

# Community Networks and the Growth of Private Enterprise in China

Ruochen Dai, Dilip Mookherjee, Kaivan Munshi and Xiaobo Zhang

July 17, 2018

# Introduction

- China has experienced the same degree of industrialization in three decades as Europe did in two centuries (Summers, 2007)
- This transformation began in the 1980's with the emergence of TVE's, and accelerated with the entry of private firms in the early 1990's
  - By 2014, there were 15 million registered private firms in China, accounting for over 90% of all registered firms and 60% of aggregate industrial production
- China is the world's largest exporter today and the world's largest or second-largest economy (Wu, 2016)

# Introduction

- Chinese growth occurred without the preconditions that are believed to be necessary for economic development; i.e. without effective legal systems or well functioning financial institutions (Allen *et al.*, 2005)
- The government compensated for some of these limitations (Long and Zhang, 2011; Wu, 2016)
- Informal mechanisms based on reputation and trust must have been at work to allow millions of (rural-born) entrepreneurs to establish and grow their business (Allen *et al.*, 2005; Song *et al.*, 2011; Peng, 2004; Greif and Tabellini, 2017)

# Introduction

- Case studies of production clusters; e.g. Fleisher *et al.* (2010) and Nee and Oppen (2012) indicate that long-established relationships among relatives and neighbors (from the rural origin) substitute for formal contracts
- We utilize comprehensive data covering the universe of registered firms over many years to identify and quantify the role played by hometown networks in the growth of private enterprise in China

# Community Networks and Private Enterprise

- Social networks or *guanxi* facilitated China's historically unprecedented rural-urban labor migration; e.g. Zhao (2003), Zhang and Li (2003), Hu (2008)
- These networks are organized around the birth county and most migrants end up living and working with *laoxiang* or “native-place fellows” (Honig, 1992; Goodman, 1995; Cai Fang, 1997; Ma and Xiang, 1998)
  - If the sending county is the domain around which migrant labor networks are organized, then it will also be the domain around which business networks supporting county-born entrepreneurs are organized

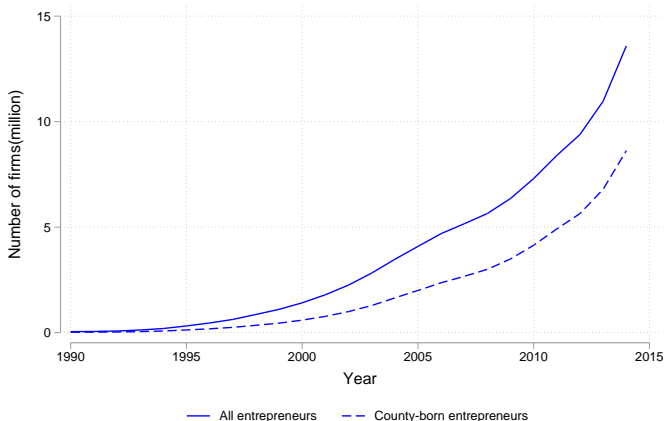
# Analysis Plan

- ① Establish that (historical) population density is a good proxy for social connectedness in a county
  - Business networks drawn from higher population density counties will sustain higher levels of mutual help - are of higher quality - regardless of where they are located
- ② Develop a theoretical model that describes the relationship between network quality and the dynamics of entry, concentration and firm size
  - Rule out competing non-network explanations
- ③ Test the predictions of the model
  - Implement direct tests of the network mechanism
- ④ Estimate the structural parameters and quantify the impact of the networks on firm entry and capital stock
  - Additional counter-factual simulations shed light on misallocation and industrial policy in economics where networks are active

# Social Connectedness and Economic Cooperation

- We proxy social connectedness by population density
  - The frequency of local social interactions is increasing in population density (spatial proximity)
  - Frequent social interactions support higher levels of economics cooperation by improving community enforcement
  - This argument is robust the definition of the community; i.e. clan vs. county, but may not hold in the city
- Empirically validate the preceding arguments with data from China Family Panel Survey [▶ Table](#)
  - Thus focus on county-born entrepreneurs, using city-born entrepreneurs as a control group

# Growth of Private Enterprise, by Birthplace of Entrepreneurs



Source: SAIC registration database.



# Table of Contents

1 Introduction

2 A Model of Network Dynamics

3 Empirical Analysis

# Population and Technology

- There are many origin counties, each with an exogenous level of social connectedness,  $p$
- An equal-sized cohort of new agents is born in each county in each period,  $t = 1, 2, \dots$ , who live forever thereafter
- Each agent is born with ability  $\omega$ ;  $\log \omega \sim U[0, 1]$
- Every cohort  $t$  agent makes a once-and-for-all occupational choice at  $t$ 
  - The choice is between a traditional non-entrepreneurial (T) sector and two business sectors,  $B_1$  and  $B_2$
  - Denote entry into sector  $B_i$  by past cohorts by  $n_{i,t-1}$

# Population and Technology

- Profit in the T sector is  $\omega^\sigma$ , where  $\sigma \in (0, 1)$
- In sector  $B_i$  at date  $t$ , the production function is

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

- where  $\alpha \in (0, 1)$  and  $A_{it} = A_0 \exp(\theta(p)n_{i,t-1})$  is Community TFP (CTFP)
- This is the first source of network complementarity in the model
- There is a fixed product price (normalized to unity) and all agents incur the same cost of capital,  $r$

# Occupation Choice

- For each origin, sector, time period, the optimal capital size  $K$  must maximize  $A\omega^{1-\alpha}K^\alpha - rK$  and thus satisfies:

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

- The resulting profit satisfies

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

- The network effect works through productivity, but it could also work through  $r$

# Occupation Choice

- A fixed fraction  $k$  of agents in every cohort has the opportunity to become an entrepreneur
- Each such agent receives an opportunity to enter one of the two business sectors
  - The probability of getting an opportunity in  $B_i$  equals the share of incumbents in that sector,  $s_{i,t-1}$
  - This is the second source of network complementarity
- Agents receiving a referral will enter sector  $B_i$  if

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

- Entry into sector  $i$  in period  $t$

$$e_{it} = k s_{i,t-1} [1 - \log \underline{\omega}_{i,t-1}]$$

# Dynamics of Entry and Concentration

## Proposition 1.

*Entry and concentration are*

- (i) increasing in  $t$  for any  $p$*
- (ii) increasing in  $p$  at any  $t$*
- (iii) increasing more steeply in  $p$  over time*
  - This result holds as long as the share of the larger sector is not too close to 1*

# Firm Size Dynamics

- Higher CTFP has two effects on the marginal entrant's initial capital
  - The direct effect raises firm size by raising firm level TFP (for a given  $\omega$ )
  - The negative selection on ability lowers firm TFP and size

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

- Noting that the average entrant has (log) ability  $\frac{1+\log \omega}{2}$

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

- Thus, average initial capital is decreasing in CTFP iff  $\sigma > \frac{1}{2}$
- Firm growth is independent of  $\omega$  and determined by changes in CTFP

# Firm Size Dynamics

## Proposition 2.

*Averaging across sectors:*

- (a) *Initial capital and ability of marginal entrants (and of average entrants if  $\sigma > \frac{1}{2}$ ) are*
  - (i) *decreasing in  $t$  for any  $p$*
  - (ii) *decreasing in  $p$  at any  $t$*
  - (iii) *decreasing more steeply in  $p$  over successive cohorts*
- (b) *The growth rate of capital of incumbent entrepreneurs of any past cohort  $t$  from  $t' - 1 (> t)$  to  $t'$  is rising in  $p$  and in  $t'$  (more steeply with higher  $p$ )*



# Alternative Explanations

- While population density may proxy for social connectedness, it could also be correlated with independent determinants of the model's outcomes
- Introduce new sources of heterogeneity at the origin, which are correlated with population density, and allow higher  $p$  origins to have better, and increasing, access to favorable destinations (locations)

# Origin Heterogeneity

- ① Population density is correlated with population, education, and traditional occupational patterns
  - These variables are associated with the stock of potential entrepreneurs, ability, and wealth
  - Can explain entry results without networks, but not negative selection or concentration
- ② Population density is associated with lower and decreasing payoffs in the traditional occupation
  - Can now explain negative selections, but not concentration or post entry growth

# Destination Heterogeneity

- The key  $s_{it}$ ,  $A_{it}$  terms are exogenously specified to match the endogenous evolution of these terms in our model
  - If each origin locates at a unique set of destinations, then the alternative models cannot be disentangled
  - In practice, firms from multiple origins locate at the same destination, so destination-time period dummies can be included in the estimating equation (this accounts for geography, government infrastructure, and agglomeration effects)
- We can add ability heterogeneity to alternative model, but this will not explain why firms from high- $p$  counties start small and then grow faster
  - Banerjee and Munshi (2004) observe the same pattern in Tirupur, but this is explained by positive selection
  - Neoclassical growth model will also generate convergence, so we control for initial capital

# Table of Contents

1 Introduction

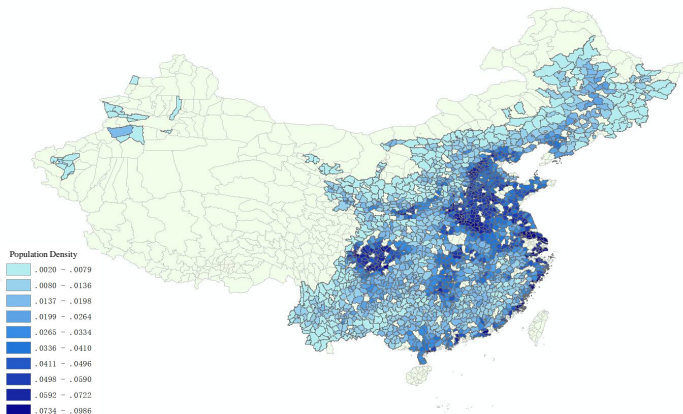
2 A Model of Network Dynamics

3 Empirical Analysis

# Firm Data

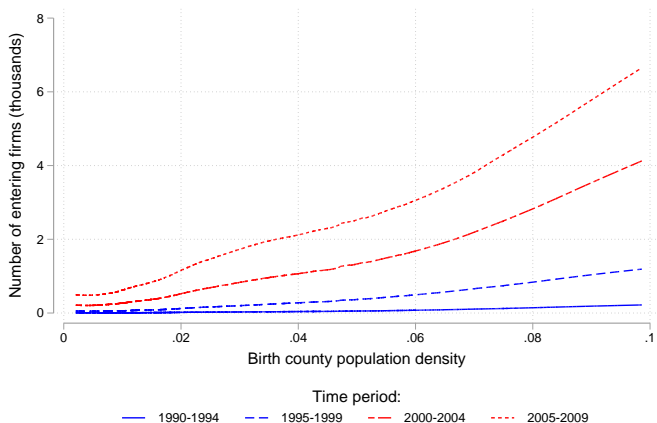
- Core analysis uses SAIC registration database
  - establishment date
  - 4-digit sector and location
  - ownership type
  - registered capital (initial and subsequent changes)
  - list of major shareholders and managers with citizenship ID
- Analysis of firm growth uses SAIC inspection database and industrial census (1995,2004,2008)

# Population Density across Counties



Source: 1982 population census.

# Evidence on Firm Entry



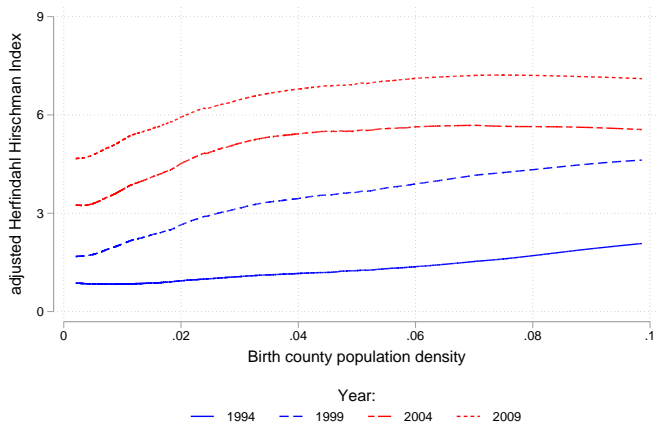
Source: SAIC registration database and 1982 population census.

# Firm Entry and Population Density

- 1 Verify that 1982 population density in the birth county has a positive and significant effect on firm entry in each time period
  - Robust to including additional county characteristics in the estimating equation
- 2 Verify that firms from higher population density counties did not select into sector-locations that received a relatively larger number of entrants from other origins in any time period
- 3 Verify that the results are robust to including sector and location dummies in each time period

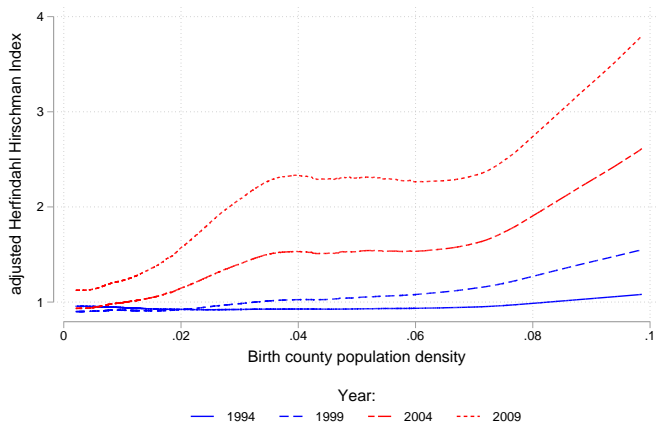


# Evidence on Sectoral Concentration



Source: SAIC registration database and the 1982 population census

# Spatial Concentration, within Sectors

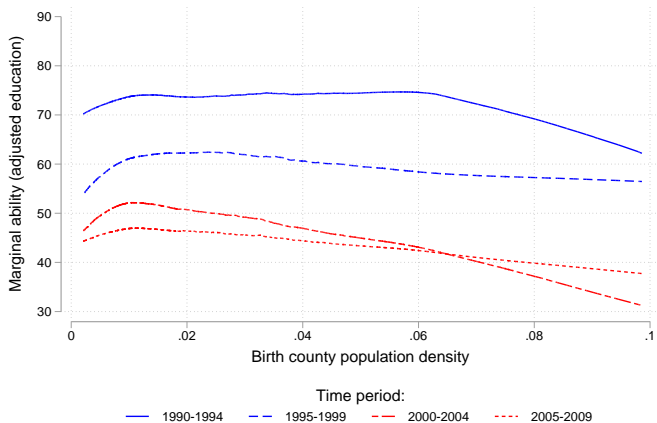


Source: SAIC registration database and 1982 population census.

# Changes over Time and Interaction Effects

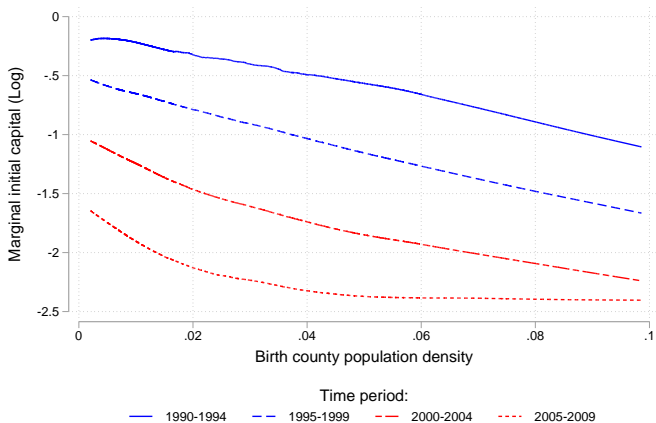
Dependent variable:	county			city district		
	number of entrants	sectoral HHI	spatial HHI	number of entrants	sectoral HHI	spatial HHI
	(1)	(2)	(3)	(4)	(5)	(6)
Time period	0.517*** (0.016)	1.717*** (0.020)	0.417*** (0.014)	0.661*** (0.026)	0.344*** (0.007)	2.359*** (0.082)
Birth place population density $\times$ time period	0.353*** (0.029)	0.151*** (0.019)	0.156*** (0.017)	0.355*** (0.041)	-0.029*** (0.004)	0.046 (0.045)
Observations	6,496	6,494	71,148	3,224	3,222	20,285

# Marginal Ability and Population Density



Source: SAIC registration database and 1982 population census.

# Marginal Initial Capital and Population Density

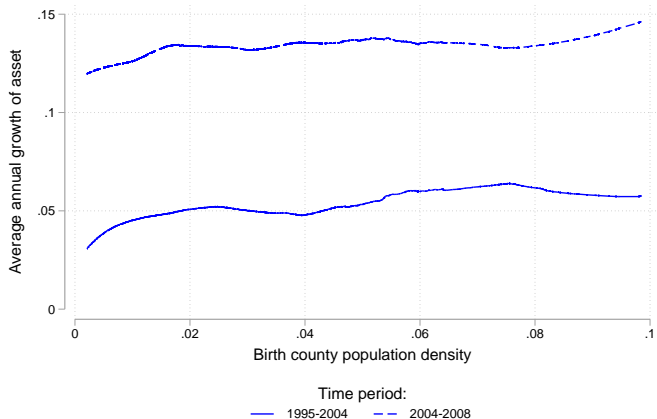


Source: SAIC registration database and 1982 population census.

# Evidence on Negative Selection

Dependent variable:	marginal ability	marginal initial capital	average initial capital	marginal initial capital	average initial capital
	(1)	(2)	(3)	(4)	(5)
Time period	-18.532*** (0.409)	-0.882*** (0.012)	-0.115*** (0.008)	-0.655*** (0.009)	-0.109*** (0.007)
Birth county population density $\times$ time period	-1.040*** (0.394)	-0.028** (0.012)	0.002 (0.008)	-0.069*** (0.010)	-0.022*** (0.007)
Mean of dependent variable	49.36	-1.744	-1.744	-1.223	-1.223
Observations	21,028	43,579	43,579	46,417	46,417
Origin-sector fixed effects	Yes	Yes	Yes	Yes	Yes
Location fixed effects	No	No	No	Yes	Yes

# Asset Growth and Population Density



Source: industrial census (1995, 2004, 2008) and 1982 population census.

# Asset Growth and Population Density

Dependent variable:	annual growth of asset					
	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 effects	No	Yes	No	Yes	No	Yes
Observations	5,517	5,664	31,234	64,258	18,701	43,470



# The Mechanism

Dependent variable: Birth place: Time period:	subsequent entrants from the birth place					
	county		county		city district	
	2000-2004	2005-2009	2000-2004	2005-2009	2000-2004	2005-2009
	(1)	(2)	(3)	(4)	(5)	(6)
Initial entrants from the birth place	7.120*** (0.711)	8.935*** (0.972)	5.239*** (1.065)	5.796*** (1.356)	7.830*** (0.959)	6.994*** (0.982)
All initial entrants at the location	0.054 (0.050)	-0.020 (0.057)	—	—	—	—
Initial entrants from the birth place $\times$ birth place population density	—	—	1.361** (0.619)	2.262** (0.991)	-0.073 (0.240)	-0.437** (0.220)
Mean of dependent variable	3.065	3.128	3.065	3.128	4.001	3.515
Observations	413,452	804,918	413,452	804,918	313,520	449,207

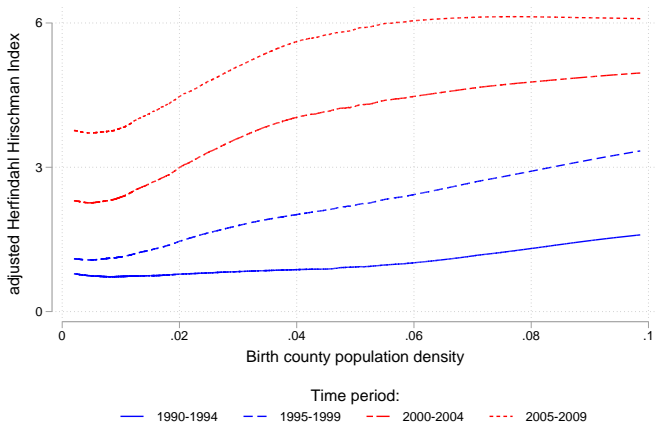
Note: all specifications include birth place-sector fixed effects.

# The Effect of Initial Entry (within clans)

Dependent variable: Time period:	subsequent entrants from the clan			
	2000-2004	2005-2009	2000-2004	2005-2009
	(1)	(2)	(3)	(4)
Initial entrants from the clan	3.275*** (0.330)	4.082*** (0.355)	2.434*** (0.315)	2.713*** (0.310)
All initial entrants at the location from the birth county	0.031*** (0.007)	0.051*** (0.012)	0.013 (0.013)	0.008 (0.019)
Initial entrants from the clan $\times$ birth county population density	—	—	0.641** (0.267)	1.071*** (0.334)
Initial entrants from the birth county $\times$ birth county population density	—	—	0.013* (0.007)	0.031** (0.014)
Mean of dependent variable	1.372	1.392	1.372	1.392
Observations	888,331	1,743,760	888,331	1,743,760

Note: all specifications include birth county-sector fixed effects.

# Clan Concentration and Population Density

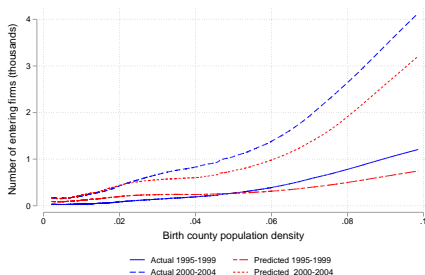


Source: Registration database and 1982 population census.

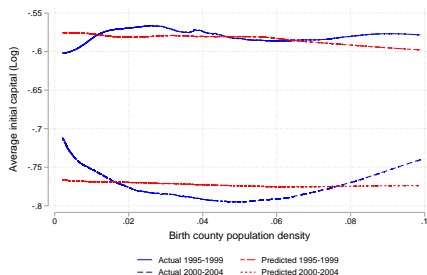
# Structural Estimation and Quantification

- Structural estimation based on the entry equation and the (average) initial capital equation
  - Estimated at the birth county-sector-time period level (1995-2004)
  - $\theta(p) = \theta p; \theta(0) = 0$
  - Allow  $\alpha$  to vary across 4 broad sectors to accommodate heterogeneity in capital requirements: 6 parameters ( $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \sigma, \theta$ )
  - Set  $A_0 = 1, r = 0.2$
- More flexible specifications allow for forward-looking behavior and sector-level spillovers

# Actual and Predicted, Entry and Initial Capital



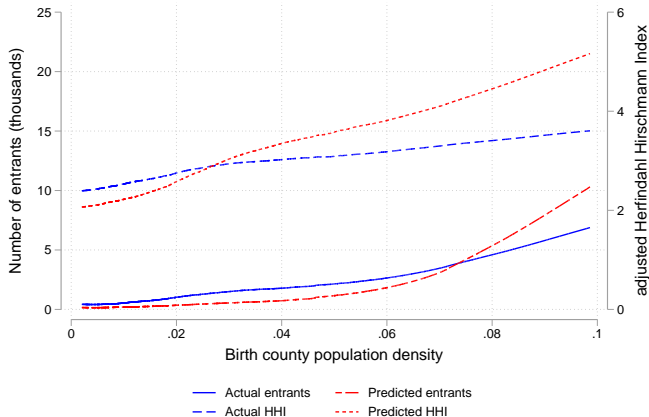
(a) Entry



(b) Initial Capital

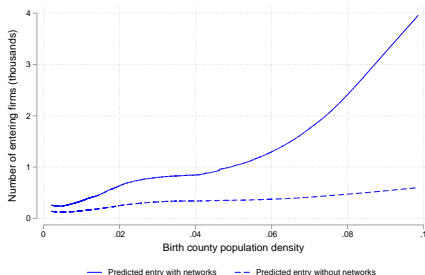
Source: SAIC registration database, model generated data, and 1982 population census.

# Out of Sample Tests – Entry and Sectoral Concentration, 2005-2009

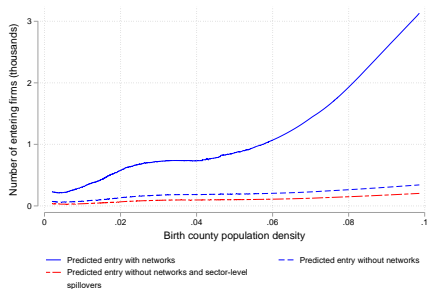


Source: SAIC registration database, model generated data, and 1982 population census.

# Counter-Factual Simulation: Effect of Community Networks on Entry



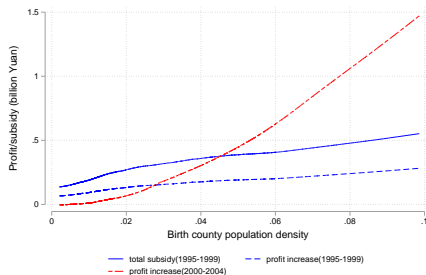
**(a)** Benchmark specification



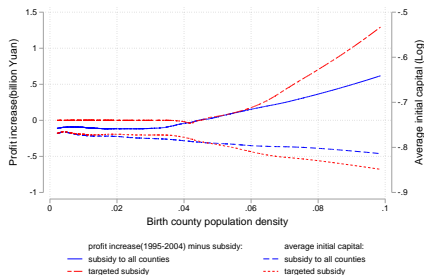
**(b)** Specification with sector-level spillovers

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

# Counter-Factual Simulation: Effect of Interest Rate Subsidy on Profits



(a) Subsidy to all counties



(b) Targeted subsidy vs. subsidy to all counties

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



# Conclusion

- There are no mark-ups in output prices or wedges in factor prices in our model, unlike the misallocation literature; e.g. Restuccia and Rogerson (2008); Hsieh and Klenow (2009)
  - Small firms and wide dispersion in firm size and productivity in our analysis are consequences of networks that substitute for missing markets, rather than inefficient taxes or regulations
  - Optimal second-best policies could entail subsidies targeting more connected communities, which would increase existing dispersion and induce even smaller firms to enter
  - More generally, we would not want to infer that one developing economy is less efficient than another because it has smaller firms or greater dispersion in firm size

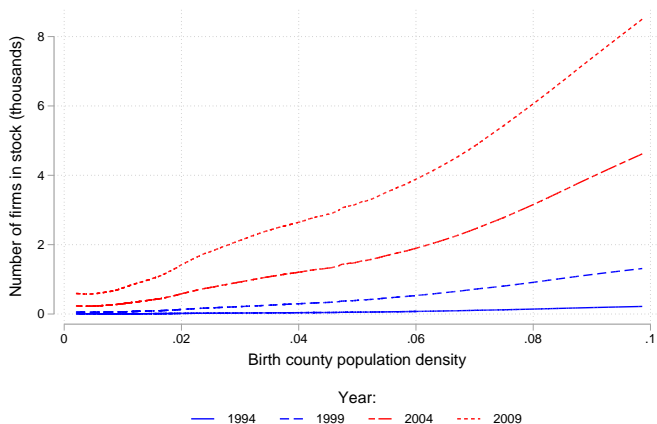
# Conclusion

- An additional implication of our network-based analysis is that subsidies should account for intra-community spillovers and individual ability
  - Existing efforts to stimulate entrepreneurship through business training programs or business plan competitions (McKenzie and Woodruff, 2014; McKenzie, 2017) do not incorporate these spillovers, potentially resulting in a substantial loss in efficiency
  - At the same time, policies that target more connected communities are likely to exacerbate existing inter-community inequality, while promoting intra-community equality, with complex distributional consequences

# Local Social Interactions, Trust and Population Density

Dependent variable:	frequency of chatting per month with local residents	whether the respondent chats most with a local resident	trust in local residents
	(1)	(2)	(3)
Population density	5.123*** (1.837)	0.075** (0.031)	0.465*** (0.160)
Constant	17.849*** (0.965)	0.229*** (0.017)	6.619*** (0.083)
Observations	8,572	20,070	19,389

# Stock of Firms and Population Density



Source: Registration database, and 1982 population census.