Online Appendix
Special Deals from Special Investors

(Not for publication)

A Registered Capital

According to China's Company Law (Article 26), registered capital is the total amount of capital all the shareholders of a limited liability company are obligated to pay into the company's accounts. The amount of registered capital is publicly available information in the company's business license and represents the maximum liability of the company's shareholders. Since early 2014 most shareholders do not actually have to transfer these funds into the company's accounts, but it is still the case that the shareholders are legally liable up until the stated registered capital. For example, if some one starts a company and claims that the company's registered capital is 1 million yuan. If the owner only pays 500 thousand yuan into the company's bank account (called paid-in capital) and the company goes bankrupt and owes a million yuan, the debtor can seize the remaining 500 thousand yuan from the owner's personal assets.

Table A.I compares a company's registered capital with other more standard metrics. Specifically, in the first panel, we merge the 2013 firm registration records with the 2013 Chinese Annual Industrial Survey which provides more financial information of industrial firms in China. The variable we take from the registration records is the firm's registered capital; the variables we take from the firm survey are the firm's reported total assets and sales. We then regress the firm's registration capital on total assets (first two columns) or sales (last two columns). For example, a regression of firm's log registered capital on its log total assets (sales) yields a coefficient of 0.93 (0.65) with a R^2 of 0.48 (0.16). If we drop the bottom 25% firms, the estimated coefficient is almost 1.

Shareholders are allowed to change (mostly increase) the registered capital of their company. To do this, they have to report to the local office of State Administration for Market Regulation. Once the application is approved, the company's business license will be changed accordingly. As a result, what we see from the 2013 firm registration records is the firm's most up-to-date registered capital by the end of 2013. We do not

Table A.I: Registered Capital vs Total Assets and Sales for Industrial Firms

	Dependent Variable: Registered Capital 2013			
	Survey 2013			
	All Firms	Assets > p(25)	All Firms	Sales > p(25)
Total Assets	0.926***	0.965***		
	(0.000)	(0.000)		
Sales			0.648***	0.745***
			(0.000)	(0.000)
# of Obs.	297,000	222,749	297,038	222,778
Adjusted R^2	0.48	0.41	0.16	0.16
	Survey 2007			
	All Firms	Assets > p(25)	All Firms	Sales > p(25)
Total Assets	0.966***	0.998***		
	(0.000)	(0.000)		
Sales			0.862***	0.922***
			(0.000)	(0.000)
# of Obs.	183,259	137,442	182,995	137,243
Adjusted R^2	0.53	0.47	0.27	0.27
	Survey 2002			
	All Firms	Assets > p(25)	All Firms	Sales > p(25)
Total Assets	0.930***	0.980***		
	(0.000)	(0.000)		
Sales			0.899***	0.967***
			(0.000)	(0.000)
# of Obs.	82,581	61,693	81,549	61,160
Adjusted R^2	0.53	0.48	0.31	0.30
	Survey 1998			
	All Firms	Assets > p(25)	All Firms	Sales > p(25)
Total Assets	0.902***	0.943***		
	(0.000)	(0.000)		
Sales			0.879***	0.952***
			(0.000)	(0.000)
# of Obs.	54,806	41,103	54,173	40,629
Adjusted R^2	0.53	0.45	0.29	0.29

Note: The unit of observation are firms in the 1998, 2002, 2007, and 2013 industrial survey matched with firms in the 2013 registration data. The dependent variable is log registered capital from the 2013 registration data. The independent variables are log total assets or log total sales from the 2013, 2007, 2002, and 1998 industrial firm surveys.

have information on the firm's historical registration capital. However, since we use registered capital as a proxy for firm size, we can check how the 2013 registration data aligns with contemporaneous data on total assets and sales in previous years for the same firm. The second through fourth panels in Table A.I show the results when we merge the 2013 firm registration records with earlier industrial surveys in 2007, 2002, and 1998. As can be seen, the regression coefficients are virtually the same as in the top panel. This suggests that using registered capital as a proxy for the firm's size in previous years is somewhat reasonable.

B Inferred Historical Data

The firm registration records include both active firms and firms which have been closed. For firms that have been closed, the records contain their exit years. By combining the information of firm's registration and exit year, we're able to infer historical data. Table A.II illustrates our approach for a hypothetical example. Suppose that there are 5 firms in the 2019 firm registration records. Three of them (A, C and E) are still active by the end of 2019, one (B) was closed in 2005 and another (D) was closed in 2015. To identify historical active firms at the end of year t, we select the firms established before or in year t and that have not been closed by the end of year t. Thus, as shown by Table A.II, at the end of 2000, active firms include A, B, C and D; at the end of 2010, active firms include A, C, D and E; while at the end of 2019, active firms include A, C and E.

Table A.II: Method for Inferring Historical Data

Firm	Registration Year	Exit Year	Active Firm in		
			2000	2010	2019
A	1985	.	Yes	Yes	Yes
B	1985	2005	Yes	–	–
C	1995	.	Yes	Yes	Yes
D	1995	2015	Yes	Yes	–
E	2005	.	–	Yes	Yes

For the sample of firms we infer as active in a given year, we then identify each firm's ultimate owners and their ownership equity shares using the information of firms' immediate shareholders in 2019 (or 2013). The implicit assumption is that the immediate shareholders of a firm is constant over time. Of course this assumption is not true for some firms, as a firm can have new shareholders, or some shareholders sell their equity, or the equity shares change. There's no solid data on the firm's historical shareholders in all years.[18] However, we do have contemporaneous registration data in 2013 and 2019. We can gauge the bias due the assumption we make that the most recent shareholder information is the same as in the past using these two datasets. Specifically, we compare the key results using data from 2013 inferred from 2019 data with those from the real 2013 data.

Table A.III summarizes the results of the comparison. The numbers of active firms are very close in inferred and real data. The number of ultimate owners in inferred data is slightly larger than that in real data. In particular our inference using the 2019 data overstates the share of the state owners in 2013. Our inference using the 2019 data also understates the share of connected private owners in 2013 by about 3 percentage points. On the other hand, the numbers of connected state owners and average downward connections per state owner are very similar in the inferred and real data.

The understatement of the share of connected private owners by the inferred data suggests that we possibly miss some connections between private owners. As in the right panel of Figure A.1, the numbers of owners with distance 1 and 2 are almost the same in the inferred and real data, but the numbers of owners with distance ≥ 3 in the inferred data are significantly smaller than those in real data. This is also confirmed by the less average downward connections per private owner (see left panel in Figure A.1).

Underestimating the historical number and share of indirectly connected private owners implies that we may overestimate the expansion and growth effect of indirectly connected private owners. We now assess the difference this makes for our estimate of the aggregate effect of the expansion of the connected sector with the "real" historical data. Specifically, we calibrate γ_d in the real historical data. Figure A.2 plots these numbers, showing both γ_d calibrated from the real 2013 data and from the 2013 data inferred from 2019 data. As we can see, γ_1 and γ_2 calibrated using inferred and real

[18] For some provinces and several years, there do exist some text records showing changes of firms' shareholders, but not for all provinces. Another problem is that there's no encrypted personal ID for individual shareholders in these records. These two shortages make us not able to recover firms' real historical shareholders from these text records.

Table A.III: Inferred vs Real 2013 Data

	Real	Inferred
# Active Firms	14,013,154	14,125,941
# Ultimate Owners	28,947,269	30,132,765
State	31,468	36,686
Private	28,915,801	30,096,079
% Registered Capital		
State	24.9%	31.0%
Private	75.2%	69.0%
# Connected State Owners	6,559	6,734
# Downward Connections per State Owner	9.5	10.2
% Registered Capital		
Connected State Owners	23.7%	29.9%
Directly Connected Private Owners	13.3%	14.1%
Indirectly Connected Private Owners	14.9%	11.1%

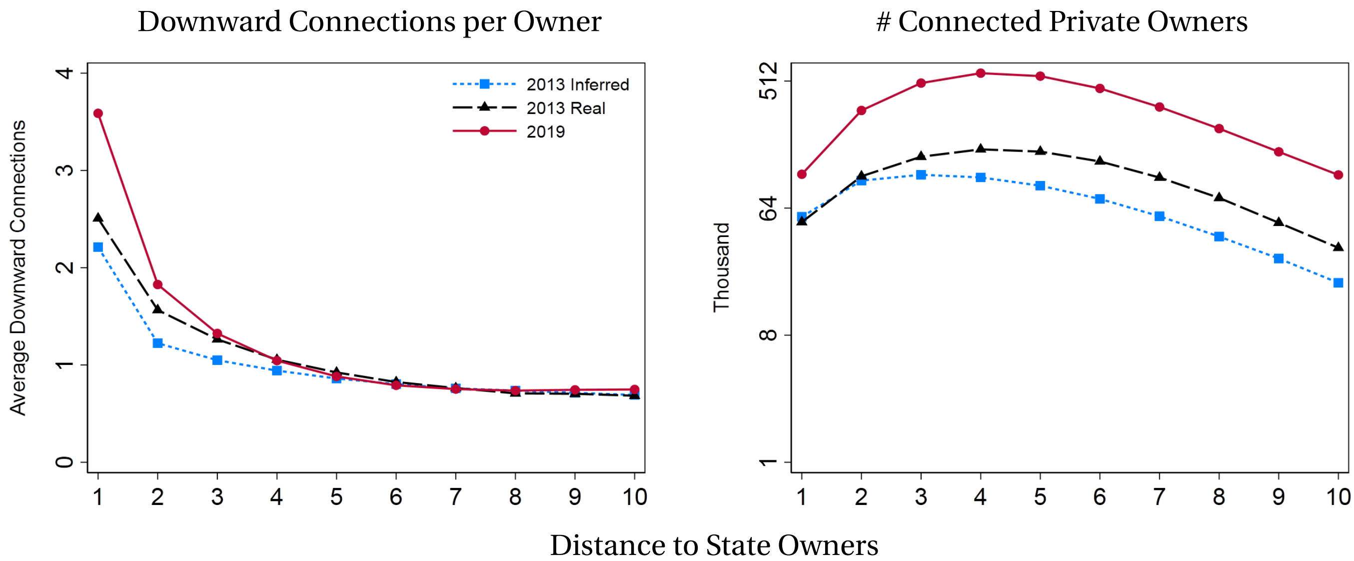
Note: "Real" uses the 2013 data. "Inferred" uses the 2019 data to calculate the statistics of firms and owners in 2013.

data are very close. But for $d = 3, 4, 5, 6$, the real data give higher γ_d. Then, as in Table 8, we measure how much the increase in γ explains economic growth by comparing γ calibrated using 2019 data with γ calibrated using inferred and real 2013 data respectively. The results are summarized in Table A.IV. When using the real historical data, the output growth caused by the increase of γ is 5.6%, just 0.1 percentage point lower than that calculated using inferred data. But the growth contributed by increase of benefit of indirect connection is smaller (2.4%) compared with that from inferred data (3.5%).

C Identification of State Owners

We identify whether a shareholder is state owner or not by its name. Specifically, we compiled a list of Chinese central, provincial, city- and county-level administrative divisions. We also compiled a list of all the departments at each level of government. The list of state owner names includes all possible combinations of division name and department name. Then we match the name of the shareholder in the registration data

Figure A.1: Connected Private Owners, Inferred vs Real 2013 Data



Note: Left panel shows the average number of downward connections per private owner by distance from state using inferred 2013 data, real 2013 data and real 2019 data. Right panel shows the total number of connected private owners by distance from state owners.

Table A.IV: Aggregate Output Growth From Connected Investors, 2013-2019: Inferred vs Real 2013 Data

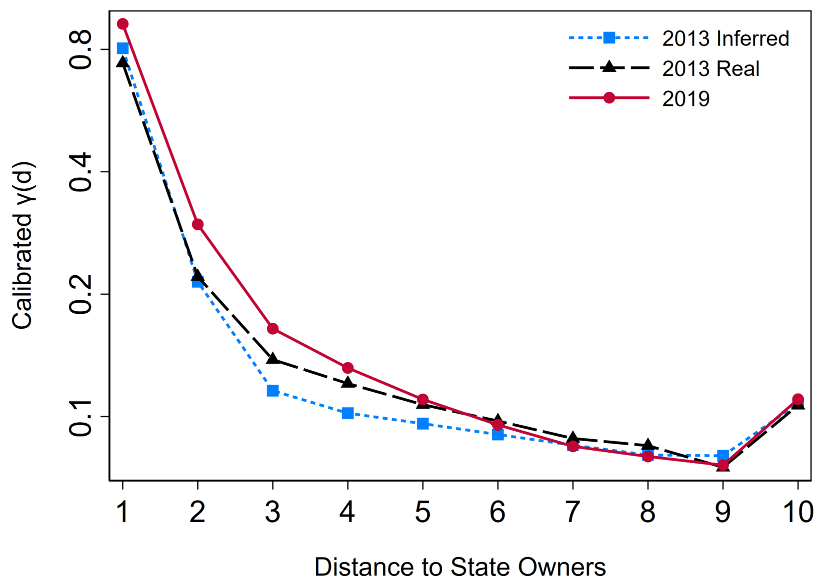
	Inferred 2013 Data	Real 2013 Data
Output Growth in Model (% per year):		
$\Delta\gamma_d, d \geq 1$	5.7%	5.6%
$\Delta\gamma_d, d = 1$ only	2.1%	3.2%
$\Delta\gamma_d, d \geq 2$ only	3.5%	2.4%
Output Growth in Model (% per year):		
Directly Connected Owners	2.4%	3.1%
Indirectly Connected Owners	4.1%	3.3%

Note: Output growth from "connected" private owners is average annual GDP growth in the model due to the change in γ we infer between 2013 (inferred and real data) and 2019.

with our list of state owners.

We treat all the departments that belong to the same level of government as one

Figure A.2: Calibrated Benefit from Connected Owners:
Inferred vs Real 2013 Data



Note: Figure plots benefit from a connected owner γ_d in 2019, inferred and real 2013 data.

state owner. For example, we treat the Department of Finance of Shandong Province and the SASAC (State-Owned Assets Supervision and Administration Commission) of Shandong Province as the same owner as both are different departments of the Shandong provincial government. However, we assume that the government of Shandong Province and the government of Jinan City (the capital city of Shandong) are two different owners. The exception to this rule are the largest state firms which we classify as separate owner, even if they are owned by the SASAC of the relevant government body. For example, although SAIC is owned by Shanghai's SASAC, we assume SAIC is a separate state owner.

D Model Proofs

Sufficient Conditions for Partially Connected Equilibrium: Denote $\{\tilde{A}^d\}_{d=1}^{\overline{d}}$ as the solution of:

$$\left[(1-\tau+\gamma_d)^\theta-(1-\tau)^\theta\right]\ A_i^\theta-\lambda\alpha\left(\frac{\tilde{N}_d}{\tilde{N}_{d-1}}\right)^{\alpha-1}=0,$$

where $\tilde{N}_d=\int_{\tilde{A}^d}^{\tilde{A}^{d-1}}f(A)dA$ and $\tilde{A}^0$ is set to $+\infty$; $\gamma_d>0$ if $d\le\overline{d}$ and $\gamma_d=0$ for all $d>\overline{d}$; and the number of private owners is normalized to 1.

If $\tilde{A}^{\overline{d}}>1$, the lower bound of private owner's TFP distribution, in the equilibrium only a subset of the private owners choose to connect (directly or indirectly) with the state.

Proof of Proposition 1: There is a unique distance d^* and TFP A^{d^*} such that there are owners connected with each distance $d\in[1,d^*]$, no owner is connected with distance $d>d^*$ and all owners with $A_i>A^{d^*}$ are connected and all owners with $A_i<A^{d^*}$ are not connected.

We will first prove the existence and uniqueness of d^* and skip to the proofs of Proposition 2 and 3 by assuming that there exist unique equilibrium. And then, we'll come back and prove the rest of Proposition 1 (existence and uniqueness of equilibrium).

We will prove that $d^*=\overline{d}$, where $\overline{d}$ satisfies $\gamma_d>0$ if $d\le\overline{d}$ and $\gamma_d=0$ for all $d>\overline{d}$. Let's denote:

$$\pi(d|A_i)=\left[(1-\tau+\gamma_d)^\theta-(1-\tau)^\theta\right]A_i^\theta,$$

which is private owner i's revenue gain from choosing $d_i=d$ compared with being unconnected.

Suppose there exist a set of private owners, whose solutions to equation (4) $d_i>\overline{d}$. As $\gamma_d=0$ for $d>\overline{d}$, these private owners have zero revenue gain from being connected, $\pi(d_i|A_i)=0$. However, the sum of their net gains from being the connected investor to other private owners net of the price they pay to their connected investors is negative, because (1) most of the price collected by them as connected investors and paid by them to their connected investors can cancel out except the price paid by owners at level $\overline{d}+1$ to owners at level $\overline{d}$; and (2) they also need to bear non-negative costs for providing the connections. As a result, the sum of these owners' total gain from being connected is negative, which means that at least some of them have negative gain and staying unconnected would make them better off. It contradicts with that $d_i > \overline{d}$ is their solution to equation (4). So we prove that the above mentioned set of private owners don't exist. Thus, if $\gamma_d = 0$, no owner will choose $d_i = d$.

We next show $N_d > 0$ if $N_{d-1} > 0$ and $\gamma_d > 0$, where N_d is the number of private owners choosing $d_i = d$ in the equilibrium. Suppose instead $N_d = 0$. This cannot be true since for any unconnected owner i, choosing $d_i = d$ would be strictly better off as $\pi(d|A_i) > 0$, the price paid to connected investor at distance $d - 1$ is zero because $N_d = 0$, and the net gain from providing downward connections is non-negative (see equation (4)). Finally, $N_0 = N_s > 0$ allows us to establish recursively that $N_d > 0$ for all d with $\gamma_d > 0$.

So we prove that there is a unique d^* which equals $\overline{d}$, that there are owners connected with each distance $d \in [1, d^*]$ and no owner is connected with distance $d > d^*$.

Proof of Proposition 2: Among the set of private owners that are connected, owners with $A_i \in [A^d, A^{d-1})$ choose to connect at distance d where $\{A^d\}_{d=0}^{d^*}$ is a strictly decreasing sequence in d.

We first prove that in the equilibrium, private owners are sorted by TFP into distance d, which means that owners with higher TFP choose relatively smaller d.

Suppose there exist owner i and j, with $A_i > A_j$, and they choose $d_i > d_j$ respectively. As owner j's solution to equation (4) is d_j, we have $U(d_j|A_j) \geq U(d_i|A_j)$. Holding other thing constant, $U(d_j|A) - U(d_i|A)$ is a strictly increasing function of A. Thus we should also have $U(d_j|A_i) \geq U(d_i|A_i)$, which contradicts with that d_i is owner i's solution to equation (4). So in the equilibrium, private owners are sorted by TFP into distance d.

As we've already proved that $N_d > 0$ for all d with $\gamma_d > 0$ in Proposition 1, we have that in the equilibrium, there exist a strictly decreasing sequence $\{A^d\}_{d=0}^{d^*}$ that owners with $A_i \in [A^d, A^{d-1})$ choose to connect at distance d.[19]

Proof of Proposition 3: The average number of downward connections per owner $\frac{N_{d+1}}{N_d}$ is decreasing in distance from the state for $d \in [0, d^* - 1]$.

In Proposition 2, we have proved that an equilibrium can be characterized by a strictly decreasing sequence $\{A^d\}_{d=0}^{d^*}$. Note that given the distribution of private owner's

[19] A^0 is set to $+\infty$.

TFP, $\{A^d\}_{d=0}^{d^*}$ also determine the sequence of number of private owners at each distance d, $\{N_d\}_{d=1}^{d^*}$. The two sequences should also satisfy $U(d|A^d) = U(d+1|A^d)$ for $d \in [1, d^*-1]$ and $U(d^*|A^{d^*}) = (1-\tau)^\theta (A^{d^*})^\theta$.[20]

Let $d = d^* - 1$, we have

$$\left[(1-\tau+\gamma_{d^*-1})^\theta - (1-\tau+\gamma_{d^*})^\theta\right](A^{d^*-1})^\theta \\ -\lambda\alpha\left(\frac{N_{d^*-1}}{N_{d^*-2}}\right)^{\alpha-1} + \lambda(\alpha-1)\left(\frac{N_{d^*}}{N_{d^*-1}}\right)^{\alpha} + \lambda\alpha\left(\frac{N_{d^*}}{N_{d^*-1}}\right)^{\alpha-1} = 0.$$

As the first term of the left-hand side is strictly positive, we should have

$$-\lambda\alpha\left(\frac{N_{d^*-1}}{N_{d^*-2}}\right)^{\alpha-1} + \lambda(\alpha-1)\left(\frac{N_{d^*}}{N_{d^*-1}}\right)^{\alpha} + \lambda\alpha\left(\frac{N_{d^*}}{N_{d^*-1}}\right)^{\alpha-1} < 0,$$

$\Rightarrow$

$$\lambda\alpha\left(\frac{N_{d^*}}{N_{d^*-1}}\right)^{\alpha-1} < \lambda\alpha\left(\frac{N_{d^*-1}}{N_{d^*-2}}\right)^{\alpha-1} - \lambda(\alpha-1)\left(\frac{N_{d^*}}{N_{d^*-1}}\right)^{\alpha} < \lambda\alpha\left(\frac{N_{d^*-1}}{N_{d^*-2}}\right)^{\alpha-1}.$$

So we shows that $\frac{N_{d^*}}{N_{d^*-1}} < \frac{N_{d^*-1}}{N_{d^*-2}}$.

For any $d \in [0, d^*-2]$, we have

$$\left[(1-\tau+\gamma_{d})^\theta - (1-\tau+\gamma_{d+1})^\theta\right](A^{d})^\theta \\ -\lambda\alpha\left(\frac{N_{d}}{N_{d-1}}\right)^{\alpha-1} + \lambda(\alpha-1)\left(\frac{N_{d+1}}{N_{d}}\right)^{\alpha} + \lambda\alpha\left(\frac{N_{d+1}}{N_{d}}\right)^{\alpha-1} - \lambda(\alpha-1)\left(\frac{N_{d+2}}{N_{d+1}}\right)^{\alpha} = 0.$$

As the first term of the left-hand side is strictly positive, we should have the rest part of the left-hand side negative, which means

[20] If $U(d|A^d) < U(d+1|A^d)$, it means owners with $A_i = A^d$ would be better off by choosing $d+1$ rather than d, which contradicts with Proposition 2. If $U(d|A^d) > U(d+1|A^d)$, it means that owners with $A_i \in [A^d - \varepsilon, A^d)$ (ε is a positive small number) would be better off by choosing d rather than $d+1$, which also contradicts with Proposition 2.

$$\lambda\alpha\left[\left(\frac{N_d}{N_{d-1}}\right)^{\alpha-1}-\left(\frac{N_{d+1}}{N_d}\right)^{\alpha-1}\right]>\lambda(\alpha-1)\left[\left(\frac{N_{d+1}}{N_d}\right)^{\alpha}-\left(\frac{N_{d+2}}{N_{d+1}}\right)^{\alpha}\right].$$

This equation says that $\frac{N_d}{N_{d-1}} > \frac{N_{d+1}}{N_d}$ if $\frac{N_{d+1}}{N_d} > \frac{N_{d+2}}{N_{d+1}}$. Since we've already proved $\frac{N_{d^*-1}}{N_{d^*-2}} > \frac{N_{d^*}}{N_{d^*-1}}$, we can recursively prove that the average number of downward connections per owner $\frac{N_{d+1}}{N_d}$ is decreasing in distance from the state.

Now we come back to prove the rest of **Proposition 1**: There exist a unique equilibrium, which can be characterized by a strictly decreasing sequence $\{A^d\}_{d=0}^{d^*}$, that all owners with $A_i > A^{d^*}$ are connected and all owners with $A_i < A^{d^*}$ are not connected.

First we prove the existence of the equilibrium.

Let's define $X = [N_1, N_2, \cdots, N_{d*}]$, which is defined on: $[0,1]^{d^*} \bigcap \sum_{i=1}^{d^*} N_i <= 1$. It is obvious a convex compact set. Then we define a function $f(X)$, which is described as below:

Step 1: Note that given the distribution of owner's TFP, $\{A^d\}_{d=0}^{d^*}$ is one-to-one mapped to $\{N_d\}_{d=1}^{d^*}$. So first get corresponding $[A^1, A^2, \cdots, A^{d^*}]$ from X.

Step 2: Holding other A^d constant, get new $A^{1\prime}$ by solving $U(1|A^{1\prime}) - U(2|A^{1\prime}) = 0$.[21] Update the sequence to $[A^{1\prime}, A^2, \cdots, A^{d^*}]$.

Step 3: Repeat step 2 for $d = 2, 3, \cdots, d^*$ and finally get $[A^{1\prime}, A^{2\prime}, \cdots, A^{d^*\prime}]$. Get $X' = [N_1', N_2', \cdots, N_{d*}']$, which is also defined on $[0,1]^{d^*} \bigcap \sum_{i=1}^{d^*} N_i <= 1$.

$X' = f(X)$ is a continuous function from a convex compact set to itself. Using Brouwer fixed-point theorem, there exist X_0 that $X_0 = f(X_0)$. By mapping X_0 to corresponding $[A^1, A^2, \cdots, A^{d^*}]$, we can get the equilibrium.

Proof of Proposition 4: A non-decreasing change in γ_d increases: (i) the number of connected owners; (ii) the maximum distance from the state d^*; (iii) the average number of downward connections per owner for some d; and (iv) aggregate output.

Note that the equilibrium can be characterized by a strictly decreasing sequence $\{A^d\}_{d=0}^{d^*}$, which is the solution of the following equation system:

[21] $U(1|A^{1\prime}) - U(2|A^{1\prime}) = 0$ has unique solution on $(A^2, +\infty)$, because (1) the left-hand side is an strictly increasing continuous function, and (2) $U(1|A^2) - U(2|A^2) = -\infty$ and $U(1|A^0) - U(2|A^0) = +\infty$.

$$
\begin{aligned}
U(d|A^d) &= U(d+1|A^d) \text{ for } d \in [1, d^*-1], \\
U(d^*|A^{d^*}) &= (1-\tau)^\theta (A^{d^*})^\theta.
\end{aligned} \tag{5}
$$

Take derivatives with respect to $\{\gamma_d\}_{d=1}^{d^*}$. We have:

$$
(YAY^T + C)X \begin{bmatrix}
\frac{\partial A^1}{\partial \gamma_1} & \frac{\partial A^1}{\partial \gamma_2} & \frac{\partial A^1}{\partial \gamma_3} & \cdots & \frac{\partial A^1}{\partial \gamma_{d^*}} \\
\frac{\partial A^2}{\partial \gamma_1} & \frac{\partial A^2}{\partial \gamma_2} & \frac{\partial A^2}{\partial \gamma_3} & \cdots & \frac{\partial A^2}{\partial \gamma_{d^*}} \\
\frac{\partial A^3}{\partial \gamma_1} & \frac{\partial A^3}{\partial \gamma_2} & \frac{\partial A^3}{\partial \gamma_3} & \cdots & \frac{\partial A^3}{\partial \gamma_{d^*}} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial A^{d^*}}{\partial \gamma_1} & \frac{\partial A^{d^*}}{\partial \gamma_2} & \frac{\partial A^{d^*}}{\partial \gamma_3} & \cdots & \frac{\partial A^{d^*}}{\partial \gamma_{d^*}}
\end{bmatrix} = B, \tag{6}
$$

where:

- $X = \text{diag}(f(A^1), f(A^2), f(A^3), \cdots, f(A^{d^*}))$. $f(x)$ is the probability density function of private owner's TFP distribution. We assume that A_i follows a Pareto distribution with scale parameter 1 and shape parameter k;
- $Y = \begin{bmatrix}
1 & -\frac{N_1+N_2}{N_1} & \frac{N_3}{N_2} & 0 & 0 & \cdots & 0 \\
0 & 1 & -\frac{N_2+N_3}{N_2} & \frac{N_4}{N_3} & 0 & \cdots & 0 \\
0 & 0 & 1 & -\frac{N_3+N_4}{N_3} & \frac{N_5}{N_4} & \cdots & 0 \\
0 & 0 & 0 & 1 & -\frac{N_4+N_5}{N_4} & \cdots & 0 \\
0 & 0 & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 1
\end{bmatrix}$;
- $A = \text{diag}(a_{11}, a_{22}, a_{33}, \cdots, a_{d^*d^*})$. $a_{ii} = \lambda\alpha(\alpha-1)\left(\frac{N_i}{N_{i-1}}\right)^{\alpha-1}\frac{1}{N_i}$;
- $B = \begin{bmatrix}
-b_{11} & b_{12} & 0 & \cdots & 0 \\
0 & -b_{22} & b_{23} & \cdots & 0 \\
0 & 0 & -b_{33} & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & -b_{d^*d^*}
\end{bmatrix}$. $b_{ij} = \theta\left(1-\tau+\gamma_j\right)^{\theta-1}(A^i)^\theta$ if $i = j$ or

$i = j - 1$;

- $C = \text{diag}(c_{11}, c_{22}, c_{33}, \cdots, c_{d^*d^*})$. $c_{ii} = \left[(1-\tau+\gamma_i)^\theta - (1-\tau+\gamma_{i+1})^\theta\right](A^i)^\theta \frac{\theta}{k} \frac{1}{\sum_{d=1}^{i} N_d}$ if $i \in [1, d^*-1]$ and $c_{d^*d^*} = \left[(1-\tau+\gamma_{d^*})^\theta - (1-\tau)^\theta\right](A^{d^*})^\theta \frac{\theta}{k} \frac{1}{\sum_{d=1}^{d^*} N_d}$.

It is useful to note that if there exist a social planner to allocate each private owner to a certain distance to the state and its goal is to maximize the total net gain in the economy:

$$\max_{\{A^d\}_{d=1}^{d^*}} : \Pi = \sum_{d=1}^{d^*} \int_{A^d}^{A^{d-1}} (1-\tau+\gamma_d)^\theta x^\theta f(x)dx + \int_{\underline{A}}^{A^{d^*}} (1-\tau)^\theta x^\theta f(x)dx$$
$$- \sum_{d=1}^{d^*} \lambda \left(\frac{N_d}{N_{d-1}}\right)^\alpha N_{d-1}.$$

The first-order conditions are

$$-f(A^d)U(d|A^d) + f(A^d)U(d+1|A^d) = 0 \text{ for } d \in [1, d^*-1],$$
$$-f(A^{d^*})U(d^*|A^{d^*}) + f(A^{d^*})((1-\tau)(A^{d^*})^\theta = 0.$$

which are identical to the equation system determining $\{A^d\}_{d=0}^{d^*}$. Further more, the Hessian matrix is $-X\left(YAY^T + C\right)X^T$, which is a negative definite matrix. As a result, the solution of the equation system, $\{A^d\}_{d=0}^{d^*}$, is also a local maximum of the social planner's problem.

Then we come back to the derivatives. Equation (6) gives:

$$\left(\frac{\partial A^i}{\partial \gamma_j}\right)_{d^* \times d^*} = X^{-1}(AY^T + Y^{-1}C)^{-1}Y^{-1}B.$$

It is easy to show that elements of matrix $Y^{-1}B$ on and above the diagonal are all negative, and those below the diagonal are all zero. If we can prove that elements of matrix $(AY^T + Y^{-1}C)^{-1}$ on and below the diagonal are all positive, we will have

$$\frac{\partial A^i}{\partial \gamma_j} < 0 \text{ if } i \geq j,$$

which means that if we increase some γ_d, all the $\{A^d, A^{d+1}, \cdots, A^{d^*}\}$ would be smaller in the new equilibrium. Then we will have that (i) number of connected owners increases, as A^{d^*} is smaller; (ii) the average number of downward connections per owner for some d becomes larger, otherwise it will contracts with (i); and (iii) aggregate output increases.[22]

If we increase $\gamma_{\overline{d}+1}$ from zero to a positive value, we will see the maximum distance from the state will becomes larger. Otherwise the maximum distance will remain unchanged.

We prove that elements of matrix $(AY^T)^{-1}$ on and below the diagonal are all positive. This proves the proposition under the condition that the matrix $(AY^T + Y^{-1}C)^{-1}$ has a positive lower-triangle.[23]

E Structural Estimation

The lower bound of private owner's productivity is set to 1. τ is set to 0.90 to match the average firm size of unconnected private owners relative to that of directly connected private owners (according to 2019 data).

We next calibrate k, the shape parameter of the productivity distribution. Denote $\overline{A}_d^{Data}$ the average productivity of private owners with $d_i = d$ in the data.[24] The idea is to back out k by matching $\overline{A}_d^{Data}$. Given $\{N_d/N_p\}$, $d = 1, 2, \cdots$, $\overline{A}_d$ will only depend on k. As will be clear below, the estimated model is constructed to exactly match $\{N_d/N_p\}$ in the data. Therefore, we can directly back out $\overline{A}_d(k, \{N_d/N_p\})$ without knowing the other parameter values. Specifically, k is calibrated by

$$\hat{k} = \arg\min_k \sum_{d=1}^{d^*} \left(\log(\overline{A}_d^{Data}) - \log(\overline{A}_d(k, \{N_d/N_p\})) \right)^2,$$

where d^* is the maximum distance to the state in the steady state, which is set to 10. The estimated k is 272.

Denote $\phi \equiv [\theta, \lambda, \alpha]$. Given $\hat{k}$ and ϕ, we can use equations (5) to back out γ_d^t for $d = 1, 2, \cdots, d^*$ and $t \in \{2000, 2010, 2019\}$ by matching N_d/N_p in the 2000, 2010 and 2019 data respectively. Denote $\gamma_d^t(\hat{k}, \phi)$ the calibrated benefit of connection associated with $\hat{k}$ and ϕ.

Given the estimated $\hat{k}$, we then estimate ϕ by

[22] The effect on aggregate output is easier to be proved in the social planner's problem.
[23] This is confirmed by extensive simulations.
[24] Since we don't have TFP data for year 2019, we use 2013 data instead.

$$\hat{\phi} = \arg\min_{\phi} \sum_{t\in\{2000,2010,2019\}} \sum_{d=1}^{d^*} \Bigg[\Big(\log(Y_{d,t}^{Data}) - \log(Y_{d,t}(\hat{k},\phi,\gamma_d^t(\hat{k},\phi))) \Big)^2 + \Big(\log(\overline{Y}_{d,t}^{Data} - \log(\overline{Y}_{d,t}(\hat{k},\phi,\gamma_d^t(\hat{k},\phi))) \Big)^2 \Bigg].$$

Here, $Y_{d,t}^{Data}$ and $\overline{Y}_{d,t}^{Data}$ are the average firm size of private owners with $d_i = d$ and the average size of private owners with $d_i = d$ for year $t \in \{2000, 2010, 2019\}$, normalized by respective values of unconnected private owners. $Y_{d,t}(\hat{k},\phi,\gamma_d^t(\hat{k},\phi))$ and $\overline{Y}_{d,t}(\hat{k},\phi,\gamma_d^t(\hat{k},\phi))$ are the corresponding values predicted by the model with parameter values of $\hat{k}$, ϕ and the associated calibrated values of $\gamma_d^t(\hat{k},\phi)$.

Figure A.3 shows the fit of the model in terms of average firm size and owner size (across all the firms of each owner) by distance to the state.

Figure A.3: Target and Estimated Moments