, Akshai et al. "Economic Activity Map of India." *Institute for Financial Management & Research (IMFR).* http://cdf.ifmr.ac.in/?project=economic-activities-map (accessed March 20, 2012).

Balakrishnan, Pulapre. *Economic Growth in India.* New Delhi: Oxford University Press, 2010.

Bolton, R. *Handbook of Regional and Urban Economics, Vol. 4, Cities and Geography.* Edited by J. Vernon Henderson and Jacques-François Thisse. Journal of Regional Science, 47: 372–377, 2004.

"Business: Unequal Opportunities: Why Entrepreneurs Thrive in Some Countries More than Others." OECD 50: Better Policies for Better Lives (June 29, 2011) http://www.oecd.org/document/8/0,3746,en˙21571361˙44315115˙48276232˙1˙1˙1˙1,00.html. (accessed October 8, 2011).

"Chasing Gazelles to Achieve Poverty Reduction Development Goals" NextBillion (December 20, 2011). http://www.nextbillion.net/blog/2011/12/20/chasing-gazelles-to-achieve-poverty-reduction-development-goals (accessed December 25, 2011).

Ciccione, Anthony and Robert E. Hall. "Productivity and the Density of Economic Activity." *National Bureau of Economic Research*. March 1996.

Desmet et al. "The Spatial Development of India. " Working Paper, unpublished (December 2011).

Desmet, K. and Rossi-Hansberg, E.. "Spatial Growth and Industry Age," *Journal of Economic Theory* 144 (2009): 2477-2502.

"Economic survey of India 2010." Ministry of Finance, Government of India. http://indiabudget.nic.in/budget2011-2012/survey.asp (accessed April 1, 2012).

Edward L. Glaeser and Janet E. Kohlhase. *Cities, Regions, and the Decline of Transport Costs*. Cambridge, MA: *National Bureau of Economic Research,* NBER Working Paper No. 9886 (July 2003). http://www.nber.org/papers/w9886 (accessed April 1, 2012).

"Entrepreneurship at a Glance 2011." OECD 50: Better Policies for Better Lives. OECD-Eurostat Entrepreneurship Indicators Programme (June 2011). http://www.oecd.org/document/0/0,3746,en˙2649˙37461˙48107008˙1˙1˙1˙37461,00.html (accessed October 8, 2011).

Fritsch, Dr. Michael and Alexandra Schroeter. "Why does the effect of new business formation differ across regions?" (January 22, 2010). Springerlink.com (accessed October 7, 2011).

Fritsch, Dr. Michael and Dr. Yvonne Schindele. "The Contribution of New Businesses to Regional Employment – An Empirical Analysis." *Economic Geography*. Clark University (2011). http://onlinelibrary.wiley.com/doi/10.1111/j.1944-8287.2011.01113.x/full (accessed October 7, 2011).

Glaeser, Edward. "Why Humanity Loves and Needs Cities." *The New York Times Online, Economix Blog* (April 23, 2010). http://economix.blogs.nytimes.com/2010/04/13/why-humanity-loves-and-needs-cities/ (accessed March 20, 2012).

Glaeser et al. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns." *Harvard University and NBER* (January 2009). http://www.people.hbs.edu/wkerr/Ellison˙Glaeser˙Kerr˙Coagglomeration˙Jan09˙Full.pdf (accessed online 3/31/12).

"The Importance of Startups in Job Creation and Job Destruction." *Kauffman Foundation,* Research Series: Firm Formation and Economic Growth (July 2010). http://www.kauffman.org/uploadedfiles/firm˙formation˙importance˙of˙startups.pdf.

"India and the Knowledge Economy: Leveraging Strengths and Opportunities." Carl Dahlman and Anuja Utz. *World Bank Institute*. New Delhi (November 9, 2006). http://siteresources.worldbank.org/KFDLP/Resources/IndiaKE01192004.pdf (accessed October 7, 2011).

"India's Economy: Losing its magic." *The Economist Online* (March 24, 2012). http://www.economist.com/node/21551061?fsrc=nlw%7Chig%7C3-22-2012%7Ceditors˙highlights (accessed March 25, 2012).

Jaison R. Abel, Ishita Dey, and Todd M. Gabe. "Productivity and the Density of Human Capital." *Federal Reserve Bank of New York Staff Reports*, no. 440 (March 2010). http://www.newyorkfed.org/research/staff˙reports/sr440.pdf (accessed March 20, 2012).

Ministry of Commerce. "Special Economic Zones in India." http://www.sezindia.nic.in/about-osi.asp Ministry of Commerce & Industry, Government of India. (accessed April 9th, 2012).

Nijman, Jan. "India's Urban Challenge." Eurasian Geography and Economics. Bellwether Publishing, Ltd., (February 22, 2012). http://bellwether.metapress.com/content/f2082021mu441r23/. (accessed March 20, 2012).

"Report on Support to SMEs in Developing Countries Through Financial Intermediaries." *Aspen Network of Development Entrepreneurs* (November 2011).

Shirish Sankhe, et al. "India's urban awakening: Building inclusive cities, sustaining economic growth." *McKinsey Global Institute* (April 2010). http://www.mckinsey.com/Insights/MGI/Research/Urbanization/Urban˙awakening˙in˙India. (accessed March 20, 2012).

Slaper, et al. "The Index of Innovation: A New Tool for Regional Analysis." *Economic Development Quarterly.* 2011.

"Support to SMEs in Developing Countries Through Financial Intermediaries." Dalberg, Global Impact Investing Network, ANDE, Cordaid, Peacock Investments, Entrepreneurial Finance Lab, and Bamboo Finance (April 2011).

Timm, Stephen. "How South Africa can Boost Support to Small Businesses: Lessons from Brazil & India." *Trade and Industrial Policy Strategies* (April 2011). http://www.tips.org.za/files/2˙small˙businesses˙in˙india˙brazil˙and˙south˙africa˙basic˙characteristics.˙pg14-19.pdf (accessed April 1, 2012).

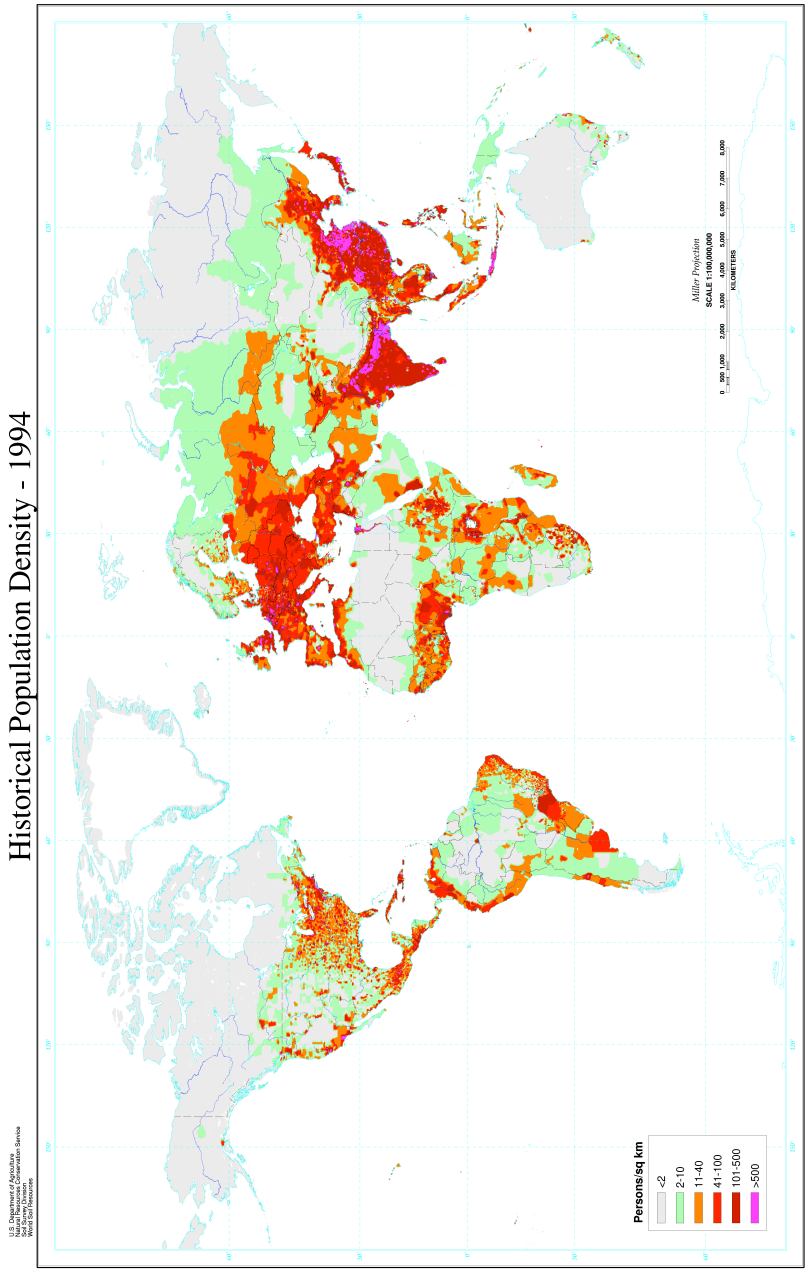
Tobler, W., V. Deichmann, J. Gottsegen, and K. Maloy. "The global demography project." Technical Report TR-95-6. *National Center for Geographic Information Analysis*, Univ. Santa Barbara, CA (1995).

Topalova, Petia. "Trade Liberalization and Firm Productivity: The Case of India." IMF Working Paper, Asia and Pacific Department (February 2004). http://www.imf.org/external/pubs/ft/wp/2004/wp0428.pdf (accessed online March 21, 2012).

"Urban Infrastructure, Housing, Basic Services and Poverty Alleviation." Government of India, Planning Commission (2008). http://planningcommission.nic.in/plans/planrel/fiveyr/11th/11˙v3/11v3˙ch11.pdf. (accessed March 20, 2012).

van Oort, Frank and Otto Raspe. "Growth of new firms and spatially bounded knowledge externalities." (December 15, 2009). *Springerlink.com* (accessed October 7, 2011).

Appendix A. Global population density map



U.S. Department of Agriculture, National Resource Conservation Service.
http://soils.usda.gov/use/worldsoils/mapindex/popden.html

Appendix B. Table of values for key variables

State	Year	Population density (people/sq km)	Labor density (workers/ sq km)	Output density (Rs crore/ sq km)	New firm density (number of firms/ sq km)	Population density 1971 (people/sq km)	Rainfall (2009)
Andhra Pradesh	1981	194.68	89.09	0.03	0.00010	158	196.3
Andhra Pradesh	1991	241.81	108.94	0.11	0.00025	158	196.3
Andhra Pradesh	2001	275.33	126.87	0.48	0.00011	158	196.3
Arunachal Pradesh	1981	7.55	3.96	0.00	0.00001	6	1174.2
Arunachal Pradesh	1991	10.32	4.78	0.01	0.00000	6	1174.2
Arunachal Pradesh	2001	13.03	5.77	0.02	0.00000	6	1174.2
Assam	1981	253.67	84.64	0.03	0.00005	186	940
Assam	1991	285.75	103.13	0.12	0.00003	186	940
Assam	2001	339.61	121.61	0.43	0.00001	186	940
New Bihar	1981	402.09	130.07	0.04	0.00001	324	742.3
New Bihar	1991	496.75	160.89	0.13	0.00002	324	742.3
New Bihar	2001	631.41	76.63	0.46	0.00000	324	742.3
Goa	1981	272.29	103.19	0.09	0.00162	215	2007.8
Goa	1991	316.05	111.56	0.28	0.00216	215	2007.8
Goa	2001	363.05	141.27	1.60	0.00054	215	2007.8
Gujarat	1981	173.89	64.80	0.03	0.00019	136	460.5
Gujarat	1991	210.73	84.79	0.12	0.00058	136	460.5
Gujarat	2001	258.11	108.44	0.47	0.00006	136	460.5
Haryana	1981	292.27	92.46	0.07	0.00016	227	122.4
Haryana	1991	372.36	115.42	0.28	0.00032	227	122.4
Haryana	2001	476.84	189.47	1.21	0.00000	227	122.4
Himachal Pradesh	1981	76.90	32.58	0.01	0.00004	62	417.5
Himachal Pradesh	1991	92.86	39.77	0.05	0.00005	62	417.5
Himachal Pradesh	2001	109.16	53.74	0.25	0.00007	62	417.5
Jammu & Kashmir	1981	26.94	11.92	0.00	0.00000	89	648
Jammu & Kashmir	1991	34.73	14.41	0.01	0.00000	89	648
Jammu & Kashmir	2001	45.31	16.89	0.06	0.00000	89	648
Karnataka	1981	193.63	77.92	0.03	0.00009	153	602.9
Karnataka	1991	234.51	98.48	0.11	0.00019	153	602.9
Karnataka	2001	274.95	122.71	0.50	0.00010	153	602.9
Kerala	1981	654.97	200.01	0.10	0.00015	549	1647.4
Kerala	1991	748.73	235.34	0.31	0.00051	549	1647.4
Kerala	2001	819.24	264.62	1.64	0.00010	549	1647.4
New Madhya Pradesh	1981	117.67	50.51	0.02	0.00000	94	879.4
New Madhya Pradesh	1991	149.24	63.91	0.06	0.00001	94	879.4
New Madhya Pradesh	2001	183.16	80.00	0.21	0.00000	94	879.4
Maharashtra	1981	204.05	86.83	0.05	0.00059	164	487.6
Maharashtra	1991	256.53	110.20	0.19	0.00075	164	487.6
Maharashtra	2001	314.42	133.80	0.71	0.00028	164	487.6

Manipur	1981	63.64	27.50	0.01	0.00000	48	283.7
Manipur	1991	82.28	34.71	0.03	0.00000	48	283.7
Manipur	2001	106.96	47.92	0.13	0.00000	48	283.7
Meghalaya	1981	59.57	27.33	0.01	0.00000	45	2354
Meghalaya	1991	79.14	33.75	0.03	0.00000	45	2354
Meghalaya	2001	102.81	43.25	0.16	0.00004	45	2354
Mizoram	1981	23.43	10.63	0.00	0.00000	16	794.2
Mizoram	1991	32.73	15.99	0.01	0.00000	16	794.2
Mizoram	2001	42.27	22.15	0.07	0.00000	16	794.2
Nagaland	1981	46.75	22.56	0.01	0.00006	31	416.1
Nagaland	1991	72.92	31.12	0.03	0.00000	31	416.1
Nagaland	2001	119.91	51.15	0.19	0.00000	31	416.1
Orissa	1981	169.36	64.37	0.02	0.00003	141	781.5
Orissa	1991	203.32	76.32	0.06	0.00005	141	781.5
Orissa	2001	235.74	91.69	0.25	0.00000	141	781.5
Punjab	1981	333.37	105.00	0.09	0.00016	269	217.4
Punjab	1991	402.70	124.34	0.33	0.00020	269	217.4
Punjab	2001	482.29	181.23	1.35	0.00004	269	217.4
Rajasthan	1981	100.11	36.65	0.01	0.00004	75	180.1
Rajasthan	1991	128.58	49.98	0.05	0.00007	75	180.1
Rajasthan	2001	165.01	69.45	0.21	0.00000	75	180.1
Sikkim	1981	44.53	21.42	0.01	0.00000	30	1427.8
Sikkim	1991	57.22	23.82	0.03	0.00000	30	1427.8
Sikkim	2001	76.10	37.06	0.12	0.00000	30	1427.8
Tamil Nadu	1981	372.20	155.30	0.06	0.00031	317	217.5
Tamil Nadu	1991	429.49	186.02	0.21	0.00073	317	217.5
Tamil Nadu	2001	477.56	214.35	1.00	0.00029	317	217.5
Tripura	1981	195.78	63.23	0.03	0.00000	148	1235.9
Tripura	1991	262.92	81.92	0.09	0.00000	148	1235.9
Tripura	2001	304.31	110.62	0.49	0.00000	148	1235.9
New Uttar Pradesh	1981	376.56	115.66	0.05	0.00003	300	633.4
New Uttar Pradesh	1991	472.51	152.16	0.17	0.00008	300	633.4
New Uttar Pradesh	2001	2937.87	194.01	0.59	0.00001	300	633.4
West Bengal	1981	614.98	185.51	0.11	0.00170	499	732.3
West Bengal	1991	767.05	246.92	0.35	0.00088	499	732.3
West Bengal	2001	903.88	332.18	1.49	0.00024	499	732.3
Andaman & Nicobar Islands	1981	22.91	8.49	0.01	0.00000	14	1428.2
Andaman & Nicobar Islands	1991	33.94	12.00	0.02	0.00000	14	1428.2
Andaman & Nicobar Islands	2001	43.16	16.49	0.11	0.00000	14	1428.2
Delhi	1981	4194.20	1349.97	1.66	0.04181	2742	192.6
Delhi	1991	6351.99	2009.44	6.91	0.08361	2742	192.6
Delhi	2001	9293.32	3064.73	37.42	0.04046	2742	192.6
Puducherry	1981	1227.64	373.98	0.34	0.00002	959	243.1
Puducherry	1991	1640.24	542.68	1.08	0.00407	959	243.1
Puducherry	2001	2031.32	716.08	7.27	0.00002	959	243.1

Appendix C. Results of labor productivity growth regressions (1981-1991, 1991-2001)

X-Var	Y-Var
1991-2001	**Log labor productivity (2001) - Log labor productivity (1991)**
Log population density (1991)	0.0456
	(0.0347)
Constant	0.575***
	(0.0305)
R-squared	0.039
Log labor density (1991)	0.0493
	(0.0376)
Constant	0.598***
	(0.0459)
R-squared	0.037
Log output density (1991)	0.0428
	(0.0332)
Constant	0.586***
	(0.0383)
R-squared	0.003
Log new firm density (1991)	0.0197
	(0.0156)
Constant	0.630***
	(0.0719)

X-Var	Y-Var
1981-1991	**Log labor productivity (1991) - Log labor productivity (1981)**
Log population density (1981)	-0.0519**
	(0.0212)
Constant	0.430***
	(0.0206)
R-squared	0.114
Log labor density (1981)	-0.0576**
	(0.0239)
Constant	0.401***
	(0.0315)
R-squared	0.078
Log output density (1981)	-0.0476**
	(0.0191)
Constant	0.396***
	(0.0323)
R-squared	0.136
Log new firm density (1981)	-0.0344*
	(0.0186)
Constant	0.313***
	(0.0871)

PLEDGE

This thesis represents my own work in accordance with University regulations.