

# THE COMMUNITY ORIGINS OF PRIVATE ENTERPRISE IN CHINA

By RUOCHEN DAI, DILIP MOOKHERJEE, KAIVAN MUNSHI AND XIAOBO ZHANG\*

*This paper identifies and quantifies the role played by birth-county-based community networks in the growth of private enterprise in China. We document that historically determined population density is positively associated with enforceable trust in Chinese counties and the resulting quality of the business networks that emerged from those counties after privatization. This motivates a model of network-based spillovers that predicts how the dynamics of firm entry, concentration, and size vary with birth county population density, independently of other factors such as government infrastructure and agglomeration effects in the locations where firms are established. 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 overall firm entry in the economy over the 1995-2004 period and the associated capital stock in 2004 would have been 11% and 12.5% lower without the rural hometown networks. Additional counter-factual simulations shed light on industrial policy.*

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 an increase in the number of State Owned Enterprises (SOE's) and then 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 (as documented in Figure 1a). 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 (see Figure 1b).

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, Qian and Qian, 2005). While the government played an important role in China's economic transformation by providing infras-

\* Dai: Central University of Finance and Economics, r.dai@cufe.edu.cn. Mookherjee: Boston University, dilipm@bu.edu. Munshi: Yale University, kaivan.munshi@yale.edu. Zhang: Peking University x.zhang@nsd.pku.edu.cn. We are grateful to numerous seminar and conference participants and discussants for their constructive comments. Research support from the Economic Development and Institutions (EDI) RA4 program, Cambridge-INET, China National Science Foundation (project number 71874008) and the Agence Nationale de la Recherche (ANR) under the EUR Project ANR-17-EURE-0010 is gratefully acknowledged. We are responsible for any errors that may remain.

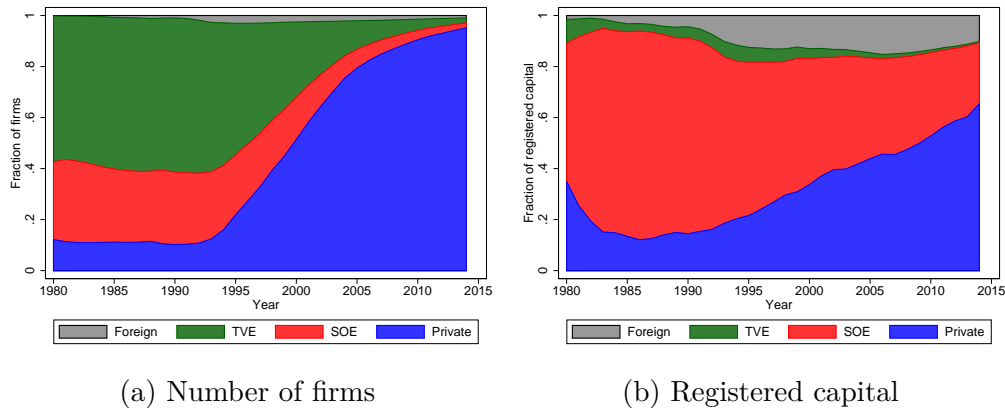


Figure 1. Distribution of Firms, by Type

Source: State Administration of Industry and Commerce (SAIC) registration database.

structure 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, Qian and Qian, 2005; Greif and Tabellini, 2017). There are many accounts of how social networks organized around the hometown supported China’s historically unprecedented rural-urban migration over the past decades; e.g. Zhao (2003), Hu (2008), complementing an older sociological literature that takes the position that ethnicity in China is defined by the native place (Honig, 1992, 1996; Goodman, 1995). Chambers of commerce that bring entrepreneurs from the same origin together (*yidi shanghui*) are commonly found in Chinese cities. Based on anecdotal accounts in the media and Chinese academia, entrepreneurs from a single rural county sometimes dominate an entire industry. While these are exceptional cases, we will use comprehensive administrative data covering the universe of registered firms to document the important role played by hometown networks in supporting entrepreneurship in China.

A longstanding economics literature describes how firms respond to the difficulty in enforcing formal contracts in developing economies by establishing relational contracts (McMillan and Woodruff, 1999), building market reputation (Banerjee and Duflo, 2000), and vertically integrating (Woodruff, 2002). More closely related to our analysis, a complementary literature has studied how firms exploit pre-determined ethnic ties to overcome market imperfections in these economies; e.g. Fafchamps (2000); Fisman (2003); Banerjee and Munshi (2004); Munshi (2011). Recent contributions to the literature on firms in developing countries have exploited sophisticated research designs to examine the determinants and the evolution of relational contracts (Macchiavello and Morjaria, 2015, forthcoming) and vertical integration (Hansman et al., forthcoming). Although much progress has been made, this new literature does not engage with the extant literature on ethnic business networks and neither does the emerging literature on networks in developing countries. The latter literature is largely focussed on

information diffusion and program targeting; e.g. Banerjee et al. (2013); Alatas et al. (2016); Banerjee et al. (2019). Two prominent papers, Karlan et al. (2009) and Jackson, Rodriguez-Barraquer and Tan (2012) do examine cooperation in networks, but not between firms. Our research advances the literatures on firms in developing countries and networks in developing countries by identifying and quantifying the contribution of community-based business networks to China’s remarkable economic transformation over a period of two decades, starting from the very onset of privatization in the early 1990’s.

We start the analysis by constructing a population-based measure of trust. Building on an older literature on norms and community enforcement and a more recent literature on cooperation in networks, discussed below, we argue that population density in rural areas is positively associated with the frequency of local social interactions, which gives rise to more connected social networks and higher levels of trust and economic cooperation. The localized trust that we focus on is distinct from the generalized trust that has received much attention in the rapidly growing economics literature on culture (Alesina and Giuliano, 2015). Localized trust is restricted to neighbors rather than the general population, and is sustained via external enforcement rather than by internalized cultural values. We provide evidence linking population density to localized (but not generalized) trust, at the country level with data from the World Values Survey (WVS) and at the county level with nationally representative data from the China Family Panel Study (CFPS).

While we uncover a positive relationship between population density and both social interactions and trust in Chinese counties, this relationship does not extend to cities. This can be explained by the offsetting effect of social heterogeneity, which is shown to be increasing in urban (but not rural) population density. The empirical analysis, linking birthplace population density to entrepreneurship, consequently focuses on county-born businessmen; their firms account for two-thirds of all registered private firms, and a comparable share of private registered capital, in China. Although the majority of county-born businessmen establish their firms outside their birth counties, our assumption is that these businessmen remain connected to their origin communities and that business networks drawn from higher population density counties will support higher levels of mutual cooperation among their members regardless of where they are located. We provide empirical support for these assumptions using comprehensive administrative data obtained from the State Administration of Industry and Commerce (SAIC) that covers the universe of registered firms. Key personnel and shareholders in firms located outside the birth county are disproportionately drawn from the origin county (the level of homophily is 40-50 times greater than what would be obtained with random assignment). Moreover, this homophily, the number of links to other firms at the destination, the fraction of these links that are with firms from the same birth county, and the strength of the network links (based on a measure of cooperation – *support* – proposed by Jackson, Rodriguez-Barraquer and Tan (2012)) are all shown to be increasing in population density.

Having established that birth county-based business networks are active and that their quality is increasing in population density, we next proceed to demonstrate that these networks have consequences for firm-level decisions and outcomes that are relevant for aggregate growth. To do this, we develop a dynamic model

that features successive cohorts of agents who are heterogeneous in their abilities and must make a choice between a traditional occupation and entrepreneurship. The payoff from entrepreneurship depends, in addition to individual ability, on the contribution of the network to productivity via mutual help.<sup>1</sup> When sector-location choice is incorporated in the model, an additional channel for network-based spillovers opens up, via a pre-entry referral process as in Chaney (2014), which increasingly directs 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, which is increasing in network quality, proxied by its exogenous determinant – origin population density – in the empirical analysis. While the increasing returns embodied in our network-based model can explain the explosive growth in the total number of firms that is observed in China’s early industrialization phase, the model generates the additional predictions that both entry and concentration (across sectors and locations) will be increasing in birth county population density at each point in time, with a slope that is increasing over time.

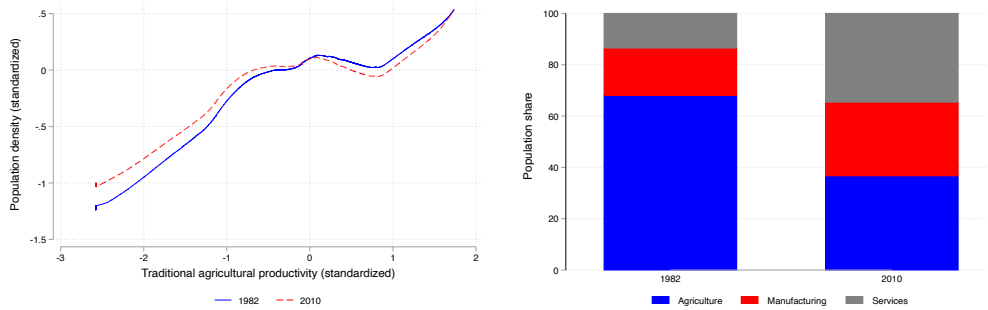
With regard to capital investment, the network-based spillovers that raise productivity over time have two conflicting effects on the initial size of the marginal entrant’s firm: The direct effect, for a given level of ability, is to increase firm size by raising the firm’s TFP. However, an increase in network size and quality 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. Once a firm has entered, its growth will be driven entirely by changes in network size in our model. Because (higher quality) networks from higher population density birth counties are growing faster, firms from those counties will start small but subsequently grow faster.

The third step in the analysis tests the predictions of the model over the 1990-2009 ‘early industrialization’ phase with the administrative data that were used for the analysis of homophily and network quality described above. These data are linked to the industrial census and the SAIC inspection database for the analysis of firm growth. An especially useful feature of the administrative data is that the birth counties of key personnel and shareholders, including the legal representative who we designate as the “entrepreneur,” can be recovered for each firm. The model generates cross-sectional and dynamic predictions for the relationship between birth county population density, which we measure in 1982 (prior to the onset of privatization) and a rich set of outcomes – firm entry, sectoral and spatial concentration, initial firm size, and firm growth – and the data match each of them.

One advantage of taking a step back and using a population characteristic – population density – as the source of forcing variation in the analysis is that we

<sup>1</sup>The network could also improve the outcomes of its members by providing cheap credit (Banerjee and Munshi, 2004). As shown below, all the results of the model would go through with this alternative role for the network and, hence, we do not attempt to disentangle these (possibly coexisting) mechanisms.

learn something more fundamental about network quality. An additional advantage is that population density, unlike more direct measures of network quality such as degree and support, is not jointly determined with outcomes of the model such as entry and concentration. Nevertheless, it is useful to examine the determinants of population density in rural counties. In our analysis, we measure population density in 1982, when the Chinese economy (after decades of stalled industrialization) was still largely agrarian. Prior to industrialization, agricultural productivity would have determined the population density that could be supported in a local area (Galor and Weil, 2000). The pre-modern economy was essentially stagnant for centuries, and while there were improvements in agricultural technology in the twentieth century, we would expect rural population densities to be largely determined by historical productivity (in the distant past). To verify this assumption, we follow Galor and Özak (2016) and construct an exogenous measure of (potential) agricultural productivity, at the county level, that is based on traditional (pre-modern) technology. As observed in Figure 2a, population density in 1982 is increasing in this measure of agricultural productivity across Chinese counties. As observed in Figure 2b, agriculture was the dominant occupation in China at this time, but the occupational structure changed entirely by 2010 (the end point of our analysis). Despite this structural transformation, and the accompanying rural-urban migration, we see that 2010 population density tracks closely with the corresponding 1982 statistic.



(a) County population density and tra- (b) Occupational structure in the Chi-  
ditional agricultural productivity nese economy

Figure 2. Determinants and Stability of Population Density

*Source:* Source: Chinese population census, FAO-GAEZ database. See Appendix A for variable construction.

Our empirical strategy is based on the idea that rural population densities, which were determined by historical agricultural conditions and which have remained stable over time, had unanticipated effects on the quality of the new business networks that emerged in a rapidly developing economy. A variant of this strategy has been employed in the literature on ethnic trading networks, which goes back to Rauch and Trindade (2002). Recent contributions to this literature; e.g. Cohen, Gurun and Malloy (2017); Burchardi, Chaney and Hassan (2019) exploit historical accidents to generate exogenous variation in ethnic

links between specific origin-destination pairs that, in turn, determine the level of trade. In our analysis, what matters is not the strength of origin-destination ties, but the quality of the origin-based networks that form at different locations, with consequences for a variety of firm-level decisions and outcomes.

The qualifier to the preceding argument is that county population density, even if it is predetermined and had unanticipated effects on network quality, could still be correlated with other factors that independently determine the outcomes of interest. This is especially relevant because pre-modern population density has been associated with the contemporaneous level of economic development (Galor and Weil, 2000). Counties with higher population densities (today and in the past) might well have been more developed historically, with long-term consequences for human capital accumulation and the structure of the local economy. Population, education, and the occupational structure in the birth county (measured in 1982) are thus included as exogenous controls in the estimating equations, but such conditioning may not account for all relevant factors. An important feature of our analysis is that we comprehensively examine alternative non-network explanations by systematically relaxing different assumptions of the model, which introduces additional sources of heterogeneity at the origin and the destination. We show that no other explanation can account for all of our results simultaneously.

As discussed below, origin heterogeneity of any sort cannot explain the increased clustering over time, in particular sectors and locations, that is predicted and documented for firms from higher population density birth counties. Nor can it explain why firms from those counties enter smaller and then grow faster. There must be some force at the destination that boosts post-entry growth. In our model, it is the endogenously evolving network, but an alternative non-network mechanism based on destination heterogeneity could also generate this result if firms from higher population density birth counties have preferred access to destinations that were exogenously growing faster for other reasons, such as improved infrastructure or destination-based productivity spillovers. To address this concern, we exploit the fact that firms from many birth counties are established in the same sector-location. Sector-time period and location-time period effects are thus included as controls in the estimating equations, without affecting the results.

Having tested and validated the network-based model, the final step in the analysis seeks to quantify the impact of these networks on aggregate firm entry and capital stocks by estimating the structural parameters of the model. Given that firms from multiple origins were established in each destination (sector-location) and that the structural equations are linear in variables, it is possible to control flexibly for local government and agglomeration effects by including destination-time period effects in the estimating equation. Consistent with the reduced form results, which find no evidence that entrepreneurs from higher population density birth counties have preferred access to faster growing destinations, the destination controls have little effect on the estimated structural parameters. Although the model is extremely parsimonious, it does a good job of matching entry and initial capital across the range of birth county population densities, during a period (1995-2004) in which the Chinese economy was growing at an explosive rate. This increases our confidence in the results of a counter-factual experiment, which estimates that overall entry in the Chinese economy over the 1995-2004 period would have declined by 11% and the capital stock in 2004 would have declined by

12.5% in the absence of the rural hometown networks.

We conclude by discussing the implications of the networks for industrial policy. There is a common view among academics and policymakers; e.g. Morrison (2019) that Chinese economic growth has been fuelled by the availability of subsidized capital and that particular firms are favored by the government. We cannot rule out the possibility that birth county networks are lobbying government official to receive cheap capital, rather than enhancing the productivity of their members. A counter-factual policy experiment that exploits network spillovers by providing a temporary credit subsidy to entering firms indicates, nevertheless, that the government would want to subsidize higher quality networks, drawn from higher population density counties, even further if its objective is to maximize total profit (surplus). These findings must, however, be placed in context. Policies that target specific communities will have complex distributional consequences and it is not obvious that they will be effective in all economies (societies). Moreover, as discussed in the final section, the long-term impact of the networks on growth and mobility is ambiguous.

### I. The Network Mechanism

The point of departure for our analysis is the assumption that localized trust in relatively sparsely populated rural areas is increasing in population density. The underlying idea is that higher population density, which is mechanically correlated with 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,  $\gamma$ , is rising in population density. A higher  $\gamma$  directly raises the degree of the social network (the number of links per capita), and indirectly also network 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  $\gamma$ . Coleman (1988) argues that network closure is a necessary condition for economic cooperation, enforced by social sanctions. Jackson, Rodriguez-Barraquer and Tan (2012) make a similar argument based on a related network property, which they refer to as “support.” A link in a network is supported if the two nodes have mutual links to a third node.<sup>2</sup> In general, the level of cooperation will be increasing in the number of mutual links that “support” a given link, which, in turn, will be increasing in  $\gamma$ .

The preceding discussion implicitly assumes that the population is socially homogeneous. Matters are more complex when the population is fragmented into

<sup>2</sup>The intuition for this measure of network cooperation is that when a link is supported, a unilateral deviation can result in the loss of multiple links. For example, consider the triad,  $i, j, k$ . Suppose that bilateral cooperation cannot be sustained, but that cooperation is possible if a deviation results in the loss of at least two links. If agent  $i$  reneges on his obligation to agent  $j$ , or vice versa, that link is severed (by assumption). Both  $i$  and  $j$  are no longer trustworthy because they are left with a single link each and, hence, agent  $k$  will sever his links to both of them. Unilateral deviation by  $i$  or  $j$  and, by the same argument, by  $k$ , results in the loss of two links, allowing cooperation to be sustained.

smaller communities. Suppose that individuals only interact within their communities and that social sanctions consequently only apply within those communities. If communities are not (perfectly) spatially segregated, then the frequency of social interactions, the effectiveness of social sanctions, and resulting enforceable trust are all decreasing in social heterogeneity. If social heterogeneity is uncorrelated with population density, as documented below in Chinese counties, then both social interactions and enforceable trust will be unambiguously increasing in population density. However, social heterogeneity could potentially be increasing in population density in cities. Most urban residents in developing countries are recent arrivals, typically from many different origins. Greater population density in an urban neighborhood might well be associated with a larger migrant presence, and a greater diversity of origins, in which case the relationship between population density and both social interactions and enforceable trust will be ambiguous.

To provide empirical support for each component of the preceding argument, we begin by estimating the relationship between trust and population density. A detailed description of the variables used for the analysis in the current section, some of which are also used in subsequent sections, is provided in Appendix A. The trust-population density relationship is not China-specific and, hence, we expect it to be observed across a wide cross-section of countries. We first measure trust with data from the World Values Survey (WVS); while the advantage of these data is that they cover many countries, one limitation is that responses from rural and urban residents cannot be distinguished. We (partially) address this limitation by only including large developing countries, with large rural populations, in the sample. Figure 3a presents a binned scatter plot describing the relationship between trust in local residents, which measures localized enforceable trust, and population density (obtained from the World Development Index) for the 31 countries in our restricted sample. This relationship is strongly positive and statistically significant (based on regression estimates not reported). Figure 3b presents a binned scatter plot describing the relationship between trust in outsiders; i.e. individuals that the respondent would meet for the first time, which measures generalized trust, and population density. No relationship can now be detected and this is also true for the companion regression estimates (not reported).

We subject the preceding results to closer scrutiny by repeating the analysis with data from the China Family Panel Studies (CFPS). The CFPS is a nationally representative, longitudinal, general social survey conducted at the individual, household, and community level that was launched in 2010, with subsequent rounds in 2012 and 2014. Different rounds of the CFPS include one-off modules on trust, social interactions, and social heterogeneity that we utilize for the analysis that follows. The advantage of the CFPS, apart from its obvious relevance for our China-based analysis of entrepreneurship is that the relationship between trust and population density can be estimated separately in counties and cities.<sup>3</sup> 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. The adult

<sup>3</sup>There are approximately 2,000 counties and 250 prefecture-level and province-level cities (which are further divided into urban districts) in China.



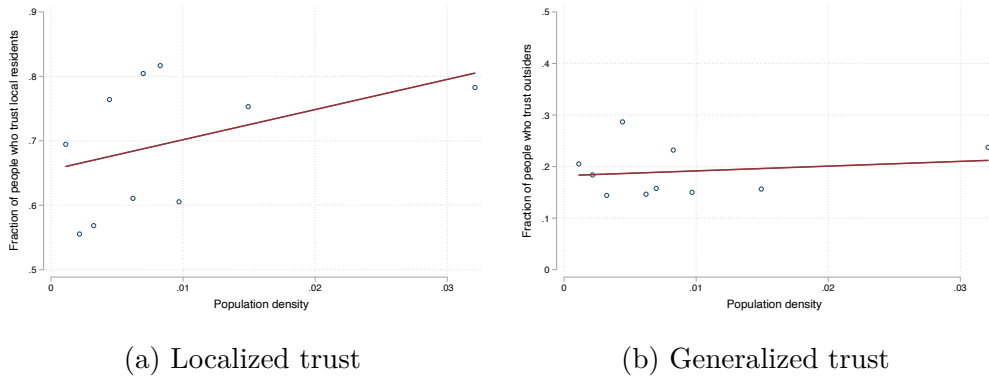


Figure 3. Trust and Population Density: Cross-Country Comparison

Source: World Values Survey and World Development Index.

individual module of the 2012 CFPS collected information on trust, which we aggregate up to the county or city level. As with the WVS, the analysis distinguishes between trust in local residents, which measures localized trust, and trust in outsiders, which measures generalized trust. Table 1 reports the estimated relationships between trust and population density, controlling for population, education, and occupational structure in the county or city (also measured in 1982). These features of the local economy, which are correlated with population density, could potentially determine trust independently. We see that the only economically and statistically significant parameter estimate in the table is the coefficient on county population density, with localized trust as the outcome. The results with the CFPS thus match the cross-country results with the WVS, but only in rural counties.<sup>4</sup>

Table 1— Trust and Population Density (China)

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	0.026*** (0.005)	0.004 (0.012)	-0.001 (0.010)	-0.006 (0.015)
Mean of dependent variable	0.883	0.289	0.876	0.252
Observations	93	93	39	39

Note: Trust measures are obtained from the adult individual module of the China Family Panel Study (2012), converted to a binary variable and then aggregated up to the county/city level. Control variables include population, education and occupation distribution in the county or city. Population density, population, education and occupation distribution are computed from the 1982 population census. Population density is measured as a Z-score. Standard errors clustered at the county or city level are reported in parentheses. \* significant at 10%, \*\* at 5%, \*\*\* at 1%.

<sup>4</sup>The binned scatter plots corresponding to Table 1, but without the controls, are reported in Appendix Figures E1a and E1b. They are broadly consistent with the estimated coefficients in the table and the cross-country relationships in Figures 3a and 3b.

In our framework, localized trust is determined by social interactions. In light of the preceding results, we would expect local social interactions to be increasing in population density in counties but not cities. Appendix Table F1 verifies that this is indeed the case, using data from the family module of the 2010 CFPS. One explanation for the weak relationship between urban population density and both social interactions and localized trust is that dense urban neighborhoods are more socially heterogeneous. We see in Appendix Table F2, using data from the community module of the 2010 CFPS, that 90% of county residents are born in the village where they currently reside, whereas less than 50% of urban residents are born in their neighborhood. Urban neighborhoods will thus be comprised of individuals from diverse origins and we see in the table that this social heterogeneity, measured by the fraction of migrants, is increasing in neighborhood population density, which, in turn, is increasing in city population density. In contrast, the fraction of migrants is uncorrelated with rural population density.

Population density is not a good measure of trust in urban areas and, based on the preceding results, this is because of its offsetting effect on social heterogeneity. Our analysis of community-based entrepreneurship, which relies on population density as an exogenous source of variation, will thus focus on county-born businessmen. Entrepreneurs in developing countries rely on each other for credit, connections to buyers and suppliers (who also provide credit) as well as for information about new technologies and markets. This type of informal support is difficult to sustain through the market mechanism, due to the inherent problem of verifying help sought and received, coupled with a weak legal environment. Cooperation is based instead on community enforcement, 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.

We use comprehensive administrative data, obtained from the State Administration of Industry and Commerce (SAIC) to provide support for the preceding assumption and to show that business networks drawn from higher population density counties are of higher quality. The SAIC database lists the key personnel and the major shareholders in each registered firm. We designate the firm’s legal representative, who typically functions as its president, chairman, or proprietor as the “entrepreneur” for the purpose of the empirical analysis and his birth county, which can be recovered from his citizenship ID, thus applies to the firm as a whole. In addition to the legal representative, the SAIC database also lists other key personnel in the firm: directors, senior managers, and “supervisors.” The last group is especially relevant for the current analysis because the supervisors are meant to provide external oversight, but can also serve as a bridge between firms. These individuals are often key internal personnel within one firm and external supervisors in another or, alternatively, supervisors in multiple firms. We define two firms (located in the same prefecture) as being linked if the same individual is listed in both of them. Firms with cross-linked personnel can support higher levels of cooperation. If these links are between firms from the same origin, then the social network at the origin can be used to support even higher levels of cooperation. Based on the analysis above, links between firms drawn from higher population density counties are more likely to be supported by their (denser) origin social networks. Given the greater value from such mutual links, we expect

firms drawn from higher population density counties to be more likely to establish cross-linkages and (conditional on creating a link) to be more likely to link with firms from the same origin.<sup>5</sup>

We begin our test of the preceding hypothesis in Table 2, Panel A, Column 1 by estimating the relationship between the fraction of firms that list individuals (other than the legal representative) and birth county population density. The sample in this table and Table 3 that follows is restricted to firms located outside their birth county since the additional objective is to establish that firms retain their ties to the origin when they are established elsewhere.<sup>6</sup> We also restrict attention to firms that were active in 2009, the end point of our dynamic analysis, to emphasize the persistence of these connections. As observed in Column 1, the coefficient on birth county population density is positive and significant. This is not because firms from those counties are larger and, therefore, more likely to list multiple individuals. Closer inspection of the data indicates that this result is driven by the fact that they are more likely to list supervisors who, as noted, can serve as a bridge between firms. Moreover, conditional on listing multiple personnel, the fraction of personnel born in the same county as the legal representative is increasing in population density in Table 2, Panel A, Column 2. Notice that this fraction, which measures birth county homophily, is on average 50 times larger than what would be obtained by random assignment of listed individuals across firms in the location (prefecture) where each firm is established.

Given that firms from higher population density counties are more likely to list multiple individuals, particularly supervisors, we expect that they will mechanically have a greater propensity to be linked to other firms in the locations where they are established. Since links between firms from higher population density counties are more likely to be supported by the origin social network, there is an additional motivation for such links (as discussed above). As observed in Table 2, Panel A, Column 3, firms from higher population density counties are indeed more likely to report having external links. Moreover, conditional on being linked, firms from those counties are more likely to be linked to firms from the same origin (established in the same location). Once again, this homophily is on average an order of magnitude (40 times) larger than what would be obtained if firms with links in each prefecture were randomly matched.

Table 2, Panel B repeats the preceding exercise using listed shareholders other than the legal representative (who is a shareholder 90% of the time). Although our model focuses on the network's role in enhancing the productivity of its members, we have noted its potentially co-existing function of channelling capital to them. One way that this can be done is through joint ownership (shared equity). We expect firms from higher population density counties to be more likely to adopt this strategy because of the higher levels of trust that can be supported by the population at the origin. We also expect these firms to be more likely to use cross-ownership to create external links with other firms from the same origin (in

<sup>5</sup>An alternative interpretation of the cross-linkages is that they compensate for low levels of trust. Given the positive association between trust and population density, this would imply that firms from higher population density counties will have fewer links, which is at odds with what we report below.

<sup>6</sup>Among the county-born legal representatives, 41% establish their firm in their birth county, 15% in their birth prefecture but outside the birth county, 15% in their birth province but outside the birth prefecture, and 29% outside their birth province.

Table 2— Homophily and Population Density

Panel A: Listed Individuals				
Dependent variable:	fraction of firms with listed individuals	fraction of listed individuals from the legal representative's birth county	fraction of firms with links in the prefecture	fraction of linked firms that are linked to a firm from the same birth county
	(1)	(2)	(3)	(4)
Birth county population density	0.029*** (0.003)	0.031*** (0.004)	0.012*** (0.001)	0.021*** (0.004)
Mean of dependent variable	0.841	0.421	0.232	0.402
Counter-factual mean	–	0.008	–	0.010
Observations	1,624	1,624	1,624	1,624
Panel B: Shareholders				
Birth county population density	0.006*** (0.002)	0.019*** (0.003)	0.004*** (0.001)	0.029*** (0.004)
Mean of dependent variable	0.763	0.332	0.209	0.339
Counter-factual mean	–	0.008	–	0.008
Observations	1,624	1,624	1,624	1,623

*Note:* The sample is restricted to firms established outside their birth counties and active in 2009. Listed individuals and shareholders do not include (are in addition to) the legal representative. Linked firms have at least one individual in common. The exception is pairs of firms that have the same legal representative (since they may not be independent entities). The counter-factual mean is based on the random assignment of listed individuals or shareholders and the random matching of linked firms in the prefectures where they are located. Control variables include population, education and occupation distribution in the birth county. Population density is measured as a Z-score. Legal representatives, listed individuals, and shareholders are 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%.

the locations where they are established) because those links are more valuable. As expected, we see in Columns 1-4 that higher birth county population density is associated with (i) a greater fraction of firms listing multiple shareholders, (ii) a higher fraction of shareholders from the legal representative's birth county in such firms, (iii) a higher fraction of firms with cross-ownership in the locations where they are established, and (iv) a higher fraction of such links that are between firms from the same origin. Once again, the observed homophily is on average an order of magnitude (40 times) larger than what would be obtained by random assignment of shareholders in Column 2 and random matching of firms with links in Column 4, within each prefecture.

The level of cooperation that can be maintained between firms that share a common listed individual or shareholder is high to begin with, but as Jackson, Rodriguez-Barraquer and Tan (2012) emphasize, it will be even higher if it is “supported” by mutually connected firms. Focusing on links between firms from the same birth county located in the same prefecture (but outside their birth

county) we observe in Table 3 that such links between firms from higher population density counties are (i) more likely to be supported, and (ii) supported by a greater number of firms, conditional on being supported. This is true regardless of whether links are defined by cross-listing (Columns 1-2) or cross-ownership (Columns 3-4). As in Jackson, Rodriguez-Barraquer and Tan (2012) we consider a variant of the network that combines both types of links (Columns 5-6) without changing the results.<sup>7</sup>

Table 3— Link Support and Population Density

Network:	listed individuals		shareholders		combined	
	fraction of links that are supported	number of supporting firms, conditional on support	fraction of links that are supported	number of supporting firms, conditional on support	fraction of links that are supported	number of supporting firms, conditional on support
Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
Birth county population density	0.016*** (0.006)	0.196*** (0.072)	0.010** (0.005)	0.018 (0.025)	0.015*** (0.006)	0.196*** (0.068)
Mean of dependent variable	0.281	1.836	0.161	1.392	0.288	1.842
Observations	1,624	1,338	1,624	1,069	1,624	1,376

*Note:* The sample is restricted to firms established outside their birth counties and active in 2009. Linked firms have at least one individual in common. The exception is pairs of firms that have the same legal representative (since they may not be legal entities).

A link (cross-listing, cross-ownership) is supported if the two nodal firms have mutual links to a third firm. Number of supporting firms is the number of mutual connections.

Control variables include population, education and occupation distribution in the birth county. Population density is measured as a Z-score.

Legal representatives, listed individuals, and shareholders are 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%.

The results in Tables 2 and 3, taken together, indicate that while all firms maintain ties with their birth county, even when located elsewhere, firms from higher population density counties display higher levels of homophily. Firms from these counties are more likely to be linked to other firms in the prefectures where they are located and these links are more likely to be with firms from the same origin.<sup>8</sup> The individual links between firms from higher population density counties will be stronger for two reasons: First, they are more likely to be supported by the social network at the origin. Second, these links are more likely to be supported at the destination. These results are robust to the spatial domain of the network. We have assumed that the network is defined at the level of the prefecture but, as observed in Appendix Tables F3 and F4, the results are largely unchanged when

<sup>7</sup>70% of listed individuals are shareholders. Thus, while there is substantial overlap in these networks, it is not complete.

<sup>8</sup>All the estimating equations in Tables 2 and 3 include county population, education, and occupational structure as controls. None of these variables display the consistent positive and significant relationship between population density and each outcome that we observe in the tables.

the links between firms are restricted to the counties where they are located; the fraction of linked firms does decline, but the level of homophily and support remain the same, as does their variation with birth county population density.

A unique feature of our network data is that all firms and all links between firms are observed. As with any empirical analysis of networks, the preceding results are obtained for specific types of links (cross-listing and cross-ownership). Nevertheless, given the consistency of the results, we expect that they will apply more generally to other (possibly informal) linkages. The business networks that emerge from higher population density counties will thus be of higher quality, and the model that follows will examine the consequences of this heterogeneity for the dynamics of firm entry, concentration, and size. Although it is tempting to use the network quality measures directly to test the model, these measures; e.g. the number of links, are determined by outcomes of the model such as entry and concentration. We will thus use an exogenous determinant of network quality – birth county population density – to test the model.

## II. The Model

### A. Population and Technology

The population is comprised of a large number of communities, distinguished by their population density,  $p$ . As discussed above, the level of trust that can be supported in a community, and the quality of the networks that are drawn from that community, is increasing in  $p$ . Each community consists of a continuum of agents, with equal sized cohorts of new agents born at successive dates  $t=1,2,\dots$ . Agents vary in individual ability  $\omega$ , which is drawn independently from a log uniform distribution on the unit interval. The ability distribution is identical across cohorts and communities.

Each agent makes a once-and-for-all choice between a traditional occupation (such as farming or wage labor) and becoming an entrepreneur. The returns to entering the traditional occupation for an agent of ability  $\omega$  is  $\omega^\sigma$  where  $\sigma \in (0, 1)$ . There are multiple destinations  $B_i, i = 1, 2..$  for entrepreneurship, with destinations denoting 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  $\omega$  selects capital size  $K$ , and has a production function

$$(1) \quad y = A_{it}\omega^{1-\alpha}K^\alpha$$

where  $\alpha \in (0, 1)$  is the capital elasticity and  $A_{it}$  denotes community TFP (CTFP); i.e. the contribution of the network to the firm's productivity.  $A_{it} = A_0(1 + h(p))^{n_{i,t-1}}$ , where  $h(p)$  denotes per-member help provided by members of the network to one another, which is increasing in  $p$ , and  $n_{i,t-1}$  measures the stock of entrepreneurs from the community who are already established in sector  $i$  by the end of period  $t - 1$  (and thus in a position to provide support to the cohort of new entrants that follows).<sup>9</sup> This specification captures the idea that

<sup>9</sup>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

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,  $n_{i,t-1}$  (for given  $h(p)$ ). 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 origin population density  $p$ ; i.e.  $h'(p) > 0$ . Letting  $\theta(p)$  denote  $\log(1 + h(p))$ , the preceding expression reduces to

$$(2) \quad A_{it} = A_0 \exp(\theta(p)n_{i,t-1})$$

where network quality,  $\theta(p)$ , is increasing in  $p$ .<sup>10</sup>

In our model, networks are based on the social origins of entrepreneurs rather than the destinations they select:  $\theta$  reflects social connectedness in the birth county and  $n_{it}$  is the number of firms from that county operating in a given destination. The implicit assumption in our formulation of the production function is that networks from different birth counties operate independently at a given destination. In the standard agglomeration model, the  $\theta$  parameter would measure exogenous destination characteristics and  $n_{it}$  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.

The dependence of CTFP,  $A_{it}$ , 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 destination  $i$ . 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.

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,

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.

<sup>10</sup>Based on the analysis in the previous section, higher quality networks increase the number of links and the level of cooperation that can be supported by a given link. We restrict attention to the second mechanism in the model, assuming for simplicity that all incumbent firms contribute to the network. We could allow the fraction of firms that contribute to the network (and are linked to entrants) to be increasing in  $p$ , but from equation (2) it is evident that this additional term would be subsumed in  $\theta(p)$ .

and would thus amplify the dynamics generated by the latter alone.<sup>11</sup> 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 sector-locations are comprised of a large number of origin county networks (as documented below) and both domestic and international market opportunities are large.

### B. Occupational Choice

To determine occupational choice, we first calculate the profits a new agent in any cohort with a given ability  $\omega$  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 (2)). The optimal capital size  $K$  must maximize  $A\omega^{1-\alpha}K^\alpha - rK$ , and thus satisfies:

$$(3) \quad \log K(\omega, A) = \log \omega + \log \phi + \frac{1}{1-\alpha} \log A - \frac{1}{1-\alpha} \log r$$

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

$$(4) \quad \log \Pi(\omega, A) = \log \omega + \log \psi + \frac{1}{1-\alpha} \log A - \frac{\alpha}{1-\alpha} \log r$$

(where  $\psi \equiv \phi^\alpha - \phi$ ).<sup>12</sup>

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. These agents will be endowed with a level of ability that exceeds a threshold  $\underline{\omega}$ :

$$(5) \quad \log \omega > \log \underline{\omega} \equiv \frac{1}{1-\sigma} \left[ \log \frac{1}{\psi} - \frac{1}{1-\alpha} \log A + \frac{\alpha}{1-\alpha} \log r \right]$$

We assume that the threshold lies in the interior of the support 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.

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

<sup>11</sup>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.

<sup>12</sup>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.



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. Others, who decided to stay out on myopic grounds, might now wish to enter when they anticipate future network growth, which would further raise the returns to entrepreneurship.

### C. Dynamics of Entry and Concentration

We make the simplifying assumption that 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 below 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 (5), 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

$$(6) \quad e_{it} = Ls_{i,t-1} + \kappa(p)N_{t-1}s_{i,t-1}^2$$

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 entrepreneurs from past cohorts from the same origin. Aggregating (6) across destinations, we obtain an expression for the dynamics of aggregate entry:

$$(7) \quad N_t - N_{t-1} \equiv E_t \equiv \sum_i e_{it} = L + \kappa(p)N_{t-1}H_{t-1}$$

where  $H_{t-1} \equiv \sum_i s_{i,t-1}^2$  denotes the Herfindahl Hirschman Index for concentration

at  $t-1$ . Equations (6,7) 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 and 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. The network complementarity associated with post-entry productivity spillovers, embodied in the  $N_{t-1}$  term in (7), is reinforced by the network complementarity associated with the referrals; i.e. the  $H_{t-1}$  term. Entry  $E_t$  will rise over time from (7) 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  $i = 1, 2$ . Without loss of generality, assume that destination 1 has a higher initial (and subsequent) presence, in which case  $H_t = s_{1t}^2 + (1 - s_{1t})^2$  is monotonically increasing in  $s_{1t}$ . Variation in  $s_{1t}$  is thus synonymous with variation in  $H_t$ . From equations (6) and (7),

$$\begin{aligned}
 s_{1t} \equiv \left[ \frac{N_t}{n_{1t}} \right]^{-1} &= \left[ \frac{L + N_{t-1} + \kappa(p)N_{t-1}H_{t-1}}{Ls_{1,t-1} + n_{1,t-1} + \kappa(p)N_{t-1}s_{1,t-1}^2} \right]^{-1} \\
 (8) \qquad \qquad \qquad &= \left[ 1 + \left( \frac{1}{s_{1,t-1}} - 1 \right) \left\{ \frac{L + N_{t-1} + \kappa(p)N_{t-1}(1 - s_{1,t-1})}{L + N_{t-1} + \kappa(p)N_{t-1}s_{1,t-1}} \right\} \right]^{-1}
 \end{aligned}$$

We use this expression, together with equation (7), which characterizes variation in entry, to derive the following result.

**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 destination  $B_1$  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.<sup>13</sup> Moreover, with two locations within any sector, the results on concentration in Proposition 2 apply across sectors or across locations within sectors.

#### D. Ability Selection and Firm Size Dynamics

Next we derive predictions concerning entrepreneurial ability and firm size. From (3), an increase in CTFP,  $A_{it}$ , increases initial capital, for a given level of entrepreneurial ability,  $\omega$ . However, we also know that network effects generate negative selection on ability: from (5), as CTFP increases over time, the threshold for entry falls, and entrepreneurs with lower ability start entering. This negative selection has a negative effect on initial capital from (3). Substituting from (5) in (3), we see that the latter effect dominates and, hence, that the initial capital of the marginal entrant is unambiguously decreasing in CTFP,  $A_{it}$ :

$$(9) \quad \log K_{it}^m = U' - \frac{\sigma}{(1-\sigma)(1-\alpha)} \log A_{it}$$

where  $U' \equiv \log \phi - \frac{1}{1-\sigma} \log \psi - \frac{1}{1-\alpha} \log r$ , and  $\log A_{it} = \log A_0 + \theta(p)n_{i,t-1}$ . The marginal entrepreneurs that enter later in time, when their networks are stronger (with higher CTFP) are thus less productive and have smaller firms.<sup>14</sup> 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  $\sigma < \frac{1}{2}$ .<sup>15</sup> To see this, observe that substituting from (5) in (3), the capital size of the average entrant satisfies:

$$(10) \quad \log K_{it}^a = W + \frac{1-2\sigma}{2(1-\alpha)(1-\sigma)} \log A_{it}$$

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

<sup>13</sup>The reason is that the expression for entry flow into sector  $c$  is modified to  $e_{ct} = \kappa L s_{c,t-1} + \kappa N_{t-1} s_{c,t-1}^2 H_{c,t-1}$ , so the term involving the quadratic term in lagged sectoral share is weighted by lagged intra-sectoral spatial concentration  $H_{c,t-1}$ .

<sup>14</sup>This result does not depend on assumptions concerning the distribution of ability. To see this, observe that expressions (3, 4) 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.

<sup>15</sup>This depends on the assumption of a log uniform distribution of ability.

for any given cohort can be shown to be rising in  $p$  and over time. Equation (3) implies that the capital at date  $t' > t$  of a cohort  $t$  entrepreneur with ability  $\omega$  is given by

$$(11) \quad \log K_{it'}^a = \log \omega + \log \phi - \frac{1}{1-\alpha} \log r + \frac{1}{1-\alpha} [\log A_{it} + \theta(p) \sum_{l=t}^{t'-1} e_{il}]$$

implying a growth rate at period  $t'$ :

$$(12) \quad \log K_{it,t'}^a - \log K_{it,t'-1}^a = \frac{1}{1-\alpha} \theta(p) e_{it'}$$

In our model, growth in incumbent firm size is independent of the entrepreneur's ability and cohort and is driven entirely by the contemporaneous growth in CTFP,  $\theta(p)e_{it'}$ .

**PROPOSITION 3:** *With two destinations:*

- (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 (5) and (9), the marginal entrant's ability and initial capital are decreasing in  $\log A_{it}$ . From (10), the average entrant's ability and initial capital are also decreasing in  $\log A_{it}$  if  $\sigma > \frac{1}{2}$ .  $\log A_{it} \equiv \log A_0 + \theta(p)N_{t-1}s_{i,t-1}$  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 (12). The result then follows from Propositions 1 and 2. Firms from high- $p$  origins enter smaller, but subsequently grow faster.<sup>16</sup>

### III. 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 may be positively associated with trust in the birth county and with network quality, it could also be correlated with other factors that

<sup>16</sup>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). In our model, which fits the Chinese data, firms with access to higher quality networks (and lower ability) grow faster. To fit the Indian data, 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 would need to dominate the network effect.

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 alternative models that are examined below.

#### A. Origin Heterogeneity

Our model assumes cohort size and the share of potential entrepreneurs,  $k$ , are constant across origin counties and cohorts. Suppose that we relax these assumptions and let  $k(p, t)$ , which now refers to the number of potential entrepreneurs, 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,  $\omega^\sigma$ , where  $\omega$  is individual ability, 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  $v_p < 0, v_t < 0, v_{pt} < 0$ .<sup>17</sup>

If higher- $p$  counties have a larger pool of entrepreneurs and this advantage is growing over time, as specified above, then our model's predictions for entry could be obtained without requiring networks to be active. Heterogeneity in  $k$  would not, however, explain the negative selection with respect to entrepreneurial ability and initial capital that is implied by our model. The specified heterogeneity in outside options,  $v(p, t)$ , will generate this negative selection and, simultaneously, our model's predictions for entry. However, origin heterogeneity, no matter how it is specified, cannot explain why firms from higher- $p$  counties are increasingly clustered over time in particular sectors and locations. Nor can it explain why firms from those counties enter smaller and then grow faster.

#### B. Destination Heterogeneity

Now consider heterogeneity in 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

<sup>17</sup>An alternative interpretation of  $v(p, t)$  is that it represents the payoff from origin-based networks operating in the traditional sector.

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 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).

This alternative model 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 preferred 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), 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 effects can then be included in the estimating equation. Conditional on these fixed effects, the network model would imply that firms from higher- $p$  origins will grow faster on average because growth is determined by changes in CTFP. In contrast, there is no relationship between firm growth and  $p$  in the alternative model once destination-time period effects 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. 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).

## IV. Testing the Model

### A. 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. We test these predictions with data over the 1990-2009 period because, as documented below, the decline in the ability threshold predicted by the model starts to weaken beyond that point.<sup>18</sup> Our analysis covers the universe of registered firms, in contrast with previous analyses of firms in China; e.g. Brandt, Van Biesebroeck and Zhang (2012), Aghion et al. (2015) which 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. The above-scale firms account for less than 15% of all private firms in the registration database in 2008. Nevertheless, with nearly 2,000 counties, the number of entering firms from a given birth county at each point in time is relatively small. Each time period in the dynamic analysis that follows thus covers a five-year window. As with the analysis of network quality, the legal representative's birth county applies to the firm as a whole in the tests of the model and all the analysis that follows.

Figure 4 reports nonparametric estimates of the relationship between the entry of firms from each birth county in each time period and 1982 population density. The entry patterns in the figure are visually consistent with the model's predictions.<sup>19</sup> Table 4, Columns 1-4 report parametric estimates corresponding to Figure 4, 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 with the analysis of trust and network quality, we include birth county population, education, and occupational structure (also measured in 1982) as covariates in the estimating equation. We will do the same in the analyses of concentration and firm size. This allows for the possibility that population density is correlated with particular county characteristics that could independently determine the opportunities for entrepreneurship (in specific sectors) and access to capital. We see in Table 4, 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 predictions (ii) and (iii) above. Formal tests of these predictions are reported later in this section.

As discussed above, both origin heterogeneity and destination heterogeneity can explain the entry results without requiring origin-based networks to be active. We

<sup>18</sup>Recall that the model only applies to the initial, rapid growth, phase of industrialization.

<sup>19</sup>Appendix Figure E2a 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 E2a that the model's predictions for the stock of firms go through as well, despite the exits. As an additional robustness test, Appendix Figure E2b 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.

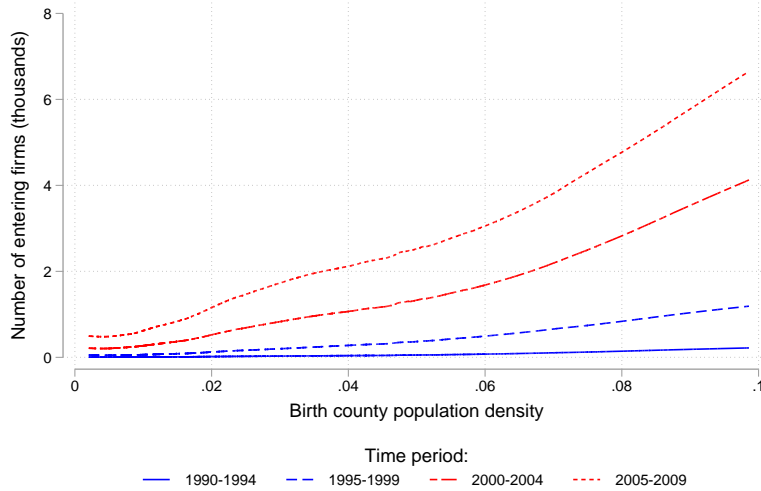


Figure 4. Firm Entry and Population Density

Source: SAIC registration database and 1982 population census.

Table 4— Firm Entry and Population Density

Dependent variable: Time period:	number of entering firms				number of entering firms			
	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*** (0.003)	0.092*** (0.013)	0.289*** (0.035)	0.448*** (0.052)	0.014** (0.006)	0.118*** (0.037)	0.382*** (0.101)	0.575*** (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 (in thousands) is measured at the birth county - time period level in Columns 1-4 and at the birth county-sector-location-time period level in Columns 5-8 (adjusted to account for variation in the number of sectors and locations across birth counties).

Control variables include population, education and occupation distribution in the birth county. Population density is converted to a Z-score.

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%.

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 (sectors and locations). 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 4, Columns 5-8 includes (two-digit) sector fixed effects and location (county or urban district) 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, at a relatively fine level, over time.<sup>20</sup> Although our analysis focuses on county-born businessmen, we place no restriction on the location of their firms; there are 3,235 counties or urban districts where firms locate in our data. Birth county population density continues to have a positive and significant effect on entry in each time period in Table 4, 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 destination 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 (receiving fewer entrants overall).

### *B. 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 5a reports nonparametric estimates of the relationship 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 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 C. 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 5b reports nonparametric estimates of the relationship between spatial concentration, within one-digit sectors, and birth county population density in five-year intervals.<sup>21</sup> Although the model assumes that all destinations are symmetric, one obvious asymmetry in practice is that moving costs are lower when the entrepreneur chooses to stay back home. We avoid this asymmetry by including all locations in the analysis of spatial concentration, measured at the county or urban district level, except for the birth location. 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. 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.

<sup>20</sup>Entry in Table 4, 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 (to be comparable with the estimates in Columns 1-4). For a given birth county, all sectors and locations that ever have entrants are included in all time periods (assigned zero entry where necessary).

<sup>21</sup>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 there are multiple entrants from each birth county in each time period.

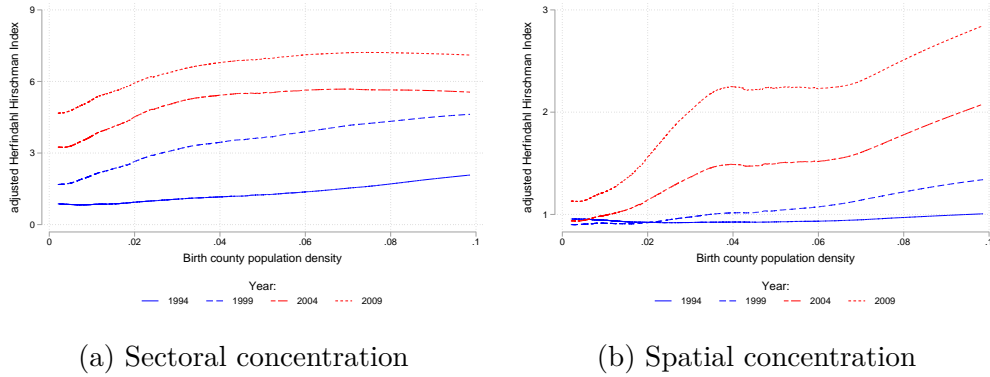


Figure 5. Concentration and Population Density

Source: SAIC registration database and 1982 population census.

Table 5 reports parametric estimates corresponding to Figure 5a and Figure 5b. The additional birth county characteristics are included in the estimating equation as usual. 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 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 5— Sectoral and Spatial Concentration and Population Density

Dependent variable:	adjusted HHI across sectors				adjusted HHI across locations			
	1994	1999	2004	2009	1994	1999	2004	2009
Year:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth county population density	0.106*** (0.023)	0.417*** (0.057)	0.444*** (0.061)	0.574*** (0.074)	0.014* (0.008)	0.041*** (0.014)	0.047* (0.027)	0.072* (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 is 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.

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. Population density is measured as a Z-score.

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 6 formally tests the model's predictions for changes in firm entry, sec-

toral 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 (an ordinal variable corresponding to successive five-year windows over the 1990-2009 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 6, 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 6, Columns 4-6.<sup>22</sup> There is no association between localized trust and population density 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 community 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.

Table 6— Entry, Concentration, and Population Density (time and interaction effects)

Dependent variable:	county			city		
	number of entrants	sectoral HHI	spatial HHI	number of entrants	sectoral HHI	spatial HHI
	(1)	(2)	(3)	(4)	(5)	(6)
Time period	0.517*** (0.016)	1.686*** (0.020)	0.506*** (0.021)	0.661*** (0.026)	0.299*** (0.008)	2.054*** (0.069)
Birth place population density × time period	0.353*** (0.029)	0.165*** (0.019)	0.134*** (0.012)	0.353*** (0.041)	-0.026*** (0.004)	0.030 (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. Sector fixed effects are also included with spatial HHI as the dependent variable. See Tables 4 and 5 for a detailed description of variable construction. Standard errors clustered at birthplace level are reported in parentheses. \* significant at 10%, \*\* at 5%, \*\*\* at 1%.

<sup>22</sup>We measure urban population density at the city rather than the disaggregated urban district level because urban districts were created after 1982. However, firm location is always measured at the urban district level.

### C. 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 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. Sectors are measured at the two-digit level and locations are defined by the county or urban district. We begin in Figure 6a 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 from the U.S. indicates that education is also a good measure of entrepreneurial ability (Levine and Rubinstein, 2017). In a developing economy, however, 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.<sup>23</sup> 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 6a that the marginal entrant's measured ability declines over time; from around the 70<sup>th</sup> percentile of his birth county-birth cohort education distribution in the 1990-1994 period to just around the 40<sup>th</sup> percentile in the 2005-2009 period. The relationship between the marginal entrant's ability and population density is also negative in each time period and grows steeper over time.<sup>24</sup> 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 6b 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

<sup>23</sup>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.

<sup>24</sup>Appendix Table F5 reports parametric estimates corresponding to Figure 6a, separately in each time period. These estimates indicate that birth county population density has a negative and significant effect on marginal ability among entering entrepreneurs at each point in time.

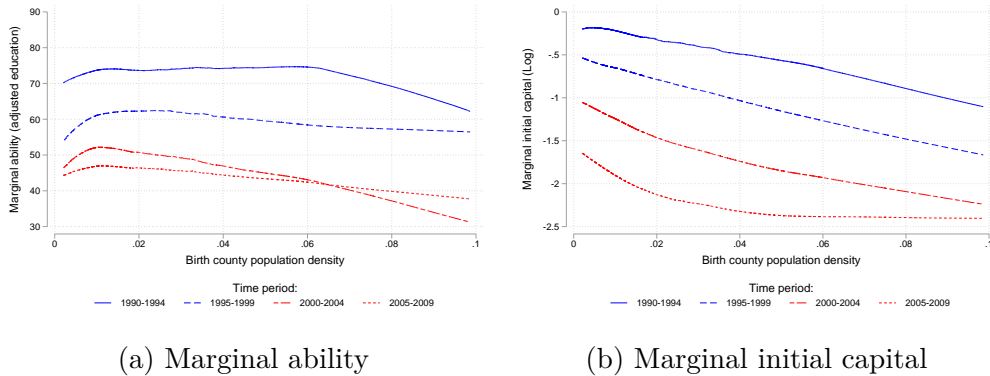


Figure 6. Marginal Ability, Marginal Initial Capital and Population Density

Source: SAIC registration database and 1982 population census.

capital is defined as the bottom one percentile of the initial capital distribution at the birth county-sector-time period level.<sup>25</sup> As predicted by the model, marginal initial capital is decreasing over time and decreasing in birth county population density in each time period.<sup>26</sup>

Notice from Figure 6b 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 7 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 6a and 6b. We see in Table 7, Column 1, which includes birth county-sector fixed effects, that the marginal entrant is drawn from lower down in his birth county-cohort education 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 7, 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

<sup>25</sup>The initial capital for a firm is determined by its initial registered capital, which can be recovered from the SAIC registration database. 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.

<sup>26</sup>Appendix Table F5 reports parametric estimates corresponding to Figure 6b, separately by time period. The population density coefficient is negative and significant in each time period.

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 7, 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.<sup>27</sup> 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.

Table 7— Evidence on Negative Selection

Dependent variable:	marginal	marginal	average	marginal	average
	ability	initial capital	initial capital	initial capital	initial capital
	(1)	(2)	(3)	(4)	(5)
Time period	-17.908*** (0.496)	-0.868*** (0.012)	-0.116*** (0.009)	-0.609*** (0.010)	-0.095*** (0.008)
Birth county population density × Time period	-0.926*** (0.351)	-0.026** (0.011)	0.002 (0.007)	-0.061*** (0.009)	-0.020*** (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).

Initial capital (in million Yuan) is measured in logs.

The marginal entrepreneur (firm) is located at the bottom one percentile of the ability (initial capital) distribution in the birth county-sector-time period (Columns 1-3) or birth county-sector-location-time period (Columns 4-5). Average initial capital is the mean of the relevant distribution.

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 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

<sup>27</sup>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-location-time period. The sample in Columns 4-5 is restricted to birth county-sector-locations 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.

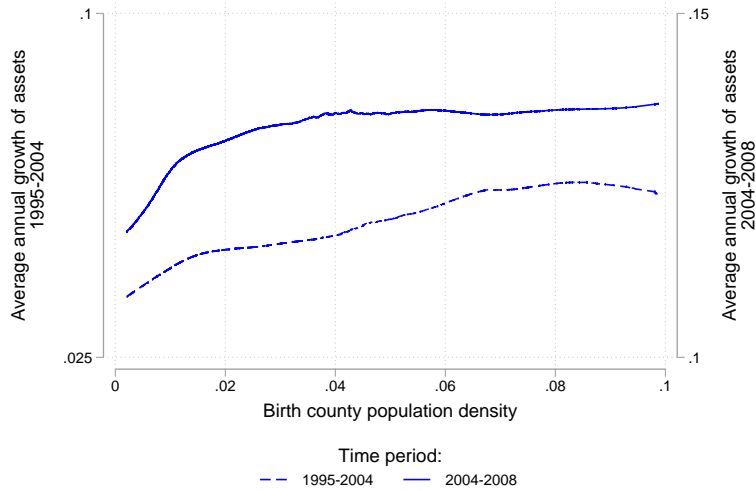


Figure 7. Asset Growth and Population Density

*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.

the analysis of firm growth, we 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 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.<sup>28</sup> Figure 7 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 8 reports parametric estimates corresponding to Figure 7. Columns 1-4 report results with industrial census data and Columns 5-6 repeat the analysis with SAIC inspection data, which include all sectors (not just manufacturing, as in the industrial census).<sup>29</sup> 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 although the growth equations are estimated in the cross-section, the analysis in Table 8 mirrors the analysis in Table 7. In the former we examine

<sup>28</sup>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 network quality and CTFP, which apply equally to all active firms from a given birth county at a given point in time.

<sup>29</sup>Data coverage for seven provinces is poor with the inspection data and these provinces are thus dropped from the analysis.

Table 8— Growth of Assets and Population Density

Dependent variable:	average annual growth of assets					
	industrial census		industrial census		inspection data	
	1995-2004		2004-2008		2004-2008	
	(1)	(2)	(3)	(4)	(5)	(6)
Birth county population density	0.006*** (0.002)	0.007* (0.004)	0.004** (0.002)	0.003** (0.001)	0.004*** (0.001)	0.002* (0.001)
Initial capital	–	0.002*** (0.000)	–	0.001*** (0.000)	–	0.001*** (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.

Population density is measured in units of 10,000 people per square km, and then converted to Z-score. 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%.

the increase in firms' assets over time, whereas in the latter we examine the corresponding decline in entering entrepreneurs' ability and initial firm capital. To test the model's predictions, we measure firm growth at the birth county-sector level in specifications with sector fixed effects (Columns 1, 3 and 5) and at the birth county-sector-location level in specifications with sector and location fixed effects (Columns 2, 4 and 6) and then estimate the (average) effect of birth county population density on these growth measures. 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 in Columns 2, 4 and 6 to allow for convergence; as discussed above, 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. Although this is not a feature of our model, it may reflect exogenous variation in access to capital across birth counties and individuals. Regardless of the specification, the consistent finding in Table 8 is that population density in the birth county has a positive and significant effect on firm growth.<sup>30</sup> Firms from high population density birth counties enter 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.

<sup>30</sup>A pooled regression (not reported) which combines industrial census data over both time periods indicates, as implied by the model, that firm growth is increasing significantly over time.



## V. 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 network to have any effect on entry and initial capital size in the model, there must be positive lagged entry. The structural estimation and the quantification of network effects that follows is thus restricted to destination-time periods with positive lagged entry. To increase the fraction of firms that are included in these destination-time periods and, hence, in the structural analysis, we now define destinations at the one-digit sector and prefecture level.

Based on the entry and initial capital equations in the model, (6) and (10), and retaining its notation (with the addition of a community,  $c$ , subscript), we estimate the following structural equations simultaneously:

$$(13) \quad e_{ci,t} = G(\alpha, \sigma, r, A_0)k_c s_{ci,t-1} + \frac{\theta}{(1-\sigma)(1-\alpha)}k_c s_{ci,t-1} \cdot pn_{ci,t-1} + u_{ci,t}$$

$$(14) \quad \log K_{ci,t}^a = J_t(\alpha, \sigma, r, A_0, f_t) + \frac{\theta(1-2\sigma)}{2(1-\sigma)(1-\alpha)}pn_{ci,t-1} + v_{ci,t}$$

The functional forms for  $G(\alpha, \sigma, r, A_0)$  and  $J_t(\alpha, \sigma, r, A_0, f_t)$  are obtained directly from equations (6) and (10), with the addition of the separable  $f_t$  term (described below) in the  $J_t$  function.  $e_{ci,t}$  measures the number of entrants and  $\log K_{ci,t}^a$  measures average initial capital (in logs) for birth county  $c$  and destination (sector-location)  $i$  in time period  $t$ . We parameterize the  $\theta(p)$  function to be increasing linearly in  $p$ :  $\theta(p) = \theta p$ , with the restriction that  $\theta(0) = 0$ . This restriction is based on the idea that as the population density goes to zero in a hypothetical county, there can be no social interactions and, hence, no enforceable trust. The network effect is thus represented by a single parameter,  $\theta$ .  $n_{ci,t-1}$  is the stock of firms from birth county  $c$  that are already established in destination  $i$  at the beginning of time period  $t$ .  $s_{ci,t-1}$  is the share of destination  $i$  in the stock of firms originating from county  $c$  at the beginning of period  $t$ . 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  $t$ .

The theoretical model assumes that the size of each cohort and the fraction of potential entrepreneurs in the cohort,  $k$ , are the same in all counties. In practice, cohort size will depend on the county population and the fraction of potential entrepreneurs will depend on the level of education in the county.  $k_c$  in equation (13) thus refers to the number of potential entrepreneurs in county  $c$ . This number is calculated from the 1990 population census, based on the characteristics of actual entrepreneurs when they established their firms. Most entrepreneurs in the SAIC database have at least high school education and relatively few were younger than 25 when their firm was established. Assuming, as in the model, that individuals must make a one-time occupational choice at the start of their

careers,  $k_c$  for each (five-year) cohort of entering entrepreneurs is thus specified to be the number of men born in county  $c$ , aged 25-29, with at least high school education, as reported in the 1990 population census.

The residual terms,  $u_{ci,t}$  and  $v_{ci,t}$ , include the effect of local government inputs, agglomeration, and sector-level spillovers on access to capital, firm productivity, and accompanying entry. The structural equations are linear in observed variables; (i)  $k_c s_{ci,t-1}$  (ii)  $k_c s_{ci,t-1} \cdot p n_{ci,t-1}$  (iii)  $p n_{ci,t-1}$  and, hence, we will be able to control flexibly for these potentially confounding effects, which could bias the structural estimates, as shown below. There are four reduced-form coefficients in equations (13) and (14). 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:  $\alpha, \sigma, r, A_0, \theta$ . The model is under-identified and we thus assign values to  $r$  and  $A_0$  prior to estimation.

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).<sup>31</sup> The productivity multiplier is set to one in all destinations; i.e.  $A_0 = 1$ . As in the model, variation in productivity across destinations (and origin counties) 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  $\alpha$  parameter, which measures the marginal returns to capital, to vary across six broad sector categories: high-tech services, wholesale and retail services, manufacturing and transportation, heavy industry (mining, electricity, and construction), non-financial services (hotels, catering, education), and finance. This increases the number of structural equations to 12, given that there are now two equations in each sector category, and the number of structural parameters to be estimated to eight;  $\alpha_1, \dots, \alpha_6, \sigma, \theta$ . The structural parameters are estimated by matching on entry and average initial capital in each birth county-destination-time period.<sup>32</sup> The model is estimated over the 1995-2004 period; i.e. over two time periods, matching the reduced form analysis with initial capital as the outcome.  $f_t$ , which adjusts capital from physical to monetary units in each

<sup>31</sup>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, Kambourov and Storesletten (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.

<sup>32</sup>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 within and across sector categories. For example, the ratio of the coefficients on  $k_c s_{ci,t-1} \cdot p n_{ci,t-1}$  and  $p n_{ci,t-1}$  in (13) and (14), 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.

time period, and which appears additively in the  $J_t$  function, is thus estimated separately in 1995-1999 (period 1) and 2000-2004 (period 2).

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-destination-time periods is minimized. Parameter estimates, with bootstrapped standard errors in parentheses, are reported in Table 9, Column 1.<sup>33</sup> The  $\sigma$  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 estimated  $\theta$  parameter is positive and highly significant. The remaining structural parameters ( $\alpha_1, \dots, \alpha_6$ ) and the capital adjustment term,  $f_t$ , which is estimated separately in each time period are not reported to preserve space.

Table 9, Column 2 reports estimates from an augmented specification that includes lagged entry in the sector-location from all origins (including urban origins). The additional  $n_{i,t-1}$  term is included additively to capture generalized agglomeration effects and to proxy for destination effects, such as government infrastructure, that would induce entry from all origins. This term, like  $n_{ci,t-1}$ , is treated as exogenous in the structural estimation, and we allow for separate  $n_{i,t-1}$  coefficients in the entry and initial capital equations. The augmented specification in Column 2 also adds time period effects to the entry equation to accommodate secular changes in the economy over time. The estimated  $\sigma$  and  $\theta$  parameters are hardly affected by these additions to the estimating equation. Given the additive structure of the estimating equations, we can control even more flexibly for destination effects by including sector-location-time period effects. Parameter estimates with this specification are reported in Table 9, Column 3. The  $\sigma$  parameter estimate is entirely unchanged, and while the  $\theta$  estimate does decline slightly, we cannot reject the hypothesis that the  $\theta$  estimates are equal across all specifications in Table 9. Consistent with the reduced form evidence, entrepreneurs from higher population density birth counties do not appear to have preferred access to superior destinations (if they did, then the  $\theta$  estimates would change when destination-time period effects are included in the estimating equation).

Figures 8a and 8b assesses the goodness of fit of the model by comparing actual and predicted values, separately for entry and initial capital, across the range of birth county population densities. The destination-time period effects in Table 9, Column 3 are not estimated and, hence, our most flexible specification cannot be used for prediction. The predicted values are thus based on the next most flexible specification, reported in Table 9, Column 2. Although there are just 14 parameters in this specification, it does a good job of predicting entry and initial capital across nearly 2,000 birth counties, in each time period, at a time when the Chinese economy was growing at an explosive rate.<sup>34</sup>

A major objective of our research is to quantify the role played by community

<sup>33</sup>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.

<sup>34</sup>The model under-predicts entry at the top of the population density distribution in Figure 8a, but this is on account of the fact that the distribution is skewed to the left (as documented in Appendix Figure E3) with a mean of 0.03. The estimation thus puts more weight on matching at lower population densities. Overall, the prediction error in total entry is 9% in the 1995-1999 period and 2% in the 2000-2004 period.

Table 9— Structural Estimates

Model:	benchmark	with agglomeration effects	with destination-time period effects
	(1)	(2)	(3)
$\sigma$	0.779 (0.001)	0.780 (0.001)	0.798 (0.020)
$\theta$	0.883 (0.274)	0.878 (0.283)	0.654 (0.106)

*Note:* all specifications include the following additional parameters (not reported);  $\alpha_1, \dots, \alpha_6$ , and  $f_t$ . The augmented specification in column 2 includes the lagged stock of firms in each destination-time period to capture generalized agglomeration effects and adds time period effects to the entry equation. The most flexible specification in column 3 includes destination-time period effects. The destination is defined by the one-digit sector and prefecture. 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. Bootstrapped standard errors in parentheses.

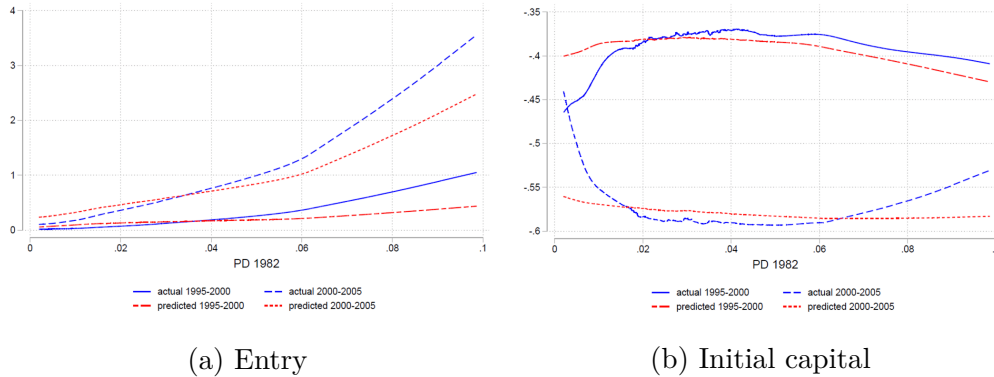


Figure 8. Actual and Predicted, Entry and Initial Capital

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

networks in the growth of private enterprise in China. This is accomplished by setting the  $\theta$  parameter to zero and then generating counter-factual entry and capital investment over the estimation period. This exercise incorporates the dynamic network multiplier effect that is a key feature of our model; exogenous initial entry in a birth county-destination induces subsequent entry, which, in turn, leads to further entry, and so on. A corresponding dynamic multiplier cannot be applied to overall entry in the destination,  $n_{i,t-1}$ , in the augmented specification reported in Table 9, Column 2 because it is outside the model; while this term can be included as a control in the estimation, it cannot be incorporated in the counter-factual simulation. The simulations that follow are thus based on the benchmark specification reported in Table 9, Column 1. The  $\sigma$  and  $\theta$  estimates are almost identical for the two specifications in any case. The counter-factual

simulation with entry as the outcome is reported in Figure 9a. 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 predicted entrants would have declined by 21.7% over the 1995-2004 period if the networks had not been active. In a related counter-factual exercise, reported in Figure 9b, the predicted total stock of capital in 2004 (taking account of the number of firms that entered, their initial capital, and the subsequent growth in their capital) would have declined by 28.5% had the networks been absent. Adjusting for the fact that the counter-factual analysis is restricted to birth county-destination-time periods with positive lagged entry and to county-born entrepreneurs, for whom the hometown networks are relevant, this amounts to a 10.8% decline in the number of entrants and a 12.5% decline in the stock of capital for the economy as a whole.<sup>35</sup>

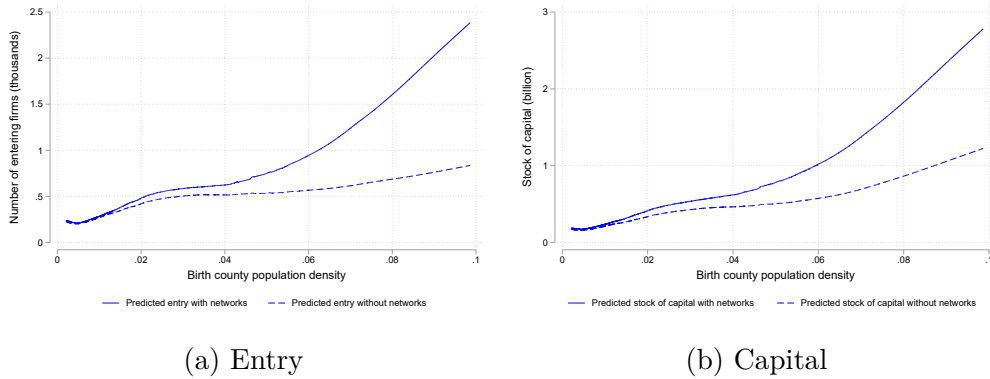


Figure 9. Counter-Factual Simulation: Effect of Community Networks on Entry and Capital

Source: Model generated data and 1982 population census.

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. Wu (2016); Morrison (2019). 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 profit of all (marginal and infra-marginal)

<sup>35</sup>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.

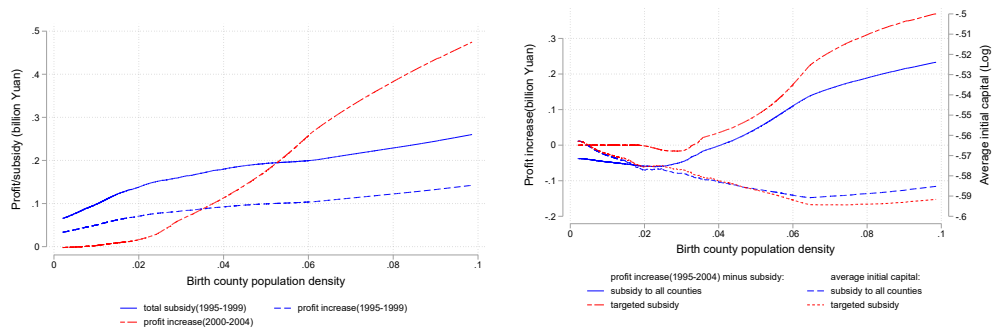
entrants. As observed in Figure 10a, the total profit increase generated by the subsidy in 1995-1999 is less than the cost to the government in all birth counties. However, the spillover effect of the one-time subsidy on profits in the subsequent 2000-2004 period is substantial (and even larger than 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 a 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 5.7% for counties above the mean population density and -46.7% for counties below the mean.

Figure 10b 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; these counties, with a population density above 0.04, lie in the top quartile of the population density distribution. To keep the total amount of the subsidy constant, the interest rate for the targeted counties is reduced to 0.12. 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). In our framework, higher quality networks either enhance the productivity of their members or provide cheaper capital. One way that they can accomplish the latter is by more effectively lobbying government officials. Even if that were the case, our analysis indicates that the optimal policy would be to subsidize the high quality networks even further.<sup>36</sup> Notice also that average initial capital, which is declining with population density, declines even more steeply with the more efficient targeted program. 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 10b).

## VI. 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 trust and network quality (proxied by population density). We validate each of these predictions over a twenty year period with unique adminis-

<sup>36</sup>The implicit assumption underlying this policy prescription is that birth county networks lack the market power to distort prices in the destinations where they are established. Based on the registration data, firms from a given origin county account for 12% of firms at the destinations (rural counties or urban districts) 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. The corresponding statistic for the capital share is 18% and both the firm-share and the capital-share are declining in birth county population density (on account of the diversity in sectors and locations where firms from different origins are established). While pricing may be non-competitive in China (see, for example, Brooks, Kaboski and Li (2016)), the origin-based networks, even those drawn from high population density counties, do not appear to be directly associated with these distortions.



(a) Subsidy to all counties      (b) Targeted subsidy vs. subsidy to all counties

Figure 10. Counter-Factual Simulation: Effect of Interest Rate Subsidy on Profits

Source: Model generated data and 1982 population census.

trative 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. Having validated the model, we estimate its structural parameters and conduct counter-factual simulations. The first simulation indicates that aggregate entry in the 1995-2004 period and total capital stock in 2004 would have been 11% and 12.5% lower in the economy as a whole in the absence of the rural hometown 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 argued to have played an important role, 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).<sup>37</sup> Chi-

<sup>37</sup>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).

nese 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 the networks. 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 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.

The third caveat concerns the unintended inefficiencies that are generated by networks and which must be considered when designing network-oriented policies. The model assumes that networks operate independently; while they may distribute resources and provide mutual help efficiently within the boundaries of the community, this implies that valuable inputs will not cross community lines. Highlighting the importance of such inefficiencies, Cai and Szeidl (2017) document that an intervention that brought Chinese entrepreneurs together had a substantial positive effect on their productivity. In addition to such static inefficiencies, networks can also generate dynamic inefficiencies. In particular, once networks are established and their members are locked into specific activities, they may be discouraged from taking advantage of new economic opportunities. Based on our research, hometown networks appear to have successfully adapted to the new economic environment that emerged with privatization in China to support occupational and spatial mobility. Recall, however, that our model only applies to early phases of industrialization. Thirty years later, the next step would be to improve product quality and shift into more remunerative exporting or to move into advanced (high-tech) sectors. The question is whether the hometown networks will successfully reorient themselves to help their members navigate this next step or whether they will hold them back. This would appear to be a fruitful area for future research, not just to shed light on the specifics of Chinese growth but also to enhance our understanding of the process of economic development.



## REFERENCES

- Aghion, Philippe, Jing Cai, Mathias Dewatripont, Luosha Du, Ann Harrison, and Patrick Legros. 2015. "Industrial Policy and Competition." *American Economic Journal: Macroeconomics*, 7(4): 1–32.
- Alatas, Vivi, Abhijit Banerjee, Arun G Chandrasekhar, Rema Hanna, and Benjamin A Olken. 2016. "Network structure and the aggregation of information: Theory and evidence from Indonesia." *American Economic Review*, 106(7): 1663–1704.
- Alesina, Alberto, and Paola Giuliano. 2015. "Culture and institutions." *Journal of Economic Literature*, 53(4): 898–944.
- Allen, F, J Qian, and M Qian. 2005. "Law, Finance, and Economic Growth in China." *Journal of Financial Economics*, 77(1): 57–116.
- Banerjee, Abhijit, and Kaivan Munshi. 2004. "How Efficiently is Capital Allocated? Evidence from the Knitted Garment Industry in Tirupur." *Review of Economic Studies*, 71(1): 19–42.
- Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. 2013. "The diffusion of microfinance." *Science*, 341(6144).
- Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. 2019. "Using gossips to spread information: Theory and evidence from two randomized controlled trials." *The Review of Economic Studies*, 86(6): 2453–2490.
- Banerjee, Abhijit V, and Esther Duflo. 2000. "Reputation effects and the limits of contracting: A study of the Indian software industry." *The Quarterly Journal of Economics*, 115(3): 989–1017.
- Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten. 2016. "Firm Entry and Regional Growth Disparities: the Effect of SOEs in China." *University of Toronto Mimeo*.
- Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2012. "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing." *Journal of Development Economics*, 97(2): 339–351.
- Brooks, Wyatt J., Joseph P. Kaboski, and Yao Amber Li. 2016. "Growth Policy, Agglomeration, and (the Lack of) Competition." *NBER Working Paper Series*.
- Burchardi, Konrad B, Thomas Chaney, and Tarek A Hassan. 2019. "Migrants, ancestors, and foreign investments." *The Review of Economic Studies*, 86(4): 1448–1486.
- Cai, Jing, and Adam Szeidl. 2017. "Interfirm Relationships and Business Performance." *Quarterly Journal of Economics*, qjx049.
- Chaney, Thomas. 2014. "The network structure of international trade." *American Economic Review*, 104(11): 3600–3634.
- Ciccone, Antonio, and Robert E. Hall. 1996. "Productivity and the Density of Economic Activity." *American Economic Review*, 86(1): 54–70.
- Cohen, Lauren, Umit G Gurun, and Christopher Malloy. 2017. "Resident networks and corporate connections: Evidence from World War II internment camps." *The Journal of Finance*, 72(1): 207–248.
- Coleman, James S. 1988. "Social Capital in the Creation of Human Capital."

- American Journal of Sociology*, 94: S95–S120.
- Ellison, Glenn.** 1994. “Cooperation in the Prisoner’s Dilemma with Anonymous Random Matching.” *Review of Economic Studies*, 61(3): 567–588.
- Fafchamps, Marcel.** 2000. “Ethnicity and credit in African manufacturing.” *Journal of Development economics*, 61(1): 205–235.
- Fisman, Raymond J.** 2003. “Ethnic ties and the provision of credit: Relationship-level evidence from African firms.” *The BE Journal of Economic Analysis & Policy*, 3(1).
- Galor, Oded, and David N Weil.** 2000. “Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond.” *American economic review*, 90(4): 806–828.
- Galor, Oded, and Ömer Özak.** 2016. “The agricultural origins of time preference.” *American Economic Review*, 106(10): 3064–3103.
- Goodman, Bryna.** 1995. *Native Place, City, and Nation: Regional Networks and Identities in Shanghai, 1853–1937*. University of California Press.
- Greif, Avner.** 1993. “Contract enforceability and economic institutions in early trade: The Maghribi traders’ coalition.” *The American economic review*, 525–548.
- Greif, Avner.** 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.
- Greif, Avner, and Guido Tabellini.** 2017. “The Clan and the Corporation: Sustaining Cooperation in China and Europe.” *Journal of Comparative Economics*, 45(1): 1–35.
- Hansman, Christopher, Jonas Hjort, Gianmarco León, and Matthieu Teachout.** forthcoming. “Vertical integration, supplier behavior, and quality upgrading among exporters.” *Journal of Political Economy*.
- Honig, Emily.** 1992. *Creating Chinese Ethnicity: Subei People in Shanghai, 1850-1980*. New Haven:Yale University Press.
- Honig, Emily.** 1996. *Regional Identity, Labor, and Ethnicity in Contemporary China*. Berkeley, California:Institute of East Asian Studies, University of California.
- Hsieh, Chang-Tai, and Peter J. Klenow.** 2009. “Misallocation and Manufacturing TFP in China and India.” *Quarterly Journal of Economics*, 124(4): 1403–1448.
- Hu, Biliang.** 2008. “People’s Mobility and Guanxi Networks: A Case Study.” *China & World Economy*, 16(5): 103–117.
- Jackson, Matthew O, Tomas Rodriguez-Barraquer, and Xu Tan.** 2012. “Social capital and social quilts: Network patterns of favor exchange.” *American Economic Review*, 102(5): 1857–97.
- Kandori, Michihiro.** 1992. “Social Norms and Community Enforcement.” *Review of Economic Studies*, 59(1): 63–80.
- Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl.** 2009. “Trust and social collateral.” *The Quarterly Journal of Economics*, 124(3): 1307–1361.
- Levine, Ross, and Yona Rubinstein.** 2017. “Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?” *Quarterly Journal of Economics*, 132(2): 963–1018.

- Lin, Justin Yifu, and Yan Wang.** 2016. *Going Beyond Aid: Development Cooperation for Structural Transformation*. Cambridge University Press.
- Long, Cheryl, and Xiaobo Zhang.** 2011. "Cluster-based Industrialization in China: Financing and Performance." *Journal of International Economics*, 84(1): 112–123.
- Macchiavello, Rocco, and Ameet Morjaria.** 2015. "The value of relationships: evidence from a supply shock to Kenyan rose exports." *American Economic Review*, 105(9): 2911–45.
- Macchiavello, Rocco, and Ameet Morjaria.** forthcoming. "Competition and relational contracts in the Rwanda coffee chain." *The Quarterly Journal of Economics*.
- McMillan, John, and Christopher Woodruff.** 1999. "Interfirm relationships and informal credit in Vietnam." *The Quarterly Journal of Economics*, 114(4): 1285–1320.
- Morrison, Wayne M.** 2019. "China's economic rise: History, trends, challenges, and implications for the United States." *Congressional Research Service*, RL33534.
- Munshi, Kaivan.** 2011. "Strength in Numbers: Networks as a Solution to Occupational Traps." *Review of Economic Studies*, 78(3): 1069–1101.
- Nee, Victor, and Sonja Oppen.** 2012. *Capitalism from Below: Markets and Institutional Change in China*. Harvard University Press.
- Peng, Yusheng.** 2004. "Kinship Networks and Entrepreneurs in China's Transitional Economy." *American Journal of Sociology*, 109(5): 1045–1074.
- Rauch, James E, and Vitor Trindade.** 2002. "Ethnic Chinese networks in international trade." *Review of Economics and Statistics*, 84(1): 116–130.
- Summers, Lawrence.** 2007. "The Rise of Asia and the Global Economy." *Research Monitor*, , (2).
- Woodruff, Christopher.** 2002. "Non-contractible investments and vertical integration in the Mexican footwear industry." *International Journal of Industrial Organization*, 20(8): 1197–1224.
- Wu, Mark.** 2016. "The China, Inc. Challenge to Global Trade Governance." *Harv. Int'l LJ*, 57: 261.
- Zhang, Junyi.** 2016. "Chinese Foreign Assistance, Explained." Retrieved from <https://www.brookings.edu/blog/order-from-chaos/2016/07/19/chinese-foreign-assistance-explained/>.
- Zhao, Yaohui.** 2003. "The Role of Migrant Networks in Labor Migration: the Case of China." *Contemporary Economic Policy*, 21(4): 500–511.

## VARIABLE CONSTRUCTION

**1. Traditional crop productivity:** The Food and Agricultural Organization (FAO) Global Agro-Ecological Zones (GAEZ) database provides estimates of potential crop yields that can be matched to any administrative unit, such as a Chinese county, for which spatial shape files are available. Galor and Özak (2016) convert the potential yields derived for low technology - rain fed agriculture to caloric production. They then average across crops to construct a Caloric Suitability Index (CSI) that they document is a good indicator of historical population density (going back many centuries). Our measure of the CSI is based on three staple crops – wheat, rice, and barley – that historically dominated (and continue to dominate) agricultural production in China. All variables in Figure 2a are standardized by subtracting their means and then dividing by their standard deviations.

**Trust (WVS):** The World Values Survey (WVS) provides the fraction of respondents for a given country in the following categories: trust completely, trust somewhat, trust not very much, trust not at all. We combine the first two categories and the latter two categories to construct a binary measure of trust. Countries with annual GDP per capita exceeding \$20,000, with an area less than  $100,000km^2$ , or with missing population data in the World Development Index are dropped from the analysis.

**Trust (CFPS):** The adult individual module of the 2012 China Family Panel Study (CFPS) measures trust as an ordinal variable, taking values from 0 to 10. To construct a binary trust measure that is consistent with the WVS, we selected a cutoff, which turns out to be 5, such that trust in neighbors and strangers obtained from the CFPS matches most closely with the corresponding statistics for China from the WVS.

**Population density:** Population density in rural counties and in cities is obtained from the 1982 census and measured in units of 10,000 people per square km. The threshold density for rural counties 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.

**Population:** Population in rural counties and cities is obtained from the 1982 census and is measured in millions.

**Education:** Education in rural counties and cities is obtained from the 1982 census and is measured by the fraction of the population that is literate.

**Occupational structure:** Occupational structure in rural counties and cities is obtained from the 1982 census and is measured by the share of workers in agriculture and manufacturing, with services the excluded category.

**Birth county:** The SAIC database provides the citizenship ID of the legal representative, the listed personnel, and the major shareholders in each registered firm. Citizenship ID's were first issued in September 1985. The first six digits of the citizenship ID reveal the birth county of individuals born after that date and the county or city of residence (in September 1985) for individuals born before

that date. Given the limited opportunities for labor migration in the 1980's 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 reveal the county of birth, with few exceptions, even for those born before September 1985.

#### 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  $\xi$  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  $\omega$  is

$$(B1) \quad P_{it}(\omega) = \omega \xi A_0^{\frac{1}{1-\alpha}} \exp(\theta p n_{i,t-1} \frac{1}{1-\alpha}) [1 + \delta \exp(\theta(p) e_{it} \frac{1}{1-\alpha})]$$

while of staying in  $T$  is

$$(B2) \quad N_{it}(\omega) = \omega^\sigma [1 + \delta]$$

The agent will enter if

$$(B3) \quad \log \omega > \frac{1}{1-\sigma} [-\log \xi - \frac{1}{1-\alpha} \log A_0 + \log(1+\delta) - \frac{1}{1-\alpha} \theta(p) n_{i,t-1} - \log\{1 + \delta \exp(\theta(p) e_{it} \frac{1}{1-\alpha})\}]$$

Define the function

$$g(e|s_{i,t-1}, n_{i,t-1}, A_{i0}) = k s_{i,t-1} \{1 + \frac{1}{1-\sigma} [\log \xi + \frac{1}{1-\alpha} \log A_0 - \log(1 + \delta) + \frac{1}{1-\alpha} \theta(p) n_{i,t-1} + \log\{1 + \delta \exp(\theta(p) e \frac{1}{1-\alpha})\}]\}$$

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

$$(B4) \quad g(e|s_{i,t-1}, n_{i,t-1}, A_{i0}) = e$$

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

$$(B5) \quad g'(e|s_{i,t-1}, n_{i,t-1}, A_{i0}) = s_{i,t-1} \frac{\delta \exp(\frac{\theta(p)e}{1-\alpha})}{1 + \delta \exp(\frac{\theta(p)e}{1-\alpha})} \frac{k\theta(p)}{(1-\alpha)(1-\sigma)}$$

Hence if

$$(B6) \quad \frac{k\theta\bar{p}}{(1-\alpha)(1-\sigma)} < 1$$

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.

#### 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 (6). 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<sup>38</sup>  $\sum_i s_{i,t-1}^2 g_{it} = \sum_i s_{i,t-1} x_{it} > 0$ , which proves that concentration is rising in  $t$ , and hence (using (7)) the same is true for  $E_t$ .

*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 (7) at  $t=1$ . Next observe that the right-hand-side of (8) 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 (7) then implies  $E_t$  (and  $N_t$ ) is rising in  $p$ . Moreover, observe that the right-hand-side of (8) is rising in  $N_{t-1}$  and in

<sup>38</sup>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.

$s_{1,t-1}$ , 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 (7)

$$(C1) \quad E_{t+1} - E_t = \kappa(p)[N_t H_t - N_{t-1} H_{t-1}] = \kappa(p)[E_t H_t + N_{t-1}(H_t - H_{t-1})]$$

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 (8) can be rewritten as

$$(C2) \quad s_{1t} - s_{1,t-1} = \kappa(p)N_{t-1} \frac{(2s_{1,t-1} - 1)(1 - s_{1,t-1})s_{1,t-1}}{(L + N_{t-1})(2 - s_{1,t-1}) + \kappa(p)N_{t-1}(s_{1,t-1}^2 + 1 - s_{1,t-1})}$$

$\kappa(p) < 1$  implies that the denominator of the right-hand-side of (C2) 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  $\kappa$ . 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 (9) 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).<sup>39</sup> A similar argument operates for ability and size of the average entrant from (10) and taking the average across destinations. Part (b) follows from averaging across destinations in (12), and applying Propositions 1 and 2.

#### 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  $\mathbf{X} = (X_1, \dots, X_k)$  has a multinomial distribution with parameters  $(n, k, p_1, \dots, p_k)$ . Applied to our context, (i)  $n$  is the total number of firms from a given birth county, (ii)  $k$  is the total number of destinations that they are allocated to, and (iii)  $p_1, \dots, p_k$  are the probabilities that a firm allocated randomly would end up in each of those destinations. We assume that there is an equal probability of choosing any destination;  $p_j = \frac{1}{k}, \forall j$ .

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

<sup>39</sup>  $N_t$  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.

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

Based on the general properties of the multinomial distribution,

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

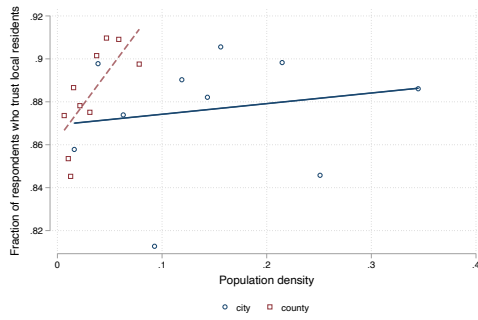
It follows that,

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

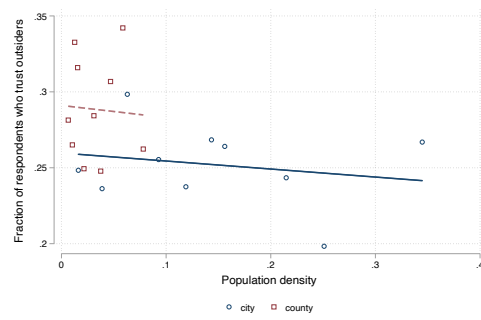
For large  $n$ ,  $E(HHI) \approx \frac{1}{k}$ . For small  $n$ ,  $E(HHI)$  is decreasing in  $n$ . We account for this by constructing a normalized HHI statistic, which is simply the unadjusted HHI, based on the observed distribution of firms across destinations, divided by  $E(HHI)$ . If firms are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



FIGURES



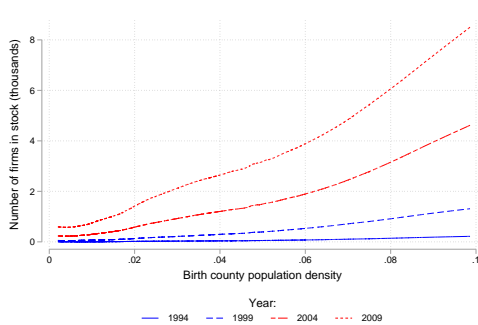
(a) Localized trust



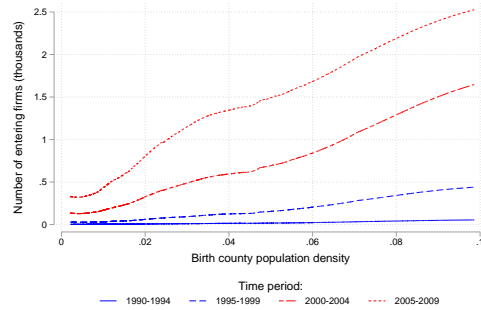
(b) Generalized trust

Figure E1. Trust and Population Density: China

Source: China Family Panel Study and 1982 population census.



(a) Stock of Firms



(b) Firms Located Outside the Birth County

Figure E2. Firm Entry and Population Density

Source: Registration Database and 1982 population census.

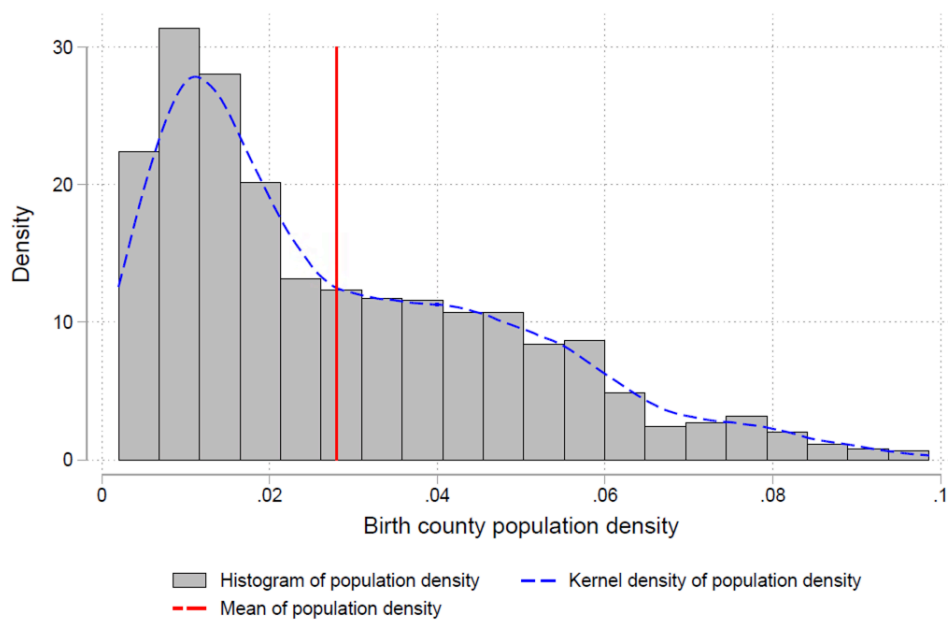


Figure E3. Distribution of Population Density

Source: 1982 population census.

TABLES

Table F1— Frequency of Local Social Interactions and Population Density

Dependent variable:	average frequency of social interactions per month with local residents			
	county		city	
	planned	unplanned	planned	unplanned
Respondent's location:	(1)	(2)	(3)	(4)
Population density	0.391 (0.415)	1.640** (0.821)	-0.774 (0.505)	-0.098 (1.184)
Mean of dependent variable	4.027	15.70	3.700	12.92
Observations	93	93	39	39

*Note:* Note: Social interactions are obtained from the family module of the China Family Panel Study (2010) and aggregated up to the county/city level. Social interactions are divided into planned interactions and unplanned interactions.

Planned interactions include group entertainment, visits to neighbors' homes, and dining together. Group entertainment includes playing *mahjong* or cards, reading newspapers, listening to the radio, or watching TV with others.

Unplanned interactions are one-on-one social meetings without other background activities.

Control variables include population, education and occupation distribution in the birth place. Population density is converted to a Z-score.

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 F2— Social Heterogeneity and Population Density

Dependent variable:	population density at neighborhood/village level		fraction of residents born locally	
	county	city	county	city
	(1)	(2)	(3)	(4)
Population density at city/county level	2.076*** (0.548)	2.763*** (1.016)	–	–
Population density at neighborhood/village level	–	–	1.550 (2.109)	-7.866*** (2.468)
Mean of dependent variable	0.112	0.794	90.16	47.90
Observations	265	134	265	134

*Note:* Information on local population density and social heterogeneity are obtained from the community module of the China Family Panel Study (2010) at the neighborhood/village level. Population density at city/county level is computed from the 1982 population census.

Columns 1-2 establish that neighborhood/village population density is increasing in city/county population density.

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%.

Table F3— Homophily and Population Density (links within the county)

Network:	listed individuals		shareholders	
Dependent variable:	fraction of firms with links in the county	fraction of linked firms that are linked to a firm from the same birth place	fraction of firms with links in the county	fraction of linked firms that are linked to a firm from the same birth place
	(1)	(2)	(3)	(4)
Birth county population density	0.005*** (0.001)	0.016*** (0.004)	-0.000 (0.001)	0.026*** (0.004)
Mean of dependent variable	0.140	0.380	0.120	0.313
Counter-factual mean	–	0.025	–	0.026
Observations	1,624	1,622	1,624	1,622

*Note:* The sample is restricted to firms established outside their birth counties and active in 2009. Listed individuals and shareholders do not include (are in addition to) the legal representative. Linked firms in the same county have at least one listed individual or shareholder in common. The exception is firms that have the same legal representative (since they may not be independent entities). The counter-factual mean is based on the random assignment of listed individuals or shareholders and the random matching of linked firms in the counties where they are located. Control variables include population, education and occupation distribution in the birth county. Population density is measured as a Z-score. Legal representatives, listed individuals and shareholders are 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 F4— Link Support and Population Density (links within the county)

Network:	listed individuals		shareholders		combined	
	fraction of links that are supported	number of supporting firms, conditional on support	fraction of links that are supported	number of supporting firms, conditional on support	fraction of links that are supported	number of supporting firms, conditional on support
Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
Birth county population density	0.018*** (0.006)	0.165*** (0.067)	0.010** (0.005)	0.004 (0.022)	0.016*** (0.006)	0.169*** (0.063)
Mean of dependent variable	0.215	1.585	0.111	1.280	0.218	1.583
Observations	1,624	1,099	1,624	724	1,624	1,159

*Note:* The sample is restricted to firms established outside their birth counties and active in 2009. Linked firms have at least one individual in common. The exception is pairs of firms that have the same legal representative (since they may not be independent entities). A link (cross-listing, cross-ownership) is supported if the two nodal firms have mutual links to a third firm. Number of supporting firms is the number of mutual connections. Control variables include population, education and occupation distribution in the birth county. Population density is converted to a Z-score. Information on legal representatives, listed individuals and shareholders 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 F5— Marginal Ability, Marginal Initial Capital and Population Density

Dependent variable:	marginal ability				marginal initial capital			
	1990-1994	1995-1999	2000-2004	2005-2009	1990-1994	1995-1999	2000-2004	2005-2009
Time period:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth county population density	-1.829 (1.369)	-2.685*** (1.103)	-3.570*** (0.807)	-1.435*** (0.755)	-0.100*** (0.030)	-0.118*** (0.026)	-0.134*** (0.017)	-0.041*** (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. Initial capital (in million Yuan) is measured in logs. The marginal entrepreneur (firm) is located at the bottom one percentile of the ability (initial capital) distribution in the birth county-sector-time period. Control variables include population, education and occupation distribution in the birth county. Population density is measured as a Z-score. Standard errors clustered at birth county level are reported in parentheses. \* significant at 10%, \*\* at 5%, \*\*\* at 1%.