The Community Origins of Private Enterprise in China*

Ruochen Dai† Dilip Mookherjee‡ Kaivan Munshi§ Xiaobo Zhang¶

February 28, 2019

Abstract

This paper identifies and quantifies the role played by birth-county-based community networks in the growth of private enterprise in China. The starting point for the analysis is the observation that population density is positively associated with local social interactions, social homogeneity, and enforceable trust in counties (but not cities). This motivates a model of network-based spillovers that predicts how the dynamics of firm entry, concentration, and firm size vary with birth county population density. The predictions of the model are validated over the 1990-2009 period with administrative data covering the universe of registered firms. Competing non-network-based explanations can explain some, but not all of the results. We subsequently estimate the structural parameters of the model and conduct counter-factual simulations, which indicate that entry and capital stock over the 1995-2004 period would have been 40% lower without community networks. Additional counter-factual simulations shed light on misallocation and industrial policy.

Keywords. Community Networks. Enforceable Trust. Entrepreneurship. Misallocation. Informal Institutions. Growth and Development.
JEL. J12. J16. D31. I3.

*We are grateful to Emily Breza, Matt Elliot, Debraj Ray, Daniel Xu, and numerous seminar and conference participants for their constructive comments. Research support from the Economic Development Initiative (EDI)and Cambridge-INET is gratefully acknowledged. We are responsible for any errors that may remain.

†Peking University dairuochenpku@gmail.com
‡Boston University dilipm@bu.edu
§University of Cambridge munshi@econ.cam.ac.uk
¶Peking University x.zhang@nsd.pku.edu.cn

1 Introduction

China has witnessed the same degree of industrialization in three decades as Europe did in two centuries (Summers, 2007). This economic transformation began in the early 1980's with the establishment of township-village enterprises (TVE's) and accelerated with the entry of private firms in the 1990's. Starting with almost no private firms in 1990, there were 15 million registered private firms in 2014, accounting for over 90% of all registered firms and 60% of aggregate industrial production. The surge in the number of private firms has had major macroeconomic consequences. China is the largest exporter in the world today and, depending on how the accounting is done, the world's largest or second-largest economy (Wu, 2016).

What is perhaps most striking about the growth of private enterprise in China is that it occurred without the preconditions generally believed to be necessary for market-based development; i.e. without effective legal systems or well functioning financial institutions (Allen *et al.*, 2005). While the government played an important role in China's economic transformation by providing infrastructure and credit (Long and Zhang, 2011; Wu, 2016), it has been argued that informal mechanisms based on reputation and trust must have been at work to allow millions of entrepreneurs, most of whom were born in rural areas, to establish and grow their businesses (Peng, 2004; Allen *et al.*, 2005; Song *et al.*, 2011; Greif and Tabellini, 2017; Zhang, 2017). Case studies of production clusters; e.g. Fleisher *et al.* (2010) and Nee and Opper (2012) indicate that long-established relationships among relatives and neighbors (from the rural origin) substitute for legally enforced contracts between firms. These case studies document a high degree of mutual reliance of firms within the cluster for exchange of intermediate goods, information, marketing connections and finance. Our analysis, which utilizes comprehensive data covering the universe of registered firms over many years, advances this line of research by identifying and quantifying the role played by informal community-based business networks in the growth of private enterprise in China.

Our analysis proceeds in four steps. First, we argue that population density in rural counties, which is mechanically correlated with spatial proximity, is positively associated with social connectedness. Spatial proximity results in more frequent social interactions that, under plausible assumptions on the matching process, give rise to more inter-connected social networks (Coleman, 1988; Jackson *et al.*, 2012). Inter-connected networks sustain greater economic cooperation via norms based on community enforcement (Greif, 1993, 1994; Greif and Tabellini, 2017). The nature of trust and cooperation we focus on is distinct from from the generalized trust that has received much attention in the rapidly growing economics literature on culture (Alesina and Giuliano, 2015). The domain of the former is restricted to local residents rather than the general population, and is sustained via external enforcement rather than by internalized cultural values. We provide evidence linking population density and localized trust with data from the population census and the China Family Panel Survey (CFPS) which covers a nationally representative sample of households: the frequency of local social interactions and trust in neighbors are both increasing in county population density, after controlling for total population, occupation structure and average education. This result holds in counties, but not cities. Moreover, even within counties, it applies only to trust in local residents rather than outsiders.

The argument that local cooperation is increasing in population density is, however, predicated on the assumption that the local population is socially homogeneous, or that the degree of homogeneity is either independent or increasing in population density. While there is a well established sociological literature that describes the role of the hometown or birth county in supporting economic activity in China; e.g. Honig (1992), Goodman (1995), it has also been claimed that the clan, a traditional kinship unit within the county, has renewed its relevance in the post-collectivist era (Peng, 2004; Greif and Tabellini, 2017). We provide evidence, using the CFPS and the population census, that clan concentration within counties (using surnames to identify clans, as in Peng (2004)) is positively correlated with population density. In contrast, social homogeneity is found to be decreasing in population density in cities, which may explain why local social interactions and localized trust are not associated with urban population densities. The second step in our analysis, which links population density to entrepreneurship, thus focuses on county-born businessmen. This is an important group to study because their firms account for two-thirds of all registered private firms (and a comparable share of total registered capital) in China. The majority of our county-born entrepreneurs establish their firms outside their birth counties, often far away. Our key assumption, which is consistent with the literature on temporary migration; e.g. Morten (2019) is that these entrepreneurs remain connected to, and can be sanctioned by, their communities. If this assumption is satisfied, business networks drawn from higher population density counties will support higher levels of mutual help among their members.

The canonical model of community-based entrepreneurship that we develop to validate the preceding hypothesis features successive cohorts of agents that make a choice between a traditional occupation (such as farming or wage labor) and becoming a private entrepreneur. Individual abilities are drawn from an i.i.d. process and affect returns to both occupations; the payoff from entrepreneurship depends additionally on what we refer to as community TFP (CTFP); i.e. the contribution of the network (defined by the set of incumbent entrepreneurs from a common birth county) via mutual help. Help provided by different network members is mutually complementary, which implies, in turn, that there are increasing returns to network size. Moreover, the level of help (for given network size) is increasing in birth county population density, a measure of 'network quality', for reasons discussed above. The model generates predictions for the dynamics of entry across successive cohorts and the consequent evolution of network size, as a function of initial entry, population density at the origin, and time. It features a network 'multiplier' where higher initial entry generates greater entry flows in later cohorts, with the multiplier increasing in population density and in time.

These predictions are tested with administrative data obtained from the State Administration of Industry and Commerce (SAIC) covering the universe of registered firms in China. The following information is available for each private firm: establishment date, sector, location, registered capital, and a list of major shareholders and managers, with their citizenship ID. The county of birth can be extracted from the citizenship ID and the firm's legal representative is designated as the "entrepreneur" for our analysis. Since the first wave of private entry in China commenced in the early 1990's, we estimate the effect of initial entry in 1990-1994 on subsequent entry, separately in 2000-2004 and 2005-2009, at the level of the birth county-sector-location. All counties and cities (urban districts) in the country are included in the set of locations in the benchmark specification, although we verify that the results are robust to excluding the birth county itself. Consistent with a dynamic network multiplier effect, we find that one additional entrant from the birth county in a particular sector-location in the initial 1990-1994 period is associated with seven additional entrants over the 2000-2004 period and nine additional entrants over the 2005-2009 period.

While these results highlight the inter-linkage between firms from the same birth county, non-network explanations for the inter-temporal correlation are also available. A stronger test that networks are active is that the initial entry effect should be increasing in birth county population density, regardless of where firms are located, and this is indeed what we find. We also find that conditional on the number of initial entrants from the birth county, the total number of initial entrants in a sector-location (aggregating across all origins) does not predict the number of subsequent entrants. This is consistent with the hypothesis that the birth county is the domain within which business networks are organized in China, rather than destination sectors or locations. The absence of a (negative) cross-community effect also goes against the view that the networks are simply competing with each other for subsidized credit from the local government, as in Bai *et al.* (2019), although our framework does not rule out the presence of such rent seeking.

We supplement these results by testing the hypothesis that the primary locus of informal cooperation within a county is the clan. Using birth county and surname to measure clan affiliation, and implementing the same test as above, we find that network spillovers are concentrated within the clan, although greater initial entry by non-clan members from the birth county also has a positive and significant (albeit smaller) effect on subsequent entry. Assuming that the spillovers are restricted to the clan, the model predicts in addition that entering entrepreneurs will be increasingly concentrated in particular clans, with a steeper increase in clan concentration over time for entrepreneurs from higher population density birth counties. The data support each of these predictions, highlighting the sociological roots of private enterprise in China.

The third step of our analysis augments the model of community-based entrepreneurship to incorporate sector-location choice and capital investment. Although the augmented model is set up to match the basic features of the canonical model, there are now two sources of network-based spillovers: (i) post-entry mutual help among networks members, as before, and (ii) a pre-entry referral process which increasingly channels entering firms from a given origin into an initially favored destination (a term we use to denote either sectors or locations). The interaction between the two types of spillovers generates dynamic increasing returns to network size in any given destination, resulting in increased sectoral and spatial (within sector) concentration over time. Concentration is also rising in origin population density at each point in time, with a slope that increases over time at early stages of the industrialization process.

With regard to capital investment, the network-based spillovers that raise CTFP over time have two conflicting effects on the initial size of the marginal entrant's firm: the direct effect, for a given level of ability, is to increase firm size by raising the firm's TFP, but an increase in CTFP also lowers the ability threshold for entry into entrepreneurship and this negative selection works in the opposite direction to lower TFP. We show that the latter effect dominates; the marginal entering firm from a higher population density birth county will be unambiguously smaller, with this negative relationship growing stronger over time as networks get larger. Under specific conditions on the model's parameters, this result is shown to hold for average initial firm size as well. This contrasts with the model's predictions for the post-entry growth in firm size. This growth is driven by changes in CTFP over time and is the same for all firms in a network at a given point in time, regardless of their cohort or the ability of the entrepreneur. Because networks from higher population density birth counties are growing faster, firms from those counties will start small but subsequently grow faster.

The predictions of the augmented model are tested over the 1990-2009 period with SAIC registration data, supplemented with the industrial census and the SAIC inspection database for the analysis of firm growth. Population, education, and the occupational structure (measured in the birth county in 1982) are included in the estimating equations as controls. Population density could, nevertheless, be correlated with other unobserved determinants of the outcomes of interest and, hence, we focus on the less easily explained *interaction* between birth county population density and time; this is similar to a difference-in-difference analysis, except that we are leveraging the dynamic implications of the model. Exploiting the fact that firms from many birth counties are established in the same sector-location, we also include sector-time period and location-time period effects in the estimating equations. This effectively controls for access to resources provided by local governments, and for the destination-based productivity spillovers emphasized in the literatures on endogenous growth (Romer, 1990, 1986; Segerstrom *et al.*, 1990; Aghion and Howitt, 1992; Jones, 1995; Segerstrom, 1998) and agglomeration (Ciccone and Hall, 1996; Ellison and Glaeser, 1997; Au and Henderson, 2006; Combes *et al.*, 2012). The model generates dynamic predictions for the relationship between birth county population density and a rich set of outcomes – firm entry, sectoral and spatial concentration, firm size – and the data match each of them. Moreover, the relationships get stronger when sector and location effects are added to the estimating equation. This indicates that entrepreneurs from higher population density birth counties are, if anything, selecting into less favorable sectors and locations. As a final check, we consider a comprehensive set of potential non-network explanations for our results, which are obtained by systematically relaxing different assumptions of the model. We show that while these explanations can explain some of the results, there is no other explanation that can account for all of them simultaneously.

There are many anecdotal examples of the role played by communities in supporting the business activities of their members, in China and elsewhere (see Munshi (2014) for a survey). There have also been some attempts; e.g. Banerjee and Munshi (2004), Munshi (2011), to estimate the impact of community networks on selection into entrepreneurship and firm outcomes in specific industries and locations. However, our analysis is the first that we are aware of to document the role played by communities in the evolution of private enterprise in an entire economy. Having tested and validated the network-based model, the final step in the analysis thus seeks to quantify the impact of these networks on aggregate firm entry and capital stocks by estimating the structural parameters of the model and then conducting counter-factual simulations. Although the model is extremely parsimonious, it does a good job of matching entry and initial capital within sectors, across the range of birth county population densities, within the sample period for the structural estimation (1995-2004) and out of sample. This increases our confidence in the results of a counter-factual experiment, which estimates that entry from county origins would have declined by as much as 64% over the 1995-2004 period (with a comparable decline in the total capital stock) had the networks not been active. As entry and capital stocks from county origins accounted for approximately two-thirds of all entry and capital in China, this amounts to an impact of approximately 40% for the entire country. Given the dynamic increasing returns that are generated by the networks, the long-term consequences of their absence would have been even more substantial.

Two stylized facts motivate a large and growing literature in macro-development: (i) the variation in marginal productivity and, hence, firm size within narrow sectors is especially wide in developing economies, and (ii) firms in those economies are small (Peters, 2016). Although a number of mechanisms can explain these facts; e.g. Caunedo (2016), Asker *et al.* (2014), Akcigit *et al.* (2016), Haltiwanger *et al.* (2018), perhaps the simplest is based on a model with mark-ups in output prices and wedges in factor prices (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009, 2014; Peters, 2016). There are no price distortions in our model. Our objective is to demonstrate that wide dispersion in firm size and productivity can be a consequence of community-based inter-firm spillovers, rather than inefficient taxes or regulations. The results of a second counter-factual policy experiment, which provides a temporary credit subsidy to entering firms, suggests that optimal second-best policies should target these subsidies to more socially connected communities, as a way of exploiting the resulting network spillovers from increased entry. This would increase increasing dispersion and the proportion of small firms even further. More generally, we would not want to infer that one developing economy is less efficient than another developing economy because it has smaller firms or greater dispersion in firm size. Indeed, these characteristics may be symptomatic of a more dynamic economy in which underlying community networks are responding more effectively to market frictions. Although this insight, and the message that policies aimed at raising growth and efficiency should incorporate intra-community spillovers over and above individual ability, would apply to all economies where community networks are active, the important qualifier is that the magnitude of the estimated network effects are specific to China and would not necessarily extend to other countries with different social structures and economic infrastructure. Moreover, policies that attempt to exploit these network effects will have complex distributional consequences; in particular, policies that target more connected communities are likely to exacerbate inter-community inequality, while promoting intra-community equity.

2 Institutional Setting

2.1 Private Enterprise in China

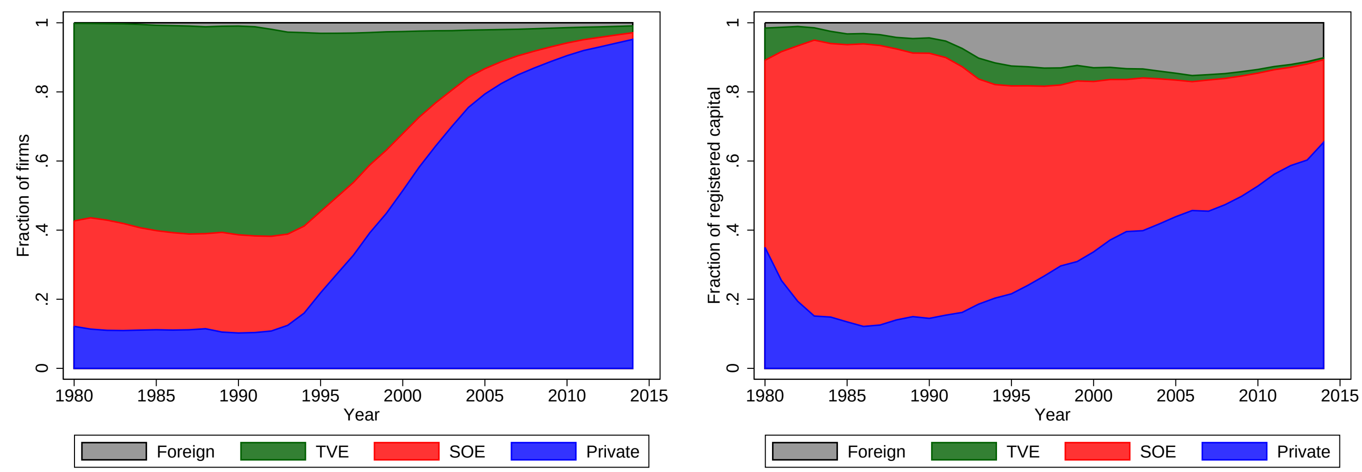
The State Administration of Industry and Commerce (SAIC) database that we utilize for much of the empirical analysis comprises the universe of registered firms in China, regardless of their size, from 1980 onwards. These firms are classified as township-village enterprises (TVE's), state owned enterprises (SOE's), foreign owned firms, and private (domestically owned) firms. Our analysis focuses on the private firms. The SAIC database lists the major shareholders and managers, with their citizenship ID, in each registered private firm. New firms enter the database each year, while a fraction of incumbents exit. We can thus trace the initial growth phase of private enterprise in China in its entirety; starting with a relatively small number of private firms in 1990, there were close to 15 million registered private firms in 2014 (and 7.3 million private firms in 2009, which will be the end point of our statistical analysis).[1] As documented in Figure 1a, private firms accounted for approximately 10% of all firms in the early 1990's. By 2014 they accounted for over 90% of all firms.[2]

[1]In contrast, previous analyses of firms in China have relied on a publicly available database of manufacturing firms with sales above a threshold level (5 million Yuan) over the shorter 1998-2008 period; e.g. Hsieh and Klenow (2009), Song *et al.* (2011), Brandt *et al.* (2012), Aghion *et al.* (2015). The above-scale firms account for less than 15% of all private firms in the registration database in 2008.

[2]Appendix Figure D.1 reports the number of private firms, TVE's, and SOE's, in each year over the 1990-2014 period. As can be seen, the number of private firms is an order of magnitude larger than the number of TVE's and SOE's by the early 2000's, with the gap continuing to grow steeply thereafter. The ownership type assigned to a given firm in the figure is based on its final type (in the event that it was privatized). However, just 5.5% of registered private firms report any change in their ownership structure, which includes changes in their ownership type. Given the dominance of the private firms in terms of their numbers, their growth evidently could not have been generated simply by privatizing existing TVE's and SOE's.

Figure 1b reports the share of total registered capital, by firm-type, over the 1980-2014 period.[3] As with their numbers, the share of registered capital held by private firms grew steeply from the early 1990's onwards and by 2014 they held 60% of total registered capital in the economy.

Figure 1. Distribution of Firms, by Type



(a) Number of Firms

(b) Registered Capital

Source: SAIC registration database.

The preceding facts suggest that private firms played an important role in China's rapid growth over the past decades. This raises the natural question about the underlying cause of this surge in private enterprise. It is generally believed that governments at the local (county), provincial, and central level played a substantial role in China's economic transformation. Local governments provided the infrastructure to support production clusters located throughout the country, which are a distinctive feature of the Chinese economy (Long and Zhang, 2011). Provincial governments and the central government supported firms by giving them subsidized credit and by aggressively promoting exports (Wu, 2016). Our interest, however, is in the role played by informal community-based institutions in supporting private entrepreneurship. Allen *et al.* (2005) argue that reputation and relationships must have substituted for missing financial institutions for China to grow so rapidly. Song *et al.* (2011) explain China's unique growth path as consequence of the fact that more productive private firms had to rely on self-financing in the absence of low cost formal finance. Case studies of Chinese production clusters; e.g. Huang *et al.* (2008), Ruan and Zhang (2009), and Fleisher *et al.* (2010) consistently find that the impetus for their formation came from within, with groups of entrepreneurs setting up firms with little external support. The involvement of local governments is found to come later, through the provision of infrastructure such as roads, markets, and quality control.

What specific institutional arrangements allowed these early entrants to come together and become private entrepreneurs? There are many accounts of the role played by social networks or *guanxi* in facilitating China's historically unprecedented rural-urban labor migration over the past decades; e.g. Zhao (2003), Zhang and Li (2003), Hu (2008). These accounts describe how migrant networks are organized around the rural hometown,

[3]The initial registered capital represents the total amount paid up by the shareholders. This amount is deposited with the SAIC, and can be used to pay the firm's operating expenses before it becomes cash flow positive. Access to bank credit is also dependent on the firm's registered capital, which is why firms will often choose to increase their registered capital over time.

complementing an older anthropological literature that takes the position that ethnicity in China is defined by the native place; e.g. Honig (1992, 1996), Goodman (1995).[4] If the sending county is the domain around which migrant labor networks are organized, then it is plausibly the natural domain around which business networks supporting county-born entrepreneurs are organized.[5]

Counties in China are divided into villages, which consist, in turn, of one or more clans or lineages (Peng, 2004; Tsai, 2007). These clans historically supported the business activities of their members, who were bound together by mutual moral obligations. It has been argued that this role has re-emerged in the post-collectivist era (Peng, 2004; Zhang, 2017; Greif and Tabellini, 2017). We remain agnostic about the boundary of the social unit from which business networks are drawn in this paper; i.e. whether it is the county or the clan. As discussed below, the county characteristic that we use as the source of forcing variation in the empirical analysis would apply to the county as a whole and to all clans within the county, and thus our results would apply in either case.

2.2 Population Density and Trust in Chinese Counties

The point of departure for our analysis is the assumption that social connectedness in Chinese counties is increasing in population density. The underlying idea is that greater spatial proximity raises the frequency of social interactions and facilitates communication with neighbors. This, in turn, helps sustain higher levels of mutual cooperation, supported by the threat of social sanctions, as argued in early papers on social norms and community enforcement (Greif, 1993, 1994; Kandori, 1992; Ellison, 1994). To make this argument more precise, consider a random graph model in which the probability that an individual is connected to any other individual in a local population is equal to γ, which is rising in spatial proximity. A higher γ directly raises the degree of the social network (the number of links per capita), and indirectly also network inter-connectedness i.e. the probability that friends of friends are linked, and so on. For example, the rate of triadic closure – the probability that any three individuals are directly linked – is increasing in γ. Coleman (1988) argues that network closure is a necessary condition for economic cooperation, enforced by social sanctions. Jackson *et al.* (2012) make a similar argument based on a related network property, which they refer to as "support." These results do not necessarily rely on the random matching assumption, and are also likely to hold if the matching process exhibits homophily.[6]

The preceding discussion implicitly assumes that the county is socially homogeneous and comprises a single community. Matters are more complex when the county population is fragmented into smaller communities. For example, suppose that each county is composed of a number of clans $j = 1, \ldots,$ where cooperation

[4]Migrants from the same rural origin move to the city in groups and most migrants end up living and working with *laoxiang* or "native-place fellows" (Fang, 1997; Ma and Xiang, 1998; Zhang and Xie, 2013). In Chinese cities, migrant-peasant enclaves are often named after a sending province, but as Ma and Xiang (1998) note, this nomenclature is misleading because the enclave typically consists of peasants from a single county or two neighboring counties.

[5]A similar argument has been made in past research in India, where the endogamous caste or *jati* is the common domain around which networks supporting rural-urban migration, business, and other functions are organized (Munshi and Rosenzweig, 2006, 2016; Munshi, 2011).

[6]Jackson and Rogers (2007) describe a model of matching in which some links are formed randomly, whereas other links are formed strategically. The latter process results in clustering or homophily, with individuals having higher degree (more links) more likely to be matched to each other. While adding a strategic element to the matching process will generally increase the inter-connectedness of the network, the relationship between population density (γ) and inter-connectedness now becomes more complex. Nevertheless, we continue to expect this relationship to be positive by a limit argument; as γ goes to zero, no one is connected and the rate of triadic closure goes to zero. As γ goes to one, everyone is connected, and the rate of triadic closure goes to one.

is sustained only within clans, not across clans. Let ζ_j denote the demographic weight of clan j, so clan composition of the county is represented by the vector $\zeta \equiv \{\zeta_j\}_j$. The likelihood of sanctions being applied by clan members is increasing in the population density p (which determines the frequency of interactions within every clan in the county), while the magnitude of the sanction is rising in the relative size of the clan (e.g., a larger clan will correspond to a higher proportion of neighbors that impose sanctions on any deviant). The total expected sanction and, hence, the level of cooperation in the clan is the product of two functions, $\nu_1(p)$ which is rising in population density p, and $\nu_2(\zeta)$ which is rising in the relative size of the clan. Aggregating across clans, average cooperation in the county equals $\nu_1(p)\sum_l \zeta_l \nu_2(\zeta_l)$. If clan composition is independent of p, then again a rise in population density is associated with higher average cooperation. However, if clan composition is correlated with p, the effect of a higher p is complicated as it depends on the nature of this correlation, as well as the curvature of the function $\nu_2(\zeta)$.[7] Nevertheless, if clan concentration is positively associated with p, and greater concentration (which implies greater social homogeneity) is associated with an increase in cooperation — as frequently observed in practice[8] — then higher population density will continue to be associated with higher average cooperation in the county.

The assumption that social homogeneity is increasing in population density is unlikely to apply to cities. Most urban residents in developing countries are recent arrivals, typically from diverse origins. Greater population density in an urban area may thus reflect greater diversity of origins, rather than higher intra-community density. The social enforcement that can be supported within long-established clans or among local residents in counties (who have been living together for generations) thus will not be necessarily sustainable in urban neighborhoods.

The preceding discussion indicates that cooperation is increasing in population density in a socially heterogeneous population if both the frequency of local social interactions and social homogeneity are increasing in population density. We use the China Family Panel Survey (CFPS), which covers a nationally representative sample of households, to provide empirical support for each of these relationships, in counties but not in cities.[9] Table 1 reports the relationship between population density and local social interactions, separately in counties and cities, based on the 2010 round of the CFPS. We divide the different types of social interactions reported in the family module of the CFPS into planned interactions; i.e. group entertainment, visits to neighbors' homes, and dining together, and unplanned interactions, which are defined as meetings or conversations without other background activities.[10] Population density in the counties and cities covered by the CFPS, and in all the analysis that follows, is computed from the 1982 population census. This is before the first wave of privatization in the early 1990's and the accompanying rural-urban migration, which could have endogenously shifted the population density in ways that are correlated with the outcomes of interest.[11] Moreover, the estimating

[7] A mean preserving increase in spread of the distribution $\{\zeta_j\}_j$ will increase (resp. decrease) average cooperation if ν_2 is convex (resp. concave).

[8] Peng (2004) and Zhang (2017) show that entrepreneurship and employment in private enterprise in China are positively associated with measures of clan concentration within villages or prefectures, after controlling for local area characteristics.

[9] There are approximately 2,000 counties and 250 prefecture-level and province-level cities (which are further divided into urban districts) in China.

[10] Group entertainment includes playing *mahjong* or cards, reading newspapers, listening to the radio, or watching TV with others.

[11] Population density is measured in units of 10,000 people per square km. The threshold density in our analysis is set at 0.002; i.e. 20 people per square km. This excludes sparsely populated regions such as Western China, Inner Mongolia, and Tibet, which are inhabited by ethnic minorities with a different culture than the majority Han Chinese.

equation in Table 1, and all the analyses in the paper that estimate the direct effect of population density, include population, education, and occupational structure in the county or city to allow for the possibility that these variables (which are correlated with population density) could be directly determining the outcomes.

Columns 1-2 and 4-5 report the relationship between population density and the frequency of each category of local social interactions, per month, for respondents residing in counties and cities, respectively. The population density coefficient is positive (and statistically significant with unplanned interactions as the dependent variable) for county residents, whereas it is negative (and statistically insignificant) for city residents. Table 1, Columns 3 and 6 provide additional support for this distinction between counties and cities by estimating the relationship between population density and a different measure of local social interactions obtained from the adult individual module of the CFPS. This measure is based on a question that ascertains who the respondent has the most unplanned social interactions with: local residents, relatives, classmates, colleagues, or others. We construct a binary variable, which indicates whether the respondent lists local residents as his most frequent interaction partners. With this measure as the dependent variable, the population density coefficient is positive and statistically significant for county residents (Column 3) and negative and statistically insignificant for city residents (Column 6).

Table 1. Frequency of Local Social Interactions and Population Density

Respondent's location:	county			city		
Dependent variable:	frequency of social interactions per month with local residents		indicates whether most unplanned interactions are with local residents	frequency of social interactions per month with local residents		indicates whether most unplanned interactions are with local residents
Type of social interactions:	planned	unplanned		planned	unplanned	
	(1)	(2)	(3)	(4)	(5)	(6)
Population density	0.343	1.635**	0.022**	-0.778*	-0.573	-0.019
	(0.422)	(0.789)	(0.010)	(0.398)	(0.834)	(0.012)
Mean of dependent variable	3.907	15.54	0.197	3.434	13.56	0.124
Observations	8,830	8,830	20,669	3,255	3,255	7,088

Source: China Family Panel Survey (2010). Columns 1-2 and 4-5 based on family module and Columns 3 and 6 based on adult individual module.
Planned interactions include group entertainment, visits to neighbors' homes, and dining together. Unplanned interactions are one-on-one social meetings without other background activities.
Population density is measured in units of 10,000 people per square km, and then converted to Z-score.
Control variables include population, education and occupation distribution in the birth place.
Population density, population, education and occupation distribution are computed from the 1982 population census.
Standard errors clustered at the county or city level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Table 2 provides an explanation for the observed differences in social interactions between counties and cities based on their distinct social structures. We begin, in Columns 1 and 3, by establishing that population density at the aggregate – city or county – level, obtained from the 1982 population census, is positively associated with population density at the local – neighborhood or village – level, obtained from CFPS (2010). Next, we compare the fraction of locally born residents in counties and cities in Columns 2 and 4, respectively. Based on the mean of the dependent variable, these fractions are very different; while around 90% of county residents are born there, this statistic is lower than 50% in cities. Moreover, the fraction of locally born residents is decreasing in local population density in cities (Column 2) but not counties (Column 4), and this increasing social heterogeneity might explain the weakly negative relationship between population density and social interactions that we observe in the cities. In contrast, Column 5 indicates that social homogeneity, appropriately defined, could even be increasing in population density in Chinese counties. If economic cooperation is organized within the clan, then the appropriate measure of social homogeneity is clan concentration. Members of a clan will share the same surname (Peng, 2004) and the CFPS (2010) provides the fraction of the village population that has the most popular surname. We see in Table 2, Column 5 that local population density has a positive and significant effect on this measure of social homogeneity. Clan affiliation is less relevant in the city, particularly among the urban born, which is why the surname information is not collected for urban residents in the CFPS.

Table 2. Social Structure and Population Density

Respondent's location:	city		county		
Dependent variable:	population density at neighborhood level	fraction of residents born locally (%)	population density at village level	fraction of residents born locally (%)	fraction of households from the largest clan (%)
	(1)	(2)	(3)	(4)	(5)
Population density at city/county level	3.266***	–	2.077***	–	–
	(0.993)		(0.548)		
Population density at neighborhood/village level	–	-7.750***	–	1.567	18.460*
		(2.591)		(2.109)	(9.608)
Mean of dependent variable	0.772	49.27	0.112	90.14	42.46
Observations	124	124	264	264	264

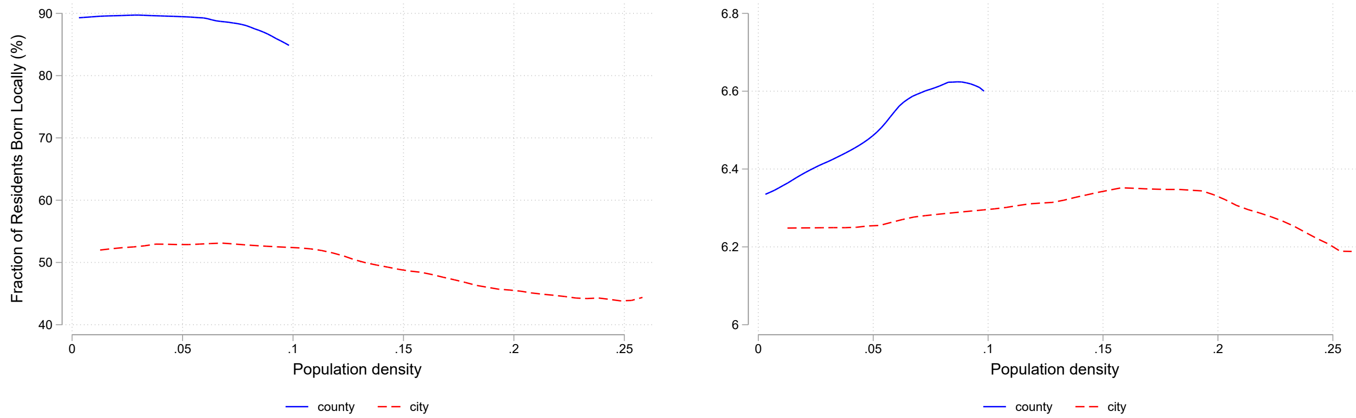
Source: China Family Panel Survey (2010).
Clan affiliation is measured by surname. Population density at neighborhood/village level is computed from the community module of the CFPS. Population density at city/county level is computed from the 1982 population census.
Population density is measured in units of 10,000 people per square km.
Robust standard errors are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Figure 2a nonparametrically estimates the relationship between the fraction of residents born locally and population density, separately in counties and cities. The additional information provided by the figure is that although the population density is greater in cities than in counties on average, there is substantial overlap in the range. The gap in the fraction of residents born locally, between counties and cities, is retained even in the region where the population densities overlap.

The preceding results suggest that trust or economic cooperation should be increasing in population density in counties, but not cities. The adult individual module of the CFPS (2012) collected information on trust, which proxies for economic cooperation. Trust is measured as an ordinal variable, taking values from 0 to 10.[12] We see in Table 3, Columns 1 and 3 that trust in local residents is increasing in population density for respondents residing in counties, but not in cities. Recall that the localized and enforceable trust that is relevant for our analysis is distinct from the generalized trust that has received much attention in the culture literature. The implicit assumption in our analysis is that generalized trust is uncorrelated with population density in counties and in cities, and this is indeed what we observe in Columns 2 and 4, with population

[12]The question on trust is designed to match the well known and frequently used question on trust in the World Values Survey (WVS). The WVS measures trust as an ordinal variable that takes values from 1 to 4. The level of trust in the WVS is assessed for the respondent's family, in his neighborhood, among people that the respondent knows personally, among people he meets for the first time, among people of another religion, and among people of another nationality. The CFPS measures trust in parents, neighbors, Americans, strangers, cadres, and doctors. Our analysis focuses on trust in neighbors or local residents and strangers; i.e. people the respondent meets for the first time (and who thus live outside the local area). These are the two categories that overlap with the WVS.

Figure 2. Fraction of Residents Born Locally, Trust and Population Density



(a) Fraction of Residents Born Locally **(b)** Trust in Local Residents

Source: China Family Panel Survey. Fraction of residents born locally is calculated from the adult individual module of China Family Panel Survey (2010) Trust measures are obtained from the adult individual module of China Family Panel Survey (2012).

density having no impact on trust in outsiders (which plausibly measures generalized trust).

Table 3. Trust and Population Density

Dependent variable:	trust in local residents	trust in outsiders	trust in local residents	trust in outsiders
Respondent's location:	county		city	
	(1)	(2)	(3)	(4)
Population density	0.225***	-0.018	-0.073	-0.014
	(0.054)	(0.079)	(0.056)	(0.064)
Mean of dependent variable	6.429	2.194	6.239	2.115
Observations	20,047	20,047	6,236	6,236

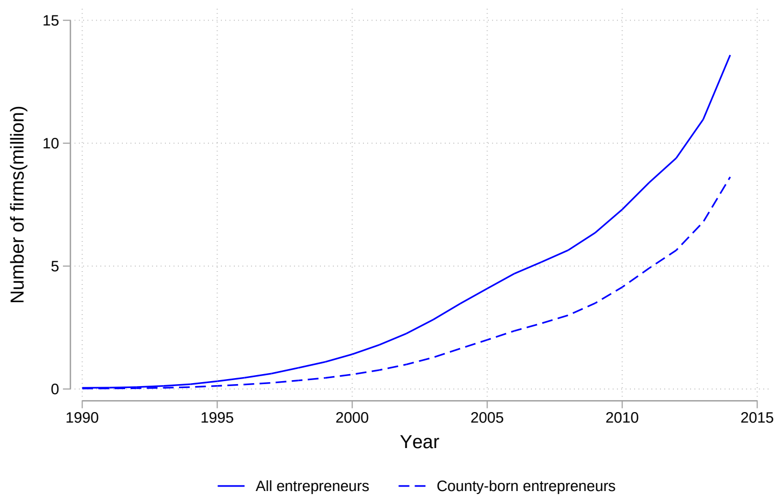
Source: Trust measures are obtained from the adult individual module of China Family Panel Survey (2012).
Trust is an ordinal variable, and takes value from 0 to 10.
Population density is measured in units of 10,000 people per square km, and then converted to Z-score.
Control variables include population, education and occupation distribution in the county or city.
Population is measured in millions and education is measured by the percent of the population that is literate.
Occupation distribution is measured by the share of workers in agriculture and industry with services the excluded category.
Population density, population, education and occupation distribution are computed from the 1982 population census.
Standard errors clustered at the county or city level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Figure 2b reports nonparametric estimates corresponding to Table 3, Columns 1-2, but without the additional covariates. Localized trust is increasing in population density in counties, but not in cities. This distinction is retained even in the region where the population densities overlap. Our analysis of community-based entrepreneurship, which uses population density as the source of forcing variation, is thus restricted to county-born entrepreneurs.[13] Figure 3 switches back to the SAIC data, describing the growth in the total number of private registered firms and in the number of firms owned by county-born entrepreneurs, respectively, over the 1990-2014 period. We see that county-born entrepreneurs made up about two-thirds of all entrepreneurs in China, with this ratio remaining quite stable over time. Firms owned by county-born entrepreneurs are just slightly smaller than the average registered firm (not reported). The contribution of these

[13]This does not imply that city-born entrepreneurs do not have access to networks; based on the evidence provided above, population density is just not a good indicator of social connectedness for them.

entrepreneurs, most of whom are first-generation businessmen, to Chinese economic growth has thus been substantial. Our objective in the analysis that follows is to identify and quantify the role played by birth county networks in supporting the explosive entry and subsequent growth of their firms.

Figure 3. Growth of Private Enterprise, by Birthplace of Entrepreneurs



Source: SAIC registration database.

2.3 Population Density and Entrepreneurship: A Canonical Model

Entrepreneurs in developing countries often rely on other entrepreneurs operating in the same sector and location for different forms of help. Chinese production clusters, for example, are characterized by a high degree of specialization by firms in specific stages of production, extensive exchange of intermediate goods, and flexible adjustment to workloads and product specifications in response to volatile market demand. Entrepreneurs also rely on each other for connections to buyers and suppliers (who often provide credit) as well as for information about new technologies and markets. Sustaining high productivity in this economic environment thus requires considerable mutual help. This is difficult to generate via market transactions, due to the inherent problem of verifying help sought and received, coupled with a weak legal environment. Cooperation is based instead on community norms, backed by social ties among the entrepreneurs in question (Nee and Opper, 2012). The key assumption in our analysis is that these social ties are based on the birth county, despite the fact that a majority of county-born entrepreneurs establish their firms elsewhere. Migrant entrepreneurs typically have close family members, such as aged parents, in the birth county, and visit it frequently. They (or close family members) thus continue to interact socially in the birth county. It follows that cooperation in the sector-location where firms are operating can be backed effectively by social sanctions imposed by the wider community in the birth county. The discussion that follows derives the relationship between population density in the birth county, network quality and, hence, the returns to entrepreneurship.

Let the size of the business network, which consists of entrepreneurs from the same origin who have already entered a particular sector-location in period t be denoted by n_t. The TFP of an entrepreneur with individual ability ω in this sector-location equals $f(\omega)A_t$ where f is an increasing function of ω and $A_t = A_0(1+h)^{n_{t-1}}$, with h denoting help provided by incumbent members of the network to one another. This specification captures the idea that help provided by network members is mutually complementary, which implies, in turn, that there are increasing returns with respect to the size of the network (for given per-member help). If the help provided by one network member to another is observed by other network members and entrepreneurs remain connected to their origin communities, then the maximal incentive compatible level of help (for given network size) will be increasing in population density and social homogeneity in the birth county. Given that social homogeneity, measured by clan concentration, is increasing in population density, it follows that the equilibrium level of help, $h(p)$, which reflects the 'quality' of the network, will be unambiguously increasing in the population density of the birth county, regardless of whether the network is drawn from the clan or the county as a whole. Hence $A_t = A_0(1 + h(p))^{n_{t-1}}$. Letting $\theta'(p)$ denote $\log(1 + h(p))$, this reduces to

$$A_t = A_0 \exp(\theta'(p) n_{t-1}) \tag{1}$$

We will use this expression for community TFP (CTFP thereafter) to derive testable predictions for business choices and outcomes when networks are active.

But before we do so, it is useful to note the difference between this specification and others more commonly used in the enodogenous growth and agglomeration literatures. In our formulation, networks are based on the social origins of entrepreneurs rather than the destinations they select: θ' reflects social connectedness in the birth county and n_t is the number of firms from that county operating in a given destination. In the standard agglomeration model, the θ' parameter would measure exogenous destination characteristics and n_t would be the total number of firms operating in that destination, irrespective of the social origin of their respective entrepreneurs. Ciccone and Hall (1996), for instance, use the number of workers per square km as a proxy for agglomeration effects in a given location. We will exploit this difference in the empirical analysis to distinguish between birth county network effects and agglomeration effects.

To derive the implications of our formulation for the dynamics of entry, consider a given origin county with social connectedness $\theta'(p)$, which is increasing in population density, p. There are equal sized cohorts of new agents born at successive dates $t = 1, 2...$ Each agent makes a once-and-for-all choice between a traditional occupation (such as farming or wage labor) and becoming an entrepreneur in the single sector-location that is available. Agents vary in individual ability ω, which is drawn independently from a log uniform distribution on the unit interval. The returns to entering the traditional occupation for an agent of ability ω is ω^σ where $0 < \sigma < 1$, while the return from entrepreneurship is $A_t\omega$, where CTFP A_t given by (1) depends on the incumbent stock n_{t-1}, i.e., the total number of entrepreneurs from previous cohorts from the same origin. The linear ability term in the production function ensures that there will be positive selection on ability into entrepreneurship. We will see below that the linearity is derived endogenously, and not by assumption, in the augmented model that includes capital in the production function. Higher ability agents thus become entrepreneurs, while low ability agents remain in the traditional occupation, with the precise threshold depending on the incumbent stock. The threshold ability, where returns to the occupations are equalized, is seen to be

$$\log \underline{\omega}_t = -\frac{1}{1-\sigma}[\log A_0 + \theta'(p) n_{t-1}] \tag{2}$$

We assume that the threshold lies in the interior of the unit interval. The model thus applies to the 'early industrialization' phase, which seems empirically relevant for the period we consider in China.[14] The fraction of cohort t agents that enter is then given by

$$e_t = 1 + \frac{1}{1-\sigma}[\log A_0 + \theta'(p)n_{t-1}] \tag{3}$$

which can be written as

$$e_t = B + \theta(p)n_{t-1} \tag{4}$$

with $B \equiv 1 + \frac{1}{1-\sigma}\log A_0 > 0$ and $\theta(p) \equiv \theta'(p)\frac{1}{1-\sigma}$.

Solving recursively, equation (4) yields the following expression for the evolution of entry flows as a function of initial entry n_0 and origin social connectedness, $\theta(p)$:

$$e_t = (B + n_0)(1 + \theta(p))^{t-1} \tag{5}$$

It follows from equation (5) that entry in any period t is increasing in initial entry. Moreover, the initial entry effect is increasing over time due to its compounding effect on future entry when networks are active. These dynamic increasing returns to initial entry result in an increase in entry over time (this is easy to verify because $\theta > 0$), which is consistent with the explosive growth in the stock of firms that we observed in Figure 3. At any point in time, the initial entry effect is increasing in $\theta(p)$ and, hence, in origin population density, p. Finally, the effect of the initial entry - population density interaction, which is positive, will grow stronger over time, once again due to the compounding network effect.[15]

2.4 Testing the Predictions of the Canonical Model

Although the model assumes that there is a single entrepreneur in each firm, in practice most registered firms consist of multiple shareholders. The SAIC database lists the major shareholders and managers, with their citizenship ID, in each registered private firm. The first six digits of the citizenship ID reveal the birth county of the individual.[16] We designate the firm's legal representative as the "entrepreneur" for the purpose of the empirical analysis and his birth county thus applies to the firm as a whole. This individual is legally responsible for the firm's liabilities and typically plays a key role in its functioning; for example, 75% of legal representatives are shareholders in their firms. Given the high degree of clustering by birth county within firms, our choice of the designated entrepreneur has little bearing on the analysis in any case. The legal representative and the largest shareholder belong to the same birth county in over 90% of firms. Even among the 58% of firms that are established outside the legal representative's birth county, as many as 74% of the listed individuals belong to his birth county.[17] This is substantially higher than the statistic that would be

[14] If $\log \underline{\omega}_t$ hits zero, then all agents will want to become entrepreneurs. Entry flows will subsequently plateau. We do not, however, observe such a slowing down in entry flows in the Chinese data in Figure 3.

[15] Our canonical model is closely related to Munshi's (2011) model of occupational choice with community networks. The important difference is that the source of forcing variation in his analysis is based on differences in outside options across communities, which is captured by the σ parameter in our model, whereas we exploit exogenous variation in social connectedness across communities, measured by the θ parameter.

[16] Citizenship ID's were first issued in September 1985 and people born after that date are given an ID at birth. Those born before that date were registered in the county or city where they resided at the time. Given the limited opportunities for labor migration in that period and the cost of moving due to the *hukou* system, almost all rural-born individuals resided in their birth-counties in 1985. The only exceptions were college students, college graduates, and soldiers, but these numbers were small. The first six digits of the citizenship ID thus reveals the county of birth, with few exceptions, even for those born before September 1985.

[17] Among the county-born legal representatives, 42% establish their firm in their birth county, 11% in their birth prefecture but outside the birth county, 18% in their birth province but outside the birth prefecture, and 29% outside their birth province.

obtained by random assignment of listed individuals in the firm's sector-location, which is just 6%, highlighting the role played by the birth county in supporting business.[18]

Table 4 reports tests of the canonical model, corresponding to (5), that estimate the effect of initial entry on subsequent entry within birth county-sector-locations. Although a single sector-location is available to all firms in the model, firms from a given birth county will in practice enter multiple sectors and multiple locations within these sectors. The implicit assumption when measuring entry within sector-locations is that birth county networks operate independently at that level. Sectors are measured at the two-digit level and locations are either counties (including the birth county) or urban districts.[19] Initial entry is measured in the 1990-1994 period, when private firms were first starting to emerge in China, and subsequent entry is measured separately in 2000-2004 and 2005-2009; the five-year window allows for sufficient entry flow at the birth county-sector-location level. The analysis is restricted to the 1990-2009 period because we will see that the decline in the ability threshold predicted by the model starts to weaken beyond that point in time. The benchmark specification in Table 4, Columns 1-2 includes, in addition, birth county-sector fixed effects and the total number of initial entrants in the sector-location from all origins to capture generalized location-based agglomeration.[20] We see that initial entry from the birth county has a positive and significant effect on subsequent entry, with this effect growing stronger over time, as predicted by the model. One additional initial entrant generates seven additional entrants in the 2000-2004 period and nine additional entrants in the 2005-2009 period.

Conditional on the number of initial entrants from the birth county, the total number of initial entrants in a given sector-location has no effect on subsequent entry from that birth county in that sector-location. This result provides empirical support for the key assumption in the model that the birth county is the domain within business networks are organized in China and that these networks operate independently. It also provides support for the assumption that individual networks cannot influence the price of the product. If the members of the network could collude (depending on their market share) or there were limits to market size, then the total number of entrants, conditional on the number of entrants from the birth county, would also be relevant.[21]

Although we assume that network spillovers increase the productivity of their members, an alternative explanation for clustering is that the networks capture economic rents. Bai *et al.* (2019), for example, describe how favored firms have superior access to capital allocated by local governments. If local government officials

[18]The listed individuals must exert effort and contribute in different ways to the firm, and the moral hazard problem that applies to the provision of help between firms will also be relevant within the firm. Using the same logic as above, cooperation within the firm will be backed by social sanctions in the origin. More effective social sanctions will, in turn, be associated with greater representation by individuals from the origin in the firm. Appendix Table E.1, Column 1 reports the estimated relationship between the fraction of listed individuals from the legal representative's birth county and its population density. The coefficient on population density is positive and significant, as expected. Column 2 reports the same relationship for legal representatives born in cities where, in contrast, the coefficient on population density is negative and significant.

[19]There are 3,235 counties or urban districts where firms locate in our data. We measure urban population density at the city rather than the dissaggregated urban district level in our analysis because urban districts were created after 1982. However, firm location is always measured at the urban district level.

[20]All locations which had a positive number of entrants by 2000-2004 and 2005-2009, respectively, for a given birth county-sector are included in the estimating equation.

[21]Appendix Table E.2 replicates Table 4 restricting attention to locations outside the birth county. Although the coefficient on the total number of initial entrants is now statistically significant, it remains positive and an order of magnitude smaller than the coefficient on the number of entrants from the birth county. If all firms from the same origin collude, but there is competition across (origin-based) networks at the destination, then we would expect to see a negative effect of entry from other origins. While pricing may be non-competitive in China (see, for example, Brooks *et al.* (2016)), the origin-based networks do not appear to be directly associated with these distortions.

Table 4. The Effect of Initial Entry on Subsequent Entry (birth place level)

Dependent variable:	subsequent entrants from the birth place					
Birth place:	county		county		city	
Time period:	2000-2004	2005-2009	2000-2004	2005-2009	2000-2004	2005-2009
	(1)	(2)	(3)	(4)	(5)	(6)
Initial entrants from the birth place	7.120***	8.935***	5.198***	5.723***	7.385***	6.668***
	(0.686)	(0.956)	(0.972)	(1.281)	(0.750)	(0.799)
All initial entrants at the destination	0.054	-0.020	–	–	–	
	(0.048)	(0.056)				
Initial entrants from the birth place × birth place population density	–	–	1.356**	2.277**	-0.016	-0.261**
			(0.564)	(0.937)	(0.107)	(0.113)
Distance to the birth place	–	–	-2.225***	-2.962***	-3.369***	-3.145***
			(0.104)	(0.106)	(0.229)	(0.162)
Mean of dependent variable	3.156	3.170	3.156	3.170	4.148	3.579
Origin-sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Location fixed effects	No	No	Yes	Yes	Yes	Yes
Observations	384,031	778,897	384,031	778,897	313,915	446,696

Note: number of entrants is measured at the birth place- sector-destination level.
Initial entry is computed over the 1990-1994 period and sectors are defined at the two-digit level.
Number of entrants is obtained from the SAIC registration database and birth place population density is computed from the 1982 population census.
Standard errors clustered at birth place-sector level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

favor entrepreneurs from their hometown, then this would explain why firms from a given birth county cluster at the same location.[22] However, while pre-existing social ties might confer a temporary advantage, resulting in additional entry, it will not persist because government officials are frequently rotated precisely to avoid such corruption. Based on data from city yearbooks and the online CV's of government officials, we estimate that the median (mean) term-length for a city mayor in China is 4 (3.7) years. Government favoritism based on pre-existing social ties, thus cannot explain the long-term effect of initial entry on subsequent entry that we observe in Table 4.

A related mechanism that might explain the inter-temporal correlation is based on the idea that stronger networks are more effective at lobbying the government official who is in place. A testable implication of network competition for scarce government credit, and resources more generally, is that greater entry from other origins into a particular location will reduce entry from a given origin into that location. We find no evidence of such negative cross-community spillover effects. While particular entrepreneurs may have favored access to government credit, the birth county networks do not appear to be facilitating this process. What they do, instead, is to mobilize capital internally. Recently available data from the Enterprise Survey for Innovation and Entrepreneurship in China (ESIEC) which uses the registration database as the sampling frame for a subset of firms indicates that initial registered capital is obtained from the following sources: self finance (76%), the owners' social network (15%), and bank loans (9%). More importantly, and in line with the productivity enhancing role of the network that is assumed in the model, 55% (62%) of the firms report that their largest stable supplier (buyer) either belongs to their social network or was referred by a member of the network.[23]

[22]Fisman *et al.* (2018) document the same type of favoritism, based on hometown ties, in the Chinese Academy of Sciences.
[23]The social network is defined to include individuals with pre-existing social ties to the entrepreneur: relatives, *laoxiang*

Although the model assumes that initial entry is exogenously determined, entrepreneurs will select particular sectors and locations in practice. If initial entrants were more likely to select into persistently favorable destinations, then a positive correlation between initial entry and subsequent entry could be obtained without requiring networks to be active. Alternatively, if entrepreneurs tend to locate at proximate destinations, then a spurious correlation could once more be obtained (because unobserved proximity is fixed over time). We account for these possibilities in Table 4, Columns 3-4 by adding location fixed effects and the distance between the birth county and the location under consideration in the estimating equation. Location fixed effects can be included because entrepreneurs from multiple origins will establish their firms in the same place, but the variable measuring initial entry from all origins will now be subsumed in the fixed effects. The additional controls will account for a wide class of omitted variables; what remains is the possibility that some factor other than distance ties birth counties to specific sector-locations. We address this possibility by testing the additional prediction of the model, based on the interaction of birth place population density and initial entry. Table 4, Columns 3-4 reports estimation results with the augmented specification, where we see that the interaction coefficient is positive and significant, and increasing from 2000-2004 to 2005-2009, as predicted by the network model. As a placebo test, we estimate the same equation in Columns 5-6 with city-born entrepreneurs. The major change in the estimated coefficients, when compared with the results obtained with county-born entrepreneurs, is that the interaction coefficient is now negative (and significant with 2005-2009 entry as the dependent variable).

We remain agnostic about the domain of the network in our analysis; i.e. whether it covers the entire birth county or the clan within the county. Given that clan concentration is increasing in population density in Chinese counties, the predictions of the model can be tested at the county level even if networks are organized within clans. Having tested the model at the county level, we now go down to the clan level to directly ascertain the domain of the network. As with the analysis using CFPS data, we focus on the county-born entrepreneurs and infer their clan affiliation from their birth county and surname. The predictions of the model are tested at the clan level by estimating the effect of initial entry in the 1990-1994 period on subsequent entry, separately in 2000-2004 and 2005-2009, within clan-sector-location. The estimating equation includes birth county-sector fixed effects as well as the total number of initial entrants from the birth county into that sector-location (to allow for the possibility that network spillovers extend beyond the clan to others from the same birth county). The coefficient on the number of initial entrants, at the level of the clan and the county, is positive and significant in Table 5, Columns 1-2, although the former is an order of magnitude larger. As predicted, the initial effects grow larger over time; i.e. from 2000-2004 to 2005-2009. Columns 3-4 interact initial entry with the population density of the birth county. The initial effects are larger for higher population density birth counties, at the level of the clan and the county, in 2000-2004 and in 2005-2009, and increasing over time, once again as predicted by the model.

The results in Table 5 indicate that spillovers are concentrated within the clan, although they do extend to a limited extent beyond its boundaries to other entrepreneurs in the same county from other clans. Assuming that the spillovers operate exclusively within the clan, our model generates independent predictions for the

(homeplace fellows), classmates, friends, army comrades, colleagues, and past business partners. These categories are not exclusive; in particular, relatives, classmates, and many individuals from the other categories will be *laoxiang*.

Table 5. The Effect of Initial Entry on Subsequent Entry (clan level)

Dependent variable:	subsequent entrants from the clan			
Time period:	2000-2004	2005-2009	2000-2004	2005-2009
	(1)	(2)	(3)	(4)
Initial entrants from the clan	3.275***	4.082***	2.422***	2.689***
	(0.325)	(0.353)	(0.304)	(0.305)
All initial entrants at the destination from the birth county	0.031***	0.051***	0.012	0.006
	(0.007)	(0.012)	(0.013)	(0.017)
Initial entrants from the clan × birth county population density	–	–	0.650**	1.092***
			(0.258)	(0.331)
Initial entrants from the birth county × birth county population density	–	–	0.012*	0.029**
			(0.007)	(0.013)
Distance to the birth county	–	–	-0.174***	-0.259***
			(0.012)	(0.015)
Mean of dependent variable	1.381	1.397	1.381	1.397
Origin-sector fixed effects	Yes	Yes	Yes	Yes
Location fixed effects	No	No	Yes	Yes
Observations	866,344	1,723,509	866,344	1,723,509

Note: number of entrants is measured at the clan-sector-location level. Clan is defined by the same surname and the same birth county. Initial entry is computed over the 1990-1994 period and sectors are defined at the two-digit level.
Number of entrants is obtained from the SAIC registration database and birth place population density is computed from the 1982 population census.
Standard errors clustered at birth place-sector level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

concentration of business activity within clans drawn from the same birth county, across counties with different population densities and over time. As discussed above, the total expected sanction and, hence, the level of cooperation in clan j can be expressed as the product of two functions: $\nu_1(p)$, which is increasing in population density p, and $\nu_2(\zeta_j)$, which is increasing in the relative size of the clan. When entrepreneurs remain connected to their birth counties, this formulation applies to the incentive compatible level of help that can be supported in the business network. The θ parameter, which measures this help, can thus be approximated by the product of two functions when it is measured at the level of the clan: $\theta_1(p)$ and $\theta_2(\zeta_j)$.[24] Now suppose that there are two clans in the county: a majority clan, accounting for a fraction $M > \frac{1}{2}$ of the population, and a minority clan accounting for a fraction $m \equiv 1 - M$ of the population. With two clans, the clan concentration among entering entrepreneurs (measured by the Herfindahl Hirschman Index) will be locally monotonically increasing in the entry share of the majority clan or, equivalently, in the ratio of its share to the share of the smaller clan. From equation (5):

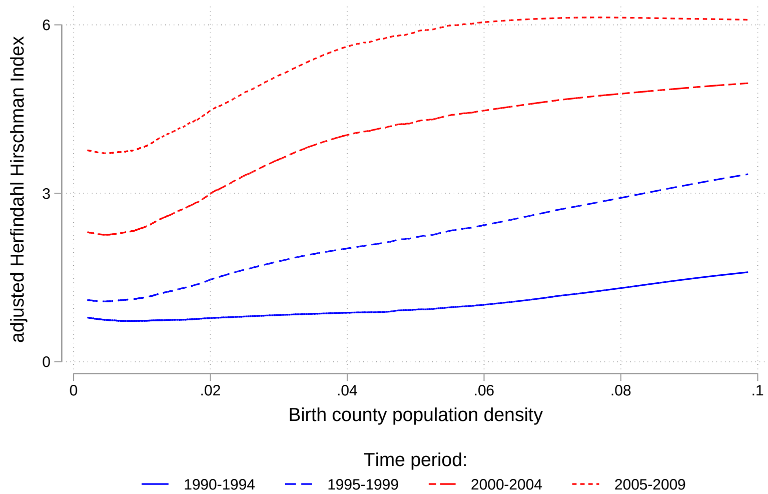
$$\frac{e_{Mt}}{e_{mt}} = \left[\frac{1+\theta_1(p)\theta_2(M(p))}{1+\theta_1(p)\theta_2(m(p))}\right]^{t-1} \tag{6}$$

We know from the analysis with CFPS data that clan concentration is increasing in p. Hence, $M(p)$ is increasing in p, while $m(p)$ is decreasing in p. We can then infer the following from equation (6): (i) Clan concentration among entering entrepreneurs is increasing over time (because the term in square brackets is greater than one). (ii) Clan concentration is increasing in population density at each point in time. (iii) The

[24]The reason why we could express the θ parameter, at the level of the county, as a function of p alone is because clan concentration is increasing in p and, hence, $\theta_2(\zeta)$ aggregated to the county level is increasing in p.

positive relationship between clan concentration and population density will grow stronger over time.

Figure 4. Clan Concentration and Population Density



Source: SAIC Registration database and 1982 population census.
Clan concentration measured by the Herfindahl Hirschman Index (HHI) across surnames of entering entrepreneurs from a given birth county, divided by the expected HHI that would be obtained by random assignment, given the number of entrants and the number of surnames at each point in time.

Table 6. Clan Concentration and Population Density

Dependent variable:	adjusted HHI across clans			
Time period:	1990-1994	1995-1999	2000-2004	2005-2009
	(1)	(2)	(3)	(4)
Birth county population density	0.072***	0.367***	0.583***	0.596***
	(0.013)	(0.029)	(0.042)	(0.054)
Mean of dependent variable	0.830	1.660	3.291	4.790
Observations	1,610	1,624	1,624	1,624

Note: clan is defined by the same surname and the same birth county.
Clan concentration measured by the Herfindahl Hirschman Index (HHI) across surnames of entering entrepreneurs from a given birth county, divided by the expected HHI that would be obtained by random assignment, given the number of surnames (clans) and the number of entrants from the birth county at each point in time.
Population density is measured in units of 10,000 people per square km, and then converted to Z-score.
Control variables include population, education and occupation distribution in the county.
Population density, population, education and occupation distribution are computed from the 1982 population census.
Robust standard errors are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Figure 4 reports the relationship between the clan concentration of entering entrepreneurs, measured by the Herfindahl Hirschman Index (HHI), and population density in their birth county at each point in time. The HHI statistic is adjusted for the fact that measured concentration could vary with the number of entrepreneurs and the number of clans just by chance, using a normalization derived in Appendix A.[25] We will use the same adjustment for all the analyses of concentration that follow. Clan concentration among entering entrepreneurs is increasing over time and increasing in birth county population density at each point in time, precisely as

[25]Previous attempts to examine the spatial distribution of production; e.g. Ellison and Glaeser (1997), Duranton and Overman (2005), have also taken account of this feature of all concentration statistics.

our model of within-clan-network-based cooperation would predict. Although it is difficult to visually assess whether the relationship is growing stronger over time, we see in Table 6, which reports parametric estimates corresponding to Figure 4 in each time period, that the coefficient on population density is positive and significant in each time period and increasing over time.[26]

3 The Augmented Model: with Sector-Location Choice and Capital Investment

In this section we extend the canonical model to incorporate sector-location choice and capital investment. One advantage of the augmented model is that it generates predictions for additional outcomes; the distribution of firms across sectors and across locations within sectors, as well as firm size. A second advantage is that once the structural parameters of the augmented model are estimated, it can be used to quantify the impact of the origin-community based networks on both aggregate firm entry and the stock of capital at a critical stage in the growth of the Chinese economy. While the predictions of the canonical model were derived with respect to initial capital, to emphasize the inter-temporal link between firms from the same birth county, we now derive robust (but less direct) predictions for changes in outcomes over time and, more importantly, for changes in the relationship between birth county population density and these outcomes over time, when community networks are active.

There are multiple destinations $B_i, i = 1, 2..$ for entrepreneurship. The destinations denote sectors or locations. For simplicity we assume these destinations are *ex ante* symmetric, except for entry at the initial date. In destination B_i at date t, an entrepreneur with ability ω selects capital size K, and has a production function

$$y = A_{it}\omega^{1-\alpha}K^{\alpha} \tag{7}$$

where $\alpha \in (0, 1)$ is the capital elasticity, and A_{it} denotes CTFP in destination i which takes the form (as explained in the previous section):[27]

$$A_{it} = A_0 \exp(\theta(p)n_{i,t-1}) \tag{8}$$

where $n_{i,t-1}$ denotes incumbent stock from the origin or birth county in destination i at the end of $t - 1$. As before, social connectedness $\theta(p)$ is increasing in population density in the birth county.

The dependence of CTFP on the size of the incumbent stock represents one source of network complementarity, reflecting gains from intra-network cooperation in improving productivity for those who have already entered sector i. We add to this a second source of network complementarity, which pertains to 'referrals' or 'access' to particular business sectors or locations. A fixed fraction $k \in (0, 1)$ of new agents in every cohort receive an opportunity to become an entrepreneur. Within this group of 'potential entrants', the fraction that get an opportunity to enter destination B_i equals $s_{i,t-1}$, the share of incumbent entrepreneurs from the origin community already in that destination. This reflects the formation of aspirations, access to information, or referrals provided by older members from the same origin community in a given destination.

[26]Pooling data from all time periods, we see in Appendix Table E.3 that the time period coefficient and the population density-time period coefficient are both positive and significant. This result is robust to restricting the sample to firms established outside their entrepreneur's birth county.

[27]The A_0 term incorporates the product price and labor productivity. Labor is not included as a variable input in the production function because it is not observed in our data. With the Cobb-Douglas specification of the production function, the optimal labor input can be derived as a function of the model's parameters and is subsumed in the A_0 term.

Apart from the decision of whether or not to enter a given destination when presented with the opportunity, an agent decides on how much capital to invest. All agents incur the same cost of capital r which is exogenous and fixed across all t and all origins. We are thus abstracting from possible network complementarities operating via internal capital markets, as in Banerjee and Munshi (2004), which arise in response to financial market imperfections. To the extent that larger and higher quality networks lower borrowing costs for their members, the resulting dynamics turn out to be very similar to those generated via productivity spillovers, and would thus amplify the dynamics generated by the latter alone.[28] We also assume a fixed price of the product, unaffected by supply from the network. This abstracts from price collusion among network members, as well as limits to market size in a competitive context. These seem plausible in the Chinese setting, where most sectors are comprised of a large number of origin county networks, and both domestic and international market opportunities are large.[29]

3.1 Occupational Choice

To determine occupational choice, we first calculate the profits a new agent in any cohort with a given ability ω expects to earn upon entering a given business destination (sector and location) when the CTFP in that destination is expected to be A. The latter is a sufficient statistic for the specific date, destination in question, and existing network size and quality (which determine CTFP as per (8)). The optimal capital size K must maximize $A\omega^{1-\alpha}K^{\alpha} - rK$, and thus satisfies:

$$\log K(\omega, A) = \log \omega + \log \phi + \frac{1}{1-\alpha} \log A - \frac{1}{1-\alpha} \log r \tag{9}$$

(where $\phi \equiv \alpha^{\frac{1}{1-\alpha}}$). The resulting profit satisfies

$$\log \Pi(\omega, A) = \log \omega + \log \psi + \frac{1}{1-\alpha} \log A - \frac{\alpha}{1-\alpha} \log r \tag{10}$$

(where $\psi \equiv \phi^{\alpha} - \phi$).[30]

Of the new agents receiving an offer, the ones that will decide to enter business are those who receive a higher profit in that destination compared to the traditional occupation.[31] These agents will be endowed with a level of ability that exceeds a threshold $\underline{\omega}$:

$$\log \omega > \log \underline{\omega} \equiv \frac{1}{1-\sigma}[\log \frac{1}{\psi} - \frac{1}{1-\alpha} \log A + \frac{\alpha}{1-\alpha} \log r] \tag{11}$$

As with the canonical model, we assume that the threshold lies in the interior of the ability distribution at the beginning of the process for each destination, and we will restrict attention to 'early phases of industrialization' when this continues to be true.

[28] We ignore the role of labor networks in the model. The owner of the firm and the workers rarely belong to the same community, even in network-based economies. The historical and contemporary experience, across the world, indicates that incumbent workers (with a reputation to maintain within their firms) are the primary source of job referrals.

[29] Based on the registration data, firms from a given origin county account for 13% of firms at the destinations where they locate, on average (within narrow two-digit sectors). This statistic is based on all entrepreneurs, including those who locate their firms in their county of birth. This limited market power explains, in part, why we failed to observe negative cross-community entry effects when testing the canonical model.

[30] If we allowed for credit networks organized around the origin county and parameterized the interest rate as $r = r_0 exp(-\eta(p)n_{i,t-1})$, then the productivity channel operating through the A term and the credit network channel would not be separately identified. Although the model is set up so that networks operate through the productivity channel, all the results that follow would go through if, instead, they operated through the credit channel.

[31] Recall that the profit in the traditional occupation was specified to be ω^{σ}, where $\sigma \in (0,1)$, in the canonical model. We retain this specification in the augmented model.

Notice that agents receiving an entrepreneurial opportunity make their decision *selfishly* and *myopically*. The former assumption implies that they ignore the consequences of their entry decisions on the profits of other agents. The latter states that they make their choice solely to maximize their date–t profits, ignoring consequences at later dates. This enables us to compute the entry dynamics recursively, simplifying the analysis considerably. If agents were more far-sighted, they would have to forecast current and future levels of entry from the same origin county, generating strategic complementarity of entry decisions within each cohort. This extension is considered in Appendix B, where entry decisions at t are based on the discounted sum of profits at t and $t+1$, rather than t alone. We show there under some natural conditions that a unique rational expectations equilibrium exists, whose comparative statics are similar to those in the simpler myopic model. If anything, the myopic model generates a conservative bias in entry decisions. This is because a network's size cannot ever decrease over time and its quality does not change, and neither do profits in the traditional sector. Those deciding to enter based on a myopic calculation would also want to enter if farsighted, while some others deciding to stay out on myopic grounds may wish to enter when they anticipate future network growth, which would further raise the returns to entrepreneurship.

3.2 Dynamics of Entry and Concentration

The different business destinations have identical 'fundamentals'. At the beginning of the process ($t = 0$), there is a small, exogenous number n_{i0} of older entrepreneurs (from cohorts preceding $t = 1$) who have already entered B_i. These represent the initial conditions for the dynamics. These historical entry levels will generically not be exactly balanced across destinations; without loss of generality suppose $n_{i0} > n_{i-1,0} > 0$ for all i. We show first that the initial imbalance across destinations will cumulate thereafter, with entrants in later cohorts increasingly locked-in to the destinations with higher initial presence.

To derive entry in subsequent cohorts, we start with the threshold condition (11), which determines the measure of agents from cohort t who would choose to enter destination B_i if they had the opportunity. Combining this with the fraction $ks_{i,t-1}$ of those agents that have an opportunity to enter, we derive the volume of entry e_{it} in cohort t into B_i as a function of the state variables $n_{i,t-1}, s_{i,t-1}$:

$$e_{it} = ks_{i,t-1}[B + C\theta(p)n_{i,t-1}]$$

where $B \equiv 1 - \frac{1}{1-\sigma}\log\frac{1}{\psi} - \frac{\alpha}{(1-\sigma)(1-\alpha)}\log r + \frac{1}{(1-\sigma)(1-\alpha)}\log A_0$ and $C \equiv \frac{1}{(1-\sigma)(1-\alpha)}$. This expression reduces further to

$$e_{it} = Ls_{i,t-1} + \kappa(p)N_{t-1}s_{i,t-1}^2 \tag{12}$$

where L denotes kB; $\kappa(p)$ denotes $Ck\theta(p)$ which is rising in p, and $N_{t-1} \equiv \sum_i n_{i,t-1}$ denotes the aggregate number of business entrepreneurs from past cohorts from the same origin. Aggregating (12) across sectors, we obtain an expression for the dynamics of aggregate entry:

$$N_t - N_{t-1} \equiv E_t \equiv \sum_i e_{it} = L + \kappa(p)N_{t-1}H_{t-1} \tag{13}$$

where $H_{t-1} \equiv \sum_i s_{i,t-1}^2$ denotes the Herfindahl Hirschman Index for concentration at $t-1$. Equations (12,13) define the dynamics of the vector $(N_t, s_{it}, i = 1, 2..)$, where $s_{it} \equiv s_{i,t-1}\frac{N_{t-1}}{N_t} + \frac{e_{it}}{N_t}$.

Proposition 1 *Concentration H_t and aggregate entry flow E_t are both rising in t.*

The proofs of this proposition, as well as subsequent propositions, are provided in Appendix C. The intuitive reason for concentration to rise over time is simple: a destination with higher incumbent stock is both more profitable and generates more opportunities for entry, so its share grows faster, reinforcing the higher initial presence. Note that the compounding effect of initial entry, due to network spillovers, which resulted in an increase in entry over time in the canonical model, is present here as well. The network complementarity associated with post-entry productivity spillovers, embodied in the N_{t-1} term in (13), is now reinforced by the additional network complementarity associated with the referrals; i.e. the H_{t-1} term. Entry E_t will rise over time from (13) if concentration is increasing over time.

The compounding network effect is stronger for firms from higher p origins, on account of the $\kappa(p)$ multiplier, so one would also expect the level of concentration and entry, and their growth over time, to be rising in p. Verifying this conjecture is more complicated, however, especially with respect to concentration. To illustrate this, consider the case of two destinations $i = 1, 2$, whereupon concentration is monotone increasing in the share of the destination with a higher initial presence. We can then focus on the dynamics of the share of this initially dominant destination, which (without loss of generality) we denote by destination 1:

$$\begin{aligned} s_{1t} \equiv [\frac{N_t}{n_{1t}}]^{-1} &= [\frac{L + N_{t-1} + \kappa(p)N_{t-1}H_{t-1}}{Ls_{1,t-1} + n_{1,t-1} + \kappa(p)N_{t-1}s^2_{1,t-1}}]^{-1} \\ &= [1 + (\frac{1}{s_{1,t-1}} - 1)\{\frac{L + N_{t-1} + \kappa(p)N_{t-1}(1 - s_{1,t-1})}{L + N_{t-1} + \kappa(p)N_{t-1}s_{1,t-1}}\}]^{-1} \end{aligned} \tag{14}$$

Proposition 2 *With two destinations:*

(a) Entry E_t and concentration H_t are rising in p (at any given t).

(b) $E_t - E_{t-1}$ and $H_t - H_{t-1}$ are both rising in p, if $\kappa(p) < 1$ for all p and the share of the larger sector at $t - 1$ is not too close to 1 (e.g., below $\frac{3}{4}$).

Part (a) confirms that the level of concentration is rising in p at any t, which in turn implies the same for entry flows. Part (b) shows that growth of entry or concentration is rising in p at 'early stages' of the industrialization process; i.e. when concentration is not too high. The qualifier is required because the share of the dominant sector B_1 is bounded above by 1. Thus, the share of the dominant destination cannot be increasing faster forever; eventually, its growth rate will flatten out as the share approaches one.

The results concerning the dynamics of concentration across destinations translate into testable predictions concerning either sectoral or spatial concentration, given that destinations correspond to sectors or locations. We partition the set of destinations into sectors, with each sector consisting of a subset of locations. Proposition 1 can then be extended to show that spatial (location) concentration within any given sector must be rising in t. The same can be shown for sectoral concentration, provided that sectors with higher initial shares are also characterized by higher intra-sectoral spatial concentration at date 0.[32] Moreover, with two locations within any sector, the results on concentration in Proposition 2 apply across sectors or across locations within sectors.

[32]The reason is that the expression for entry flow into sector c is modified to $e_{ct} = \kappa Ls_{c,t-1} + \kappa N_{t-1}s^2_{c,t-1}H_{c,t-1}$, so the term involving the quadratic term in lagged sectoral share is weighted by lagged intra-sectoral spatial concentration $H_{c,t-1}$.

3.3 Ability Selection and Firm Size Dynamics

Next we derive predictions concerning entrepreneurial ability and firm size. We first show that network effects generate negative selection on ability: from (11), as CTFP increases over time, the threshold for entry falls, and entrepreneurs with lower ability start entering. This negative selection has consequences for firm size. Substituting from (11) in (9), initial capital of the marginal entrant is decreasing in CTFP, A_{it}:

$$\log K_{it}^{m} = U' - \frac{\sigma}{(1-\sigma)(1-\alpha)} \log A_{it} \tag{15}$$

where $U' \equiv \log\phi - \frac{1}{1-\sigma}\log\psi - \frac{1}{1-\alpha}\log r$, and $\log A_{it} = \log A_0 + \theta(p)N_{t-1}s_{i,t-1}$. The negative selection on individual ability that accompanies a stronger network (with higher CTFP) outweighs its productivity benefit. The marginal entrepreneurs that enter are thus less productive and enter with smaller firm sizes.[33] The same argument applies to comparisons at any given t across different p origins: marginal entrants from higher p origins, with stronger networks, enter with smaller firm sizes. If $\sigma \in (\frac{1}{2}, 1)$ this is true also for the average entrant: firms from high p origins enter with smaller initial capital on average, with the opposite result holding if $\sigma < \frac{1}{2}$.[34] To see this, observe that substituting from (11) in (9), the capital size of the average entrant satisfies:

$$\log K_{it}^{a} = W + \frac{1-2\sigma}{2(1-\alpha)(1-\sigma)} \log A_{it} \tag{16}$$

where $W \equiv \log\phi + \frac{1}{2} + \frac{1}{2(1-\sigma)}\log\frac{1}{\psi} - \frac{2-\alpha-2\sigma}{2(1-\alpha)(1-\sigma)}\log r$. All firms face the same cost of capital and there are no mark-ups in our model. The preponderance of small and seemingly unproductive firms often noted in developing countries, which is typically attributed to wedges in factor prices and mark-ups in output price in the misallocation literature, may instead just be a manifestation of strong network effects! Our model implies that their own productivity understates their contribution via spillovers to their network.

In contrast to the results for initial capital, post-entry growth rates of firm size for any given cohort can be shown to be rising in p and over time. (9) implies that the capital at date $t' > t$ of a cohort t entrepreneur with ability ω is given by

$$\log K_{itt'}^{a} = \log\omega + \log\phi - \frac{1}{1-\alpha}\log r + \frac{1}{1-\alpha}[\log A_{it} + \theta(p)\sum_{l=t}^{t'-1} e_{il}] \tag{17}$$

implying a growth rate at period t':

$$\log K_{it,t'}^{a} - \log K_{it,t'-1}^{a} = \frac{1}{1-\alpha}\theta(p)e_{it'} \tag{18}$$

In our model, growth in incumbent firm size is independent of the entrepreneur's ability and cohort and is driven entirely by contemporaneous changes in CTFP. Once we average across destinations, these changes in CTFP are measured by aggregate entry flows of firms from the origin.

Proposition 3 *With two destinations:*

[33] This result does not depend on assumptions concerning the distribution of ability. To see this, observe that expressions (9, 10) show that capital size and entrepreneurial profit depend on individual ability and CTFP in exactly the same way. The marginal entrepreneur must be of lower ability when CTFP is higher, and must be indifferent between the traditional occupation and entrepreneurship. Profits will thus be lower in the traditional occupation for an agent with lower ability. So the same is true for entrepreneurial profit, and hence for capital size.

[34] This depends on the assumption of a log uniform distribution of ability.

(a) Averaging across destinations, ability and initial capital of marginal entrants (also of average entrants if $\sigma > \frac{1}{2}$) is decreasing in t (for any given p) and in p (for any given t), and decreasing more steeply in p across successive cohorts.

(b) Averaging across destinations, the growth rate of capital of incumbent entrepreneurs of any past cohort t from $t' - 1(> t)$ to t' is rising in t' and in p (more steeply over time).

From (11) and (15), the marginal entrant's ability and initial capital are decreasing in $\log A_{it}$. From (16), the average entrant's ability and initial capital are also decreasing in $\log A_{it}$ if $\sigma > \frac{1}{2}$. $\log A_{it} \equiv \log A_0 + \theta(p)n_{i,t-1}$ is increasing in N_{t-1} when it is averaged across destinations. From Propositions 1 and 2 we know that E_t and, hence, N_t is increasing in t (for any p), increasing in p (for any t), and increasing more steeply in p over time. Hence, part (a) of Proposition 3 follows immediately. A similar argument can be used to establish part (b). Averaging across destinations, $e_{it'}$ is replaced by $E_{t'}$ on the right hand side of equation (18). The result then follows from Propositions 1 and 2. Firms from high-p origins start smaller, but subsequently grow faster.[35] This dual prediction will be especially helpful in distinguishing our network-based model from alternative models, as discussed in the next section.

3.4 Alternative Explanations

To what extent do the preceding results rely on network spillovers? Could they be obtained, instead, by relaxing different assumptions of our model, while shutting down the network component? These questions are relevant because although population density is plausibly associated with social connectedness in the birth county and, hence, network quality, it could also be correlated with other factors that independently determine the dynamics of entry, concentration, and firm size. The discussion that follows systematically examines this possibility by introducing new sources of (possibly time-varying) heterogeneity at the origin, which are, in turn, correlated with population density, and by allowing firms from different origins to have favorable access to destinations of varying quality. Our model treats sectors and locations interchangeably. Because locational heterogeneity is an important alternative that we must consider, entrepreneurs choose between locations (which we refer to as destinations for expositional convenience) rather than sectors in the models that follow.

3.4.1 Origin Heterogeneity

Our model assumes that the stock of potential entrepreneurs, k, is constant across origin counties and cohorts. Suppose that we relax this assumption and let $k(p,t)$ be a twice differentiable function satisfying $k_p > 0, k_t > 0, k_{pt} > 0$. This could be because higher population density counties simply have larger populations that are growing relatively fast over time or because their residents have greater wealth or preferred access to finance, which facilitate entry into business. An additional source of origin heterogeneity could be in payoffs in the traditional occupation across counties. Our model assumes that the payoff, ω^σ, where ω is individual ability,

[35] In a related paper, Banerjee and Munshi (2004) find that outsiders in Tirupur's garment cluster, who face a higher cost of capital because they have weaker local credit networks, start with smaller firms and then grow faster (because they are positively selected on ability). To explain Banerjee and Munshi's findings, our model would need to be augmented to allow firm growth to be increasing in the entrepreneur's ability, with the additional condition that the ability effect needs to dominate the network effect (which is stronger for the insiders).

is the same in all counties and constant over time. However, the payoff could be lower in higher population density counties because there is a larger population for a given amount of resources (such as agricultural land). It is also possible that this population pressure is increasing over time. We allow for this possibility by representing the payoff in the traditional sector by $\omega^\sigma v(p,t)$, where $v(p,t)$ is a twice differentiable function satisfying $v_p < 0, v_t < 0, v_{pt} < 0$.[36]

Abstract for the time being from heterogeneity across destinations, so the TFP of any entrepreneur with ability ω at any destination at time t is $\omega^{1-\alpha}A_t$, with A_t growing exogenously over time. Moreover, an exogenous share s_i of potential entrepreneurs at each origin have the opportunity to enter any given destination. Owing to the absence of network effects, neither A_t nor s_i depend on p.

The ability threshold for entry into destination i from an origin with population density p in this alternative model would equal

$$\log \underline{\omega}_i(p;t) = \frac{1}{1-\sigma}\left[\log\frac{1}{\psi} + \frac{\alpha}{1-\alpha}\log r - \frac{1}{1-\alpha}\log A_t + \log v(p,t)\right] \tag{19}$$

while the expression for entry flows is:

$$e_i(p,t) = s_i k(p,t)\left[1 + Z + \frac{1}{(1-\alpha)(1-\sigma)}\log A_t - \frac{1}{1-\sigma}\log v(p,t)\right] \tag{20}$$

where $Z \equiv \frac{1}{1-\sigma}\log\psi + \frac{\alpha}{(1-\alpha)(1-\sigma)}\log r$. The entry flows will be rising in p and in t, and the slope with respect to p will be rising in t. The alternative model can thus generate our model's predictions for entry. However, the share of different destinations will be constant and independent of p. In order to obtain the same predictions for spatial concentration generated by the network model, the shares s_i of different destinations would have to (exogenously) depend on p and t in a way that exactly delivers these results. Although we do not explicitly incorporate sectors in the alternative model, it would similarly need to be augmented to exactly match our model's predictions for the dynamics of sectoral concentration.

With $\sigma \in (\frac{1}{2}, 1)$ the initial capital of entrants would fall over time due to the increase in A_t and the decline in $v(p,t)$, and would also be falling in p (more steeply over time) due to the $v(p,t)$ term. However, post-entry growth of firm size would be driven entirely by rising productivity at the destinations, A_t, which does not vary with the origins of entrepreneurs. Hence this model would not generate result (b) of Proposition 3 concerning post-entry growth in firm size across origin counties.

3.4.2 Destination Heterogeneity

Now consider the implications of varying productivity levels and growth rates across destinations. This could reflect the effect of geography, support provided by local governments (through credit and infrastructure), or agglomeration spillovers. The latter depend on the total number of firms at a destination, regardless of their origin. Let A_{it} denote productivity at destination i at t, which does not vary with the origins of entrepreneurs in the absence of network effects. Suppose in addition that high p origins have better, and increasing, access to the faster growing destinations. For instance, if there are two destinations and productivity at destination

[36] An alternative interpretation of $v(p,t)$ is that it represents the payoff from origin-based networks operating in the traditional sector.

1 is higher and growing faster than at destination 2, then the share $s_1(p,t)$ is increasing in t and in p (more steeply over time). The expressions for entry thresholds and entry flows are now

$$\log \underline{\omega}_i(p;t) = \frac{1}{1-\sigma}\left[\log\frac{1}{\psi} + \frac{\alpha}{1-\alpha}\log r - \frac{1}{1-\alpha}\log A_{it} + \log v(p,t)\right] \tag{21}$$

$$e_i(p,t) = s_i(p,t)k(p,t)\left[1 + Z + \frac{1}{(1-\alpha)(1-\sigma)}\log A_{it} - \frac{1}{1-\sigma}\log v(p,t)\right] \tag{22}$$

This model which incorporates both origin and destination heterogeneity would generate the same predictions as Proposition 2 for entry and concentration. There would be greater total entry from high p origins owing to the origin heterogeneity, coupled with greater access to the faster growing destination. Concentration would rise over time for entrepreneurs from all origins, owing to faster entry growth into destination 1. This would be more pronounced for the high p origins, so concentration would rise in p and $p * t$. Entry thresholds from high-p origins would be lower due to higher A_{it} (averaged across destinations) or lower $v(p,t)$, so the initial capital size result in part (a) of Proposition 3 would also go through. The average rate of growth of firm size (where we average across destinations) would be higher for high-p origins, owing to their preferred access to the faster growing destination.

The alternative model specified above can generate the predictions of our model relating to the dynamics of entry, concentration, and firm size because the key $s_i(p,t)$, A_{it} terms are exogenously specified to match the endogenous evolution of these terms in our model. If firms from each origin locate at a unique set of destinations, then our network-based model would not be distinguishable from the alternative model with destination heterogeneity. In practice, however, firms from multiple origins will locate at the same destination. Destination-time period dummies can then be included in the estimating equation. Conditional on these dummies, the network model would imply that firms from higher-p origins will grow faster on average because their growth is driven by changes in CTFP. In contrast, there is no relationship between firm growth and p in the alternative model once destination-time period dummies are included because there is no longer any variation within destinations.

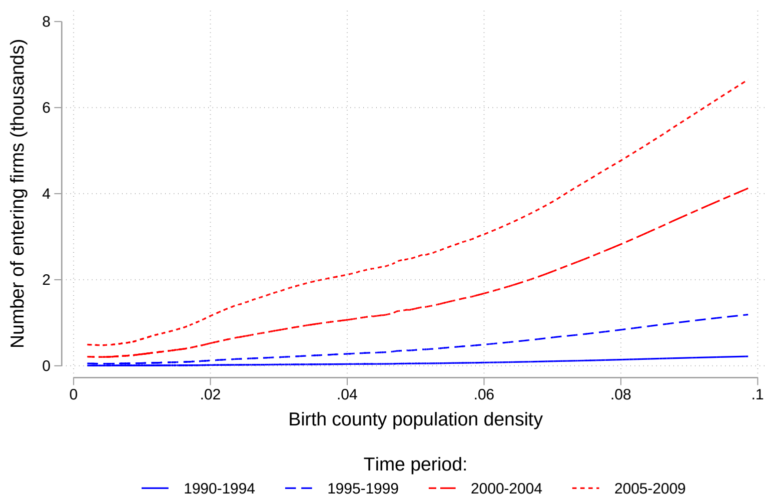
One way to incorporate heterogeneity within destinations, without networks, would be to allow firm growth to vary with the entrepreneur's ability; this is not a feature of our model. A positive relationship between p in the origin county and firm growth would then be obtained even within destination-time periods if entrepreneurs from higher p origins have higher ability on average. However, this model would not explain why firms from higher p origins, with higher ability, nevertheless have lower initial capital. An alternative model that may be considered, is that entrepreneurs do not have access to external credit and have to be entirely self-financing (Song *et al.*, 2011). Suppose that for some reason entrepreneurs from high p counties have a higher shadow cost of capital, so entering firms start with lower capital size, and thereafter grow faster owing to convergence forces akin to those in the Ramsey-Solow neoclassical growth model. This model would not be able to explain the positive relationship between population density and either entry or concentration; high p origins ought then to be associated with smaller entry flows. Nor would it be able to explain why the positive relationship between firm size growth and population density is robust to controlling for initial capital size (as shown below).

4 Testing the Augmented Model

4.1 Evidence on Firm Entry

The model predicts that firm entry is (i) increasing in origin county population density at each point in time, (ii) increasing over time, and (iii) increasing more steeply in population density over time. This is a statement about the flow of firms rather than the stock. As with the tests of the canonical model, we will thus measure entry in five-year windows from 1990 to 2009 to ensure that there is a sufficient flow of firms in each time period. Figure 5 reports nonparametric estimates of the relationship between the entry of firms from each birth county in each time period and 1982 population density.[37] The entry patterns in the figure are visually consistent with the model's predictions.[38]

Figure 5. Firm Entry and Population Density



Source: SAIC registration database and 1982 population census.

Table 7, Columns 1-4 report parametric estimates corresponding to Figure 5, separately by time period. This allows us to statistically validate the prediction that entry is increasing in birth county population density at each point in time. As noted, all analyses that estimate the direct effect of birth county population density will include population, education, and occupational structure (also measured at the county level in 1982) in the estimating equation. This allows for the possibility that population density is correlated with county characteristics that independently determine the outcomes of interest. We see in Table 7, Columns 1-4 that the population density coefficient is positive and significant in each time period. Notice also that the mean of the dependent variable and the population density coefficient are increasing across time periods, in line with

[37]The advantage of using a predetermined measure of social connectedness in all of the analysis is that it avoids the possibility that changes in population density in later time periods are generated by endogenously determined network-based migration. Nevertheless, as documented in Appendix Figure D.2 using successive rounds of the population census, the ranking of counties with respect to population density is invariant over time.

[38]Appendix Figure D.3a reports the corresponding nonparametric relationship between population density in the birth county and the stock of firms (measured at the end of each time period). The predictions of the model apply to both firm entry; i.e. the flow and the stock of firms. In practice, however, the stock will also take account of exits, which play no role in the model. We see in Figure D.3a that the model's predictions for the stock of firms go through as well, despite the exits. As an additional robustness test, Appendix Figure D.3b reports the nonparametric relationship between population density in the birth county and firm entry, restricting attention to firms that locate outside the birth county. Although the entry result is based on all locations, we see that the predictions of the model hold up with this reduced sample of locations as well.

predictions (ii) and (iii) above. Formal tests of these predictions are reported later in this section.

Table 7. Firm Entry and Population Density

Dependent variable:	number of entering firms				number of entering firms			
Time period:	1990-1994	1995-1999	2000-2004	2005-2009	1990-1994	1995-1999	2000-2004	2005-2009
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth county population density	0.013***	0.092***	0.289***	0.448***	0.014**	0.118***	0.382***	0.575***
	(0.003)	(0.013)	(0.035)	(0.052)	(0.006)	(0.037)	(0.101)	(0.126)
Mean of dependent variable	0.0306	0.208	0.787	1.560	0.0725	0.483	1.673	3.024
Sector fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Location fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Observations	1,624	1,624	1,624	1,624	1,085,169	1,085,169	1,085,169	1,085,169

Note: number of entering firms from each birth county in each time period is measured in thousands.
Population density is measured in units of 10,000 people per square km, and then converted to Z-score.
In Columns 5-8, for a given birth county, all sectors and locations that ever have entrants are included in all time periods (assigned zero entry where necessary). To adjust for differences in the number of sectors and locations across birth counties, the number of entrants is multiplied by the number of sectors × the number of locations.
Control variables include population, education and occupation distribution in the birth county.
Population is measured in millions and education is measured by the percent of the population that is literate.
Occupation distribution is measured as the share of workers in the birth county in agriculture and industry. Service is the excluded category.
Number of firms is obtained from the SAIC registration database and population density, population, education and the occupation distribution are derived from the 1982 population census.
Robust standard errors are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

As discussed above, both origin heterogeneity and destination heterogeneity can explain the entry results without requiring origin-based networks to be active. We account for important elements of origin heterogeneity by including additional birth county characteristics in the estimating equation. We next account for the possibility that entrepreneurs born in higher population density birth counties have access to faster growing destinations. Given that firms from multiple origin counties will enter each destination, we can flexibly accommodate this possibility by including destination fixed effects in the estimating equation. The estimating equation in Table 7, Columns 5-8 includes sector fixed effects and location fixed effects together with the birth county characteristics. This equation is estimated separately in each time period and so the fixed effects capture the changing fortunes of sectors and locations over time.[39] Birth county population density continues to have a positive and significant effect on entry in each time period in Table 7, Columns 5-8. A comparison of the results obtained with the benchmark specification in Columns 1-4 and the augmented specification in Columns 5-8 indicates that the inclusion of the fixed effects actually increases the point estimates. This tells us that entrepreneurs born in high population density counties are selecting sectors and locations that are less advantageous (receive fewer entrants overall).

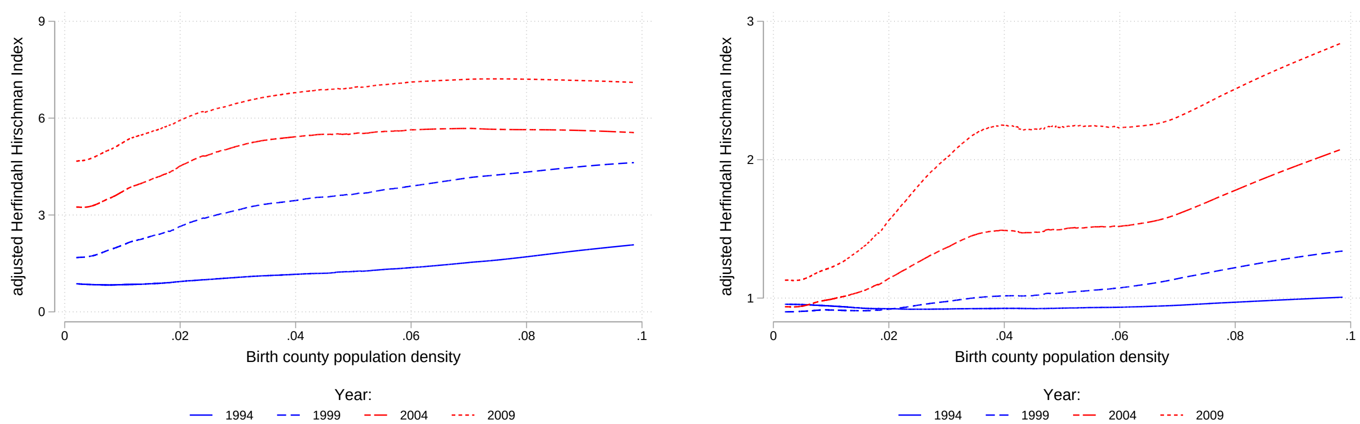
4.2 Evidence on Concentration

The model predicts that the concentration of the stock of firms, measured by the Herfindahl Hirschman Index (HHI) across destinations, is (i) increasing in birth county population density at each point in time, (ii) increasing over time, and (iii) increasing more steeply in population density over time. Destinations are defined by sectors or by locations within sectors. Figure 6a reports nonparametric estimates of the relationship

[39]Entry in Table 7, Columns 5-8 is measured at the birth county-sector-location level in each time period. The number of entrants is thus multiplied by the county-specific product of the number of sectors and the number of locations so that the dependent variable reflects entry at the level of the county.

between sectoral concentration at the two-digit level and 1982 population density in the birth county in five-year intervals from 1994 to 2009. The HHI is based on the stock of existing firms (net of exits) and, as with the analysis of clan concentration, is adjusted for the fact that measured concentration could vary with the number of firms and the number of sectors just by chance, using a normalization derived in Appendix A. The adjusted HHI is evidently increasing in population density at each point in time and increasing over time, although it is difficult to visually assess whether the slope of the relationship gets steeper over time. Figure 6b reports nonparametric estimates of the relationship between spatial concentration, within one-digit sectors, and birth county population density in five-year intervals. Although the model assumes that all destinations are symmetric, one obvious asymmetry in practice is that transportation costs are lower when the entrepreneur chooses to stay back home. We avoid this asymmetry by including all locations in the analysis, measured at the county or urban district level, except for the birth county. As with the analysis of sectoral concentration, the spatial concentration within each sector for a given birth county is based on the stock of firms (net of exits) and is adjusted for the number of firms and the number of external destinations, which would generate variation in the measured HHI just by chance.[40] Matching the predictions of the model, the spatial HHI is evidently (i) increasing in birth county population density in each time period, (ii) increasing over time, and (iii) increasing more steeply over time.

Figure 6. Concentration and Population Density



(a) Sectoral Concentration **(b)** Spatial Concentration

Source: SAIC registration database and 1982 population census.
Sectoral concentration is measured by the Herfindahl Hirschman Index (HHI) across two-digit sectors, divided by the expected HHI that would be obtained by random assignment, given the stock of firms and the number of sectors at each point in time.
Spatial concentration, within one-digit sectors, is measured by the Herfindahl Hirschman Index (HHI) across destination locations (outside the birth county) divided by the expected HHI that would be obtained by random assignment, given the stock of firms and number of destination locations at each point in time.

Table 8 reports parametric estimates corresponding to Figure 6a and Figure 6b. The usual birth county characteristics are included in the estimating equation. Sector fixed effects are also included in the estimating equation with spatial concentration (within sectors) as the dependent variable to allow for the possibility that concentration varies independently across sectors (possibly due to the nature of the production technology and

[40] We measure spatial concentration within one-digit rather than two-digit sectors to allow for a sufficient flow of firms across locations. To maintain consistency across time periods, we only include birth county-sectors that have multiple entrants in all time periods. This is not a constraint in the sectoral analysis because all birth counties have multiple entrants in each time period.

the associated gains from agglomeration). Population density in the birth county has a positive and significant effect on (adjusted) sectoral and spatial concentration at each point in time. The mean of the dependent variable and the population density coefficient are increasing over time, in line with predictions (ii) and (iii), which we test formally below.

Table 8. Sectoral and Spatial Concentration and Population Density

Dependent variable:	adjusted HHI across sectors				adjusted HHI across locations			
Year:	1994	1999	2004	2009	1994	1999	2004	2009
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth county population density	0.106***	0.417***	0.444***	0.574***	0.014*	0.041***	0.047*	0.072*
	(0.023)	(0.057)	(0.061)	(0.074)	(0.008)	(0.014)	(0.027)	(0.040)
Mean of dependent variable	1.039	2.839	4.622	6.065	0.936	1.010	1.295	1.777
Observations	1,622	1,624	1,624	1,624	5,450	15,076	23,727	26,769

Note: sectoral concentration measured across two-digit sectors and spatial concentration, within one-digit sectors, is measured across destination locations (outside the birth county). Concentration statistics are adjusted for expected concentration due to random assignment.
Population density is measured in units of 10,000 people per square km, and then converted to Z-score.
Sector fixed effects included in the regression with spatial HHI as the dependent variable.
Control variables include population, education and occupation distribution in the birth county.
Number of firms is obtained from the SAIC registration database and population density, population, education and the occupation distribution are derived from the 1982 population census.
Robust standard errors are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Table 9 formally tests the model's predictions for changes in firm entry, sectoral concentration, and spatial concentration over time and across birth county population density over time. Data from all time periods are pooled and the estimating equation now includes birth county population density, time period, and the interaction of these variables. Since the cross-sectional relationship with population density in each time period has been previously reported with each outcome, we only report the coefficient on the time period variable and the interaction coefficient. Restricting the sample to county-born entrepreneurs in Table 9, Columns 1-3, the time period coefficient and the interaction coefficient are positive and significant with the number of entrants, sectoral concentration, and spatial concentration as the dependent variables, as predicted by the model. As a placebo test, we restrict the sample to entrepreneurs born in cities in Table 9, Columns 4-6. Population density is not positively associated with social connectedness in cities and thus we do not expect to find support for the model's predictions with this set of entrepreneurs. The time period coefficient and the interaction coefficient are both positive and significant with entry as the dependent variable but, as discussed, many alternative models can generate this result without a role for birth county networks. The model's predictions for concentration are less easy to explain away. Reassuringly, the interaction coefficient for the city-born entrepreneurs is *negative* and significant with sectoral concentration as the dependent variable and statistically indistinguishable from zero (at conventional levels) with spatial concentration as the dependent variable, contrary to the predictions of our model.

4.3 Evidence on Firm Size

The model predicts that the ability and the initial capital of the marginal entrant is (i) decreasing in birth county population density at each point in time, (ii) decreasing over time, and (iii) decreasing more steeply in

Table 9. Entry, Concentration, and Population Density (time and interaction effects)

Birth place:	county			city		
Dependent variable:	number of entrants	sectoral HHI	spatial HHI	number of entrants	sectoral HHI	spatial HHI
	(1)	(2)	(3)	(4)	(5)	(6)
Time period	0.517***	1.686***	0.506***	0.661***	0.299***	2.054***
	(0.016)	(0.020)	(0.021)	(0.026)	(0.008)	(0.069)
Birth place population density × time period	0.353***	0.165***	0.134***	0.353***	-0.026***	0.030
	(0.029)	(0.019)	(0.012)	(0.041)	(0.004)	(0.026)
Observations	6,496	6,494	71,022	3,284	3,283	21,046

Note: the estimating equation includes, in addition, birthplace population density and a constant term.
Number of entering firms from each birth place in each time period is measured in thousands.
Sectoral concentration measured by Herfindahl Hirschman Index (HHI) across two-digit sectors divided by the expected HHI that would be obtained by random assignment.
Spatial concentration, within one-digit sectors, is measured by the Herfindahl Hirschman Index (HHI) across destination locations (outside the birth county) divided by the expected HHI that would be obtained by random assignment. Sector fixed effects are included in this specification.
Population density is measured in units of 10,000 people per square km, and then converted to Z-score.
Time period is an ordinal variable taking value from 1 to 4 corresponding to successive five-year time windows over the 1990-2009 period.
Number of entrants and concentration statistics are derived from the SAIC registration database and population density is derived from the 1982 population census.
Standard errors clustered at birthplace level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

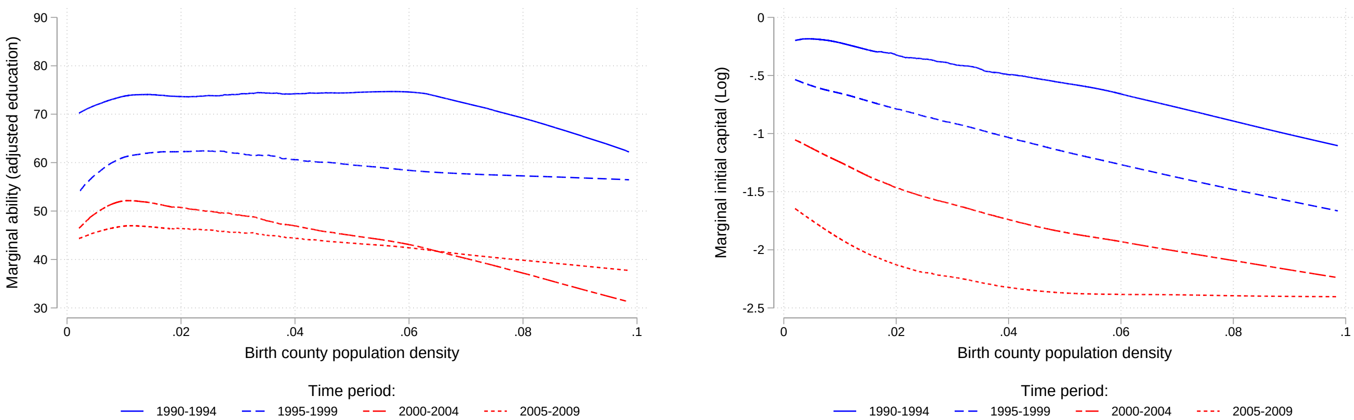
population density over time. If the negative selection on ability that accompanies a stronger network dominates the positive productivity effect of that network for inframarginal firms, then the preceding predictions apply to average initial capital as well. However, only the positive network productivity effects are relevant for post-entry growth rates of firm size.

To test the model, we measure marginal ability and initial capital at the level of the birth county-sector or birth county-sector-location in each time period and then estimate the (average) effect of birth county population density and its interaction with time on these variables. We begin in Figure 7a by nonparametrically estimating the relationship between a measure of ability, based on education, of the marginal entrepreneur in each birth county-sector-time period and population density in the birth county. It is standard practice to proxy for ability with education, and recent evidence indicates that education is also a good measure of entrepreneurial ability (Levine and Rubinstein, 2017). In a developing economy, the level of education will vary across birth cohorts and in the cross-section (across birth counties) for the same level of ability, depending on the supply of schooling. Our measure of ability is thus the entrepreneur's percentile rank in his birth county-birth cohort education distribution.[41] The marginal entrant is the entrepreneur who is placed at the bottom one percentile of the ability distribution among entering entrepreneurs in each birth county-sector-time period. We see in Figure 7a that the marginal entrant's measured ability declines over time; from around the 70^{th} percentile of his birth county-birth cohort education distribution in the 1990-1994 period to just around the 40^{th} percentile in the 2005-2009 period. The relationship between the marginal entrant's ability and population

[41] The education distribution is constructed in each county for birth cohorts from 1920 to 1989 in five-year intervals, based on data from the 2000 population census. Each entrepreneur is assigned to a birth cohort interval based on his birth year, which is available from the registration database, and his position in the relevant education distribution is determined on the basis of his education, which is also obtained from the registration database. The coverage for the education variable is not complete in the SAIC registration database, with a significant minority of entrepreneurs not reporting this information. This has no bearing on the complementary analysis of firm size, which includes all registered firms.

density is also negative in each time period and grows steeper over time.[42] Notice, however, that there is a bottoming out by the last, 2005-2009, period. Our model is only designed to capture firm dynamics up to this point, which is why the empirical analysis does not extend beyond 2009. For the dynamic analysis of negative selection that follows, and for the structural estimation, the analysis period will be restricted even further to the 1990-2004 period.

Figure 7. Marginal Ability, Marginal Initial Capital and Population Density



(a) Marginal Ability

(b) Marginal Initial Capital

Source: SAIC registration database and 1982 population census.
The entrepreneur's ability is measured by his percentile rank in his birth county- birth cohort education distribution (obtained from the 2000 population census). The marginal entrepreneur is defined as the individual at the bottom one percentile of the ability distribution among entering entrepreneurs in a given birth county-sector-time period.
Marginal initial capital is defined as the bottom one percentile of the initial capital distribution at the birth county-sector-time period level.

Figure 7b reports complementary nonparametric estimates of the relationship between marginal initial capital, measured in logs, and 1982 population density in the birth county in five-year windows over the 1990-2009 period. Marginal initial capital is defined as the bottom one percentile of the initial capital distribution at the birth county-sector-time period level.[43] As predicted by the model, marginal initial capital is decreasing over time and decreasing in birth county population density in each time period.[44]

Notice from Figure 7b that the decline in initial capital with birth county population density does not grow steeper over time (as implied by the model). One reason why this might be the case is because marginal initial capital within birth county-sector-time periods is effectively averaged across sectors in the figure. Although this is not a feature of our model, the capital requirement will vary across sectors, and this must be accounted for in the empirical analysis. Table 10 allows for this by studying the change in the ability of entering entrepreneurs and their capital investments over time, within birth county-sectors. The analysis is restricted to the 1990-2004 period because our measure of marginal ability and initial capital both bottom out (and flatten out) in the 2005-2009 period in Figures 7a and 7b. We see in Table 10, Column 1, which includes birth county-sector fixed effects, that the marginal entrant is drawn from lower down in his birth county-cohort education

[42] Appendix Table E.4 reports parametric estimates corresponding to Figure 7a, separately in each time period. These estimates indicate that birth county population density has a negative and significant effect on marginal education among entering entrepreneurs at each point in time.

[43] The initial capital for a firm is determined by its initial registered capital, which can be recovered from the SAIC registration database.

[44] Appendix Table E.4 reports parametric estimates corresponding to Figure 7b, separately by time period. The population density coefficient is negative and significant in each time period.

distribution over time and that this decline in our measure of ability is significantly steeper for entrants from higher population density counties, as predicted by the model. Table 10, Columns 2-3 use the distribution of initial capital (in logs) in each entering cohort of firms, in five-year windows over the 1990-2004 period, to identify the marginal entrant (the bottom one percentile) and the average entrant by birth county-sector. Including birth county-sector fixed effects in the estimating equation, we see that both the marginal entrant's initial capital and the average entrant's initial capital are decreasing significantly over time. Although the coefficient on the time period-birth county population density interaction is also negative and significant with the marginal entrant's initial capital as the dependent variable, the interaction coefficient is positive (albeit small in magnitude and statistically insignificant) with average initial capital as the dependent variable.

The analysis of firm size thus far has not accounted for location choices, and the possibility that variation in these choices across birth counties could be driving the results. Table 10, Columns 4-5 thus includes location fixed effects, in addition to birth county-sector fixed effects in the estimating equation. Initial capital is now measured at the birth county-sector-location level in each time period.[45] Both marginal initial capital and average initial capital are declining significantly over time, as in Columns 2-3. Moreover, the coefficient on the time period-birth county population density interaction is now negative and significant with both dependent variables, as predicted by the model. As with the analysis of firm entry, accounting for location effects only strengthens our results, indicating that entrepreneurs from higher population density birth counties selected less advantageous locations where firms had lower access to capital (within sector) on average.

Table 10. Evidence on Negative Selection

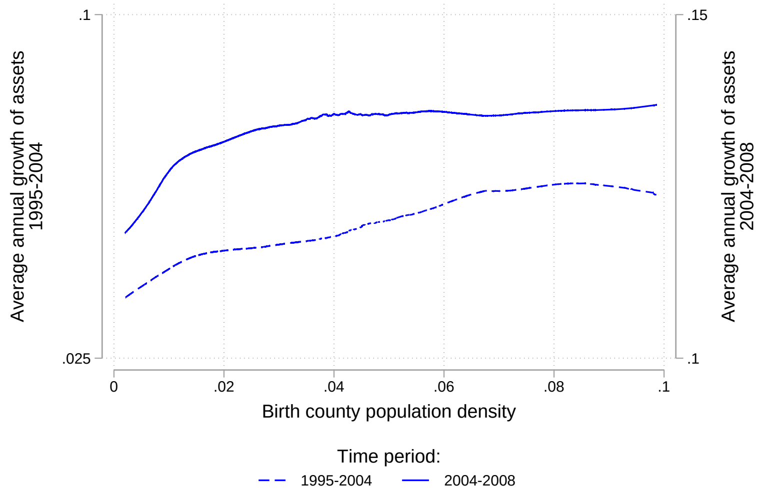
Dependent variable:	marginal ability	marginal initial capital	average initial capital	marginal initial capital	average initial capital
	(1)	(2)	(3)	(4)	(5)
Time period	-17.908***	-0.868***	-0.116***	-0.609***	-0.095***
	(0.496)	(0.012)	(0.009)	(0.010)	(0.008)
Birth county population density × Time period	-0.926***	-0.026**	0.002	-0.061***	-0.020***
	(0.351)	(0.011)	(0.007)	(0.009)	(0.006)
Mean of dependent variable	49.36	-1.744	-0.401	-1.223	-0.374
Origin-sector fixed effects	Yes	Yes	Yes	Yes	Yes
Location fixed effects	No	No	No	Yes	Yes
Observations	21,028	43,578	43,578	46,417	46,417

Note: The entrepreneur's ability is measured by his percentile rank in his birth county- birth cohort education distribution (obtained from the 2000 population census). The marginal entrepreneur is defined as the individual at the bottom one percentile of the ability distribution among entering entrepreneurs in a given birth county-sector-time period.
Initial capital (in million Yuan) is measured in logs. Marginal initial capital defined by the bottom one percentile of the initial capital distribution at the birth county-sector-time period level or the birth county-sector-location-time period level (when location fixed effects are included). Average initial capital is the mean of the distribution.
Time period is an ordinal variable taking value from 1 to 3 corresponding to successive five-year time windows over the 1990-2004 period. Education and initial capital are obtained from the SAIC registration database and birth county population density is derived from the 1982 population census.
Standard errors clustered at birth county-sector level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Conditional on having entered, the model predicts that firms from higher population density counties will

[45]The marginal entrant's initial capital and the average entrant's initial capital are now based on the distribution of capital in each birth county-sector-destination-time period. The sample in Columns 4-5 is restricted to birth county-sector-destinations with entrants in the initial period. Similarly, the sample in Columns 2-3 is restricted to birth county-sectors with entrants in the initial period.

Figure 8. Asset Growth and Population Density



Source: Industrial census (1995, 2004, 2008) and 1982 population census.
Firm-level average annual growth of assets is averaged up to the birth county-sector level in each time period.

grow faster. Although the registration database is well suited to examine entry, concentration, and initial capital investments, it is less suitable for analyses of capital growth. Registered capital does change, but given that these changes are self-reported and involve substantial administrative costs, it will not track perfectly with changes in the firm's assets over time. For the analysis of firm growth, we thus turn (separately) to the industrial census, which was conducted in 1995, 2004, and 2008 and the SAIC's inspection database, which includes annual firm-level information on assets and sales and which has reasonable coverage from 2004 onwards. We thus compute the average annual growth rate over the 1995-2004 and 2004-2008 periods with the industrial census and, to be consistent, over the 2004-2008 period with the inspection data.[46] Figure 8 reports asset growth, separately in the 1995-2004 period and the 2004-2008 period, based on the industrial census. The average annual growth of assets is increasing in population density in each time period and increasing over time, as predicted by the model, in contrast with the patterns that we observe in the data for initial firm size.

Table 11 reports parametric estimates corresponding to Figure 8. Since growth rates can only be computed at two points in time with the industrial census data and the inspection data cover a relatively short period of time, we focus on the cross-sectional predictions of the model. Note that the growth in firm size is computed by differencing the level over time and, hence, the cross-sectional relationship between population density and growth is effectively the change in the relationship between population density and the level of firm size over time. To test the model's predictions with industrial census data, we measure firm growth at the birth county-sector level in Table 11, Columns 1 and 3 and at the birth county-sector-location level in Table 11, Columns 2 and 4 and then estimate the (average) effect of birth county population density on these growth measures. The estimating equation in Columns 1 and 3 includes sector fixed effects, while the estimating

[46]The average annual growth between period t and t' is computed as the difference in log assets in t' and t divided by $t' - t$. Although there are no exits in the model, this is a feature of the data. In practice, firms with low profit levels – the young and the less able – are more likely to exit. This selective exit, based on the profit level, does not bias our estimates because growth rates in the model are determined entirely by changes in CTFP that apply equally to all active firms in the network.

Table 11. Growth of Assets and Population Density

Dependent variable:	average annual growth of assets					
Source:	industrial census		industrial census		inspection data	
Time period:	1995-2004		2004-2008		2004-2008	
	(1)	(2)	(3)	(4)	(5)	(6)
Birth county population density	0.006***	0.007*	0.004**	0.003**	0.004***	0.002*
	(0.002)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)
Initial capital	–	0.002***	–	0.001***	–	0.001***
		(0.000)		(0.000)		(0.000)
Mean of dependent variable	0.0528	0.0557	0.133	0.136	0.106	0.110
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Location fixed effect	No	Yes	No	Yes	No	Yes
Observations	5,517	5,664	31,234	64,258	18,701	43,470

Note: firm-level average annual growth of assets is averaged up to the birth county-sector level in specifications with sector fixed effects and to the birth county-sector-location level in specifications with sector fixed effects and location fixed effects.
Initial capital (in million Yuan) obtained from the SAIC registration database and birth county population density is derived from the 1982 population census.
Standard errors clustered at birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

equation in Columns 2 and 4 includes sector fixed effects, location fixed effects, and the firm's initial capital. The fixed effects account for exogenous variation in firm growth across sectors and locations, which is not a feature of our model. The firm's initial capital is included to allow for convergence; recall that one alternative explanation for why firms from high population density birth counties start small and then grow faster is mechanical convergence (with initial size being accidentally determined). What we observe, instead, is that firms that are larger to begin with, subsequently grow faster. Table 11, Columns 5-6 repeats the analysis with SAIC inspection data, which include all sectors (not just manufacturing, as in the industrial census).[47] The consistent finding across specifications is that population density in the birth county has a positive and significant effect on firm growth.[48]

The estimating equation in Table 11 is specified to be consistent with the estimating equation in Table 10. This allows for an appropriate comparison of initial size and firm growth. The main finding is that firms from high population density birth counties start small but subsequently grow faster, after accounting for sector and location effects. As discussed, this result is especially useful in distinguishing our model from alternative non-network explanations.

5 Structural Estimation

Having validated the model, we next proceed to estimate its structural parameters. This will allow us to quantify the contribution of the community networks to the growth in the number of firms and the capital stock at the aggregate level. For the purpose of the structural estimation, we measure destinations by sectors in order to restrict the dimensionality of the necessary computations. Two equations, with entry and average initial capital in each birth county-sector-time period as the dependent variables, need to be estimated. Besides

[47]Data coverage for seven provinces is poor with the inspection data and these provinces are thus dropped from the analysis.
[48]A pooled regression (not reported) which combines industrial census data over both time periods indicates, in addition, that firm growth is increasing significantly over time. While our model can explain this result, it must be augmented, perhaps by introducing credit constraints, to explain why initial capital has a positive and significant effect on firm growth.

the benchmark version of the model with myopic decision-making, we also estimate a variant (presented in Appendix B) where entrepreneurs consider profits in the current period and the subsequent period when making their entry decision.

Based on the corresponding equations in the myopic version of the model, (12) and (16), and retaining its notation, we estimate the following structural equations:

$$e_{ci,t} = G(\alpha, \sigma, r, A_0)k_c s_{ci,t-1} + \frac{\theta}{(1-\sigma)(1-\alpha)} k_c s_{ci,t-1} \cdot pn_{ci,t-1} + u_{ci,t} \tag{23}$$

$$logK^a_{ci,t} = J_t(\alpha, \sigma, r, A_0, f_t) + \frac{\theta(1-2\sigma)}{2(1-\sigma)(1-\alpha)} pn_{ci,t-1} + v_{ci,t} \tag{24}$$

$e_{ci,t}$ measures the number of entrants and $logK^a_{ci,t}$ measures average initial capital (in logs) for birth county c and sector i in time period t. We parameterize the $\theta(p)$ function to be increasing linearly in p: $\theta(p) = \theta p$, with the restriction that $\theta(0) = 0$. The network effect is thus represented by a single parameter, θ. $n_{ci,t-1}$ is the stock of firms from that birth county that are already established in that sector at the beginning of the time period. $s_{ci,t-1}$ denotes the share of sector i in the stock of firms originating from county c at $t-1$. Capital is measured in the model in physical units, whereas in the data it is measured in monetary units. The mapping from physical units to monetary units changes over time owing to changes in the price of capital goods. This is especially relevant in the structural estimation because the objective is to match predicted and actual firm size in each time period. f_t thus represents the price of capital goods in period t.

k_c measures the number of potential entrepreneurs from the birth county. The theoretical model assumed k_c was equal across birth counties and time periods. In practice, the number of potential entrepreneurs will depend on the size of the population and the level of education in the county. The number of potential entrepreneurs in each birth county is calculated from the 1990 population census, based on the characteristics of actual entrepreneurs when they established their firms. We see in Appendix Figure D.4 that most entrepreneurs in the SAIC database have at least high school education and that most were aged 25-44 when their firm was established. k_c is thus specified to be the number of men born in county c, aged 25-44, with at least high school education, as reported in the 1990 population census.

The residual terms, $u_{ci,t}$ and $v_{ci,t}$, measure the effect of local government inputs, agglomeration, and sector-level spillovers on access to capital, firm productivity, and accompanying entry. We will see, with an augmented specification, that adding sector-level spillovers to the structural equations does not affect the estimated parameters, indicating that this component of $u_{ci,t}$, $v_{ci,t}$ is uncorrelated with birth county population density, p. Recall, however, from the tests of firm entry and firm size that entrepreneurs from higher p counties appear to be selecting into less advantageous locations. The implied negative correlation between $u_{ci,t}$, $v_{ci,t}$ and p would bias the network spillover parameter, θ, downward and, hence, provide a conservative estimate of the network effects.

The structural equations are linear in observed variables; (i) $k_c s_{ci,t-1}$ (ii) $k_c s_{ci,t-1} \cdot pn_{ci,t-1}$ (iii) $pn_{ci,t-1}$, with four reduced-form coefficients.[49] One of these coefficients, J_t, cannot be used to identify the structural parameters because f_t is unobserved. This leaves three reduced-form coefficients and five structural parameters:

[49]The functional forms for $G(\alpha, \sigma, r, A_0)$ and $J_t(\alpha, \sigma, r, A_0, f_t)$ are obtained directly from (12) and (16), with the addition of the separable f_t term in the J_t function.

$\alpha, \sigma, r, A_0, \theta$. We noted in Section 3 that the productivity channel and the credit channel for the network effect cannot be separately identified. Although the model is parameterized to allow networks to increase productivity, we remain agnostic about the specific channel through which the networks operate. For the structural estimation, we specify that the network operates through the productivity channel as in the model, setting r to 0.2 (which is in line with estimates of the average interest rate faced by Chinese firms).[50] The productivity multiplier is set to one in all sectors; i.e. $A_0 = 1$. As in the model, variation in productivity across sectors (and origin counties) in the benchmark specification is generated entirely by the network effect; $exp(\theta p \cdot n_{ci,t-1})$. In addition to the factors included above in the residual terms, we thus also abstract from variation in product prices and labor productivity. The objective will be to assess how well our parsimonious model is able to match the observed dynamics of entry and firm size.

To accommodate differences in the capital requirement across sectors, we do, however, allow the α parameter, which measures the marginal returns to capital, to vary across four broad sector categories: high-tech services, wholesale and retail services, manufacturing and transportation, and heavy industry (mining, electricity, and construction). This increases the number of structural equations to eight, given that there are now two equations in each sector category, and the number of structural parameters to be estimated to six; $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \sigma, \theta$. The structural parameters are estimated by matching on entry and average initial capital in each birth county-sector-time period.[51] Initial entry in each birth county-sector is based on the number of entrants in 1990-1994 and, if there is no entry in that time period, on the number of entrants in 1995-1999. Sectors are defined at the one-digit level in the structural estimation to ensure that there is positive initial entry in (almost) all birth county-sectors and the model is estimated over the 1995-2004 period; i.e. over two time periods.

To estimate the structural parameters, we search for the set of parameters that minimize the distance between the actual and the predicted entry and average initial capital; i.e. for which the sum of squared errors over all birth county-sector-time periods is minimized. Parameter estimates, with bootstrapped standard errors in parentheses, are reported for the benchmark model in Table 12, Column 1.[52] The σ coefficient lies between 0.5 and 1, satisfying the condition, derived in the model, which ensures that average initial capital is decreasing in birth county population density. The adjustment from physical capital to capital in monetary units, f_t, appears additively in the J_t function and, thus, can be estimated separately in 1995-1999 (period 1) and 2000-2004 (period 2). Table 12, Column 2 reports estimates with the forward looking model, derived in Appendix B. Entry must now be derived as the solution to a nonlinear equation, satisfying a fixed point condition, in each birth county-sector-time period.[53] Notice that the θ parameter declines substantially when we allow for

[50] In our model, r, is the sum of the real interest rate and the depreciation rate. Hsieh and Klenow (2009) assume that the real interest rate is 0.05 in an economy, such as the U.S., with perfect financial markets and that the depreciation rate is 0.05. Using the same production function as Hsieh-Klenow and data from the Chinese industrial census, Brandt *et al.* (2016) estimate the real interest rate to be 0.15 in 1995 and 2004 and 0.18 in 2008. We thus set r to 0.2.

[51] Although the number of reduced form coefficients now exceeds the number of structural parameters, the model places additional restrictions on the reduced form coefficients that must hold across sector categories. For example, the ratio of the coefficients on $k_c s_{ci,t-1} \cdot pn_{ci,t-1}$ and $pn_{ci,t-1}$ in (23) and (24), respectively, must be $2/(1-2\sigma)$ in each sector category. The identification of the structural parameters is now more difficult to assess analytically and, hence, we verified that the parameters continue to be (just) identified by estimating the model with different values of r. The point estimates of the structural parameters do change in response, but the predicted entry and average initial capital (for each value of p) remain unchanged.

[52] When matching on entry and initial capital, we weight the error term by the reciprocal of the (bootstrapped) standard deviation of the mean of each variable. The unweighted estimates are very similar to what we report in the table.

[53] The discount factor per year, δ is set to 0.8 when estimating the model with foresight. Because one time period is five years,

forward looking behavior, while remaining statistically and quantitatively significant. This is because potential entrants require less of a "push" from the network when they take account of future benefits. Table 12, Column 3 reports estimates with an augmented model that allows for spillovers at the sector level, regardless of the county of birth. This captures spillovers of the sort considered in endogenous growth models, resulting from diffusion of R&D across firms in any given sector. An additional $n_{i,t-1}$ term is thus included in equations (23) and (24). Although the coefficient on $n_{i,t-1}$, λ, is very precisely estimated, notice that the θ coefficient is very similar in magnitude in Columns 1 and 3.

Table 12. Structural Estimates

Model:	benchmark	forward looking	with sector-level spillovers
σ	0.775	0.776	0.777
	(0.001)	(0.000)	(0.001)
θ	0.395	0.125	0.359
	(0.054)	(0.004)	(0.011)
λ	–	–	0.00003
			(0.000002)
α_1	0.137	0.119	0.138
	(0.009)	(0.001)	(0.007)
α_2	0.189	0.179	0.194
	(0.020)	(0.001)	(0.020)
α_3	0.192	0.177	0.197
	(0.011)	(0.003)	(0.016)
α_4	0.236	0.226	0.240
	(0.004)	(0.002)	(0.004)
f_1	0.223	0.233	0.216
	(0.004)	(0.003)	(0.016)
f_2	0.189	0.208	0.184
	(0.005)	(0.003)	(0.012)

Note: the parameters are estimated by matching on entry and average initial capital (in logs), measured at the birth county-sector-time period level. When minimizing the sum of squared errors, the error term is weighted by the reciprocal of the bootstrapped standard deviation of the mean of each variable.
The model is estimated over two time periods, 1995-1999 and 2000-2004, taking entry in the first period, 1990-1994, as given.
Sectors are defined at the one-digit level when measuring entry and capital, but the α parameter is estimated at the aggregate sector level: (1) new technological services; (2) wholesale, retail and business service; (3) manufacturing and transportation; (4) heavy industry (mining, electricity, and construction)
The discount factor per year δ is set to 0.8 in the forward looking model.
Bootstrapped standard errors in parentheses.
Number of entrants and average initial capital are computed from the SAIC registration database and birth county population density is computed from the 1982 population census.

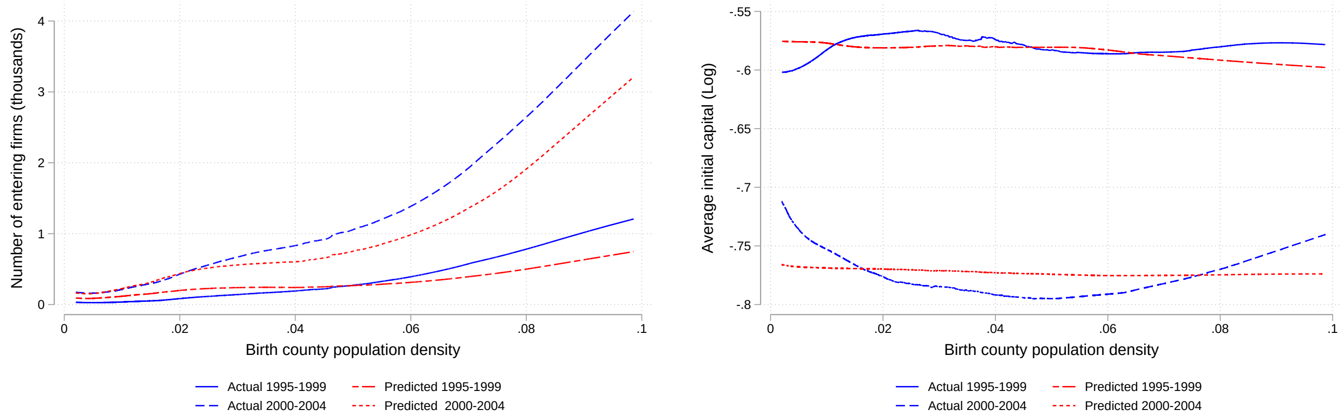
Figure 9a assesses the goodness of fit of the benchmark model by comparing actual and predicted entry across birth counties in each time period.[54] The model that we estimate is extremely parsimonious, with just six parameters. Nevertheless, it does a good job of predicting entry across nearly 2,000 birth counties and over ten years. Figure 9b repeats this exercise with average initial capital and, once again, we see that the model predicts variation across birth counties fairly accurately. Once the capital price adjustment factor is included in each time period, note that actual and predicted average initial capital will match on levels by construction.

To formally test the goodness of fit of the model, we would want to compare predicted and actual outcomes (entry and average initial capital) at each level of population density, p. Given the large number of counties, with distinct values of p, we compare, instead, the estimated population density coefficient with actual and predicted data (generated by each of the three models in Table 12). The dependent variable is either entry or

this works out to $0.8^{2.5} = 0.56$.

[54]Birth county-sectors where initial entry commences in 1995-1999 are not included in the comparison for the 1995-1999 period.

Figure 9. Actual and Predicted, Entry and Initial Capital



(a) Entry

(b) Initial Capital

Source: SAIC registration database, model generated data, and 1982 population census.

average initial capital.[55] Since the structural estimates are based on two time periods, the estimating equation includes birth county population density and a binary time period variable, which takes the value one in 2000-2004 and zero in 1995-1999. The estimated population density coefficients in Appendix Table E.5 are statistically indistinguishable between actual data and data generated by the benchmark model. Moreover, the benchmark model does a better job of matching the population density coefficient estimated with actual data than both the forward looking model and the model that allows for additional sector-level spillovers.

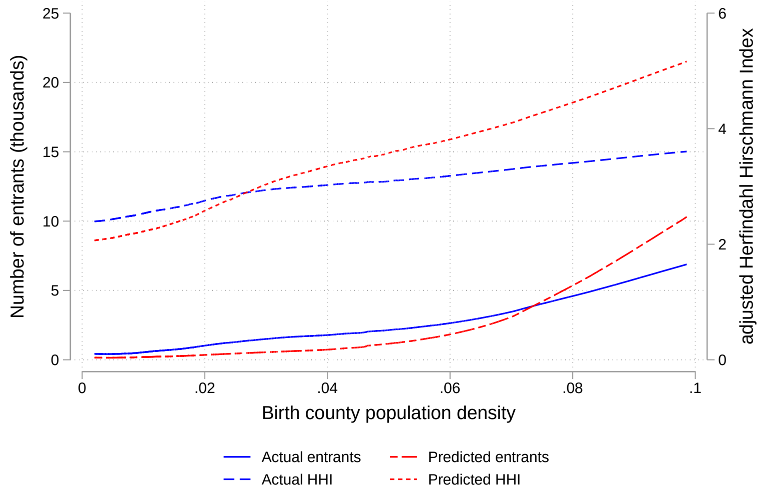
An independent test of the structural model is to assess its out of sample predictions. The model is estimated over the 1995-1999 and 2000-2004 periods. Figure 10 compares actual and predicted entry and sectoral concentration in the subsequent 2005-2009 period.[56] The model does a good job of predicting the level of firm entry and sectoral concentration on average, although the predicted slope with respect to birth county population density is higher than in the data.

With a single θ parameter common to all birth counties and sectors in the model, cross-sectional variation is generated by differences in population density and initial entry alone. Nevertheless, the model is able to match the data quite well across counties and sectors, even out of sample. The estimated parameters can thus be used for counter-factual simulations. A major objective of our research is to quantify the role played by community networks in the growth of private enterprise in China. This is accomplished by setting the θ parameter to zero and then generating counter-factual entry and capital investment over the sample period. The results of this exercise, with the benchmark model, are reported in Figure 11a. It is evident from the figure that the number of entrants would have been substantially reduced in the absence of community networks, particularly in higher population density birth counties. Based on our estimates, the total number of entrants would have declined by 64% over the 1995-2004 period if the networks had not been active. In a related counter-factual exercise, the total stock of capital in 2004 (taking account of the number of firms that entered,

[55]Entry is measured at the birth county-time period level and average initial capital is measured at the birth county-sector-time period level to be consistent with the tests of the model above.

[56]We cannot test the model's ability to predict initial capital beyond the sample period because the mapping from physical capital to capital in monetary units is unavailable.

Figure 10. Out of Sample Tests - Entry and Sectoral Concentration, 2005-2009



Source: SAIC registration database, model generated data, and 1982 population census.
Number of entrants and sectoral concentration computed at the birth county level in 2005-2009 with actual and model generated data.

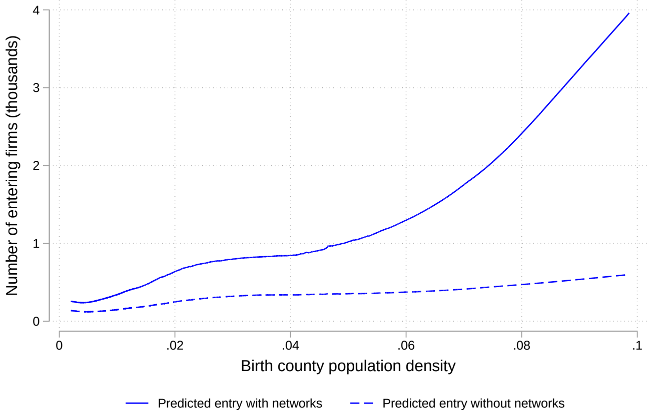
their initial capital, and the subsequent growth in their capital) would have declined by 65% had the networks been absent. Adjusting for the fact that our analysis is restricted to county-born entrepreneurs, for whom the hometown networks are relevant, this amounts to a 40% decline in the number of entrants and the stock of capital for the country as a whole.[57]

The preceding numbers are broadly in line with the results from a simple quantification exercise in which we regress entry in each birth county-sector-location-time period on a full set of birth county dummies, location dummies, and the interaction of these dummies with time. Although this exercise does not explicitly quantify the birth county network effect, the advantage of the less structured approach is that we can account flexibly for location effects. Based on this model, 45% of the predicted variation in entry over the 1995-2004 period can be explained by variation across birth counties. In an independent robustness test, Figure 11b repeats the counter-factual entry analysis with the augmented model that allows for sector-level spillovers. Notice that these spillovers have almost no impact on entry, in contrast with the substantial impact that we estimate for the birth county-sector spillovers. This indicates that origin-based networks constitute the main source of spillovers in China, rather than the sector-based spillovers considered in the endogenous growth literature.

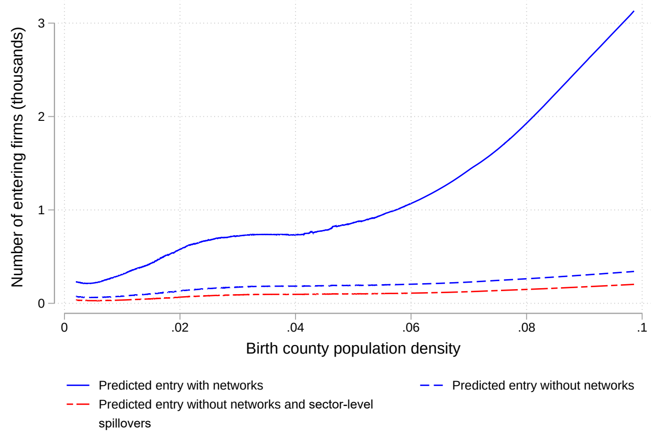
An important objective of industrial policy in any developing economy is to stimulate entrepreneurship. It has been claimed that the government played a critical role in accelerating China's growth by providing firms with subsidized credit; e.g. Song *et al.* (2011), Wu (2016). In the absence of a market-based allocation mechanism, a natural question to ask is which firms should have been targeted for the subsidy. To answer this question, we examine a counter-factual policy experiment in which all entering firms in the 1995-1999 period received credit at an interest rate of 0.15; i.e. with a subsidy of 0.05. This subsidy would have had two effects; it would have induced additional firms to enter at the margin and it would have increased the

[57] Government infrastructure and prices remain fixed in the counter-factual simulation. If the network were shut down and the number of firms declined, then output (input) prices would increase (decrease). The resulting increase in profits would encourage some additional firms to enter. In contrast, if government infrastructure and the growth of the networks are complementary, then the removal of the networks would reduce the infrastructure level, generating a further decline in the number of firms in the counter-factual scenario.

Figure 11. Counter-Factual Simulation: Effect of Community Networks on Entry



(a) Benchmark quantification



(b) Quantification with sector-level spillovers

Source: Model generated data and 1982 population census.

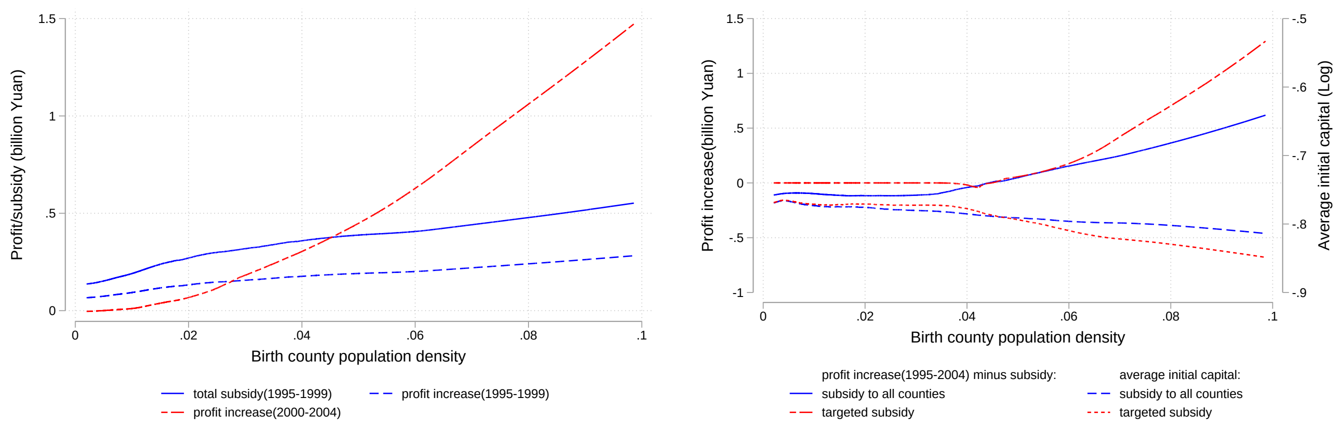
profit of all (marginal and infra-marginal) entrants. As observed in Figure 12a, the total profit increase generated by the subsidy in 1995-1999 is less than the cost to the government in all birth counties. However, the spillover effect of the one-time subsidy on profits in the subsequent 2000-2004 period is substantial (and even larger than the direct effect on profits in high population density counties). This is because the credit subsidy induces additional entry during 1995-1999, which through the compounding network effect generates large profit increases in the more socially connected counties in 2000-2004. With an annual discount factor of 0.8, the return on the subsidy, based on the additional (discounted) profits that were generated over the 1995-1999 and 2000-2004 periods minus the cost of the subsidy, would have been 12% for countries above the mean population density and -45% for counties below the mean.

Figure 12b reports the impact of an alternative government program, which only gives the subsidy to those origin counties who would have increased their aggregate discounted profits over the 1995-2004 period by more than the amount of subsidy they received in the preceding counter-factual experiment. To keep the total amount of the subsidy constant, the interest rate for the targeted counties is now reduced to 0.11. The increase in profits minus the subsidy received is reported across the population density distribution in the figure, both for the original subsidy scheme and for the targeted subsidy scheme. As can be seen, the targeted program does strictly better if the government's objective is to maximize total profit (less the subsidy cost). Notice also that average initial capital, which is declining with population density, declines even more steeply with the more efficient targeted program.[58] A distinguishing feature of our network-based mechanism is that efficiency-enhancing policies could actually result in even smaller firms and even greater dispersion in firm size in equilibrium (as observed in Figure 12b). The qualifier here is that this argument only applies if the networks increase productivity. If the networks simply capture rents (cheap capital), and higher population density birth counties have stronger networks, then a model based on community networks will still explain

[58]In our analysis, the marginal value product of capital does not vary across firms by construction. If we had used credit market imperfections to motivate network formation instead, then a efficiency-enhancing policy that exploited network spillovers would have increased the dispersion in marginal productivity within sectors (due to variation in interest rates across networks) as well. Haltiwanger *et al.* (2018) also make the point that an increase in efficiency could increase the dispersion in TFP, but with a different mechanism.

the stylized facts, but the normative interpretation will be very different. In particular, the preponderance of small firms and wide variation in firm size will now be indicative of a misallocation. Recall, however, that we found no evidence of rent seeking behavior, and the associated competition between community networks, in our tests of the canonical model.

Figure 12. Counter-Factual Simulation: Effect of Interest Rate Subsidy on Profits



(a) Subsidy to all counties

(b) Targeted subsidy vs. subsidy to all counties

Source:Model generated data and 1982 population census.

6 Conclusion

In this paper, we identify and quantify the role played by community networks, organized around the birth county, in the growth of private enterprise in China. The model that we develop generates predictions for the dynamics of firm entry, sectoral and spatial concentration, and firm size across birth counties with different levels of social connectedness (measured by population density) when networks are active. We validate each of these predictions over a twenty year period with unique administrative data that covers the universe of registered firms and provides information on entrepreneurs' birth counties. The rich set of results that we obtain, taken together, allow us to rule out alternative non-network based explanations. Additional results provide direct support for the network channel, indicating that spillovers occur within the birth county and in particular within clans within the county. Having validated the model, we estimate its structural parameters and conduct counter-factual simulations. The first simulation indicates that aggregate entry and private investment in China would have been 40% lower over the 1995-2004 period in the absence of the community networks. While the contribution of these informal institutions to Chinese growth has thus been substantial, in line with the anecdotal evidence, this still leaves room for other factors that have been associated with this process, such as government policies, infrastructure, and finance; high saving rates and foreign investment inflows; and the opening of the world market to Chinese exports.

The substantial inter-firm spillovers that we document are unlikely to be fully anticipated or internalized by individual entrepreneurs. This creates scope for industrial policies to stimulate private investment, and this is the subject of our second counter-factual simulation. This experiment, which simulates the effect of a one-time credit subsidy, shows that the optimal strategy to maximize total profits would be to target entrants from higher population density birth counties in order to take advantage of the larger resulting network externalities. There are, however, a number of caveats to such a policy prescription. First, a policy that places weight on both social affiliation and individual merit will only be effective in a population where community networks are already active or have the potential to be activated, and this will depend on the underlying social structure. In particular, the Chinese development experience will not be replicated in other countries by simply providing infrastructure and credit. This is relevant for Chinese overseas development assistance policy, which has largely focussed on infrastructure construction and industrial development (Zhang, 2016).[59] Chinese development assistance has grown exponentially in recent years (Lin and Wang, 2016), but our analysis indicates that the expected returns will only be realized if community networks in the recipient countries evolve in parallel with the infrastructure construction, just as they did in China.

The second caveat concerns the normative consequences of such networks. Our analysis has primarily focused on their positive or descriptive consequences. With regard to efficiency, credit subsidies targeted at particular communities are justified if the resulting network spillovers increase firm productivity (and profits). The potential downside is that communities will start to lobby the government for access to cheap capital; in the extreme case, such rent seeking without accompanying productivity gains could worsen existing distortions in the economy. Moreover, there are important consequences for inequality that need to be considered. By bringing in less able entrepreneurs at the margin, community networks are redistributive within their populations. However, a policy that targets individuals from more socially connected populations to take advantage of the positive externalities that their stronger networks provide will only exacerbate existing inequalities across communities. Given the dynamic increasing returns generated by the networks, these inequalities will persist and, if anything, worsen over time. Absent other redistributive mechanisms, any policy that attempts to exploit network externalities must pay attention to the potentially enduring consequences for inter-community inequality.

[59]This policy is explicitly motivated by the Chinese domestic experience, and the belief that infrastructure construction is the key to development (see, for example, China's second Africa policy paper; Xinhua, December 4, 2015).

References

Aghion, P., Cai, J., Dewatripont, M., Du, L., Harrison, A. and Legros, P. (2015). Industrial Policy and Competition. *American Economic Journal: Macroeconomics*, **7** (4), 1–32.

— and Howitt, P. (1992). A Model of Growth through Creative Destruction. *Econometrica*, **60** (2), 323–351.

Akcigit, U., Alp, H. and Peters, M. (2016). Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries. *NBER Working Paper Series*.

Alesina, A. and Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, **53** (4), 898–944.

Allen, F., Qian, J. and Qian, M. (2005). Law, Finance, and Economic Growth in China. *Journal of Financial Economics*, **77** (1), 57–116.

Asker, J., Collard-Wexler, A. and De Loecker, J. (2014). Dynamic Inputs and Resource (Mis) Allocation. *Journal of Political Economy*, **122** (5), 1013–1063.

Au, C.-C. and Henderson, J. V. (2006). Are Chinese Cities Too Small? *Review of Economic Studies*, **73** (3), 549–576.

Bai, C.-e., Hsieh, C.-T. and Song, Z. (2019). Special deals with chinese characteristics.

Banerjee, A. and Munshi, K. (2004). How Efficiently is Capital Allocated? Evidence from the Knitted Garment Industry in Tirupur. *Review of Economic Studies*, **71** (1), 19–42.

Brandt, L., Kambourov, G. and Storesletten, K. (2016). Firm Entry and Regional Growth Disparities: the Effect of SOEs in China. *University of Toronto Mimeo*.

—, Van Biesebroeck, J. and Zhang, Y. (2012). Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing. *Journal of Development Economics*, **97** (2), 339–351.

Brooks, W. J., Kaboski, J. P. and Li, Y. A. (2016). Growth Policy, Agglomeration, and (the Lack of) Competition. *NBER Working Paper Series*.

Caunedo, J. (2016). Industry Dynamics, Investment and Uncertainty. *Working Paper*.

Ciccone, A. and Hall, R. E. (1996). Productivity and the Density of Economic Activity. *American Economic Review*, **86** (1), 54–70.

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, **94**, S95–S120.

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D. and Roux, S. (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection. *Econometrica*, **80** (6), 2543–2594.

Duranton, G. and Overman, H. G. (2005). Testing for Localization Using Micro-Geographic Data. *Review of Economic Studies*, **72** (4), 1077–1106.

Ellison, G. (1994). Cooperation in the Prisoner's Dilemma with Anonymous Random Matching. *Review of Economic Studies*, **61** (3), 567–588.

— and Glaeser, E. L. (1997). Geographic Concentration in US Manufacturing Industries: a Dartboard Approach. *Journal of Political Economy*, **105** (5), 889–927.

Fang, C. (1997). Recent Trends of Migration and Population in Shandong. *Working Papers on Chinese Politics, Economy and Society*, (10).

Fisman, R., Shi, J., Wang, Y. and Xu, R. (2018). Social Ties and Favoritism in Chinese Science. *Journal of Political Economy*, **126** (3), 1134–1171.

Fleisher, B., Hu, D., McGuire, W. and Zhang, X. (2010). The Evolution of an Industrial Cluster in China. *China Economic Review*, **21** (3), 456–469.

Goodman, B. (1995). *Native Place, City, and Nation: Regional Networks and Identities in Shanghai, 1853–1937*. University of California Press.

Greif, A. (1993). Contract enforceability and economic institutions in early trade: The maghribi traders' coalition. *The American economic review*, pp. 525–548.

— (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of political economy*, **102** (5), 912–950.

— and Tabellini, G. (2017). The Clan and the Corporation: Sustaining Cooperation in China and Europe. *Journal of Comparative Economics*, **45** (1), 1–35.

Haltiwanger, J., Kulick, R. and Syverson, C. (2018). *Misallocation measures: The distortion that ate the residual*. Tech. rep., National Bureau of Economic Research.

Honig, E. (1992). *Creating Chinese Ethnicity: Subei People in Shanghai, 1850-1980*. New Haven: Yale University Press.

— (1996). *Regional Identity, Labor, and Ethnicity in Contemporary China*. Berkeley, California: Institute of East Asian Studies, University of California.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics*, **124** (4), 1403–1448.

— and — (2014). The Life Cycle of Plants in India and Mexico. *Quarterly Journal of Economics*, **129** (3), 1035–1084.

Hu, B. (2008). People's Mobility and Guanxi Networks: A Case Study. *China & World Economy*, **16** (5), 103–117.

Huang, Z., Zhang, X. and Zhu, Y. (2008). The Role of Clustering in Rural Industrialization: A Case Study of the Footwear Industry in Wenzhou. *China Economic Review*, **19** (3), 409–420.

Jackson, M. O., Rodriguez-Barraquer, T. and Tan, X. (2012). Social capital and social quilts: Network patterns of favor exchange. *American Economic Review*, **102** (5), 1857–97.

— and Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? *American Economic Review*, **97** (3), 890–915.

Jones, C. I. (1995). R & d-Based Models of Economic Growth. *Journal of Political Economy*, **103** (4), 759–784.

Kandori, M. (1992). Social Norms and Community Enforcement. *Review of Economic Studies*, **59** (1), 63–80.

Levine, R. and Rubinstein, Y. (2017). Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More? *Quarterly Journal of Economics*, **132** (2), 963–1018.

Lin, J. Y. and Wang, Y. (2016). *Going Beyond Aid: Development Cooperation for Structural Transformation*. Cambridge University Press.

Long, C. and Zhang, X. (2011). Cluster-based Industrialization in China: Financing and Performance. *Journal of International Economics*, **84** (1), 112–123.

Ma, L. J. C. and Xiang, B. (1998). Native Place, Migration and the Emergence of Peasant Enclaves in Beijing. *The China Quarterly*, **155**, 546.

Morten, M. (2019). Temporary migration and endogenous risk sharing in village india. *Journal of Political Economy*, **127** (1), 000–000.

Munshi, K. (2011). Strength in Numbers: Networks as a Solution to Occupational Traps. *Review of Economic Studies*, **78** (3), 1069–1101.

— (2014). Community Networks and the Process of Development. *Journal of Economic Perspectives*, **28** (4), 49–76.

— and Rosenzweig, M. (2006). Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy. *American Economic Review*, **96** (4), 1225–1252.

— and — (2016). Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap. *American Economic Review*, **106** (1), 46–98.

Nee, V. and Opper, S. (2012). *Capitalism from Below: Markets and Institutional Change in China*. Harvard University Press.

Peng, Y. (2004). Kinship Networks and Entrepreneurs in China's Transitional Economy. *American Journal of Sociology*, **109** (5), 1045–1074.

Peters, M. (2016). Heterogeneous Markups, Growth and Endogenous Misallocation. *Working Paper*.

Restuccia, D. and Rogerson, R. (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic Dynamics*, **11** (4), 707–720.

Romer, P. M. (1986). Increasing Returns and Long-run Growth. *Journal of Political Economy*, **94** (5), 1002–1037.

— (1990). Endogenous Technological Change. *Journal of Political Economy*, **98** (5, Part 2), S71–S102.

Ruan, J. and Zhang, X. (2009). Finance and Cluster-based Industrial Development in China. *Economic Development and Cultural Change*, **58** (1), 143–164.

Segerstrom, P. S. (1998). Endogenous Growth without Scale Effects. *American Economic Review*, pp. 1290–1310.

—, Anant, T. C. and Dinopoulos, E. (1990). A Schumpeterian Model of the Product Life Cycle. *American Economic Review*, pp. 1077–1091.

Song, Z., Storesletten, K. and Zilibotti, F. (2011). Growing Like China. *American Economic Review*, **101** (1), 196–233.

Summers, L. (2007). The Rise of Asia and the Global Economy. *Research Monitor*, (2).

Tsai, L. L. (2007). Solidary Groups, Informal Accountability, and Local Public Goods Provision in Rural China. *The American Political Science Review*, **101** (2), 355–372.

Wu, M. (2016). The China, Inc. Challenge to Global Trade Governance. *Harv. Int'l LJ*, **57**, 261.

Zhang, C. (2017). Culture and the Economy: Clan, Entrepreneurship, and Development of the Private Sector in China. *SSRN Scholarly Paper*, (ID 2865105).

— and Xie, Y. (2013). Place of Origin and Labour Market Outcomes Among Migrant Workers in Urban China. *Urban Studies*, **50** (14), 3011–3026.

Zhang, J. (2016). Chinese Foreign Assistance, Explained. Retrieved from https://www.brookings.edu/blog/order-from-chaos/2016/07/19/chinese-foreign-assistance-explained/.

Zhang, X. and Li, G. (2003). Does Guanxi Matter to Nonfarm Employment? *Journal of Comparative Economics*, **31** (2), 315–331.

Zhao, Y. (2003). The Role of Migrant Networks in Labor Migration: the Case of China. *Contemporary Economic Policy*, **21** (4), 500–511.

Appendix A. Derivation of the Adjusted HHI

Suppose that there are n trials, that each outcome j from the set of k possible outcomes has an independent probability of occurring p_j, and that the random variable X_j is the number of occurrences of outcome j. Then the multivariate random variable $\mathbf{X} = (X_1, \cdots, X_k)$ has a multinomial distribution with parameters $(n, k, p_1, \cdots, p_k)$. Applied to our context, (i) n is the total number of entrepreneurs or firms from a given birth county, (ii) k is the total number of clans or destinations that they are allocated to, and (iii) $p_1, \cdots, p_k$ are the probabilities that an entrepreneur or firm allocated randomly would end up in each of those clans or destinations. We assume that there is an equal probability of choosing any clan or destination; $p_j = \frac{1}{k}, \forall j$.

The expected HHI when firms make decisions independently can be expressed as,

$$E(HHI) = E\left(\frac{1}{n^2}\sum_{i=1}^{k} X_i^2\right) = E\left(\frac{1}{n^2}\mathbf{X}^T\mathbf{X}\right).$$

Based on the general properties of the multinomial distribution,

$$E(HHI) = \frac{1}{n^2}\left([E(\mathbf{X})]^T E(\mathbf{X}) + tr[cov(\mathbf{X})]\right).$$

It follows that,

$$E(HHI) = \frac{1}{n^2}\left(k\left(\frac{n}{k}\right)^2 + k\left[n\frac{1}{k}\left(1-\frac{1}{k}\right)\right]\right) = \frac{1}{k} + \frac{1}{n}\frac{k-1}{k}.$$

For large n, $E(HHI) \approx \frac{1}{k}$. For small n, $E(HHI)$ is decreasing in n. We account for this by constructing a normalized HHI statistic, which is simply the unadjusted HHI, based on the observed distribution of entrepreneurs or firms across clans or destinations, divided by $E(HHI)$. If entrepreneurs or firms are allocated randomly, then the adjusted HHI will be close to one, providing a useful benchmark for this statistic.

Appendix B. Entry with Foresight

Consider the consequences of allowing entrepreneurs to look ahead and incorporate profits they would expect to make after the first period they enter. We suppose cohort t agents look ahead one additional period, i.e., make their entry decision based on anticipated present value profits in periods t and $t+1$. The equilibrium can no longer be computed recursively, owing to the need for entrants to coordinate their expectations of entry decisions of one another. We shall consider equilibria where these expectations are fulfilled. We continue to assume that incumbents are committed to their previous entry decisions.

Let ξ denote $\psi r^{-\frac{\alpha}{1-\alpha}}$, and $\delta \in (0,1)$ denote the common discount factor of agents. Then expected present value of entering B_i at t for a cohort t agent of ability ω is

$$P_{it}(\omega) = \omega\xi A_0^{\frac{1}{1-\alpha}} \exp(\theta p n_{i,t-1}\frac{1}{1-\alpha})[1 + \delta\exp(\theta(p)e_{it}\frac{1}{1-\alpha})] \tag{25}$$

while of staying in T is

$$N_{it}(\omega) = \omega^{\sigma}[1+\delta] \tag{26}$$

The agent will enter if

$$\log\omega > \frac{1}{1-\sigma}[-\log\xi - \frac{1}{1-\alpha}\log A_0 + \log(1+\delta) - \frac{1}{1-\alpha}\theta(p)n_{i,t-1} - \log\{1+\delta\exp(\theta(p)e_{it}\frac{1}{1-\alpha})\}] \tag{27}$$

Define the function

$$\begin{aligned} g(e|s_{i,t-1}, n_{i,t-1}, A_{i0}) &= ks_{i,t-1}\{1 + \tfrac{1}{1-\sigma}[\log \xi + \tfrac{1}{1-\alpha} \log A_0 - \log(1+\delta) \\ &\quad + \tfrac{1}{1-\alpha}\theta(p)n_{i,t-1} + \log\{1 + \delta \exp(\theta(p)e\tfrac{1}{1-\alpha})\}]\} \end{aligned}$$

Then equilibrium entry decisions form a fixed point of this function, i.e., $e_{it} = e(s_{i,t-1}, n_{i,t-1}, A_{i0})$ solves

$$g(e|s_{i,t-1}, n_{i,t-1}, A_{i0}) = e \tag{28}$$

The intercept of this function is exactly the entry that results in the myopic equilbrium with $\delta = 0$. The function is increasing in e, with a slope

$$g'(e|s_{i,t-1}, n_{i,t-1}, A_{i0}) = s_{i,t-1} \frac{\delta \exp(\frac{\theta(p)e}{1-\alpha})}{1 + \delta \exp(\frac{\theta(p)e}{1-\alpha})} \frac{k\theta(p)}{(1-\alpha)(1-\sigma)} \tag{29}$$

Hence if

$$\frac{k\theta\bar{p}}{(1-\alpha)(1-\sigma)} < 1 \tag{30}$$

where $\bar{p}$ is an upper bound for p, an equilibrium exists and is unique. Computing the equilibrium is easy, as it involves solving for fixed points of a contracting mapping defined recursively by past entry decisions. It can be easily verified that entry is rising in $s_{i,t-1}, \theta(p)$ and $n_{i,t-1}$, just as in the myopic entry case.

Appendix C: Proofs

Proof of Proposition 1: We first prove that $s_{it} > s_{i-1,t}$ for all t. Suppose this is true at $t-1$: then $s_{i,t-1}$ is rising in i. Denote the growth rate of destination i share: $g_{it} \equiv \frac{s_{it} - s_{i,t-1}}{s_{i,t-1}} = \frac{N_{t-1}}{N_t} + [L + \kappa(p)N_{t-1}]s_{i,t-1} - 1$, upon using (12). Hence g_{it} is rising in $s_{i,t-1}$ and therefore in i, implying $s_{it} > s_{i-1,t}$. So shares are ordered for all cohorts exactly as they are in cohort 0. Also note that all destinations have positive shares in all cohorts, and growth rates cannot be zero at any t for all destinations.

Since $H_t \equiv \sum_i s_{it}^2 = \sum_i s_{i,t-1}^2(1+g_{it})^2 = \sum_i s_{i,t-1}^2 + \sum_i s_{i,t-1}^2 g_{it}^2 + 2\sum_i s_{i,t-1}^2 g_{it}$, it follows that

$$\begin{aligned} H_t - H_{t-1} &= \sum_i s_{i,t-1}^2 g_{it}^2 + 2\sum_i s_{i,t-1}^2 g_{it} \\ &> 2\sum_i s_{i,t-1}^2 g_{it} \\ &> 0 \end{aligned}$$

where the first inequality follows from the fact that all sector shares are positive and growth rates are not all zero. The second inequality follows from observing that: (i) if we define $x_{it} \equiv s_{i,t-1}g_{it} = s_{it} - s_{i,t-1}$ then $\sum_i x_{it} = 0$; (ii) $\sum_i s_{i,t-1} = 1$, and (iii) x_{it} and $s_{i,t-1}$ are both increasing in i, as explained above. Hence by a standard argument[60] $\sum_i s_{i,t-1}^2 g_{it} = \sum_i s_{i,t-1}x_{it} > 0$, which proves that concentration is rising in t, and hence (using (13)) the same is true for E_t.

[60] The distribution across destinations first order stochastically dominates the uniform distribution, in which $s_{i,t-1}$ is the same for all i, and the expected value of x under the uniform distribution equals zero. Hence the expected value of x must be positive.

Proof of Proposition 2: The increase in E_t, N_t, H_t with t follows from Proposition 1. So consider how a higher p alters the dynamics, given initial conditions. We claim that it raises aggregate entry E_t (and hence N_t) as well as H_t at every date t. This follows from an inductive argument. Observe first that it must be true for E_t (and N_t) at $t = 1$, given the initial conditions N_0, H_0, upon applying equation (13) at $t = 1$. Next observe that the right-hand-side of (14) is rising in p, given any N_{t-1} and $s_{1,t-1} > \frac{1}{2}$. Hence s_{11} must be rising in p, given the initial conditions. So the result holds for H_t at $t = 1$. Next suppose it holds until some date $t-1$, i.e., N_{t-1} and H_{t-1} are rising in p. Equation (13) then implies E_t (and N_t) is rising in p. Moreover, observe that the right-hand-side of (14) is rising in N_{t-1} and in $s_{1,t-1}$, given p and $s_{1,t-1} > \frac{1}{2}$. The share s_{1t} will then be increasing in p because it is increasing in $s_{1,t-1}$, N_{t-1} and $\kappa(p)$ respectively. Induction now ensures this will be true at every t. This establishes part (a) of Proposition 2.

Turn now to part (b). Taking first differences of (13)

$$E_{t+1} - E_t = \kappa(p)[N_t H_t - N_{t-1} H_{t-1}] = \kappa(p)[E_t H_t + N_{t-1}(H_t - H_{t-1})] \tag{31}$$

Since $\kappa, E_t, H_t, N_{t-1}$ are all rising in p, the result would hold for entry if it were also true for concentration (i.e., $H_t - H_{t-1}$ is rising in p). A sufficient condition for this to hold is that it is true for s_{1t}: i.e., if $s_{1,t} - s_{1,t-1}$ is rising in p (since $H_t - H_{t-1} = 2(s_{1t} - s_{1,t-1})(s_{1t} + s_{1,t-1} - 1)$, and we have already shown that $s_{1t}, s_{1,t-1}$ are rising in p).

Now observe that (14) can be rewritten as

$$s_{1t} - s_{1,t-1} = \kappa(p) N_{t-1} \frac{(2s_{1,t-1} - 1)(1 - s_{1,t-1}) s_{1,t-1}}{(L + N_{t-1})(2 - s_{1,t-1}) + \kappa(p) N_{t-1}(s_{1,t-1}^2 + 1 - s_{1,t-1})} \tag{32}$$

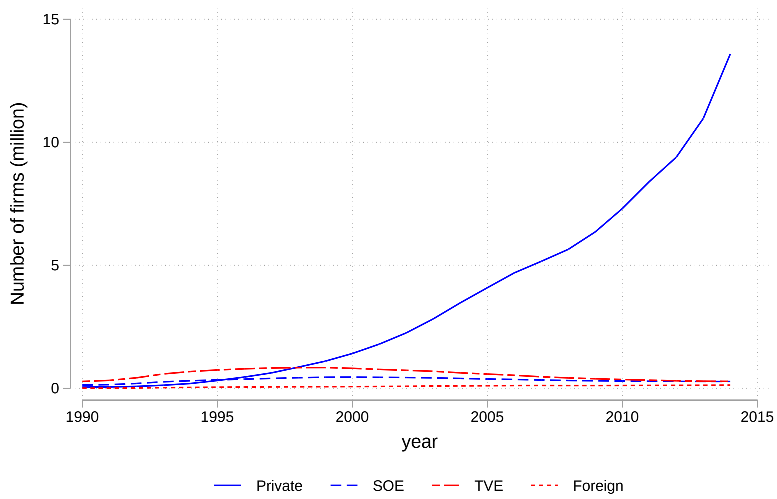
$\kappa(p) < 1$ implies that the denominator of the right-hand-side of (32) is decreasing in $s_{1,t-1}$. And the numerator is increasing in $s_{1,t-1}$ if $s_{1,t-1} < \frac{3}{4}$ (since this implies $s_{1,t-1}(1 - s_{1,t-1}) > \frac{1}{6}$). Then $s_{1t} - s_{1,t-1}$ is rising in $s_{1,t-1}$, as well as in N_{t-1} and κ. Part (b) then follows from the fact that $s_{1,t-1}, N_{t-1}$ are rising in p.

Proof of Proposition 3: To verify (a), observe that averaging (15) across destinations (and noting that $\sum_i s_{i,t-1} = 1$), initial capital of the marginal entrant is decreasing in t, p, and $p * t$ because $\theta(p)$ is increasing in p and N_{t-1} is increasing in t and p (more steeply over time).[61] A similar argument operates for ability and size of the average entrant from (16) and taking the average across destinations. Part (b) follows from averaging across destinations in (18), and applying Propositions 1 and 2.

[61] N_t is increasing in t (for any given p) and in p (for any given t) from Propositions 1 and 2. $N_t - N_{t-1} \equiv E_t$, which is increasing in p from Proposition 2, hence, the cross-partial derivative of N_t with respect to p and t is positive.

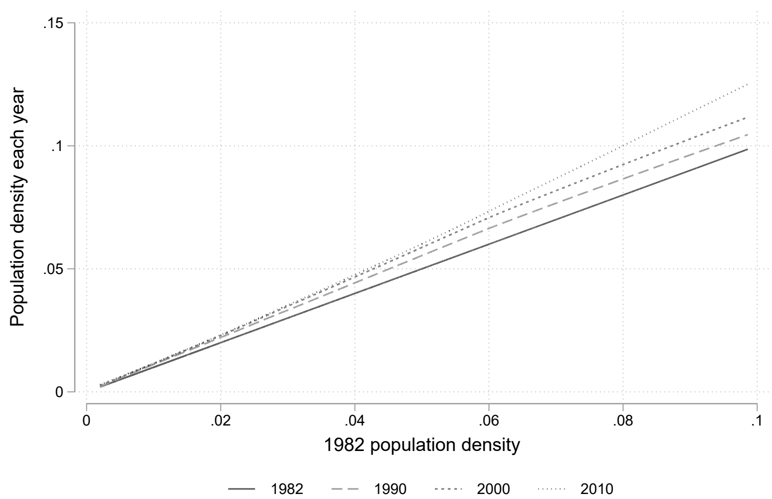
Appendix D. Figures

Figure D.1. Number of Firms, By Type



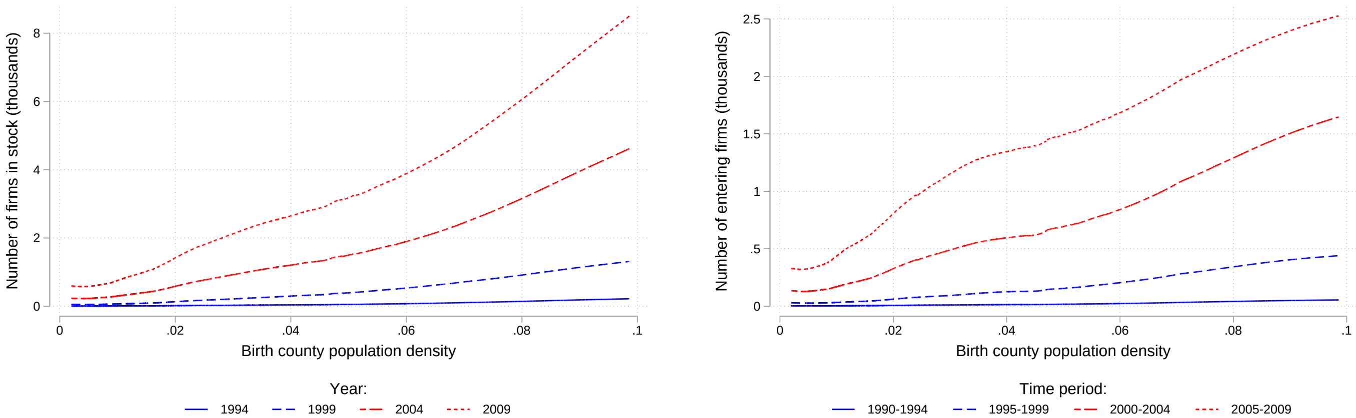
Source: SAIC Registration Database.

Figure D.2. Population Density over Time



Source: Registration Database and 1982 population census.

Figure D.3. Firm Entry and Population Density

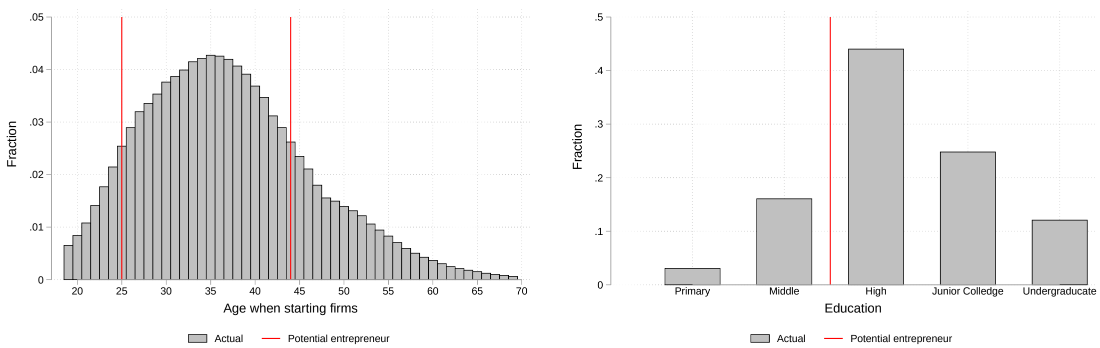


(a) Stock of Firms

(b) Firms Located Outside the Birth County

Source: Registration Database and 1982 population census.

Figure D.4. Education and Age Distribution of Entrepreneurs



(a) Age

(b) Education

Source: SAIC registration database.

Appendix E. Tables

Table E.1. Composition of Listed Individuals in the Firm, by Birth Place Population Density

Dependent variable:	fraction of listed individuals from the legal representative's birth county	
Birth place:	couty	city
	(1)	(2)
Birth place population density	0.010***	-0.017***
	(0.002)	(0.001)
Mean fraction	0.741	0.606
Counter-factual fraction with random assignment	0.0560	0.0916
Observations	490,273	245,302

Source: State Administration of Industry and Commerce (SAIC) registration database.
Sample restricted to firm's operating outside their legal representative's birth county.
Counter-factual fraction with random assignment for such a firm is measured by the number of listed individuals from the legal representative's birth county in its sector-location divided by the total number of listed individuals (among all firms whose legal representatives were born outside their birth county) in that sector-location. The sector is measured at the 2-digit level and location is the county or urban district.
"Listed individuals" includes major investors and top managers, but excludes the legal representative.
Population density in the county or city based on the 1982 population census.
Population density is measured in units of 10,000 people per square km, and then converted to Z-score.
Standard errors clustered at the county or city level reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Table E.2. The Effect of Initial Entry on Subsequent Entry (for firms located outside the birth county)

Dependent variable:	subsequent entrants from the birth place					
Birth place:	county		county		city	
Time period:	2000-2004	2005-2009	2000-2004	2005-2009	2000-2004	2005-2009
	(1)	(2)	(3)	(4)	(5)	(6)
Initial entrants from the birth place	5.481***	8.273***	3.303***	5.883**	6.813***	5.152***
	(1.044)	(1.948)	(1.197)	(2.470)	(0.506)	(0.507)
All initial entrants at the destination	0.294***	0.442***	–	–	–	–
	(0.036)	(0.047)				
Initial entrants from the birth place × birth place population density	–	–	1.820***	1.880*	0.081	-0.083
			(0.507)	(0.966)	(0.087)	(0.080)
Distance to the birth place	–	–	-1.458***	-1.723***	-2.831***	-2.609***
			(0.077)	(0.069)	(0.155)	(0.101)
Mean of dependent variable	1.964	2.124	1.964	2.124	3.382	2.926
Origin-sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Location fixed effects	No	No	Yes	Yes	Yes	Yes
Observations	334,843	713,935	334,843	713,935	289,708	420,138

Note: number of entrants outside the birth place is measured at the (two-digit) sector-destination level.
Initial entry is derived over the 1990-1994 period.
Number of entrants is obtained from the SAIC registration database and birth county population density is derived from the 1982 population census.
Standard errors clustered at birth county-sector level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Table E.3. Clan Concentration and Population Density (time and interaction effects)

Dependent variable:	clan concentration	
Sample:	all firms	firms located outside the birth county
	(1)	(2)
Time period	1352***	1.027***
	(0.014)	(0.012)
Birth county population density × time period	0.254***	0.238***
	(0.015)	(0.014)
Observations	6,482	6,300

Note: the estimating equation includes, in addition, birthplace population density and a constant term.
Clan is defined by the same surname and the same birth county.
Clan concentration measured by Herfindahl Hirschman Index (HHI) across clans divided by the expected HHI that would be obtained by random assignment.
Population density is measured in units of 10,000 people per square km, and then converted to Z-score.
Time period is an ordinal variable taking value from 1 to 4 corresponding to successive five-year time windows over the 1990-2009 period.
Clan concentration is derived from the SAIC registration database and population density is derived from the 1982 population census.
Standard errors clustered at birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Table E.4. Marginal Ability, Marginal Initial Capital and Population Density

Dependent variable:	marginal ability				marginal initial capital			
Time period:	1990-1994	1995-1999	2000-2004	2005-2009	1990-1994	1995-1999	2000-2004	2005-2009
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth county population density	-1.829	-2.685***	-3.570***	-1.435***	-0.100***	-0.118***	-0.134***	-0.041***
	(1.369)	(1.103)	(0.807)	(0.755)	(0.030)	(0.026)	(0.017)	(0.010)
Mean of dependent variable	66.07	52,47	40.79	40.15	-0.803	-1.093	-1.669	-2.225
Observations	4,079	6,595	10,354	11,137	15,601	46,877	83,276	99,877

Note: the entrepreneur's ability is measured by his percentile rank in his birth county-birth cohort education distribution.
The marginal entrepreneur is defined as the individual at the bottom one percentile of the ability distribution among entering entrepreneurs in a given birth county-sector-time period. Marginal initial capital defined as the bottom one percentile of the initial capital distribution at the birth county-sector-time period level.
Control variables include population, education and occupation distribution in the birth county.
Standard errors clustered at birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Table E.5. Estimates based on Actual and Model Generated Data

Dependent variable:	number of entrants				average initial capital			
Data:	actual	model generated (myopic)	model generated (forward looking)	model generated (with sector-level spillovers)	actual	model generated (myopic)	model generated (forward looking)	model generated (with sector-level spillovers)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth county population density	0.335***	0.215***	0.109***	0.171***	-0.010**	-0.002**	-0.003**	-0.003**
	(0.034)	(0.030)	(0.016)	(0.025)	(0.005)	(0.001)	(0.001)	(0.001)
Time period	0.476***	0.305***	0.095***	0.369***	-0.202***	-0.191***	-0.151***	-0.191***
	(0.020)	(0.021)	(0.011)	(0.019)	(0.009)	(0.002)	(0.003)	(0.002)
Observations	3,032	3,032	3,032	3,032	18,401	18,401	18,401	18,401

Note: Estimates based on two time periods, 1995-1999 and 2000-2004.
Firm entry (in thousands) measured at the birth county-time period level. Initial capital (in million Yuan) is measured in logs. Average initial capital is the average of the initial capital distribution at the birth county-sector-time period level.
Time period is a binary variable taking the value one for the 2000-2004 period and zero for the 1995-1999 period.
Actual firm entry and initial capital are obtained from the SAIC registration database and birth county population density is derived from the 1982 population census.
Standard errors clustered at the birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.