Regional Studies
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cres20

Spatial Determinants of Entrepreneurship in India

Ejaz Ghani[a], William R. Kerr[b] & Stephen O'Connell[c]

[a] The World Bank, 1818 H Street, NW, Washington, DC 20433, USA. Email:

[b] Harvard University, Bank of Finland, and National Bureau of Economic Research (NBER), Rock Center 212, Harvard Business School, Boston, MA 02163, USA

[c] The World Bank and City University of New York (CUNY) Graduate Center, 365 Fifth Avenue, New York, NY 10016, USA. Email:

Published online: 03 Dec 2013.

To cite this article: Ejaz Ghani, William R. Kerr & Stephen O'Connell (2014) Spatial Determinants of Entrepreneurship in India, Regional Studies, 48:6, 1071-1089, DOI: 10.1080/00343404.2013.839869

To link to this article: http://dx.doi.org/10.1080/00343404.2013.839869

Regional Studies, 2014
Vol. 48, No. 6, 1071–1089, http://dx.doi.org/10.1080/00343404.2013.839869

Spatial Determinants of Entrepreneurship in India

EJAZ GHANI*, WILLIAM R. KERR† and STEPHEN O'CONNELL‡
*The World Bank, 1818 H Street, NW, Washington, DC 20433, USA. Email: Eghani@worldbank.org
†Harvard University, Bank of Finland, and National Bureau of Economic Research (NBER), Rock Center 212, Harvard Business School, Boston, MA 02163, USA. Email: wkerr@hbs.edu
‡The World Bank and City University of New York (CUNY) Graduate Center, 365 Fifth Avenue, New York, NY 10016, USA. Email: soconnell@gc.cuny.edu

(Received May 2012: in revised form August 2013)

GHANI E., KERR W. R. and O'CONNELL S. Spatial determinants of entrepreneurship in India, *Regional Studies*. The spatial determinants of entrepreneurship in India in the manufacturing and services sectors are analysed. Among general district traits, the quality of the physical infrastructure and workforce education are the strongest predictors of entry, with labour laws and household banking access also playing important roles. Extensive evidence is also found of agglomeration economies among manufacturing industries. In particular, supportive incumbent industrial structures for input and output markets are strongly linked to higher establishment entry rates. In comparison with the United States, regional conditions in India play a stronger relative role for the spatial patterns of entrepreneurship compared with incumbent industry locations.

Entrepreneurship Agglomeration Development India South Asia

GHANI E., KERR W. R. and O'CONNELL S. 印度创业精神的空间决定因素，*区域研究*。本研究分析印度在製造业与服务业部门中，企业创业精神的空间决定因素。在一般的行政区特徵中，实质基础建设的质量与劳动力教育水平最能有效预测企业进入，而劳动法规与家户取得银行业务的渠道亦扮演了重要的角色。製造业中亦发现了聚集经济的大量证据。特别是投入与产出市场的当前支持性产业结构，与较高的创立进入率显着相关。与美国相较而言，印度的区域条件与目前的产业地点相较之下，在创业的空间模式上扮演了相对重要的角色。

创业精神 聚集 发展 印度 南亚

GHANI E., KERR W. R. et O'CONNELL S. Les déterminants spatiaux de l'esprit d'entreprise en Inde, *Regional Studies*. On analyse les déterminants spatiaux de l'esprit d'entreprise en Inde dans les secteurs de la fabrication et des services. Parmi les caractéristiques générales des districts, la qualité de l'infrastructure physique et de la formation professionnelle sont les meilleurs indicateurs de l'entrée, alors que le droit du travail et l'accès des ménages aux services bancaires jouent également un rôle important. Il s'avère aussi de nombreuses preuves des économies d'agglomération dans le secteur de la fabrication. En particulier, les structures d'appui industrielles établies pour les marchés amont et aval sont fortement liées à des taux d'entrée plus élevés des entreprises. Par rapport aux États-Unis, les conditions régionales en Inde jouent un rôle relatif plus fort pour ce qui est des structures spatiales de l'esprit d'entreprise par comparaison avec les emplacements industriels établis.

Esprit d'entreprise Agglomération Développement Inde Asie du Sud

GHANI E., KERR W. R. und O'CONNELL S. Räumliche Determinanten des Unternehmertums in Indien, *Regional Studies*. Wiranalysieren die räumlichen Determinanten des Unternehmertums im Produktions- und Dienstleistungssektor von Indien. Unter den generellen Merkmalen der Bezirke sind die Qualität der physischen Infrastruktur sowie der Bildungsgrad der Arbeitnehmer die stärksten Prädiktoren für Firmengründungen; die Arbeitsgesetze und die Verfügbarkeit von Haushaltsbanken spielen ebenfalls eine wichtige Rolle. Darüber hinaus finden wir unter den produzierenden Branchen umfangreiche Belege für Agglomerationsökonomien. Insbesondere besteht eine starke Verbindung zwischen den vorhandenen unterstützenden Branchenstrukturen für Input-und Outputmärkte und einem höheren Anteil an Unternehmensgründungen. Bei einem Vergleich mit den vorhandenen Branchenstandorten spielen die regionalen Bedingungen in Indien verglichen mit den USA eine größere relative Rolle für die räumlichen Muster des Unternehmertums.

Unternehmertum Agglomeration Entwicklung Indien Südasien

GHANI E., KERR W. R. y O'CONNELL S. Determinantes espaciales del empresariado en India, *Regional Studies*. En este artículo analizamos los determinantes espaciales del empresariado en los sectores de producción y servicios de India. Entre las características generales de las comarcas, la calidad de la infraestructura física y la educación de la mano de obra son los determinantes más importantes para predecir la creación de empresas, siendo las leyes laborales y la disponibilidad de servicios bancarios a hogares

también factores muy importantes. Observamos asimismo pruebas extensas de economías de aglomeración entre las industrias de producción. En particular, las estructuras industriales establecidas y de apoyo para los mercados de insumos y productos están muy vinculadas a índices más altos en la creación de nuevas empresas. En comparación con los Estados Unidos, las condiciones regionales en India desempeñan un papel relativo más importante en los patrones espaciales del empresariado en comparación con las ubicaciones industriales establecidas.

Empresariado Aglomeración Desarrollo India Asia meridional

JEL classifications: L26, M13, R10, R12

INTRODUCTION

Many policy-makers want to encourage entrepreneurship given its perceived role in economic growth and development.[1] The importance of this factor has led to extensive recent research on regional traits associated with entrepreneurship. Multiple studies consider advanced economies, but there is very little empirical evidence for developing countries. This lack of research hampers the effectiveness of policy: for example, the roles that education or infrastructure play in entry in the United States may be quite different from a setting where illiteracy and lack of roads and sanitation continue to hamper development.

AUDRETSCH *et al.* (2012) emphasize the local nature of entrepreneurship determinants. These questions are investigated for manufacturing and services in India in the present paper. Within these two industry groups, the organized and unorganized sectors are also compared. The traits of districts that systematically predict stronger entry levels are quantified. Several important themes emerge from the study. First, education levels and local infrastructure access are the most prominent local traits linked to entrepreneurship across all sectors. Second, local industrial conditions – the links that form across industries within a district – play an even stronger role in predicting entry within specific district–industries than the general district-level traits. Finally, in comparison with the United States, it is found that India's economic geography is still taking shape. At such an early point and with industrial structures not entrenched, there is room for policy to have substantial impact by shaping where industries plant their roots.

The study makes several contributions to the literature. It is among the first to quantify the spatial determinants of entrepreneurship in India. Moreover, it moves beyond manufacturing to consider services, and compares the organized and unorganized sectors. The latter analyses of the unorganized sector are among the most important contributions given the limited study of the informal economy previously and its substantial importance for India and other developing economies. More broadly, it is among the first studies to apply the incumbent industrial structures frameworks of GLAESER and KERR (2009) to a developing economy, providing insights into how agglomeration economies resemble and differ from each other. More research on agglomeration economies and entrepreneurship in developing countries is important for urban and development economics going forward.[2]

Identifying local conditions that encourage entrepreneurship and acting upon them is essential to foster economic growth. Fig. 1 shows that entrepreneurship rates are lower in South Asia than what its stage of development would suggest. Effective entrepreneurship will play a key role in job growth for India, the development of a strong manufacturing base (FERNANDES and PAKES, 2010), and the transition of people out of subsistence living and the informal sector. KHANNA (2008) emphasizes entrepreneurship for India's future, and reallocation can help close India's productivity gap (e.g., HSIEH and KLENOW, 2009).[3]

SPATIAL ENTREPRENEURSHIP RATES IN INDIA

Entrepreneurship is measured as the presence of young establishments. The primary measure, which can consistently be observed across all the datasets, is whether an establishment is less than three years old. For the organized manufacturing sector, establishments in their first year of existence can also be measured, and very similar results are found with this approach. Incumbent establishments, which are used to model existing activity in the district–industry, are firms that are three or more years old. Entry measures are principally defined through employment in young establishments, and counts of entering establishments are looked at in robustness checks.[4]

Establishment-level surveys of manufacturing and service enterprises carried out by the Government of India are employed. The manufacturing data are taken from surveys conducted in fiscal years 2005–06; services sector data come from 2001–02. While these surveys were conducted over two fiscal years, this paper refers to the initial year only. An unpublished appendix (which is available from the authors upon request) lists data sources and years employed, and additional information included in it is described below. NATARAJ (2009), KATHURIA *et al.* (2010), HASAN and JANDOC (2010), and DEHEJIA and PANAGARIYA (2010) provide detailed overviews of similar databases.

The distinction between organized and unorganized sectors relates to establishment size. In manufacturing, the organized sector is comprised of establishments with more than ten workers if the establishment uses

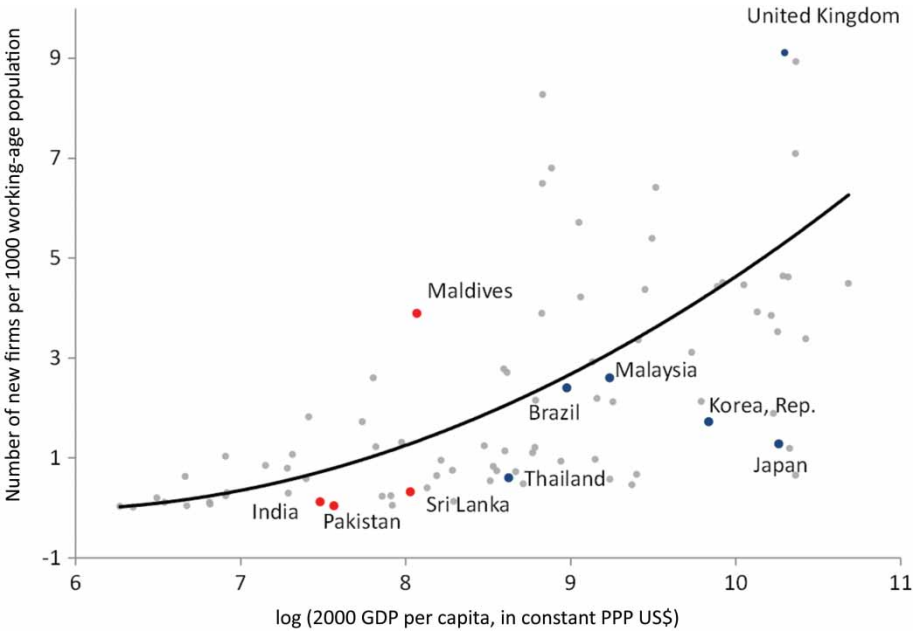

Fig. 1. *Business registration density, 2008*
Note: Countries designated as offshore tax shelters are excluded. Eighty-seven countries are shown
Sources: World Bank Group Entrepreneurship Survey, 2010; and World Development Indicators, 2010

electricity. If the establishment does not use electricity, the threshold is 20 workers or more. These establishments are required to register under the Factories Act of 1948. The unorganized manufacturing sector is, by default, comprised of establishments that fall outside the scope of the Factories Act.

Service establishments, regardless of size or other characteristics, are not required to register and thus are all officially unorganized. There are various approaches to differentiate comparably small-scale, autonomous establishments from the larger employers that constitute the organized sector, as generally defined. Services establishments with fewer than five workers and/or listed as an 'own-account enterprise' (OAE) are assigned to the unorganized sector. OAE enterprises are firms that do not employ any hired worker on a regular basis. The choice of five employees as the size cut-off recognizes that average establishment size in services is significantly smaller than in manufacturing. Using this demarcation, the organized sector makes up approximately 25% of employment in both manufacturing and services.

The organized manufacturing sector is surveyed by the Central Statistical Organisation (CSO) every year through the Annual Survey of Industries (ASI), while unorganized manufacturing and services establishments are separately surveyed by the National Sample Survey Organisation (NSSO) at approximately five-year intervals. Establishments are surveyed with state and four-digit National Industry Classification (NIC) stratification. For organized manufacturing, the business register described above forms the basis for the sampling frame. Establishments are notified if they fall into the sampled frame and are required by law to complete and return the survey questionnaire; the CSO investigates cases of non-response (typically closed plants). For the services and unorganized manufacturing sector, India's Economic Census comprises the basis for the sampling frame and stratification procedures. Establishments falling into the sample are then surveyed by government enumerators.[5]

The survey years used are the most recent data by sector for which the young establishment identifiers are recorded. The provided sample weights are used to construct population-level estimates of total establishments and employment by district and three-digit NIC industry. Employment is formally defined as 'persons engaged' and includes working owners, family and casual labour, and salaried employees.

Districts are administrative subdivisions of Indian states or territories. Currently there are approximately 630 districts spread across 35 states/union territories. Districts with a population less than 1 million (based on the 2001 Census) or with fewer than 50 establishments sampled are excluded. These small districts are excluded because limited sampling makes the data of limited value for the study (given that district–industry conditions that separate young and incumbent establishments need to be evaluated). States that experienced ongoing conflict and political turmoil during the period of study are also excluded. After these adjustments, the resulting sample retains districts in 20 major states that include more than 94% of Indian employment in both manufacturing and services.

Table 1 provides descriptive statistics; Figs 2 and 3 show spatial entry patterns; and the unpublished

Table 1. *Local industrial conditions for Indian entrepreneurship*

	Mean	SD	Median	Minimum	Maximum	Mean	SD	Median	Minimum	Maximum
District traits										
District population (*n*)	2 972 828	1 731 997	3 207 232	1 021 573	13 900 000					
District population density (persons/km^2)	810	2477	480	35	24 963					
Share of the population with a graduate education (%)	5.9	2.7	6.2	1.7	19.3					
Demographic dividend for a district (age profile)	1.32	0.26	1.41	0.92	2.12					
Index of infrastructure quality for a district	2.93	0.76	3.34	0.00	4.00					
Strength of household banking environment	0.35	0.13	0.38	0.09	0.73					
Stringency of labour laws: adjustments (state level)	0.69	0.84	0.00	0.00	3.00					
Stringency of labour laws: disputes (state level)	–0.41	1.24	0.00	–3.00	3.00					
Proximity to India's ten largest cities (minimum driving)	446	240	396	0	1020					
Consumption per capita (year 2005 US$ at purchasing power parity (PPP))	680	186	625	352	1397					
	Organized sector					Unorganized sector				
Industrial traits – manufacturing										
Total employment in district–industry	1383	5020	337	2	215 611	4517	15 389	831	1	422 193
Start-up employment in district–industry	151	788	0	0	28 576	553	2938	0	0	96 647
Labour market strength	0.09	0.13	0.05	0.00	0.97	0.09	0.11	0.04	0.00	0.97
Input/supplier strength	–1.64	0.25	–1.69	–2.00	–0.05	–1.71	0.24	–1.76	–2.00	–0.05
Chinitz index of small suppliers	0.48	1.33	0.25	0.00	45.52	n.a.	n.a.	n.a.	n.a.	n.a.
Output/customer strength (×10 for presentation)	0.01	0.02	0.03	0.00	7.64	0.01	0.02	0.00	0.00	0.97
Industrial traits – services										
Total employment in district–industry	1761	5892	400	3	173 293	2885	8145	376	1	195 863
Start-up employment in district–industry	268	1429	0	0	47 048	502	1581	46	0	50 243

Notes: Descriptive statistics are based on the Annual Survey of Industries and the National Sample Survey, various rounds.
n.a., Not applicable; SD, standard deviation.

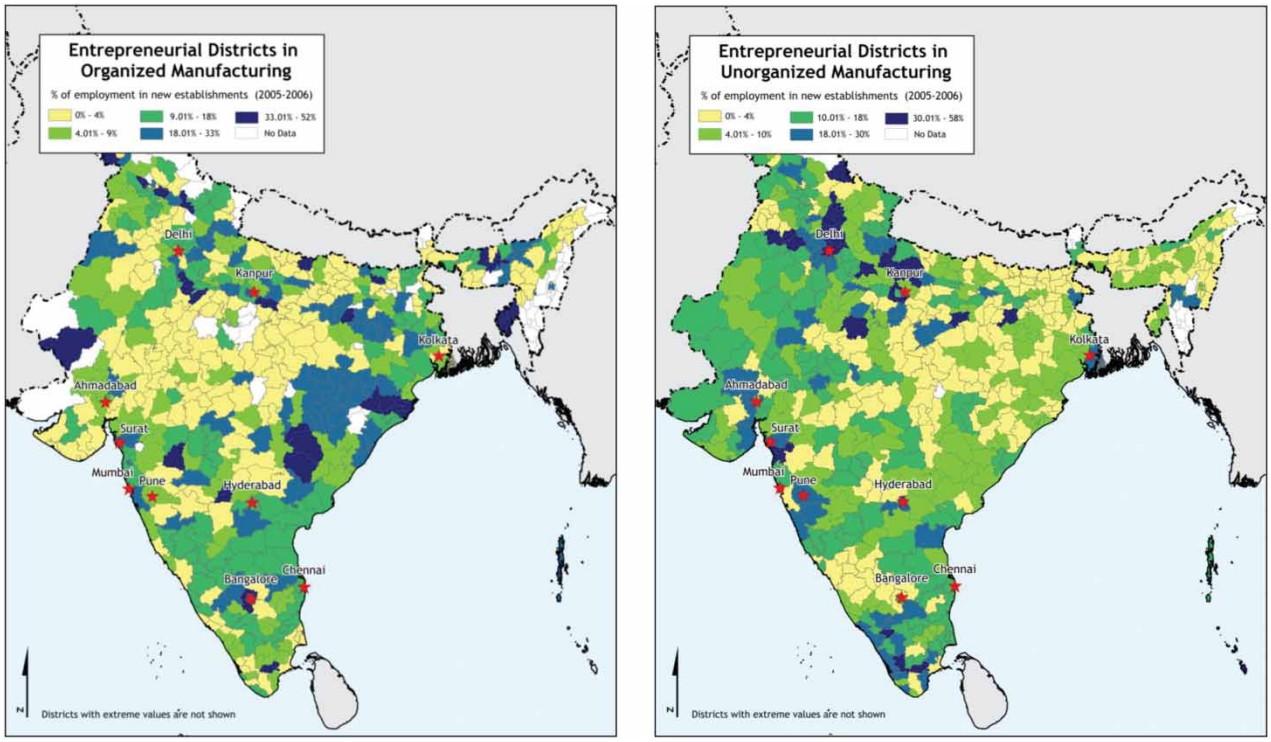

Fig. 2. Indian manufacturing entry rates, 2005–06

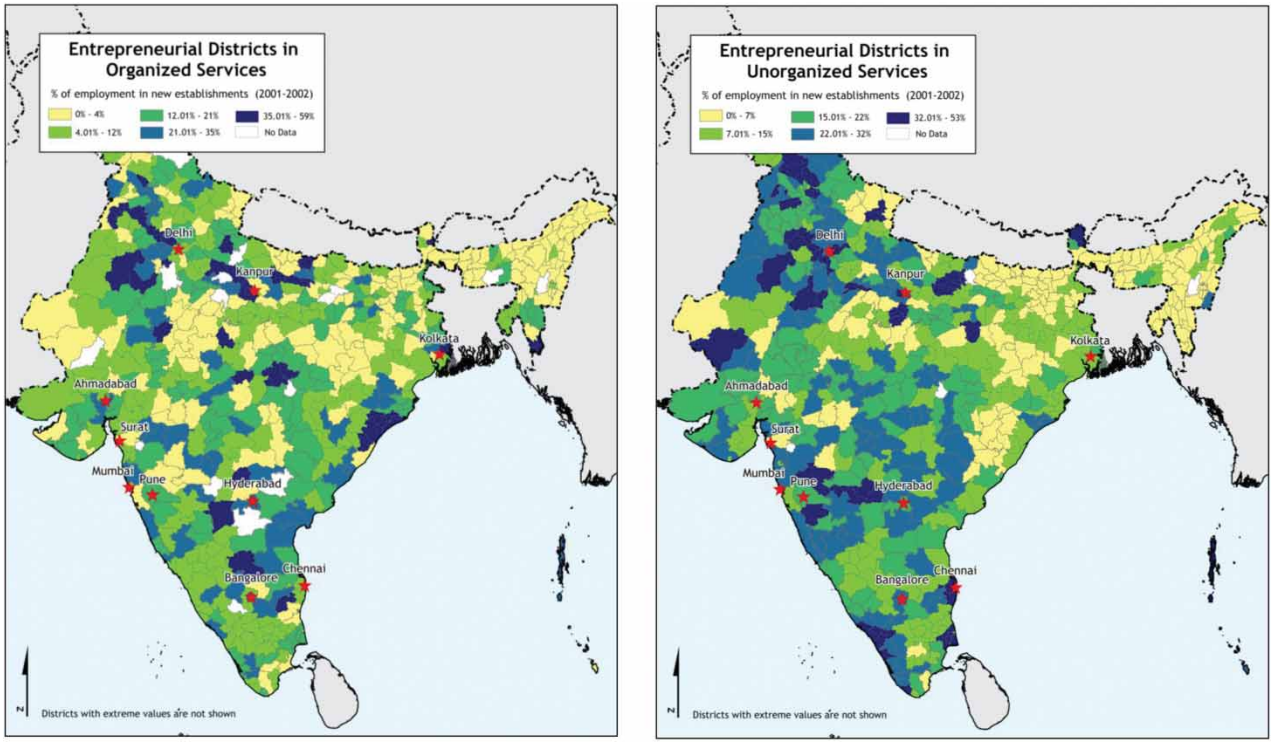

Fig. 3. Indian services entry rates, 2001–02

appendix offers additional tabulations by state. Entry rates, as a weighted average across all states, are 15% and 12% for organized and unorganized manufacturing, respectively. The entry rates are 20% for organized services and 17% for unorganized services. The spatial entry rates for organized and unorganized sectors have –0.2 and 0.3 correlations across states for manufacturing and services, respectively. VAN STEL *et al.* (2007)

emphasize the need to measure entry determinants separately across different types of entrepreneurs.

DETERMINANTS OF ENTREPRENEURSHIP

This section now describes the spatial and industrial factors that are used to predict entrepreneurship. General traits of the district that affect all entrepreneurs, regardless of industry, are first considered. These traits include both baseline features that are longstanding and slow to adjust, like the population distribution, and factors that are more directly influenced by policy-makers, such as education and infrastructure, recognizing that deep change in education and infrastructure also takes a long time to accomplish. Second, recent research stresses the central importance of heterogeneity across industries as well as regions for explaining start-up rates (e.g., FRITSCH and FALCK, 2007; GLAESER and KERR, 2009). The second category thus develops industry-specific conditions that yield this heterogeneity within regional experiences.

District-level conditions

The initial explanatory measures focus on the basic traits of districts. It is essential to understand the effect of local area traits on entrepreneurship, especially given the disproportionate degree to which entrepreneurs found businesses in their home areas (e.g., FIGUEIREDO *et al.*, 2002; MICHELACCI and SILVA, 2007). Population is first controlled for to provide a natural baseline of economic activity (e.g., consumer markets, general availability of workers). The district's age structure, measured as the ratio of the working-age population to the non-working-age population, is then considered given that the propensity to start new firms changes over the lifetimes of individuals, and the age structure of a region often connects to local entry rates (e.g., BÖNTE *et al.*, 2009; DELFMANN *et al.*, 2013). The age profile is often called the demographic dividend in the Indian context.

Third, a measure of population density is included. Unlike the clear positive predictions for the first two factors, the prediction for population density is ambiguous as it brings higher wages and land rents alongside greater market opportunities. Density has also been linked to stronger knowledge flows, and AUDRETSCH and FRITSCH (1994) use density as one source of convexity in local production that links to entry rates. Many studies link higher population density to reduced manufacturing entry rates, especially for larger plants that use established production techniques and seek to minimize costs. DURANTON and PUGA (2001) provide a formal theoretical model of this process. Ultimately, these multiple forces suggest an uncertain theoretical role for population density in explaining Indian entry rates.

Beyond these basic demographics, five primary traits of districts are considered: education of the local labour force, quality of the local physical infrastructure, access or travel time to major Indian cities, stringency of labour laws, and household banking conditions. These traits are motivated by theoretical models of entrepreneurship and their perceived importance to India's development, and other traits and their relationships to these variables are discussed below. Unless otherwise noted, these traits are taken from the 2001 Population Census.

Several studies link entrepreneurship to educated workforces in the United States (e.g., DOMS *et al.*, 2010; GLAESER *et al.*, 2010), often with the underlying conceptual model that entrepreneurship requires a degree of creativity and handling of many tasks and ambiguous circumstances that education prepares one for (e.g., the model of LAZEAR, 2005). Entrepreneurs may also benefit from the specific development of basic business skills. On the other hand, REYNOLDS *et al.* (1994) do not find this relationship holds within every country, and GLAESER and KERR (2009) find limited evidence for a link between education and US manufacturing entrepreneurship. Thus, the literature is again ambiguous. Clarifying education's role for India is very important, as many local policy-makers stress developing the human capital of their workforces, and India is no different (AMIN and MATTOO, 2008). The general education level of a district is measured by the percentage of adults with a graduate (post-secondary) degree. The results are robust to alternative definitions such as the percentage of adults with higher secondary education.

The second trait is the physical infrastructure level of the district. Basic services such as electricity are essential for all businesses, but new entrants can be particularly dependent upon local infrastructure (e.g., established firms are better able to provision their own electricity if need be, which is quite common in India). AGHION *et al.* (2012) provide a recent theoretical model. Entrepreneurship is likely to benefit from greater infrastructure so long as the tax burden imposed to provide the infrastructure is not too high. Many observers cite upgrading India's infrastructure as a critical step towards economic growth, and the Indian government has set aside substantial investment funds. The population census documents the number of villages in a district with telecommunications access, electricity access, paved roads and safe drinking water. The percentage of villages that have infrastructure access within a district is calculated and a sum is made across the four measures to create a continuous composite metric that ranges from zero (no infrastructure access) to four (full access).

India's economy is undergoing dramatic structural changes (DESMET *et al.*, 2011). From a starting point in the 1980s when the government used licensing to promote industrial location in regions that were developing slowly, the economic geography of India has been in flux as firms and new entrants shift spatially (e.g., FERNANDES and SHARMA, 2011). One feature that is important for a district in this transformation is its link to major cities. A measure from LALL *et al.*

(2011) of the driving time to the nearest of India's ten largest cities is thus included as a measure of physical connectivity and across-district infrastructure.

Local labour regulations are next modelled using state-level policy variation. Several studies link labour regulations in India to slower economic progress (e.g., Besley and Burgess, 2004; Aghion *et al.*, 2008), and Bozkaya and Kerr (2013) provide a theory model where tighter labour laws suppress entry. This effect may occur through reduced likelihood of both wanting to start a new firm or opening new facilities from a desire to avoid regulations. There may also be reduced 'push' into entrepreneurship with more protected employment positions. A composite labour regulations index by state is created from the measures constructed by Ahsan and Pages (2007).

The final measure is the strength of the household banking environment, reflecting the large literature on financial constraints and entrepreneurship, with Evans and Jovanovic (1989) being a seminal model. The percentage of households that have banking services by district is measured. This measure is likely to be particularly reflective of financing environments for unorganized sector activity.

Local industrial traits

Recent research emphasizes how local entrepreneurship varies substantially across industries, and the second set of metrics quantifies how suitable the local industrial environment is for a particular industry. The first trait is the overall employment in a district–industry for incumbent firms. This is important given that entrepreneurs often leave incumbents to start their own companies (e.g., Klepper, 2010; Falck *et al.*, 2008). From this baseline, metrics are further developed that unite the broad distribution of industry employments in districts with the extent to which industries interact through the traditional agglomeration rationales (e.g., Marshall, 1920; Duranton and Puga, 2004; Rosenthal and Strange, 2004). These forces are considered within the manufacturing sector, and these conditions are modelled through incumbent firms that predate the birth of the young businesses that are modelled in the outcome variables.[6]

The first agglomeration rationale is that proximity to customers and suppliers reduces transportation costs and thereby increases productivity. This reduction in shipping costs is the core agglomerative force of New Economic Geography theory (e.g., Fujita *et al.*, 1999). Where customers and suppliers are geographically separate, firms trade-off distances. The extent to which districts contain potential customers and suppliers for a new entrepreneur is measured. This section begins with an input–output table for India developed by the CSO. $Input_{i\leftarrow k}$ is defined as the share of industry i's inputs that come from industry k; and $Output_{i\rightarrow k}$ is defined as the share of industry i's outputs that go to industry k. These measures run from zero (no input or output purchasing relationship exists) to one (full dependency on the paired industry).

The quality of a district d is summarized in terms of its input flows for an industry i as:

$$Input_{di} = -\sum_{k=1,\ldots,I} \text{abs}\left(Input_{i\leftarrow k} - E_{dk}/E_d\right)$$

where I indexes industries. This measure aggregates absolute deviations between the proportions of industrial inputs required by industry i and district d's actual industrial composition, with E representing employment. The measure is mostly orthogonal to district size, which is considered separately, and a negative value is taken so that the metric ranges between negative two (i.e., no inputs available) and zero (i.e., all inputs are available in the local market in precise proportions). This metric assumes that firms have limited ability to substitute across material inputs in their production processes.

To capture the relative strength of output relationships, a consolidated metric is also defined:

$$Output_{di} = \sum_{k=1,\ldots,I} E_{dk}/E_d \cdot Output_{i\rightarrow k}$$

This metric multiplies the national share of industry i's output sales that go to industry k with the fraction of industry k's employment in district d. By summing across industries, a weighted average of the strength of local industrial sales opportunities for industry i in the focal market d is taken. This $Output_{di}$ measure takes on higher values with greater sales opportunities. It allows greater substitution across customer industries than the design built into the input metric, and its robustness to several design variants was tested.

Moving from material inputs, entrepreneurship is quite likely to be driven by the availability of a suitable labour force (e.g., the model of Combes and Duranton, 2006). While education and demographics are informative about the suitability of the local labour force, these aggregate traits miss the very specialized nature of many occupations. The working paper summarizes theories as to why specialized workers and firms agglomerate together and provides extended references. Unlike studies of advanced economies, India lacks the data to model direct occupational flows between industries. Greenstone *et al.* (2010) calculate from the Current Population Survey the rate at which workers move between industries in the United States. Using their measure of labour similarity for two industries, the present paper defines:

$$Labor_{di} = \sum_{k=1,\ldots,I} E_{dk}/E_d \cdot Mobility_{i\leftarrow k}$$

This metric is a weighted average of the labour similarity of industries to the focal industry i, with the weights

being each industry's share of employment in the local district. The metric is again by construction mostly orthogonal to district size.

These metrics condense large and diverse industrial structures for cities into manageable statistics of local industrial conditions. The advantages and limitations in their design are further discussed in the working paper. Perhaps the most important issue is that these district conditions do not capture interactions with neighbouring districts, but factor and product markets can be wider than a local area. The average size of an Indian district is about the same as two US counties at 5500 km^2.

The section finally turns to a special issue regarding local firm size distribution, building upon a literature that traces back to at least the work of JOHNSON and CATHCART (1979). FRITSCH and FALCK (2007) and PARKER (2009b) emphasize the strong degree to which an industrial base populated with small firms is associated with higher entrepreneurship rates. FRITSCH and FALCK (2007) note that the relationship could descend from a greater entrepreneurial culture (HOFSTEDE, 2001; BOSCHMA and FRITSCH, 2007; FALCK *et al.*, 2011),[7] better training for entrepreneurs due to them having worked in small businesses, or perhaps a reflection of the local industry's minimum efficient plant size. PARKER (2009b) emphasizes a self-selection role by entrepreneurs. For the organized manufacturing sector, the inclusion of a measure of the local small firm share (fewer than 40 employees) is tested in estimations. While there are many reasons to believe that this pattern in advanced countries will carry over to India, there are also reasons to be doubtful. For example, Indian labour laws and size regulations have long suppressed average firm size in India compared with its peers, perhaps weakening this robust relationship evident elsewhere.

A specific variant of this effect related to customer/supplier industries is also measured. CHINITZ (1961) observes that entrepreneurs often find it difficult to work with large, vertically integrated suppliers. The entrepreneur's order sizes are too small, and often the entrepreneur's needs are non-standard. Empirical studies for the United States find the Chinitz effect very important in local start-up conditions. The Chinitz effect – as distinct from the general conditions captured in $Input_{di}$ – is quantified through a metric that essentially calculates the average firm size in a district in industries that typically supply a given industry i:

$$Chinitz_{di} = \sum_{k=1,\ldots,I} Firms_{dk}/E_d \cdot Input_{i \leftarrow k}$$

Higher values of the $Chinitz_{di}$ metric indicate better supplier conditions for entrepreneurs.

ESTIMATION APPROACH

Factors related to entry are characterized through cross-sectional regressions at the district–industry level of India. This level of variation allows analysis of both district-level determinants and the underlying heterogeneity for entrants across industries due to incumbent industrial structures. Following the above literature and conceptual notes, these specifications take the form:

$$\ln\left(Entry_{di}\right) = \eta_i + \beta \cdot X_d + \gamma \cdot Z_{di} + \varepsilon_{di}$$

The dependent variable is the log measure of entry employment by district–industry. The sample includes the district–industry observations in which positive incumbent employment exists. The observation count thus differs across manufacturing and services and for organized and unorganized sectors. Many of the explanatory variables, such as incumbent district–industry employment, are also in log values so that the coefficients estimate proportionate responses. Non-log variables are transformed to have unit standard deviation for interpretation, estimations are weighted by an interaction of log industry size with log district population, and standard errors are clustered by district to reflect the multiple mappings of district-level variables across local industries.[8]

A vector of industry fixed effects η_i is included in the estimations. These fixed effects control for systematic differences across industries in their entrepreneurship rates, competition levels, average plant sizes, and similar. As FRITSCH and FALCK (2007) demonstrate, isolating spatial variations from these industry-level traits is very important. Also, the metrics of local industrial conditions utilize both fixed traits of industries (e.g., the input–output relationships, labour flows) and the distribution of industries within a district. The inclusion of industry fixed effects controls for these fixed industry-level traits except to the extent that they interact with the local industrial structure.

The vectors X_d and Z_{di} contain district and district–industry traits, respectively. The estimation approach balances several objectives. First, given that there has been so little work on India, the aim is to provide a sufficiently broad analysis to highlight where major correlations lie in the data. In doing so, one does not want to be too parsimonious in the specifications, but the analysis should not be overloaded. The set of metrics provides a good depiction of the Indian entrepreneurial landscape, motivated by theory, and the robustness section and the unpublished appendix discuss many additional factors considered when forming this baseline.

It must be emphasized that this work measures partial correlations in the data, rather than causal parameters, reflective of the initial enquiry. In all cases, local traits are predetermined for the entrepreneurship that was measured as the outcome variable. This provides some confidence against reverse causality, and including lagged entry rates as a control variable is further tested. A second concern is omitted factors that are highly correlated with the regressors, making interpretation

difficult. For example, in the baseline model, education may capture the quality of the local workforce that entrepreneurs employ, the strength of the local pool of potential entrepreneurs and/or stronger local consumer demand. Some specific checks along these lines (e.g., controlling for consumption per capita) are provided, but there will be a natural limit against checking every feasible concern. These issues are further discussed below.

EMPIRICAL RESULTS

Table 2a considers organized manufacturing. Column (1) includes just district populations, district–industry employments and industry fixed effects. The existing district–industry employment strongly shapes the spatial location of entry: a 10% increase in incumbent employment raises entry employment by around 2%. In addition, a district's population increases entry rates with an elasticity of 0.5. Higher-order population terms are not found to be statistically significant or economically important. The adjusted R^2 value for this estimation is quite modest at 0.13.

It is useful to compare these results with those evident in the United States for two reasons. First, the United States' advanced economy – and policy environment which has relatively fewer distortions – provides a useful idea of what entrepreneurship and local conditions might look like at the frontier. This is not to say that India will necessarily look like the United States when it reaches current levels of US development, just as entrepreneurship rates differ across advanced economies today. Nevertheless, in terms of broad regularities, it is very helpful to compare the India statistics against a country like the United States to provide perspective. A very well-known example in this regard is the Hsieh and Klenow (2009) comparison of the misallocation of production across plants in India and the United States. Second, and from an academic perspective, there is a growing body of evidence and intuition on how the US economy functions with respect to entrepreneurship. The extent to which the study can identify where the Indian experience resembles or differs from the US experience provides a reasonable starting point for ascertaining which lessons from the US studies can be applied to the Indian context. The conclusions section below describes some of these lessons that do or do not apply, and hopefully this paper provides a touchstone for identifying whether lessons from future studies made of the United States or other advanced economies should be taken into account when thinking about the Indian context.

Glaeser and Kerr (2009) estimate a related specification for the United States that uses long-term employment for a city–industry as the key explanatory variable. If the estimation is adjusted to match their technique more closely, an elasticity of 0.8 is obtained that is very similar to their 0.7 elasticity. While this elasticity is comparable, the R^2 value for this estimation remains quite modest at 0.29, much lower than the R^2 value of 0.80 for Glaeser and Kerr (2009). There are likely several factors behind this lower explanatory power for India, including data differences, estimations at the district versus city level, and similar. These natural differences between the Indian and US data limit perfect comparison, but the datasets are believed to be sufficiently similar to make some basic inference. Most important, it is clear that many industries within India's manufacturing sector are at a much earlier development stage than those in the United States, where the manufacturing sector is instead shrinking. Thus, while existing patterns of industrial activity explain the similarity of spatial distribution of entrepreneurship in India and the United States, India has much more variation in outcomes, which are characterized further below. Fernandes and Sharma (2011) also study these variations with respect to policy deregulations. Kathuria (2011) provides a broader exploratory framework.

Column (2) includes the district-level traits. Three factors stand out as discouraging entrepreneurship in organized manufacturing: high population density, strict labour regulations and the greater distance to one of India's ten biggest cities. The first pattern has been observed in many settings and reflects large manufacturers seeking cheaper environments. The second pattern connects with earlier studies of India that argue that strict labour laws reduce economic growth. These policies are associated with reduced entry even after conditioning on district–industry size. The final factor highlights that while manufacturers avoid the high costs of urban areas, they also avoid the most remote areas of India in favour of settings that are relatively near to large population centres, are likely to access customers directly or connect to shipping routes. On the other hand, the education of a district's workforce is linked to higher entry rates. The elasticity that is estimated here is stronger than that found in comparable US estimations.

Column (3) introduces district–industry traits. The roles of input and output markets are exceptionally strong with elasticities of 0.4–0.5. Both the labour market and Chinitz measures have positive coefficients. The decline in the main effect of incumbent employment suggests that these four new metrics capture the positive effects of local clusters on entry.

Column (4) shows quite similar results if one further controls for consumption per capita, per the discussion in the above section. This control, along with the population metrics, suggests that demand-side factors are not solely responsible for the positive roles that are seen for metrics such as education.

Column (5) finds similar results when examining the log count of entering establishments, with the Chinitz metric being more prominent. The paper will return

Table 2a. *District entrepreneurship estimations – organized manufacturing*

	Base estimation (1)	District traits (2)	Full estimation (3)	Adding consumption (4)	Using log entry count (5)
Log of incumbent employment in district–industry	0.229+++	0.186+++	–0.028	–0.030	0.032+
	(0.043)	(0.040)	(0.048)	(0.047)	(0.018)
Log of district population	0.531+++	0.483+++	0.475+++	0.482+++	0.216+++
	(0.179)	(0.155)	(0.156)	(0.161)	(0.056)
District traits					
Log of district population density		–0.569+++	–0.563+++	–0.562+++	–0.197+++
		(0.088)	(0.080)	(0.079)	(0.029)
Share of the population with graduate education		0.211+	0.235++	0.230++	0.078+
		(0.110)	(0.107)	(0.111)	(0.042)
Demographic dividend for a district (age profiles)		0.605	0.567	0.535	0.271
		(0.458)	(0.446)	(0.468)	(0.177)
Index of infrastructure quality for a district		0.018	0.096	0.086	0.015
		(0.100)	(0.094)	(0.097)	(0.038)
Strength of household banking environment		0.143	0.095	0.085	0.027
		(0.104)	(0.100)	(0.106)	(0.036)
Stringency of labour laws in a district's state		–0.210+++	–0.161++	–0.157++	–0.095+++
		(0.070)	(0.064)	(0.065)	(0.023)
Log travel time to the closest large city		–0.275+++	–0.241+++	–0.237+++	–0.091+++
		(0.090)	(0.083)	(0.083)	(0.031)
Log per capita consumption				0.152	
				(0.505)	
Local industrial conditions by incumbent firms					
Labour market strength for district–industry			0.161	0.164	0.026
			(0.102)	(0.102)	(0.041)
Inputs/supplier strength for district–industry			0.485+++	0.485+++	0.154+++
			(0.098)	(0.098)	(0.043)
Outputs/customer strength for district–industry			0.388+++	0.387+++	0.167+++
			(0.140)	(0.140)	(0.057)
Chinitz small suppliers metric for district–industry			0.279	0.279	0.337+++
			(0.213)	(0.212)	(0.129)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	4843	4843	4843	4843	4843
Adjusted R^2	0.128	0.166	0.218	0.218	0.279

Notes: The dependent variable is log entry employment by district–industry,

Estimations quantify the relationship between district–industry employment in new establishments and local conditions. District-level traits are taken from the 2001 Census. Industrial conditions are calculated from 2005–06 using incumbent establishments in the district–industry. Labour regulations are a composite of adjustment and disputes laws. Estimations weight observations by an interaction of district size and industry size, include industry fixed effects and cluster standard errors by district. Non-logarithm variables are transformed to have unit standard deviation for interpretation.

+++, ++ and + indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table 2b. *District entrepreneurship estimations – unorganized manufacturing*

	Base estimation (1)	District traits (2)	Full estimation (3)	Adding consumption (4)	Using log entry count (5)
Log of incumbent employment in district–industry	0.163+++	0.123+++	–0.075++	–0.078+++	–0.040
	(0.031)	(0.029)	(0.029)	(0.029)	(0.026)
Log of district population	1.051+++	0.878+++	1.010+++	1.025+++	0.866+++
	(0.161)	(0.157)	(0.160)	(0.153)	(0.138)
District traits					
Log of district population density		–0.019	–0.044	–0.042	–0.044
		(0.070)	(0.068)	(0.073)	(0.057)
Share of the population with graduate education		–0.002	–0.026	–0.079	–0.046
		(0.080)	(0.084)	(0.087)	(0.074)
Demographic dividend for a district (age profiles)		0.954+++	1.053+++	0.770++	0.798+++
		(0.326)	(0.330)	(0.326)	(0.285)
Index of infrastructure quality for a district		0.386+++	0.365+++	0.259++	0.325+++
		(0.096)	(0.097)	(0.104)	(0.086)
Strength of household banking environment		0.222+++	0.211+++	0.152+	0.193+++
		(0.080)	(0.080)	(0.082)	(0.071)
Stringency of labour laws in a district's state		–0.007	0.000	0.020	0.030
		(0.069)	(0.069)	(0.066)	(0.062)
Log travel time to closest large city		–0.004	0.009	0.029	0.017
		(0.069)	(0.074)	(0.074)	(0.065)
Log per capita consumption				1.191+++	
				(0.365)	
Local industrial conditions by incumbent firms					
Labour market strength for district–industry			0.263+++	0.271+++	0.228+++
			(0.075)	(0.075)	(0.067)
Inputs/supplier strength for district–industry			0.553+++	0.542+++	0.504+++
			(0.107)	(0.108)	(0.096)
Outputs/customer strength for district–industry			0.291+++	0.292+++	0.246+++
			(0.050)	(0.051)	(0.044)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	6451	6451	6451	6451	6451
Adjusted R^2	0.195	0.233	0.264	0.267	0.294

Notes: The dependent variable is log entry employment by district–industry.
See also Table 2a.

to this difference below when analysing the entrant size distribution.

Across these columns of Table 2a, the R^2 value increases from 0.13 to almost 0.30. While still modest, this growth in explanatory power due to modelling regional conditions is more substantial than that evident in the work of GLAESER and KERR (2009) for the United States. This pattern highlights the greater relative importance of existing district conditions relative to incumbent positioning for explaining entrepreneurship in India, which will be returned to in the conclusions.

Table 2b considers unorganized manufacturing, and several differences exist when compared with Table 2a. First, the local population plays a much greater role, with approximately unit elasticity. Entrepreneurship in the unorganized sector is much more proportionate to local market sizes than in the organized sector. This theme is also evident in the independence of entry from local population density or travel time to a major city, the stronger relationship of entry to the age profile of the district, and the higher R^2 values in columns (1) and (2). Unorganized manufacturing clearly conforms much more closely to the overall contours of India's economic geography than does organized manufacturing.

The other two district traits that are associated with strong entry rates are the strength of local, within-district physical infrastructure and the strength of local household banking environments. This contrasts with organized manufacturing entry, where education stands out. An intuitive explanation, which will also be reflected in the services estimations, is that these patterns and their differences reflect the factors on which each sector depends most. Organized manufacturing establishments have broader resources that reduce dependency on local infrastructure and household finance. Likewise, it is reasonable to believe that the unorganized sector depends less on educated workers than the organized sector. While intuitive, these results should be viewed as partial correlations until they can be rigorously confirmed in future research.

Again evidence is found for agglomeration economies within the unorganized manufacturing sector. The framework is similar to Table 2a except that the Chinitz effect is not considered, since by definition the unorganized sector is comprised of small firms. Partly as a consequence of this, the inputs metric is relatively stronger in these estimations. The initial gap in explanatory power between the organized and unorganized sectors that was evident in columns (1) and (2) is diminished in the complete estimations in columns (3) to (5).

Table 3 considers organized and unorganized services entry. The contrast to organized manufacturing is again quite intriguing. First, overall district population is as important as it was for unorganized manufacturing, with elasticity greater than one. Similarly, the R^2 value grows to 0.20 and 0.47 with just the parsimonious set of explanatory factors in columns (1) and (5), respectively. The R^2 value using the GLAESER and KERR (2009) approach for organized services is 0.30. Also similar to unorganized manufacturing, population density and travel time to major cities are not important in the multivariate setting, while the district's age profile does contribute to higher entry levels.

To recap, education and infrastructure matter the most among district traits. Education is generally more important, with particular relevance to organized sectors. Physical infrastructure has particular relevance to the unorganized sectors of the economy. The strength of the household banking sector is again also very important in the unorganized sectors of the economy. These channels provide three of the main ways that policy-makers can influence the spatial distribution of entry.

The role of the existing incumbent employment by district–industry for services is weak in Table 3, likely suggesting that Marshallian economies are weaker in services. Unreported estimations further model Marshallian interactions in the services sector similar to manufacturing. These results are also weak, at most suggesting a small role for labour market interactions. However, the authors hesitate to interpret this difference strongly as the weak results may be due to the application of concepts and metrics originally designed for manufacturers to the service sector.

Table 4 provides some extensions for organized manufacturing. Following the discussion in the third section, column (1) first includes the small firm incumbent share control. Including this control sharpens the earlier results further, including making the Chinitz effect more robust. Evidence for the general small firm effects outlined by FRITSCH and FALCK (2007) and PARKER (2009b), as well as the Chinitz effect, are thus found.

Columns (2) to (5) break out entrants by their sizes; and Table 5 provides a broader depiction of the entrant size distribution. Starting with Table 5, panel (A) presents the full entrant distribution that includes the organized and unorganized sectors. The complete distribution across both sectors looks broadly similar to other environments. For example, 98% of entering establishments have fewer than ten employees, and only 0.09% of entering establishments have more than 100 workers. In terms of employment shares, 76% of employment in entering establishments is contained in establishments with fewer than ten employees, versus 9.5% in those entering with more than 100. Panel (B) isolates the organized sector, and within this group the largest entrant size category contains 5.5% of establishments and 53% of employment. The district-level variation is also consistent around these traits.

Thus, the unorganized sector accounts for most entrants and employments, and includes plants that are by definition very small. The larger plants included in

Table 3. District entrepreneurship estimations – services

	Organized services				Unorganized services			
	Base estimation (1)	District traits (2)	Adding consumption (3)	Using log entry count (4)	Base estimation (5)	District traits (6)	Adding consumption (7)	Using log entry count (8)
Log of incumbent employment in district–industry	–0.003	–0.104+++	–0.105+++	–0.054++	0.094+++	0.037+	0.037+	0.037+
	(0.038)	(0.033)	(0.033)	(0.023)	(0.024)	(0.021)	(0.021)	(0.021)
Log of district population	1.278+++	1.023+++	1.023+++	0.711+++	1.213+++	1.113+++	1.113+++	1.113+++
	(0.148)	(0.135)	(0.133)	(0.092)	(0.107)	(0.111)	(0.108)	(0.111)
Log of district population density		–0.014	–0.013	–0.028		–0.097+	–0.096+	–0.097+
		(0.086)	(0.087)	(0.056)		(0.057)	(0.058)	(0.057)
Share of the population with graduate education		0.348+++	0.333+++	0.230+++		0.179+++	0.160++	0.179+++
		(0.085)	(0.088)	(0.059)		(0.068)	(0.070)	(0.068)
Demographic dividend for a district (age profiles)		0.548+	0.469	0.329		0.574++	0.465++	0.574++
		(0.331)	(0.349)	(0.230)		(0.229)	(0.235)	(0.229)
Index of infrastructure quality for a district		0.339+++	0.315+++	0.242+++		0.420+++	0.378+++	0.420+++
		(0.096)	(0.106)	(0.067)		(0.068)	(0.074)	(0.068)
Strength of household banking environment		0.174++	0.159+	0.108+		0.323+++	0.302+++	0.323+++
		(0.087)	(0.088)	(0.060)		(0.068)	(0.069)	(0.068)
Stringency of labour laws in a district's state		–0.117+	–0.112+	–0.076+		–0.154+++	–0.146+++	–0.154+++
		(0.067)	(0.067)	(0.046)		(0.048)	(0.048)	(0.048)
Log travel time to closest large city		–0.011	–0.007	–0.021		0.048	0.056	0.048
		(0.054)	(0.054)	(0.037)		(0.051)	(0.050)	(0.051)
Log per capita consumption			0.295				0.454	
			(0.369)				(0.291)	
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	3340	3340	3340	3340	6552	6552	6552	6552
Adjusted R^2	0.201	0.252	0.253	0.252	0.471	0.536	0.536	0.536

Notes: The dependent variable is log entry employment by district–industry.
See also Table 2a.

Table 4. Extended district entrepreneurship estimations – organized manufacturing

		Entering establishment employment of:					
	Including small firm share (1)	10–19 (2)	20–39 (3)	40–99 (4)	100+ (5)	One year entrants (6)	Including lagged entry (7)
Log of incumbent employment in district–industry	0.534+++	0.229+++	0.277+++	0.260+++	0.274+++	0.407+++	–0.082+
	(0.057)	(0.032)	(0.036)	(0.045)	(0.049)	(0.048)	(0.049)
Log of district population	0.358++	0.192++	0.266+++	0.178++	0.099	0.210	0.433+++
	(0.143)	(0.084)	(0.075)	(0.085)	(0.099)	(0.133)	(0.152)
District traits							
Log of district population density	–0.453+++	–0.169+++	–0.160+++	–0.249+++	–0.281+++	–0.343+++	–0.521+++
	(0.069)	(0.038)	(0.034)	(0.043)	(0.062)	(0.053)	(0.071)
Share of the population with graduate education	0.229++	0.107+	0.089	0.066	0.118++	0.184++	0.235++
	(0.099)	(0.060)	(0.055)	(0.051)	(0.054)	(0.086)	(0.106)
Demographic dividend for a district (age profiles)	0.392	0.185	0.340	0.388	0.049	0.196	0.503
	(0.410)	(0.255)	(0.240)	(0.244)	(0.309)	(0.335)	(0.449)
Index of infrastructure quality for a district	0.011	–0.017	–0.042	–0.095	–0.033	–0.104	0.082
	(0.085)	(0.061)	(0.047)	(0.058)	(0.063)	(0.070)	(0.086)
Strength of household banking environment	0.055	–0.002	0.017	0.058	0.090	0.187+++	0.061
	(0.085)	(0.049)	(0.045)	(0.060)	(0.061)	(0.070)	(0.098)
Stringency of labour laws in a district's state	–0.171+++	–0.094++	–0.145+++	–0.107+++	–0.036	–0.130++	–0.139++
	(0.060)	(0.037)	(0.037)	(0.038)	(0.047)	(0.059)	(0.060)
Log travel time to closest large city	–0.183+++	–0.067	–0.064+	–0.121+++	–0.113++	–0.139++	–0.202++
	(0.070)	(0.041)	(0.035)	(0.035)	(0.056)	(0.054)	(0.078)
Local industrial conditions by incumbent firms							
Labour market strength for district–industry	0.034	–0.151++	–0.004	0.048	0.195++	–0.036	0.186+
	(0.099)	(0.066)	(0.068)	(0.074)	(0.082)	(0.087)	(0.103)
Inputs/supplier strength for district–industry	0.204++	0.108+	0.064	0.049	0.059	0.050	0.429+++
	(0.086)	(0.056)	(0.069)	(0.068)	(0.072)	(0.076)	(0.100)
Outputs/customer strength for district–industry	0.230++	0.111++	0.159++	0.247+++	0.275+++	0.235+++	0.364+++
	(0.115)	(0.053)	(0.067)	(0.090)	(0.105)	(0.088)	(0.129)
Chinitz small suppliers metric for district–industry	0.429++	0.530+++	0.368++	0.150	–0.119	0.124	0.221
	(0.209)	(0.184)	(0.158)	(0.155)	(0.139)	(0.156)	(0.214)
Share of small incumbent firms in the district–industry	0.651+++	0.447+++	0.409+++	0.254+++	0.055	0.169+++	
	(0.115)	(0.060)	(0.068)	(0.072)	(0.085)	(0.034)	
Lagged organized manufacturing entry rate for district–industry							0.205+++
							(0.026)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4843	4843	4843	4843	4843	4843	4843
Adjusted R^2	0.169	0.179	0.192	0.196	0.197	0.246	0.245

Notes: The dependent variable is log entry employment by district–industry indicated in the column header.
See also Table 2a. Column (7) includes an unreported dummy variable for zero entry in the lagged period.

Table 5. Distribution of entrant employments and plant counts across size categories

	Employment in entering establishments (%)			Counts of entering establishments (%)		
	India as a whole (1)	District-level mean (2)	District-level standard deviation (3)	India as a whole (4)	District-level mean (5)	District-level standard deviation (6)
(A) Organized and unorganized sectors of manufacturing						
0–4 employees	59.11	68.80	26.85	92.18	92.79	11.35
5–9 employees	16.57	10.46	15.00	5.83	4.76	9.67
10–19 employees	8.85	6.15	10.15	1.57	1.65	4.01
20–39 employees	2.11	2.34	4.33	0.18	0.31	0.78
40–99 employees	3.83	4.46	8.82	0.14	0.33	1.43
100+ employees	9.52	7.80	14.01	0.09	0.16	0.49
(B) Organized sector only of manufacturing						
10–19 employees	36.42	33.55	32.63	79.33	57.81	32.35
20–39 employees	8.69	14.39	21.04	9.18	18.28	22.00
40–99 employees	15.74	18.76	23.30	7.16	14.69	19.87
100+ employees	39.15	27.37	29.93	4.33	9.23	16.40

Note: The distribution of entrant employments and establishment counts across the establishment size distribution is documented. Columns (1) and (4) provide statistics for India as a whole. The district mean and standard deviation columns summarize the unweighted variation at the district level.

the organized sector are still skewed towards the smaller end of the size distribution (e.g., 79% have 10–19 employees) but the largest plants with more than 100 employees have 53% employment share. The definitions of entrants discussed in the second section highlight that the data include new firm formation, but also some elements of new establishments opening in a district. The former dominate the unorganized sector, given its small establishment sizes, while the latter become increasingly important in the larger size categories of the entrant size distribution for the organized sector. This makes a separation very useful, as household banking conditions, for example, may matter less for the organized sector than the labour laws present in India.

Returning to Table 4, the heterogeneity across the entrant size distribution is fascinating and confirms many underlying theories and intuitions advanced above. Small entrants in the organized sector follow existing populations much more, similar to the unorganized sector (shown in Table 2b), while larger entrants in the organized sector are less tied to local demand and avoid places with high population density. The small business and Chinitz effects are much more important for small entrants in the organized sector, while labour markets and industrial output conditions are more critical for large entrants. Column (6) shows fairly similar results when using one-year entrants, with the main differences being a greater emphasis on local banking conditions than local input markets. Column (7) likewise displays broadly similar results when instead controlling for lagged entry rates.

The unpublished appendix provides additional robustness checks on these results: excluding sample weights, including additional covariates such as the female population share and local religious affiliations (e.g., MACK *et al.*, 2013), and clustering standard errors by state. The authors have also tested controls for a district's caste population (IYER *et al.*, 2011), conflict, trade levels and general development levels (leading/lagging designations at the state and district level). These additional controls do not substantively affect the results presented, and the more parsimonious specification is maintained to mirror other work from outside of India. The main specifications are also robust to controlling for incumbent firm counts or value added rather than employment. The unpublished appendix also provides additional work regarding the local industrial traits. Similar results are obtained when district fixed effects are included in the estimations, or when changes in industrial conditions from 1989 to 2005 are used partially to address omitted variable bias concerns.

CONCLUSIONS

Entrepreneurship can be an important factor for economic growth, and India has historically had low entry

rates for the formation of new businesses. This condition is starting to improve, and further growth in effective entrepreneurship is an important stepping stone in India's continued development. This paper explores the spatial determinants of local entrepreneurship for Indian manufacturing and services. Its analysis provides an important baseline for understanding what is important in India's developing economy, both as a first step for policy advice and as a guide to additional research efforts.[9] This foundation also serves a broader academic interest of comparing India's patterns with those of other economies like the United States.

At the district level, the strongest evidence points to the roles that local education levels and physical infrastructure quality play in promoting entry. Evidence is also found that strict labour regulations discourage entrepreneurship, and better household banking environments are associated with higher entry in the unorganized sector. Policy-makers wishing to encourage entrepreneurship in their local areas have several policy levers that can be exploited: investment in both people and places is an easy call for policy-makers, while reducing unnecessary regulations and restrictions is also warranted. This raises the importance of correct policy design for local areas, and it provides a nice testing ground for future work on agglomeration and urban economies. In particular, further research surrounding the time dimensions to entrepreneurship's role in the local economy (e.g., FRITSCH and MUELLER, 2004) for India might be particularly attractive given the rapid pace of the country's transformation.

Research in regional science has also stressed the heterogeneity in entry across industries within a local area. Extensive evidence is also found here that the incumbent compositions of local industries influence new entry rates at the district–industry level within manufacturing. This influence is through traditional Marshallian agglomeration economies, the small firm effect that has been observed in many countries, and the CHINITZ (1961) effect that emphasizes small suppliers. This evidence on localized agglomeration economies and entry is among the first in a developing economy for this growing literature.

Moving to comparative reflections, the similarities between the patterns observed for India and those in the United States are surprisingly large. For example, the strength of the small firm and Chinitz effects were a surprise given that many accounts of India describe how its firm size distribution has been artificially compressed. It could be imagined that the positive channels for entrepreneurship described for advanced economies are greatly diminished when the size distribution is being partially set by the government. Yet, these patterns are comparable. This general comparability is very important as it suggests a substantial degree of portability in the insights derived here in studying advanced economies (e.g., CHATTERJI *et al.*, 2013) to developing and emerging situations.[10]

The differences in the patterns between India and the United States are also instructive and provide important caveats and boundaries on this portability. First, the role and importance of education and physical infrastructure are higher in India than in comparable US studies. By contrast, other dimensions such as population density and regional age structures behave very similarly. The conjecture is that the spatial variation in the latter dimensions within India more closely resembles the variation in advanced economies, and so the same underlying economic forces operate comparably. On the other hand, many parts of India struggle with illiteracy and lack of paved roads, which are not issues on which regional comparisons from the United States can provide insights. Therefore, the important nuance to the broad comparability and portability noted above is that researchers and policy-makers need to contemplate carefully whether the variations utilized in earlier studies are reflective of the variations with which they are dealing.

A second point of comparison with the United States is very striking. While coefficient elasticities are often similar in magnitude, a very striking difference between the present work and that of GLAESER and KERR (2009) is that this paper can generally only account for about one-third of the spatial variation that the US-focused study could. It is posited that a large portion of this gap is due to India being at a much earlier stage of development, especially with the industrial landscape still adjusting to the deregulations of the 1980s and 1990s (e.g., FERNANDES and SHARMA, 2011). District traits and local conditions take on a much greater importance, vis-à-vis incumbent employment distributions, with the economy in transition. At such an early point and with industrial structures not entrenched, local policies and traits can have profound and lasting impacts by shaping where industries plant their roots. These key differences between developing and advanced economies are worthy subjects for further research.

Acknowledgments – The authors thank Ahmad Ahsan, Mehtabul Azam, Rajeev Dehejia, Gilles Duranton, Arti Grover, Michael Fritsch, Lakshmi Iyer, Henry Jewell, Arvind Panagariya, David Storey, Hyoung Gun Wang and three anonymous referees for helpful comments on this work. They thank The World Bank's South Asia Labor Flagship team for providing the primary datasets used in this paper. The authors are particularly indebted to Shanthi Nataraj for sharing her wisdom regarding the industrial survey data. The views expressed in this paper are those of the authors alone and not of any institution with which they may be associated.

Funding – Funding for this project was provided by The World Bank and Multi-Donor Trade Trust Fund.

NOTES

1. High rates of local entrepreneurship are linked to stronger subsequent job growth for regions in several countries (e.g., FRITSCH, 2008; GHANI *et al.,* 2011; GLAESER *et al.,* 2012). MUELLER *et al.* (2008) caution, however, about sweeping statements given the substantial heterogeneity in the British experience, where the job growth of regions depended strongly on the types of entrepreneurs entering and the initial conditions of the regions. BAUMOL (1990) also highlights how the positive or negative role of entrepreneurship depends upon the incentives in society.
2. In contemporaneous work, MUKIM (2011) examines spatial entry patterns for India's unorganized sector. The working paper version of this article discusses similarities and differences between the studies. Other related work includes: DRUCKER and FESER (2007, 2012), ACS and VARGA (2005), ARDAGNA and LUSARDI (2008), ROSENTHAL and STRANGE (2010), DELGADO *et al.* (2010), and CALÁ *et al.* (2013).
3. PARKER (2009a) provides a complete review of the entrepreneurship literature. STOREY (1994) and STOREY and GREENE (2010) give an overview of small businesses and their connections to entrepreneurship specifically. DEICHMANN *et al.* (2008) survey prior work on firm locations in developing economies.
4. The data combine single-unit start-ups with expansion facilities of multi-unit firms. One can, to some degree, separate the entry of multi-unit firms within organized manufacturing, although this distinction is not comprehensively available for all plants. With the splits available, very similar results are found when modelling single-unit entry rates. These splits are not possible for the unorganized sectors and services. A major development limitation for India is the growth and replication of successful initial businesses (e.g., HSIEH and KLENOW, 2009). From this perspective, many policy-makers are equally concerned about encouraging entry of expansion establishments. The working paper version of this article provides an extended discussion about the measures of entrepreneurship and alternative approaches. The paper also returns to this discussion when considering the entrant size distribution.
5. The sampling frame for the organized sector depends on the business register, and a concern might exist that firms indirectly sample out if they select a size so as to avoid registration. As panel data are lacking, corrections like DISNEY *et al.* (2003) cannot be taken. Absent a correlation with one of the explanatory variables, this measurement error will primarily be for the outcome variables and thus it will not bias the estimates. With respect to the explanatory variables, the same covariates with the unorganized sector are also studied. No evidence of this type of gaming behaviour is observed when comparing results for the two sectors.
6. This approach is used by GLAESER and KERR (2009), JOFRE-MONSENY *et al.* (2011), DAUTH (2011), and MUKIM (2011). It follows upon the co-agglomeration work of ELLISON *et al.* (2010).
7. Culture and social capital aspects are taken up by FRITSCH and WYRWICH (2013), KIBLER *et al.* (2013), and WESTLUND *et al.* (2013).
8. A value of less than one entering employee on average is recoded as one entering employee. This maintains a consistent sample size, and the distinction between zero and one employee for a district–industry is not economically meaningful. These cells can be excluded without impacting the results.
9. For example, GHANI *et al.* (2012) extend the distance to major city work by considering the development of the Golden Quadrangle highway system in India and its impact on districts (e.g., DATTA, 2011).
10. ROSENTHAL and STRANGE (2012) and GHANI *et al.* (2013) identify similar features between the United States and India in the spatial sorting patterns of female entrepreneurs.

REFERENCES

ACS Z. and VARGA A. (2005) Entrepreneurship, agglomeration and technological change, *Small Business Economics* **24**, 323–334.

AGHION P., AKCIGIT U., CAGE J. and KERR W. (2012) *Taxation, Corruption, and Growth*. Mimeo. Harvard University.

AGHION P., BURGESS R., REDDING S. and ZILIBOTTI F. (2008) The unequal effects of liberalization: evidence from dismantling the license *raj* in India, *American Economic Review* **98**, 1397–1412.

AHSAN A. and PAGES C. (2007) *Are All Labor Regulations Equal? Assessing the Effects of Job Security, Labor Dispute and Contract Labor Laws In India*. World Bank Working Paper Number 4259. The World Bank, Washington, DC.

AMIN M. and MATTOO A. (2008) *Human Capital and the Changing Structure of the Indian Economy*. World Bank Working Paper Number 4576. The World Bank, Washington, DC.

ARDAGNA S. and LUSARDI A. (2008) *Explaining International Differences in Entrepreneurship: The Role of Individual Characteristics and Regulatory Constraints*. NBER Working Paper Number 14012. National Bureau of Economic Research, Cambridge, MA.

AUDRETSCH D. and FRITSCH M. (1994) The geography of firm births in Germany, *Regional Studies* **28**, 359–365.

AUDRETSCH D., FELDMAN M., HEBLICH S. and FALCK O. (2012) Local entrepreneurship in context, *Regional Studies* **46(3)**, 379–389.

BAUMOL W. (1990) Entrepreneurship: productive, unproductive, and destructive, *Journal of Political Economy* **98**, 893–921.

BESLEY T. and BURGESS R. (2004) Can labor regulation hinder economic performance? Evidence from India, *Quarterly Journal of Economics* **119**, 91–134.

BÖNTE W., FALCK O. and HEBLICH S. (2009) The impact of regional age structure on entrepreneurship, *Economic Geography* **85**, 269–287.

BOSCHMA R. and FRITSCH M. (2007) *Creative Class and Regional Growth – Empirical Evidence from Eight European Countries*. Jena Working Paper Number 2007-066.

BOZKAYA A. and KERR W. (Forthcoming 2013) Labor regulations and European venture capital, *Journal of Economics and Management Strategy*.

Calá C., Arauzo-Carod J. and Manjón Antolín M. (2013) *Regional Determinants of Firm Entry in a Developing Country*. Universitat Rovira i Virgili Working Paper Number 2072/203159.
Chatterji A., Glaeser E. and Kerr W. (2013) *Clusters of Entrepreneurship and Innovation*. NBER Working Paper Number 19013. National Bureau of Economic Research (NBER), Cambridge, MA.
Chinitz B. (1961) Contrasts in agglomeration: New York and Pittsburgh, *American Economic Review* **51**, 279–289.
Combes P. and Duranton G. (2006) Labour pooling, labour poaching, and spatial clustering, *Regional Science and Urban Economics* **36**, 1–28.
Datta S. (2011) The impact of improved highways on Indian firms, *Journal of Development Economics* **99**, 46–57.
Dauth W. (2011) *The Mysteries of the Trade: Interindustry Spillovers in Cities*. Institute for Employment Research Discussion Paper Number 201015.
Dehejia R. and Panagariya A. (2010) *Services Growth in India: A Look Inside the Black Box*. Columbia University School of International and Public Affairs Working Paper Number 4444.
Deichmann U., Lall S., Redding S. and Venables A. (2008) Industrial location in developing countries, *World Bank Observer* **23**, 219–246.
Delfmann H., Koster S., McCann P. and van Dijk J. (2013) *Population Change and New Firm Formation in Urban and Rural Regions*. Mimeo. University of Groningen.
Delgado M., Porter M. and Stern S. (2010) Clusters and entrepreneurship, *Journal of Economic Geography* **10**, 495–518.
Desmet K., Ghani E., O'Connell S. and Rossi-Hansberg E. (2011) *The Spatial Development of India*. World Bank Policy Research Working Paper Number 6060.
Disney R., Haskel J. and Heden Y. (2003) Entry, exit and establishment survival in UK manufacturing, *Journal of Industrial Economics* **51**, 91–112.
Doms M., Lewis E. and Robb A. (2010) Local labor force education, new business characteristics, and firm performance, *Journal of Urban Economics* **67**, 61–77.
Drucker J. and Feser E. (2007) *Regional Industrial Dominance, Agglomeration Economies, and Manufacturing Plant Productivity*. Center for Economic Studies (US Census Bureau) Working Paper Number 07-31.
Drucker J. and Feser E. (2012) Regional industrial structure and agglomeration economies: an analysis of productivity in three manufacturing industries, *Regional Science and Urban Economics* **42**, 1–14.
Duranton G. and Puga D. (2001) Nursery cities: urban diversity, process innovation, and the life cycle of products, *American Economic Review* **91**, 1454–1477.
Duranton G. and Puga D. (2004) Micro-foundations of urban agglomeration economies, in Henderson V. and François Thisse J. (Eds) *Handbook of Regional and Urban Economics*, Vol. 4, pp. 2063–2117. North-Holland, Amsterdam.
Ellison G., Glaeser E. and Kerr W. (2010) What causes industry agglomeration? Evidence from coagglomeration patterns, *American Economic Review* **100**, 1195–1213.
Evans D. and Jovanovic B. (1989) An estimated model of entrepreneurial choice under liquidity constraints, *Journal of Political Economy* **97**, 808–827.
Falck O., Fritsch M. and Heblich S. (2008) *The Apple Does Not Fall From the Tree: Location of Start-Ups Relative to Incumbents*. CESifo Working Paper Number 2486. Munich.
Falck O., Fritsch M. and Heblich S. (2011) The Phantom of the Opera: cultural amenities, human capital, and regional economic growth, *Labour Economics* **18**, 755–766.
Fernandes A. and Pakes A. (2010) *Factor Utilization in Indian Manufacturing: A Look at The World Bank Investment Climate Survey Data*. NBER Working Paper Number 14178. National Bureau of Economic Research (NBER), Cambridge, MA.
Fernandes A. and Sharma G. (2011) *Together We Stand? Agglomeration in Indian Manufacturing*. World Bank Policy Research Working Paper Number 6062.
Figueiredo O., Guimaraes P. and Woodward D. (2002) Home-field advantage: location decisions of Portuguese entrepreneurs, *Journal of Urban Economics* **52**, 341–361.
Fritsch M. (2008) How does new business formation affect regional development?, *Small Business Economics* **30**, 1–14.
Fritsch M. and Falck O. (2007) New industry formation by industry over space and time: a multidimensional analysis, *Regional Studies* **41**, 157–172.
Fritsch M. and Mueller P. (2004) Effects of new business formation on regional development over time, *Regional Studies* **38**, 961–975.
Fritsch M. and Wyrwich M. (2013) *The Long Persistence of Regional Entrepreneurship Culture: Germany 1925–2005*. Jena Economic Research Papers Number 2012-036.
Fujita M., Krugman P. and Venables A. (1999) *The Spatial Economy: Cities, Regions and International Trade*. MIT Press, Cambridge, MA.
Ghani E., Grover Goswami A. and Kerr W. (2012) *Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing*. Harvard Business School Working Paper Number 13-040.
Ghani E., Kerr W. and O'Connell S. (2011) Promoting entrepreneurship, growth, and job creation, in Ghani E. (Ed.) *Reshaping Tomorrow*, pp. 166–199. Oxford University Press, Oxford.
Ghani E., Kerr W. and O'Connell S. (2013) Local industrial structures and female entrepreneurship in India, *Journal of Economic Geography* **13(6)**, 929–964.
Glaeser E. and Kerr W. (2009) Local industrial conditions and entrepreneurship: how much of the spatial distribution can we explain?, *Journal of Economics and Management Strategy* **18**, 623–663.
Glaeser E., Kerr W. and Ponzetto G. (2010) Clusters of entrepreneurship, *Journal of Urban Economics* **67**, 150–168.

Glaeser E., Pekkala Kerr S. and Kerr W. (2012) *Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines*. NBER Working Paper Number 18333. National Bureau of Economic Research (NBER), Cambridge, MA.

Greenstone M., Hornbeck R. and Moretti E. (2010) Identifying agglomeration spillovers: evidence from winners and losers of large plant openings, *Journal of Political Economy* **118**, 536–598.

Hasan R. and Jandoc K. (2010) *The Distribution of Firm Size in India: What Can Survey Data Tell Us?* Asian Development Bank (ADB) Economics Working Paper Number 213.

Hofstede G. (2001) *Culture and Organizations*. HarperCollins, London.

Hsieh C. and Klenow P. (2009) Misallocation and manufacturing TFP in China and India, *Quarterly Journal of Economics* **124**, 1403–1448.

Iyer L., Khanna T. and Varshney A. (2011) *Caste and Entrepreneurship in India*. Harvard Business School Working Paper Number 12-028.

Jofre-Monseny J., Marín-López R. and Viladecans-Marsal E. (2011) The mechanisms of agglomeration: evidence from the effect of inter-industry relations on the location of new firms, *Journal of Urban Economics* **70(2–3)**, 61–74.

Johnson P. and Cathcart D. G. (1979) The founders of new manufacturing firms: a note on the size of their 'incubator's plants, *Journal of Industrial Economics* **28**, 219–224.

Kathuria V. (2011) *What Causes Agglomeration? Policy or Infrastructure – A Study of Indian Manufacturing Industry*. eSocialSciences Working Paper Number 4473.

Kathuria V., Natarajan S., Raj R. and Sen K. (2010) *Organized Versus Unorganized Manufacturing Performance in India in the Post-Reform Period*. MPRA Working Paper Number 20317. Munich Personal RePEc Archive (MPRA), Munich.

Khanna T. (2008) *Billions of Entrepreneurs: How China and India are Reshaping their Futures – And Yours*. Harvard University Press, Boston, MA.

Kibler E., Kautonen T. and Fink M. (2013) *Regional Social Legitimacy of Entrepreneurship: Implications for Entrepreneurial Intention and Start-up Behaviour*. Mimeo. University of Turku.

Klepper S. (2010) The origin and growth of industry clusters: the making of Silicon Valley and Detroit, *Journal of Urban Economics* **67**, 15–32.

Lall S., Wang H. and Deichmann U. (2011) *Infrastructure and City Competitiveness in India*. World Institute for Development Economic Research Working Paper Number 2010/22.

Lazear E. (2005) Entrepreneurship, *Journal of Labor Economics* **23**, 649–680.

Mack E., Faggian A. and Stolarick K. (2013) *Does Religion Stifle Entrepreneurial Activity?* Mimeo. Arizona State University.

Marshall A. (1920) *Principles of Economics*. Macmillan, London.

Michelacci C. and Silva O. (2007) Why so many local entrepreneurs?, *Review of Economics and Statistics* **89**, 615–633.

Mueller P., van Stel A. and Storey D. (2008) The effects of new firm formation on regional development over time: the case of Great Britain, *Small Business Economics* **30**, 59–71.

Mukim M. (2011) *Industry and the Urge to Cluster: A Study of the Informal Sector in India*. Discussion Paper Number 0072. Spatial Economics Research Centre, London.

Nataraj S. (2011) The impact of trade liberalization on productivity and firm size: evidence from India's formal and informal manufacturing sectors, *Journal of International Economics* **85(2)**, 292–301.

Parker S. (2009a) *The Economics of Entrepreneurship*. Cambridge University Press, Cambridge.

Parker S. (2009b) Why do small firms produce the entrepreneurs?, *Journal of Socio-Economics* **38**, 484–494.

Reynolds P., Storey D. and Westhead P. (1994) Regional characteristics affecting entrepreneurship: a cross-national comparison, in *Frontiers of Entrepreneurship Research*, pp. 550–564. Babson College, Babson Park, MA.

Rosenthal S. and Strange W. (2004) Evidence on the nature and sources of agglomeration economies, in Henderson V. and François Thisse J. (Eds) *Handbook of Regional and Urban Economics*, Vol. 4, pp. 2119–2171. North-Holland, Amsterdam.

Rosenthal S. and Strange W. (2010) Small establishments/big effects: agglomeration, industrial organization and entrepreneurship, in Glaeser E. (Ed.) *Agglomeration Economics*, pp. 277–302. University of Chicago Press, Chicago, IL.

Rosenthal S. and Strange W. (2012) Female entrepreneurship, agglomeration, and a new spatial mismatch, *Review of Economics and Statistics* **94**, 764–788.

Storey D. (1994) *Understanding the Small Business Sector*. Thomson Learning, London.

Storey D. and Greene F. (2010) *Small Business and Entrepreneurship*. Prentice Hall, Englewood Cliffs, NJ.

Van Stel A., Storey D. and Thurik R. (2007) The effect of business regulations on nascent and young business entrepreneurship, *Regional Studies* **28**, 171–186.

Westlund H., Larsson J. and Olsson A. (2013) *Startups and Local Social Capital in the Municipalities of Sweden*. European Regional Science Association Conference Papers Number ersa12p91.