Special Deals from Special Investors

The Rise of State-Connected Private Owners in China

Chong-En Bai
Tsinghua University

Chang-Tai Hsieh
University of Chicago and NBER

Zheng Song
Chinese University of Hong Kong

Xin Wang *
Chinese University of Hong Kong

August 8, 2021

Abstract

We use administrative registration records with information on the owners of *all* Chinese firms to document their connections through equity investments. We document a hierarchy of private owners: the largest private owners have direct equity investments from state-owned firms, the next largest private owners have equity investments from private owners that themselves have equity ties with state owners, and the smallest private owners do not have any ties with state owners. The network of "state-connected" private owners has expanded over the last two decades. The share of registered capital of private owners with state-connected investors increased by almost 20 percentage points between 2000 and 2019, driven by two trends. First, state owners have increased their investments in joint ventures with private owners. Second, private owners with equity ties to state owners also increasingly invest in joint ventures with other (smaller) private owners. The expansion in the number of state-connected private owners *may* have increased aggregate output of the private sector by 2.5% a year between 2000 and 2019.

*We are grateful to Emanuele Colonelli, Yasheng Huang, Ernest Liu, Yingyi Qian, Aleh Tsyvinski, Wei Xiong and Xiaodong Zhu for helpful discussions. Jin Han and Yanzun Yang provided excellent research assistance. Zheng Song acknowledges financial support from the Research Grant Council of Hong Kong.

1. Introduction

In the late 1990s, a young Chinese auto manufacturer called Chery found itself up against what seemed like an unsurmountable obstacle.[1] Chery was successfully producing low-priced knockoffs of the Volkswagen Jetta, but it did not have a license to make cars. It had appealed to Chinese central planners multiple times for the necessary license, but the authorities were adamant that companies such as Chery were not part of their plan for China's automobile industry. The goal of the Chinese authorities at the time was to consolidate production around a small number of state-led giants such as Shanghai Automobile and First Auto Works. Small companies such as Chery that would compete with the industrial giants were forbidden.

In desperation Chery turned to Shanghai Automobile. It struck a deal where the state-owned giant took a 20% equity stake in Chery. Legally, this made Chery a "subsidiary" of Shanghai Automobile, which enabled Chery to get a car license from the Chinese authorities. Shanghai Automobile eventually sold its 20% equity stake back to Chery, which has gone on since then to become the largest exporter of cars and the 4th largest car producer in China in the late 2000s.

The role played by Shanghai Automobile in Chery's growth is an example of the critical role of what the Chinese call a "politically-connected investor" or a "protective umbrella" in enabling firms to grow. In this paper we use administrative registration data on the universe of Chinese firms from 2000 to 2019 to document the importance of "connected" investors such as Shanghai Automobile in the growth of Chinese private owners over the last two decades. A key feature of the registration data is that it identifies the owners of the universe of Chinese firms. We use this ownership information to identify firms with equity investments from state-owned firms or private owners with equity ties to state-owned firms.

This ownership information reveals two key facts. First, there is a clear hierarchy of private owners with respect to the closeness of their equity links with state owners. In 2019 state owners had equity stakes in the firms of more than one hundred thousand private owners. These private owners are the largest in China and also hold equity in the companies of other, typically smaller, private owners. In turn, these private owners

[1] This account of Chery is from Dunne (2011).

also invest in other, even smaller, private owners and so on. At the very bottom of the hierarchy, up to forty steps away from the state owners at the top, are the owners that do not invest in other owners. The very smallest private owners thus do not have any equity ties, direct or indirect, with state owners.

Our second finding is that the hierarchy of private owners with connected investors is relatively recent phenomena. In 2000 private owners with connected investors only accounted for about 14.1% of registered capital. By 2019, private owners with connected investors owned about 33.5% of all registered capital in China. The 19.4 percentage points increase in the share of connected private owners from 2000 to 2019 accounts for almost all of the increase in the share of *all* private owners over this period.

The growth of this hierarchy of connected owners is driven, in a proximate sense, by two related trends. First, conditional on investing in private owners, state owners on average had investments with less than 4 distinct private owners in 2000. By 2019, the average state owner had projects with 14 distinct private owners. The result is that the number of private owners pursuing joint ventures with state owners increased from about 45 thousand in 2000 to around 130 thousand by 2019.

Second, private owners associated with the state also now undertake more investments with other private owners. For example, the 45 thousand private owners pursuing joint ventures with state owners in 2000 themselves had joint ventures with less than 1 other private owner on average in that year. In 2019, the 130 thousand private owners directly connected with state owners were themselves the "connected investor" for more than 3 other private owners on average. The result is that number of private owners that the directly connected private owners invested in increased from 35 thousand in 2000 to more than 300 thousand by 2019. This effect is particularly dramatic for connected owners distant from the state. In 2000 for example, there were just around 4 thousand owners six or more steps away from the state. By 2019, there were more than *1.5 million* such owners.

By 2019 the net effect of the increase in connected private owners, and the growth of such owners after they became connected with a "connected investor," was that the assets of connected private owners accounted for 33.5% of total assets in China, or about 44% of total assets of all private owners. At the same time, the share of connected state owners at the "top of the food chain" of the connected sector, is merely 22.5%.

This is because politically connected investors are rarely the controlling shareholders. In the case of Chery, Shanghai Auto's stake was 20%. For the average private owner with joint ventures with state owners, the share of state owners was around 30% by cash flow rights and 35% by control rights in 2019.

We then filter these facts through the lens of a simple model where connected investors reduce "frictions" faced by private owners. In the model, an increase in the benefits provided by connected investors increases the number of connections per connected investor, total number of connected owners, share of connected owners in the economy, and aggregate output. We calibrate the increase in the benefits provided by connected owners from data on the number of connections made by each connected investor. We then filter this number through the model to estimate the contribution of the expansion of connected private owners to aggregate output. We find that this mechanism can explain a 2.5% annual growth in aggregate output of the private sector between 2000 and 2019.

This paper builds on multiple bodies of work. First, the closest predecessors of this paper are Bai et al. (2019) and Huang (2008). Bai et al. (2019) highlights the importance of informal institutions in the form of "special deals" by local governments in enabling private firms to grow; Huang (2008) argues that state-connected agents in China frequently get special deals. This paper focuses on a specific type of special deal that takes the form of connected *investors*, including private individuals that are connected to state owners, taking equity stakes in firms of private owners.

Second, there is a vast literature quantifying the economic effect of state ownership. Evidence from privatization episodes in many countries, including Mexico (La Porta and Lopez-de Silanes (1999)), Russia (Barberis et al. (1996)), and Eastern Europe (Frydman et al. (1999)), shows that state-owned firms are less efficient and that privatization generally results in gains in aggregate efficiency. The evidence from China also suggests that state-owned firms are less efficient and cause distortions.[2] Moreover, the massive exit and privatization of the smaller state firms in the late 1990s and early 2000s led to

[2]See Hsieh and Song (2015) and Brandt et al. (2020) for misallocation and entry barriers caused by state-owned firms.

modest gains in aggregate efficiency.[3] This paper shows that although firms that remain state owned in China are inefficient, they also increasingly invest in joint ventures with private owners. As a result, the fastest growing sector in China are the state connected private owners that are neither purely state-owned nor fully private. Furthermore, it is possible that the joint ventures of state owners, which are difficult to discern without data on the owners of all Chinese firms, may have resulted in large gains in aggregate efficiency.

Third, a large literature following La Porta et al. (1999) measures ownership concentration around the world and its correlates, including protection of investor rights, legal origin, and labor laws.[4] This paper builds on this literature in two respects. First, we identify the owners of *all* firms in China, not just the publicly listed firms. Second, we focus on equity links between owners in the form of joint ventures and show the prevalence of such alliances among all Chinese firms.

Fourth, there is by now a large literature on production networks and how shocks propagate through these networks. The network of owners in China we document is analogous to a production network where connected owners provide benefits to other owners through joint ventures. Perhaps the two most closely related papers in this literature are Baquee and Farhi (2020) and Liu (2019). Like us, they estimate the aggregate effect of micro-economic shocks in a network. Our paper differs in that the linkages in our model are endogenous. Private owners choose their location in the network, and their choices in turns change the benefits and the costs of a given location in the network. We therefore do not adopt Baquee and Farhi (2020) and Liu (2019)'s "sufficient statistic" approach but instead calibrate the model by choosing parameters that fit the observed (endogenous) network.

Our work is also related to Acemoglu and Azar (2020) and Taschereau-Dumouchel (2020) who also model how connections in a network endogenously respond to shocks to productivity or friction and how the effect of a connection change can be propagated and amplified through the endogenous network. Their mechanism for network formation is, however, different from ours. Our model is built on the span of control and

[3]See Hsieh and Song (2015) for details of the "Grasp the large, let go of the small" campaign in which the large state owned firms were consolidated into large conglomerates ("grasped") and smaller state firms were closed or sold ("let go"). See also Lardy (2014) on the falling share of the state sector over this period.

[4]See Aminadav and Papaioannou (2020) for a recent paper and a review of this literature.

the benefits conditional on distance to the state. And our focus is on how connections between state and private owners and between private owners respond to the benefits from connecting to the state.

Finally, four recent papers use the same registration records to explore the growth of China's private sector. Dai et al. (2019) identify community origins of entrepreneurs and investigate how the origin-based connections affect firm entry. Shi et al. (2020) finds a causal relationship between transfer of local government leaders and inter-regional investment flows. Brandt et al. (2019) show that a growing fraction of firms are started by serial entrepreneurs. Allen et al. (2019) construct a firm-to-firm equity investment network and estimate the effects of the firm's network. Our specific focus is on equity links between state and private owners, and between private owners with such equity ties and other private owners. We document a hierarchy of owners that transmits the benefit of special deals from state owners at the top to millions of private owners. We also document a rapid expansion of the hierarchy and argue that it may be a crucial force behind the growth of the private sector in China. Finally, Chen and Kung (2019) show that firms with owners that have personal ties to Politburo members get a substantial discount on land purchases.

The paper proceeds as follows. The next section describes the firm registration data. We then use the case of the East Hope Group and Shanghai Automobile to illustrate the importance of connected owners. The following section presents six key facts about connected owners. We then present a model of connected investors and use the model to calibrate the effect of a change in the value of becoming connected. The last section concludes.

2. Chinese Firm Registration Data

We use the firm registration records of the State Administration for Market Regulation. *All* Chinese firms are legally obligated to register with this body.[5] The data are the registration records for all firms, including those that have been closed, with the following information for each firm: registration year, exit year (if the firm has been closed), location, industry, total registered capital, and the firm's *immediate* owners and

[5] We exclude the self-employed (*Ge Ti Hu*).

the registered capital share of each owner.

The registration records identify the *immediate* owners of each firm. The immediate owners can be an individual person (identified with an encrypted personal ID), another firm (a "legal-person" owner), or other private organizations, including the publicly traded shares of publicly listed companies. A legal person can be Chinese or foreign. An important feature of the Chinese registration law is that all Chinese legal person owners also have to be registered and thus also appear in the data. We can therefore also identify the immediate owners of all Chinese legal-person owners. However, the registration data has no information on the owners of foreign legal persons or the identity of owners of the publicly traded shares of a listed company. We supplement the registration data with information on the largest 10 shareholders of listed companies.

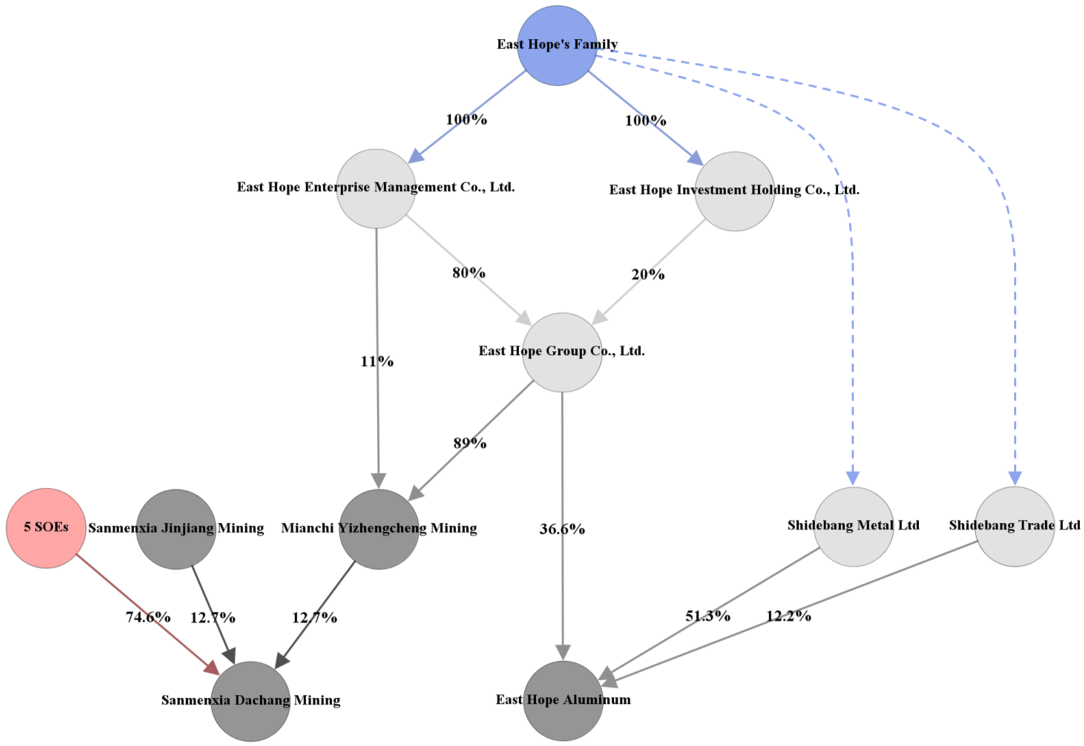
The majority of the legal-person owners are other firms, but we know from several case studies that some of them are holding shells. Take, for example, the East Hope Group, a large conglomerate with multiple companies in the heavy metals and animal food distribution and processing industries. The two dark circles at the bottom of Figure 1 represent two companies of the East Hope Group. East Hope Aluminum is one of the largest alumina producers in China. Dachang Mining is a bauxite prospecting company and a business services provider. The circles directly linked to the two firms represent their immediate owners. We distinguish different types of owners by color – light gray for suspected holding shells, dark gray for "real" private companies, red for state-owned firms, and blue for individual owners.

East Hope Aluminum is wholly owned by the family of the founder of the East Hope Group. This family, which we call "East Hope's family," owns East Hope Aluminum through five holding shells.[6] The immediate owners of East Hope Aluminum are three companies, two of which are registered in Hong Kong and one in China. We do not know for sure, but we have circumstantial evidence that the two Hong Kong holding companies are fully owned by East Hope's family.[7] As for the domestic holding shell East Hope Group Ltd., its immediate owners are two other holding shells, East Hope

[6]We define "East Hope's family" as the founder of East Hope, his wife, and his son.

[7]The two Hong Kong holding companies are Shidebang Metal Ltd. and Shidebang Trade Ltd. We identify them by an announcement made by Mingsheng Bank (`http://stock.finance.sina.com.cn/hkstock/go/CompanyNoticeDetail/code/01988/aid/488702.html`). More generally, we may fail to identify some companies of a domestic owner because of their ownership through firms registered outside China.

Figure 1: Owners of East Hope Aluminum and Dachang Mining



Note: East Hope Aluminum and Dachang Mining are the two dark gray circles at the bottom of the figure. The other circles represent the owners of East Hope Aluminum and Dachang Mining. Dark gray circles represent "real" private companies, light grey (suspected) holding shells, red for state-owned firms, and blue for individual owners.

Investment Holding Ltd. and East Hope Enterprise Management Ltd. In the registration data, these two holding shells are fully owned by East Hope's family, which we represent by the blue circle at the top of Figure 1.

Dachang Mining is a joint venture between East Hope's family and six other companies. The immediate owners of Dachang Mining are five state-owned firms (denoted by the red circle) and two private companies. One of the private companies is Mianchi Yizhengcheng Mining which is fully owned by East Hope's family through a sequence of holding shells. The other private company is Sanmenxia Jinjiang Mining which is fully owned by a large private conglomerate called the Hangzhou Jinjiang Group.

In this paper we focus on the ownership links between the *ultimate* owners. That is, we work through the ownership chain to identify each firm's *ultimate* owners, which can be state-owned firms, private individuals, foreign legal persons, or other private organizations (see Appendix D for the detailed procedure). The ultimate owner of East Hope Aluminum is East Hope's family. The ultimate owners of Dachang Mining are East Hope's family and the ultimate owners of the five state-owned firms and the Hangzhou Jinjiang Group. Hereafter we use the term "owner" to refer to a firm's *ultimate* owner.

The only economic information in the data is the firm's registered capital. China's company law stipulates that a firm's owners need to pay a fixed amount into the company's account when the firm is established. These funds, known as registration capital, represent the maximum liability of the owners in the event of a default and is viewed as a signal of the company's financial resources. Chinese law stipulates the minimum amount of registered capital for firms in certain sectors, but generally registered capital is determined by the amount of real business the firm needs to undertake. Therefore, registered capital of holding shells is minimal and a poor measure of the value of its assets, but it is a reasonable proxy of the assets and value-added of a "real" firm.[8] For each owner, we calculate the sum of registered capital of all the firms in which the owner has an equity stake weighted by her equity share in each firm, which we call the "owner's capital."[9] We also identify the controlling shareholder of each firm and assign the registered capital of the firm to the controlling shareholder (see Appendix E for details).

We have access to the registration records in 2013 and 2019.[10] The two versions provide registration information for all active and exited firms by the end of 2013 and 2019 respectively. We use the 2019 records to identify the owners (and their share of registered capital) of firms that are active in 2019, and the 2013 records to do the same for firms active in 2000 and 2010. Specifically, for a given year prior to 2013, we assume

[8]Appendix A provides more detail on what registered capital measures. We use the data from the Chinese Annual Industrial Survey with information on sales and total assets of industrial firms to check the correlation between sales, total assets and registered capital in the registration data, after matching the firms in the two data-sets.

[9]See Appendix D for details. This measure of an owner's total registered capital includes the registered capital of the holding shell companies. Ideally we should exclude holding shell companies from the firms owned by each owner, but we can not identify all the holding shells in the data. In Appendix A we check the bias due to the holding shells by comparing total registration capital of industrial firms in the Industrial Survey that belong to a given owner with total registered capital in the registration data, where the latter includes the registration capital of the intermediate owners.

[10]The registration records were kept with local offices of State Administration of Industry and Commerce (SAIC, later integrated into SAMR) until the late 2000s, when SAIC started to build a national database. To our best knowledge, the 2013 version is the earliest version of the data set with good quality.

a firm is active if it was established prior to that year and also had not exited by that year. We then infer the owners and their ownership share for each active firm from the information in 2013. The assumption is that the immediate owners of a firm are constant over time.[11] In Appendix B, we measure the error in this assumption by comparing the ownership in 2013 inferred from the 2019 data with the ownership measured directly from the 2013 data. The effects on our main findings are small and will be discussed in Section 4.

Table 1: Number of Firms and Ultimate Owners, 2000-2019

	Firms	Owners				
		All	State	Individuals	Foreign	Other
2000	4,320	5,540	257	4,635	43	606
2010	9,670	19,411	105	18,791	89	427
2019	37,546	62,887	63	62,188	140	496

Note: Table shows the number of firms and owners (in thousands) in each year in the registration data. Other includes private organizations and publicly traded shares of listed companies.

Table 1 shows the number of active firms and owners from 2000 to 2019 inferred using this procedure. The table also separately shows state owners, private individuals, foreign legal persons, and other owners.[12] Table 1 shows that almost all of the growth in the number of owners from 2000 to 2019 comes from the increase in the individual owners.

Table 2 shows the share of registration capital of each type of owner. Table 2 shows the well-known rise in the share of private individuals and the corresponding decline of the state sector. At the same time, the share of foreign legal persons has essentially not changed. This last fact suggests that the extent to which ownership is increasingly hidden behind foreign holding shells is likely to be small.

[11]The registration records show firm's most up-to-date information of immediate shareholders, which could be changed since the establishment of the firm. Although there are some text records of these changes, our understanding is that these records are incomplete, especially for earlier years. See Appendix B for more details.

[12]State owners are firms wholly and directly owned by all levels of the Chinese government (central, provincial, city- and county-level governments). Appendix C provides more details on how we identify state owners. "Other" owners are the publicly held shares of listed companies and private organizations.

Table 2: Registered Capital Share of Ultimate Owners, 2000-2019

	State	Individuals	Foreign	Other
2000	45.3%	22.1%	7.3%	25.4%
2010	35.4%	38.6%	8.9%	17.1%
2019	23.3%	63.0%	8.2%	6.5%

Note: Table shows registered capital share of state owners, individual owners, foreign legal person owners, and other private organizations.

3. Owners of East Hope and Shanghai Automobile

Earlier in the paper, we use SAIC's investment in Chery to illustrate the role of connected investors in Chery's growth. Empirically this shows up in the registration data as an ownership link between SAIC and the other owners of Chery. We now illustrate the prevalence of such links by examining SAIC and the East Hope Group in more detail.

The East Hope Group, originally a seller of animal feed in Sichuan province, started to expand into heavy metals in the early 2000s. East Hope Aluminum, the company we described earlier, was created in 2003 as a joint venture between East Hope and Huanghe Aluminum and Electricity, a state firm owned by the city of Sanmenxia (Henan Province). Huanghe initially owned 24% of East Hope Aluminum but sold its share to East Hope in 2006. The East Hope Group has created several other companies as joint ventures with the Sanmenxia local government, such as Dachang Mining in 2009.

More generally, East Hope has used joint ventures with local governments to enter into new markets. For example, East Hope expanded its animal feed business outside of its home province of Sichuan through joint ventures with local state-owned enterprises. Two examples are joint ventures with a county-level grain bureau in Henan province and with a local state-owned animal feed producer in Anhui province. East Hope has also used joint ventures to enter into new industries, as illustrated by the cases of East Hope Aluminum and Dachang Mining. Two additional examples are two joint-venture of East Hope in Chongqing and Inner Mongolia with local state-owned firms in the coal mining industry.

Table 3 uses the registration records to identify all the ultimate owners involved in

Table 3: Owners of the East Hope Group in 2019

	East Hope's Family	State Owners	Private Owners
# of Owners	1	14	9
Firms per Owner	236	443	188
East Hope's Joint Ventures	26	14	12
Capital per Owner	26.5	80.1	5.9

Note: East Hope Group defined as firms where the founder of East Hope, wife, or son own at least a 10% equity share. State owners and private owners own at least 10% of one company in the East Hope Group. Capital is registered capital owned by each owner in billion yuan.

joint ventures with the East Hope Group in 2019. We define the East Hope Group as the collection of companies where East Hope's family owns at least a 10% equity share.[13] The first column in Table 3 shows that the East Hope Group consists of 236 firms of which 210 are wholly owned by East Hope's family and 26 are joint ventures with other owners.

The second and third columns in Table 3 focus on the 14 state owners and 9 private owners operating joint ventures with East Hope's family. The 14 state owners operate on average 443 companies with total registered capital of 80 billion yuan. The registered capital of state owners linked with East Hope's family is around three times as large as that of the East Hope's family itself. These state owners are the connected owners of the East Hope Group and are significantly larger than East Hope itself.

This is not the case for the private owners that operate joint ventures with East Hope's family. For these owners, East Hope's family is itself the "connected" investor. These private owners are significantly *smaller* than the businesses owned by East Hope's family, with a total registered capital averaging 6 billion yuan which is about 22% of the registered capital of the businesses owned by East Hope's family.

Table 4 provides the same information for the SAIC Group, where the Group is defined as the collection of firms where SAIC owns at least a 10% equity share. Different from East Hope's family, SAIC is a state owner. The first column in Table 4 shows that

[13]We define East Hope's family as the founder of East Hope, his wife, his son and several overseas holding shells, which to the best of our knowledge, are owned by the three individuals.

Table 4: Owners of Shanghai Automobile in 2019

	SAIC	State Owners	Anbang, VW, GM	Private Owners
# of Owners	1	71	10	197
Firms per Owner	815	707	19.4	29.1
SAIC Joint Ventures	539	218	57	313
Capital per Owner	183	277	10.9	1.4

Note: Shanghai Automobile defined as collection of companies where SAIC owns at least a 10% equity share. State owners, Anbang, VW, GM, and other private owners own at least one company in Shanghai Automobile. Anbang, GM, and VW refer to the private owners (10 in total) of Anbang, Shanghai-GM and Shanghai-VW. Capital is registered capital owned by each owner in billion yuan.

there are 815 companies in the SAIC Group, and that SAIC's registered capital in these companies totals 183 billion Yuan. Among these 815 companies, 276 companies are wholly owned by SAIC and 539 are joint ventures with other 71 state owners and 207 private owners. The state owners directly connected to SAIC are of comparable size to SAIC as measured by the number of firms and total registered capital. Among the private owners, ten of them are linked with SAIC via 57 joint ventures with Anbang Insurance, Shanghai-VW, and Shanghai-GM (two of the ten owners are GM and VW). These ten owners are significantly larger than the remaining 197 private owners. On average the registered capital of these 197 private owners is 1.4 billion yuan, which is less than 1% of SAIC's registered capital.

Remember that East Hope's family has joint ventures with a number of private owners that are themselves even smaller than the East Hope Group. The same is true for the private owners directly connected to SAIC. In this case, the registration data indicates the 197 private owners directly connected to SAIC (excluding the owners of Anbang, VW, and GM) also operate joint ventures with 1000 private owners, who do not have joint ventures with other state owners. These owners are significantly smaller, with an average of 9.5 firms and total registered capital of 0.21 billion yuan. Because these owners are connected to SAIC through their joint ventures with the 197 private owners with direct ties with SAIC, we will say that these 1000 private owners are "indirectly connected" with a state owner (SAIC).

An example of an owner indirectly connected to SAIC is an entrepreneur (whom we call Mr. X) who established a car dealership in Yantai (Shangdong Province) in 2010. This company was a joint venture with two private owners with joint ventures with SAIC. Before 2010 Mr. X's only company was a car dealership in Hunan Province, but after creating his Yantai company, Mr. X opened 23 new car dealerships in Guizhou, Hebei, Jiangsu, Shanghai and Heilongjiang and even started a car parts manufacturer. By 2019, Mr. X owned 26 companies with a registered capital of 134 million Yuan. This evidence is only suggestive but the timing of these business developments suggests that the indirect ties Mr. X formed with private owner directly connected with SAIC in 2010 when he set up the car dealership in Yantai may have been an important factor behind the expansion of his business. In this sense, the private owners directly connected to SAIC were the connected investors that enabled Mr. X to grow beyond his original car dealership in Hunan, in the same way that SAIC was the connected investor for Chery that made it possible for the company to get their critical license.

We take three messages from these cases. First is the prevalence of equity links with multiple owners in SAIC and the East Hope Group. SAIC has joint ventures with 71 state owners and 207 private owners; the East Hope Group has joint ventures with 14 state owners and 9 private owners.

Second, there is a clear hierarchy of owners in terms of size and the number of connections they have with other owners. At the very top are state owners that are the key connected investor for many private owners. These state owners are large and undertake investments with a large number of private owners such as East Hope's family. These private owners form the next tier of owners and are themselves connected investors for other private owners such as Mr. X in the case of SAIC. Compared to their state-owned investors, these private owners are smaller and are connected to a smaller number of private owners. The next tier after that are owners such as Mr. X that are even smaller.

Third, connected investors have equity ties in only a subset of the businesses of their partners. State owners in the East Hope Group, for example, are involved in only 14 of the 236 companies in East Hope and SAIC has equity stakes in 57 of the 110 companies in Anbang, VW, and GM. SAIC also has equity in 313 of the 4,777 companies of the other private owners they have direct ties with.[14] Guided by the insights from the case studies, we will explore the whole equity network for all connected owners in the next section.

4. Connected State and Private Owners

In this section we show that the ownership links between state and private owners documented in the case studies of the East Hope Group and SAIC hold more generally across all ultimate owners in the Chinese economy. Private owners are referred to as private individuals, foreign legal persons, and other private organizations. We will use the following definitions:

- **"Directly Connected" Private Owners**: Private owners that own at least 10% of a joint venture with state owners, like East Hope's family in the case of the East Hope Group. The state owner also has to own at least 10% of the joint venture.

- **"Indirectly Connected" Private Owners**: Private owners whose only connection with a state owner is through a joint venture with another private owner, where the two parties each have at least a 10% share in the joint venture. Mr. X in the case of SAIC is an indirectly connected private owner.

- **"Distance" to the State**: The minimum number of private owners between the private owner and the state (including the private owner herself). The distance of owner i is $d_i = \min_{j \in O_i}\{d_j\} + 1$ where O_i is the set of owners directly connected to owner i. Distance = 1 for directly connected private owners such as East Hope's family. Mr. X's distance from the state is 2 because his tie with SAIC is only through the two private owners (with distance = 1) operating joint ventures with SAIC.

- **"Downward" Connections/Connected Investor**: Consider two connected owners A and B that have a joint venture together. If B is more distant from the state than A, then A has a "downward" connection with B and A is B's connected investor.

We summarize the ownership links between state and private owners as six facts.

Fact 1: Large owners are "connected"

Table 5: Connected Owners Among Top Owners, 2019

	Top 100	Top 1,000	Top 10,000	Top 100,000
State Owners	74	481	2,719	8,216
Connected	74	474	2,568	6,209
Private Owners	26	519	7,281	91,784
Directly Connected	18	334	3,566	19,217
Indirectly Connected	4	72	1,683	34,984

Note: Table shows number of top state and private owners among each group of top owners, where the size of an owner is measured by the sum of its registered capital in all the firms it owns. Connected state owners have joint ventures with a private owner. Directly connected private owners have joint ventures with a state owner. Indirectly connected private owners have a joint venture with another private owner that has a connection with a state owner.

Table 5 shows the number of connected owners among the top 100 owners ranked by total registered capital in 2019. Every single one of the 74 state owners among this top 100 have joint ventures with private owners. Among the 26 private owners in the top 100, 18 are directly connected to state owners and 4 are indirectly connected. The distinction between state and private in China becomes blurry when it comes to the largest Chinese owners. Large private owners are deeply connected to the state, and large state owners have deep ties with private owners.

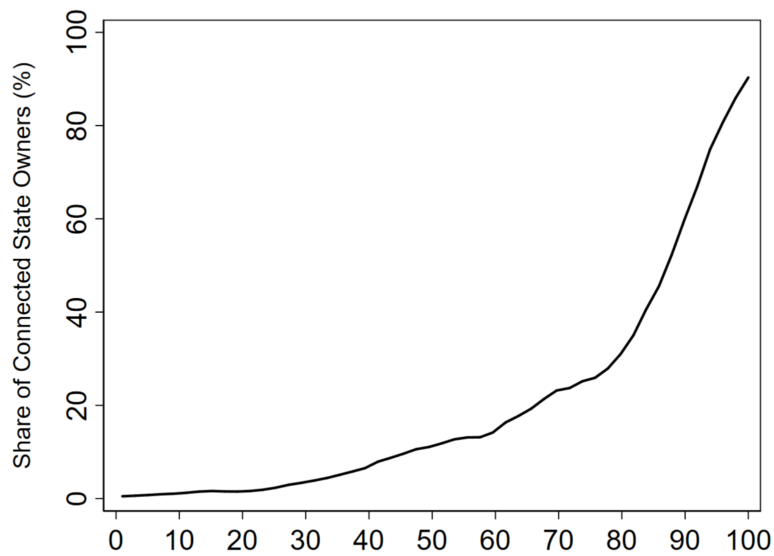
Table 5 also shows that the prevalence of these ties falls among smaller owners. Among the 92 thousand largest private owners in 2019, 19 thousand are directly connected and 35 thousand are indirectly connected to state owners. Figure 2 focuses on state owners and shows the share of state owners operating joint ventures with private owners as a function of the registered capital (in percentiles) of the state owner in 2019. Less than 20% of the bottom half of state owners have joint ventures with private owners but more than 60% of the top 10% of state owners are connected with private owners.

Fact 2: The position of a private owner in the hierarchy of connected owners is correlated with their size and number of downward connections

The left panel in Figure 3 shows the registered capital of connected private owners (rel-

[14]These private owners also own joint ventures with each other. Thus, the total number of companies is smaller than the average firms per owner multiplied by the number of private owners in Table 4.

Figure 2: State Owners with JVs with Private Owners by Size, 2019



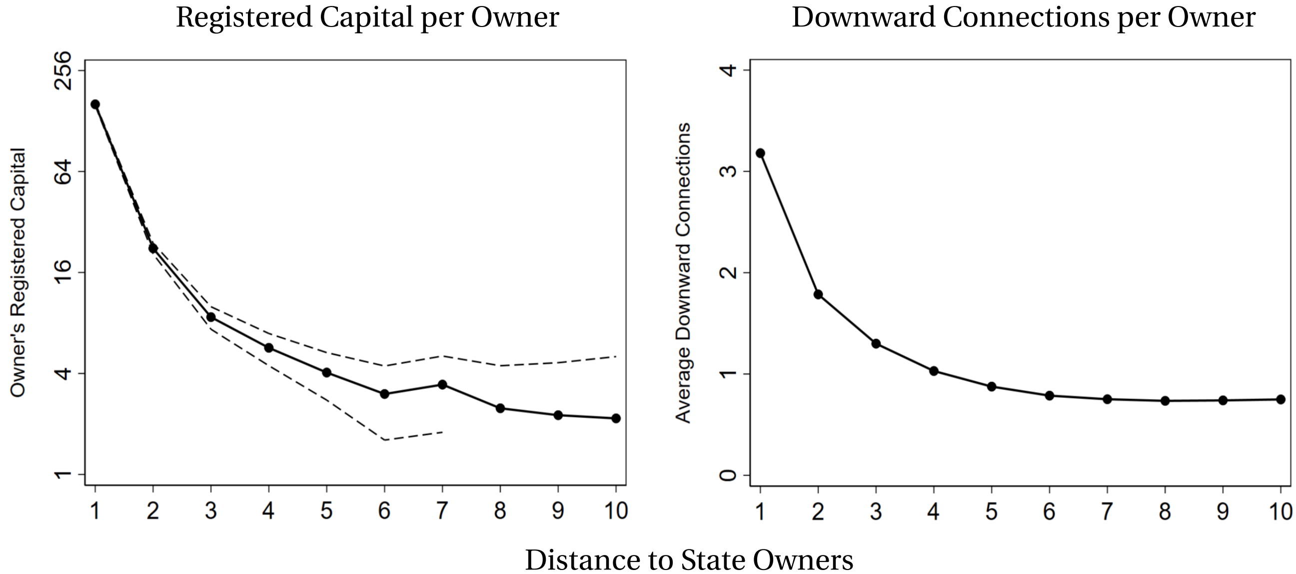
Note: Figure shows share of state owners in 2019 operating joint ventures with private owners by percentiles of total registered capital of the state owner.

ative to unconnected private owners) as a function of the owner's distance to the state. Private owners directly connected to state owners (distance = 1) are around 160 times larger (measured by total registered capital) than unconnected private owners. The gap in registered capital falls as the distance of the owner from the state gets larger. The right panel shows the average number of downward connections per owner as a function of the owner's distance to the state. We call the number of downward connections of an owner as their "span". The figure shows that private owners closer to the state have a larger span compared to owners further away from the state. Private owners directly connected with the state have joint ventures with more than 3 other private owners on average, while owners five steps away from state owners have less than one downward link on average.[15]

Fact 3: Connected investors are not controlling shareholders

[15]We also calculate the average Eigenvector centrality conditional on distance to the state and each owner's closeness centrality in the largest connected subnetwork (which is almost identical to the sector of connected state and private owners). Both measures are monotonically decreasing in distance to the state.

Figure 3: Capital and Downward Connections of Private Owners, 2019



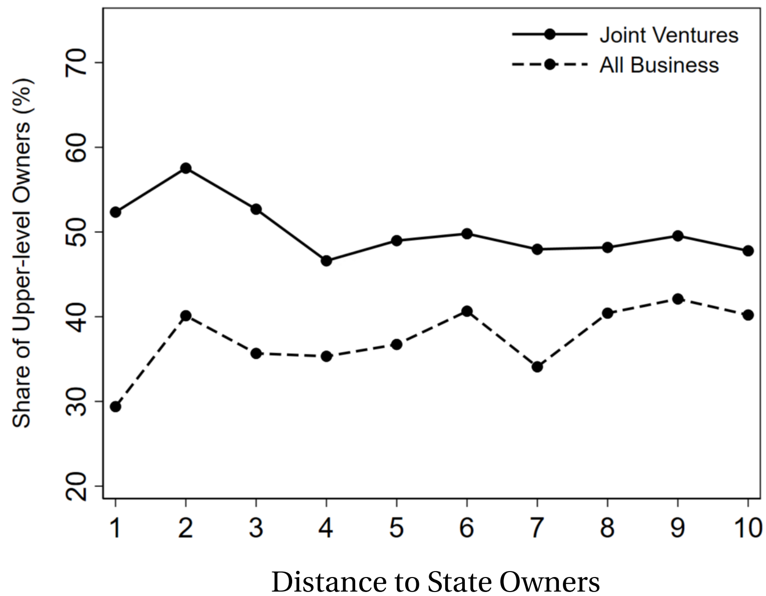
Note: Left panel shows the ratio of the average registered capital of connected private owners to the average registered capital of unconnected private owners by distance to the state (dashed lines represent 95% confidence intervals). Right panel shows the average number of downward connections of connected private owners by distance to the state.

We can also see the connected investors' share of total registered capital of the affiliated owner. The solid line in Figure 4 shows the equity share of the connected investor in the joint ventures with its downward owner. The connected investor typically owns 50% of the joint venture. The dashed line shows the investment of the connected investor as a share of the total registered capital of its partner, taking into account all the businesses of the partner. For private owners directly tied to the state (distance = 1), this share is only about 30%. For private owners more distant from the state, the share of the upward owner rarely exceeds 40%. So a private owner that takes on connected investors typically is the majority shareholder.

Fact 4: The number of private owners connected to the state has increased

The number of private owners directly connected to state owners is the product of the number of state owners that undertake investments with private owners and the number of private owners each state owner invests in. Column 1 in Table 6 shows

Figure 4: Share of Connected Investors in Private Owner's Capital, 2019



Note: The solid line shows the equity share of the connected investor in the joint venture with its downward owner. The dash line shows the ratio of the registered capital of the joint ventures owned by the connected investor to *total* registration capital of the downward owner. For two connected owners, the connected investor is closer to the state compared to the downward owner.

Table 6: Expansion of State Investment in Private Owners, 2000-2019

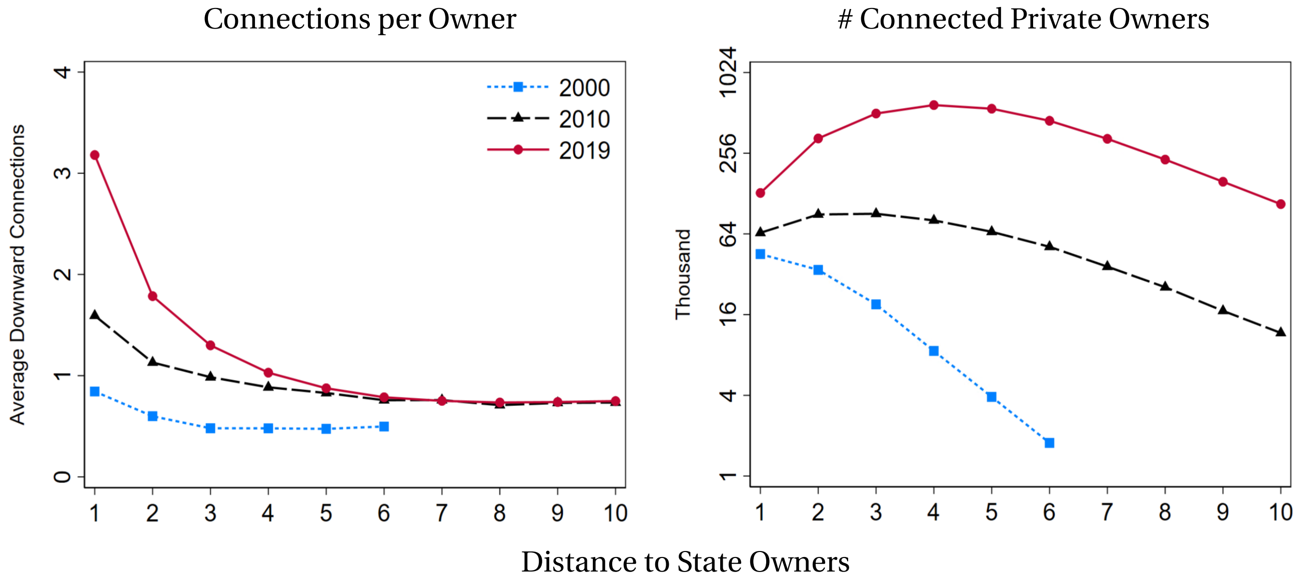
	# Connected State Owners	Connections per Owner
2000	14.4	3.5
2010	11.4	6.9
2019	12.8	14.1

Note: Column 1 shows the number of state owners (in thousands) operating joint ventures with private owners. Column 2 shows the average number of private owners a state owner has joint ventures with, conditional on investing with a private owner.

the former and column 2 the latter. The number of state owners investing in private owners fell slightly from 2000 to 2019. On the other hand, conditional on investing with a private owner, the average state owner was connected with less than 4 private owners in 2000. By 2019 the average state owner had joint ventures with 14 private owners.

Therefore the more than three-fold increase in the number of private owners directly connected to state owners, which is shown on the left panel in figure 5, is entirely due to the increase in the span of the connected state owners.

Figure 5: Increase in Number of Connected Private Owners, 2000-2019



Note: Left panel shows the average number of downward connections per private owner by distance to the state in 2000, 2010, and 2019. Right panel shows the total number of connected private owners (in thousands) by distance to the state in 2000, 2010, and 2019. We omit the observations for distance ≥ 7 in 2000 since there are less than a thousand private owners connected at these distances in 2000.

The right panel also shows that the number of private owners indirectly connected to the state also increased dramatically. This effect is particularly dramatic for owners very distant from the state. The number of owners with distance ≥ 6 increased from around 4 thousand in 2000 to more than 1.5 million by 2019. The huge increase in the number of indirectly connected owners is driven, in a proximate sense, by the significant increase in the number of private owners directly connected to the state and by the increase in the number of downward connections per private owner. The latter is shown in the left panel in Figure 5.

Fact 5: Private owners grow faster after they get connected

Here we show estimates of "event" studies to measure the effect that becoming connected has on a private owner. Specifically, we estimate the following empirical model for the panel of owners between 2000 and 2019:

$$y_{i,t} = \sum_{\tau=-14}^{\tau=15} \theta_\tau \text{Direct}_{i,t-\tau} + \sum_{\tau=-14}^{\tau=15} \beta_\tau \text{Indirect}_{i,t-\tau} + \mu_i + \lambda_t + \varepsilon_{i,t}, \tag{1}$$

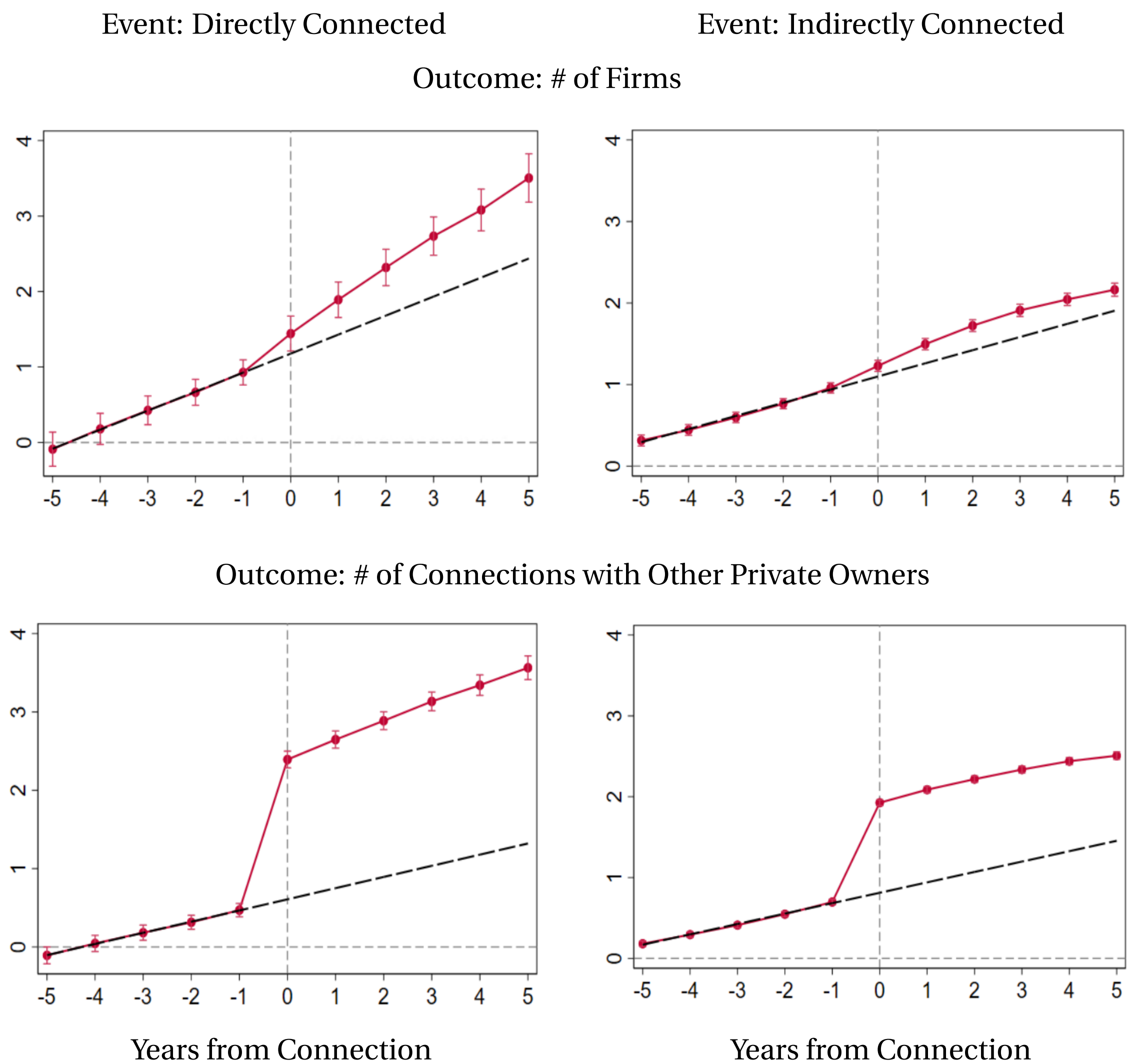
where $y_{i,t}$ is a measure of owner i's businesses; μ_i and λ_t denote owner and year fixed effects; $\text{Direct}_{i,t}$ is an indicator variable for an owner that creates a joint venture with a state owner at t; and $\text{Indirect}_{i,t}$ is an indicator variable for an owner that becomes indirectly connected (distance = 2) to the state at t. The control group are private owners who are never connected to the state during 2000-2019. The coefficients of interest are θ_τ and β_τ which summarize the values of y in the year τ before and after the "event" (becoming connected).

The first row of Figure 6 shows the results for the number of firms owned by the owner, while Figure 7 shows the results for the number of provinces the owner operates in (row 1), and number of 2-digit industries the owner operates in (row 2). We use the 2019 data to infer historical 2000-2018 data. The first column shows the estimates of θ_τ for owners that become directly connected to state owners; the second column shows the estimates of β_τ for owners that become indirectly connected to the state. The coefficient estimates and the standard errors are shown in red; the black dashed line shows the pre-trend.

These figures deliver three messages. First, there is clearly a pre-trend in all the outcome variables for owners that become connected. We interpret this as saying that owners that were growing quickly are more likely to become connected. Second, there is a clear change in the trend for all three outcomes once the owner becomes connected. Third, the magnitude of the change in the trend is larger for owners that become directly connected compared to owners that become indirectly connected.

To the extent that being connected with the state increases the size and scope of a private owner's business, as indicated by the figures, whether the connection is led by the state or private owner does not matter. However, it remains unclear whether such increases are beneficial to the private owner. For example, what state investors grab from their connected private partners can exceed the gains they generate. This hypothesis can be tested by examining if being connected to the state attracts more

Figure 6: Effect of Getting "Connected" on # Firms and # of Connections with Other Private Owners



Note: Figure plots the average # of firms (row 1) and number of connections with other private owners (row 2) of owners that become connected to state owners (column 1) or to private owners that are themselves connected to state owners (column 2) before and after the owner became connected. X-axis is number of years from the year the owner became connected (year 0). Dashed line shows the pre-connection trend.

private owners to connect with the newly connected owner. The idea is that if the state investor extends a grabbing hand instead of a helping hand, her investment in a private owner would discourage other private owners from connecting with her investee. We estimate equation (1) using the number of connections with other private owners as the independent variable. The coefficient estimates are plotted in the second row of Figure 6, which shows that the number of private owners connected to an initially unconnected private owner increases after she becomes directly connected to the state. The increase is not only quantitatively big, but highly persistent, which is inconsistent with the grabbing hand hypothesis.

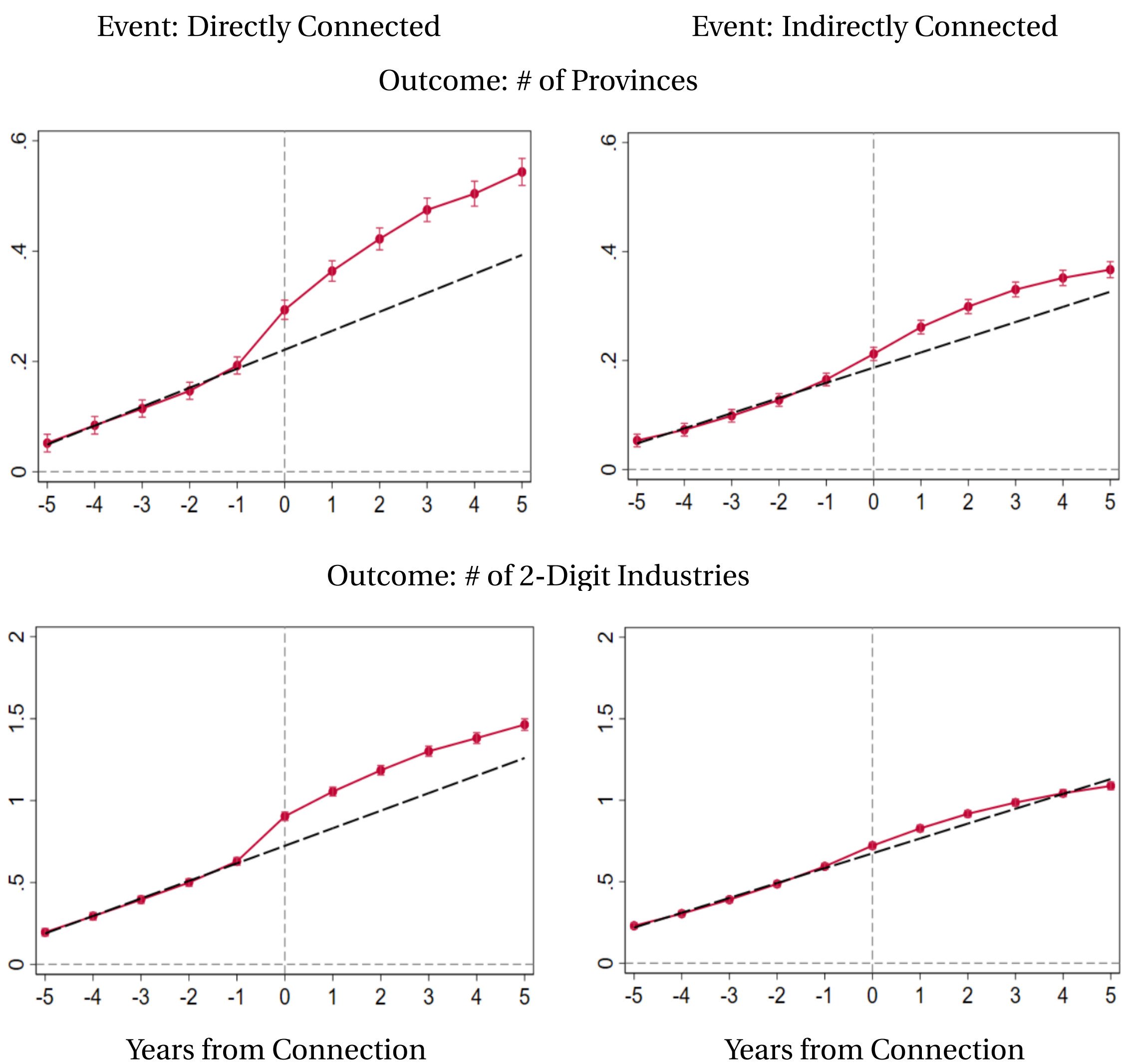
Fact 6: *Connected* private owners account for almost all the increased share of private owners

A central fact about China's growth is the decline in the size of the state sector and the expansion of the private sector. Table 7 (column 1 in the upper panel) shows that, in the registration data, the share of private owners in total registered capital increased by 22 percentage points between 2000 and 2019. Here we document that the growth of private owners is mostly due to the expansion of *state-connected* private owners. The share of private owners directly connected to state owners increased from 9.9% in 2000 to 15.3% by 2019. Meanwhile, the share of private owners indirectly connected to state owners rose from 4.2% in 2000 to 18.2% by 2019. The share of all connected private owners, including directly and indirectly connected owners, increased by 19.4 percentage points between 2000 and 2019. Note that the share of all private owners shown in the first column increased by 22 percentage points during this period. In other words, the share of unconnected private owners increased just by 2.6%.

At the same time, the share of state owners has declined, even among the state owners that have invested in private owners. The last column in Table 7 shows that the registered capital share of connected state owners fell by 16 percentage points between 2000 and 2019. Of course, this does not indicate that connected state owners have shrunk in absolute terms. It is just that the growth of the private owners enabled by state owners is much larger than the growth of the state owners that facilitated this process in the first place.

The shares of connected private owners in Table 7 should be interpreted with caution. Perhaps the most important reason is that the immediate owners of a firm can change over time. In the case of Chery and the East Hope Group, state investors played an important role in the early years of these companies, but the state owners eventually

Figure 7: Effect of Becoming "Connected" on # Provinces and Products



Note: Figure plots the average # of provinces (row 1) and 2-digit industries (row 2) of owners that become connected to state owners (column 1) or to private owners that are themselves connected to state owners (column 2) before and after the owner became connected. X-axis is number of years from the year the owner became connected (year 0). Dashed line is the pre-trend.

sold their equity share. In Appendix B we use the 2013 registration data and show that this effect may be important. Specifically, the registration capital of connected owners (as a share of total registration capital in 2013) in the 2013 data is about 6.4 percentage points higher than the registration capital share of the same owners in the 2019 data. Put differently, many connected investors in 2013 sold their equity stake by

Table 7: Share of Connected Owners in Registered Capital, 2000-2019

	All	Connected Private		Connected
	Private	Direct	Indirect	State
		Cash Flow Rights		
2000	54.7%	9.9%	4.2%	38.5%
2010	64.6%	13.9%	10.1%	33.0%
2019	76.7%	15.3%	18.2%	22.5%
		Control Rights		
2000	52.9%	8.8%	4.1%	40.0%
2010	62.7%	13.3%	9.6%	34.9%
2019	76.1%	16.3%	17.4%	23.1%

Note: Private owners are defined as individuals, foreign legal persons and other private organizations. Directly connected private owners have joint ventures with state owners. Indirectly connected private owners are linked to the state through another private owner. "All private" includes unconnected and connected private owners. Connected state owners have joint ventures with private owners. Control rights assigns all the registered capital of a firm to the controlling shareholder. See Appendix E for details.

2019. Therefore, connected investors may be more important than suggested by the contemporaneous share of connected owners.[16]

The lower panel of Table 7 shows the share of *control rights* of connected owners (see Appendix E for details). The share of control rights of connected owners is almost identical to the share of the same owners in the cash flow rights. In particular, the share of state owners in control rights is essentially the same as their share of the cash flow rights.[17]

5. A Model of Connected Investors

In this section, we build a simple model to highlight the mechanisms that are behind the stylized facts we documented in the previous section. There are two key features of the model. The first is the idea that connected investors allow a private owner to

[16]The appendix also uses alternative data to check the share of state owners (Appendix C) and whether registered capital is a reasonable proxy for sales and assets (Appendix F).

[17]We reproduce the left panel of Figure 3 and Figure 4 by *control rights*. The results are similar.

grow. The second is that a private owner who becomes connected also has the ability to provide assistance to other private owners and thus undertakes joint ventures with them. This second feature of the model is key in "explaining" the hierarchy of connected owners seen in the data.

5.1. Model Economy

Consider an economy with firms owned by two types of owners, state and private. The firm owned by owner i produces a homogeneous good according to:

$$Y_i = A_i \, L_i^{\beta},$$

where $\beta < 1$ and L_i denotes the resources employed by the firm. We represent the distortion due to "bad institutions" as an output tax $1 - \Gamma_i \in [0, 1]$. We assume $\Gamma_i = 1$ for state owners (no tax) and $\Gamma_i \leq 1$ for private owners. The owner's profit is $\Gamma_i Y_i - wL_i$, where w is the cost of L. Profit maximizing output is given by:

$$Y_i = \left(\frac{\beta \, \Gamma_i}{w} \right)^{\frac{\beta}{1-\beta}} A_i^{\frac{1}{1-\beta}}. \tag{2}$$

Output and profits are increasing in A_i and Γ_i.

We make four assumptions. First, we assume Γ_i is only a function of the private owner's distance from the state owner. So $\Gamma_i = \Gamma_d \geq \underline{\Gamma}$, where d is the owner's distance to the state, and $\Gamma_i = \underline{\Gamma}$ if the owner is not connected. We assume Γ_d is non-negative and strictly decreasing in d. We will call Γ_d the benefits from connection, although strictly speaking it comes from distortions caused by poor institutions.

Second, the distance from the state is chosen endogenously by the private owner depending on the benefits and costs of becoming connected at each distance d. Firm output and profits are increasing in firm TFP so the return from a closer connection to the state (larger Γ_d) is larger for high TFP private owners. As a consequence, higher TFP owners will sort into closer connections to the state if the cost of being connected is not related to its TFP.

Third, the cost of a connected investor providing connections to n owners is $(1 -$

$\beta)\left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\beta}}\,\lambda\, n^{\delta}$.[18] We assume $\delta > 1$ so the cost is convex in n. The connected investor charges a price for each connection she provides. Since the cost of providing connections is the same for all owners at a given distance, the price of a connection is the same and given by the standard markup over the marginal cost:

$$p_d = (1-\beta)\left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\beta}}\,\lambda\,\delta\, n_d^{\delta-1}, \tag{3}$$

where n_d is the number of connections per connected investor at distance d and referred to as the "span" of a connected investor. In equilibrium, the span is the same for all owners at each distance and equal to the ratio of the number of owners connected at distance $d+1$ to the number of owners that sort into connections at distance d. More connected owners at distance d (an increase in N_d) *lowers* the marginal cost of a connection at d, and more entry into the connected sector at distance $d+1$ (an increase in N_{d+1}) *raises* the marginal cost of a connection, both with elasticity $\delta - 1 > 0$.

Fourth, once a private owner becomes connected, she can also provide benefits to other private owners, and the price she receives from each connection is also given by (3).

Putting this together, when an owner with TFP A_i is connected at distance d, her profits are proportional to:

$$\pi\left[A_i\,|\text{choose } d\right] \propto \left(\Gamma_d\, A_i\right)^{\frac{1}{1-\beta}} + \lambda\,(\delta-1)\, n_d^{\delta} - \lambda\,\delta\, n_{d-1}^{\delta-1}. \tag{4}$$

The first term are profits from production, which is increasing in A_i and Γ_d. The second term is the net gain she gets from being the connected investor to other private owners, which is increasing in n_d (equal to $\frac{N_{d+1}}{N_d}$ in equilibrium): more "downstream" owners choosing to become connected to distance-d owners increases profits of the latter. The third term is the price she pays to her connected investor. More "upstream" owners choosing to become connected lower n_{d-1} (equal to $\frac{N_d}{N_{d-1}}$ in equilibrium) and the price of becoming connected at distance d. The last two forces imply that profits at distance d are also a function of profits at other distances when the number of "upstream" and "downstream" owners are themselves endogenous to upstream and downstream prof-

[18]The cost is scaled by $w^{\frac{-\beta}{1-\beta}}$ so that changes in the factor price has the same effect on cost as on the firm's profits. This makes the choice of the distance of a connection independent of w.

its.

5.2. Equilibrium

The equilibrium in this model is defined by the following set of conditions:

1. Conditional on being connected at distance d private owners choose output and n_d to maximize profits, taking as given the factor cost w and the price of a connection at distance d (given by (3)).

2. Conditional on choosing to become connected, private owners choose a distance d that provides them with the largest net profits (given by (4)), taking as given the price of a connection at distance $d-1$ (given by (3)).

3. Private owners choose to become connected if maximum profits from being connected is greater than their profits from remaining unconnected.

4. The span of an owner n_d is equal to the ratio of the number of owners that choose to become connected at distance $d+1$ to the number of owners that choose to become connected at distance d.

5. The price of the resource w clears aggregate demand and supply of L.

Assuming the equilibrium exists, we summarize its characteristics as three propositions. The proofs for all the propositions are in Appendix G.

Proposition 1. *There are private owners connected with each distance $d \in [1, \bar{d}]$, while no owners are connected with distance $d > \bar{d}$. There exists a cutoff TFP $A_{\bar{d}}$ defined as*

$$A_{\bar{d}} \equiv \left(\frac{\lambda \, \delta \, n_{\bar{d}-1}^{\delta-1}}{\Gamma_{\bar{d}}^{\frac{1}{1-\beta}} - \underline{\Gamma}^{\frac{1}{1-\beta}}} \right)^{1-\beta},$$

where all owners with $A_i \geq A_{\bar{d}}$ are connected and all owners with $A_i < A_{\bar{d}}$ are unconnected.

It is easy to show that profits from being connected are larger than the profits from being unconnected for owners with TFP $A_i > A_{\bar{d}}$, and vice versa. This proposition thus captures the empirical observation (Facts 1 and 2) that connected owners are larger than unconnected ones.

Proposition 2. *For all distances $d \leq \bar{d}$, there exists a cutoff TFP A_d defined as*

$$A_d \equiv \left[\frac{\lambda \left(\delta \, n_{d-1}^{\delta-1} - (\delta - 1) \, n_d^{\delta} \right) - \lambda \left(\delta \, n_d^{\delta-1} - (\delta - 1) \, n_{d+1}^{\delta} \right)}{\Gamma_d^{\frac{1}{1-\beta}} - \Gamma_{d+1}^{\frac{1}{1-\beta}}} \right]^{1-\beta},$$

where owners with $A_i \in [A_d, A_{d-1})$ choose to connect at distance d and where $\{A_d\}_{d=1}^{\bar{d}}$ is a strictly decreasing sequence in d.

The positive sorting of private owners comes from the assumption that connections closer to the state deliver greater benefits and that this gain is increasing in owner's TFP. This is consistent with our empirical finding that owners closer to the state are larger (Fact 2).

Proposition 3. *The span of an owner $n_d = \frac{N_{d+1}}{N_d}$ is decreasing in distance from the state for $d \in [0, \bar{d} - 1]$.*

This proposition comes from the assumption that Γ_d is decreasing in d and the cost of providing downward connections is convex in the span. Connected investors more distant from the state provide less benefits to its downstream owners and thus choose to provide fewer connections. This is consistent with the fact that the span of an owner decreases with distance from the state (Fact 2).

5.3. Equilibrium Effects of Changes in Connection Benefits

We now investigate the mechanisms through which changes in connection benefits Γ_d affect the hierarchy of connected private owners. Characterizing the equilibrium response to a change in Γ_d is difficult because the TFP cutoffs and the span at *all* distances potentially change in equilibrium. We can, however, characterize the equilibrium response to a change in Γ_d in the case where the TFP distribution is "sufficiently" compact, by which we mean that the dispersion of TFP is small enough such that changes in Γ_d do not change the TFP cutoffs and only changes the average number of connections.[19] When this is the case, we can state the following propositions about the equilibrium effects of a change in Γ_d. In the next section of the paper, we will empirically estimate the equilibrium effects of changes in connection benefits in the data without imposing this distributional assumption.

Proposition 4. *If the distribution of private owners' TFP is "sufficiently compact," the elasticity of the number of owners connected at distance $j \le \bar{d}$, N_j, with respect to Γ_d is:*

$$\frac{\Delta \ln N_j}{\Delta \ln \Gamma_d} = \sum_{i=0}^{\min\{j-1, d-1\}} \frac{\Delta \ln n_i}{\Delta \ln \Gamma_d} \quad \text{for any } j, \quad d \in [1, \bar{d}] \tag{5}$$

where

$$\frac{\Delta \ln n_i}{\Delta \ln \Gamma_d} = \begin{cases} \rho \, \Gamma_d^{\frac{1}{1-\beta}} \, n_i^{-(\delta-1)} \prod_{k=i+1}^{d-1} n_k & \text{if } 0 \le i < d-1 \\ \rho \, \Gamma_d^{\frac{1}{1-\beta}} \, n_i^{-(\delta-1)} & \text{if } i = d-1 \\ 0 & \text{if } d \le i \le \bar{d}-1 \end{cases} \tag{6}$$

and $\rho \equiv (\min\{A_i\})^{\frac{1}{1-\beta}} \, [(1-\beta) \, \lambda \, \delta \, (\delta - 1)]^{-1}$.

Proposition 4 states that the number of connected owners at each distance j with respect to Γ_d is the *sum* of the elasticity of the span of all the owners closer to the state ($i \le j-1$), where the elasticity of the span is positive for all upstream owners ($i \le d-1$) and zero otherwise. The elasticity of the span at all distances is non-negative so the number of connected owners at every distance $j \le \bar{d}$ increases when Γ_d rises. This proposition thus suggests the large increase in the number of connected owners (Fact 4) can be "explained" by an increase in the benefits of being connected.

The number of connected owners at a given distance j will increase when Γ_d increases, even at distances $j \neq d$ that do not directly benefit from the increase in Γ_d. The number of upstream owners rises because more owners want to be connected at distance d, and the price the upstream owner receives from providing a connection increases. In (4), the second term increases for the upstream owners. In equilibrium,

[19]See details and the proof in Appendix G. The distribution of TFP $F_s(x)$ is "sufficiently compact" when $\lim_{s\to\infty} F_s(x) = 1$ for $x > \min\{A_i\}$, $\lim_{s\to\infty} F_s(x) = 0$ for $x < \min\{A_i\}$, $\min\{A_i\} > 0$, and $F'' < 0$.

the number of connected owners at a given distance rises until the marginal owner is indifferent between being connected at that distance and another distance.

The number of downstream owners ($j > d-1$) also increases, albeit for a different reason. This effect is driven by the increase in the number of upstream owners providing connections, which lowers the price of a connection. In (4), this effect is seen in the third term, which falls at a given distance. Again, in equilibrium, the number of connected downstream owners rises until the marginal owner is once again indifferent between being connected at that distance and another distance.

Proposition 5. *If the distribution of private owners' TFP is "sufficiently compact," the share of output of connected private owners increases in Γ_d for any d.*

Aggregate output $\mathcal{Y}$ (including output by state owners) is given by:

$$\mathcal{Y} = \left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\beta}} \left(\underbrace{\sum_{d=1}^{\bar{d}} N_d \left(\Gamma_d^{\beta} \tilde{A}_d\right)^{\frac{1}{1-\beta}}}_{\text{Connected Owners}} + \underbrace{\left(N - \sum_{d=1}^{\bar{d}} N_d\right) \left(\underline{\Gamma}^{\beta} \tilde{A}_u\right)^{\frac{1}{1-\beta}}}_{\text{Unconnected Owners}} + \underbrace{N_0 \tilde{A}_0^{\frac{1}{1-\beta}}}_{\text{State Owners}} \right), \tag{7}$$

where $\tilde{A}_d$, $\tilde{A}_u$, and $\tilde{A}_0$ denote the TFP of the representative owner connected at distance d, unconnected owners, and state owners, respectively; N and N_0 are the total number of private and state owners, respectively.[20] w is determined by $\mathcal{L}$, the aggregate supply of resources (see Appendix G). The first term in (7) is the contribution from connected private owners to aggregate output; the second term is the contribution of unconnected private owners; and the third term is the contribution of state owners (distance 0).

The expression for aggregate output in (7) suggests two channels through which an increase in connection benefits Γ_d raises aggregate output under the assumption of "sufficiently " compact distribution of owner's TFP. First, the increase in Γ_d increases output for the owners connected at distance d. We call this the *direct* effect of connection benefits on output. Second, from (5), an increase in Γ_d increases the number of connected owners. This also increases aggregate output because more owners benefit

[20]Representative TFP of owners connected at distance d is defined as $\tilde{A}_d \equiv \left(\sum_{i \in d} A_i^{\frac{1}{1-\beta}} \Big/ N_d\right)^{1-\beta}$. Representative TFP of unconnected and state owners are defined similarly.

from being connected, even if Γ_j with $j \neq d$ has not changed at the specific distance they are connected. We will call this the "indirect" effect of connection benefits and provide the details in Appendix G. These two mechanisms also increase the share of output from the connected sector relative to the unconnected sector, so this can explain the increase in the share of connected private owners (Fact 6) in the data.

The effect of Γ_d on aggregate output is, however, ambiguous. It is possible that some of the benefit that connected owners get improves or worsens the efficiency of resource allocation. Alternatively, one may model "bad institutions" by assuming Γ_d as part of firm TFP. Since the profits of an owner at distance d remain the same, all the above theoretical results are intact. Yet, a higher Γ_d that increases productivity, rather than reduces distortion, will unambiguously increase aggregate output. See Appendix H for details. We do not have much evidence to distinguish distortion and productivity in Γ_d. So, we leave this question for future work. Nevertheless, the next section will provide tentative estimates of changes in Γ_d in different scenarios of Γ_d. The aggregate implications, especially the effects of the estimated changes in Γ_d on the private-sector output growth, turn out to be quantitatively similar, regardless of whether Γ_d is distortion or TFP.

6. Equilibrium Effect of Expansion of Connected Investors

In this section, we filter the data on the expansion of connected investors through the lens of the model laid out in the previous section. We first describe how we estimate the key parameters of the model, primarily the benefits of connection Γ. We then measure the effect of the estimated change in Γ on aggregate output and the share of the private sector.

6.1. Model Calibration

We make two extensions to estimate the model. First, we add capital, K, as another input factor and reinterpret L as labor. The output elasticity of capital and labor are denoted by α and β, respectively. Second, we assume an owner can create multiple firms, where the marginal cost of owning m firms is $m^{\frac{1}{\theta-1}}$. With these two extensions, total profit-maximizing output of the owner $\bar{Y}_i$ is:

$$\bar{Y}_i = (1-\alpha-\beta)^{\theta-1} \left[\left(\frac{\alpha}{r}\right)^{\alpha} \left(\frac{\beta}{w}\right)^{\beta} \right]^{\frac{\theta}{1-\alpha-\beta}} \Gamma_i^{\frac{\theta}{1-\alpha-\beta}-1} A_i^{\frac{\theta}{1-\alpha-\beta}}. \tag{8}$$

These extensions change the elasticity of output with respect to Γ_i from $\frac{1}{1-\beta} - 1$ in (2) to $\frac{\theta}{1-\alpha-\beta} - 1$.

The model is summarized by six parameters $\{\alpha, \beta, \theta, \lambda, \delta, \underline{\Gamma}\}$ and two forcing variables $\{A_i, \Gamma_i\}$. We assume the output elasticity parameters are $\alpha = 0.4$ and $\beta = 0.4$ and set $\underline{\Gamma}$ to 0.41 from the relative size of firms owned by unconnected private owners compared with firms owned by state owners. We fit a Pareto distribution to the average firm TFP of owners for each distance to the state.[21].

We then estimate $\frac{\theta}{1-\alpha-\beta}$, $\lambda^{\frac{1-\alpha-\beta}{\theta}}$, and δ by targeting two sets of moments: (1) the average registered capital of firms owned by an owner at d (relative to unconnected owners); (2) the average registered capital of an owner at d (also relative to unconnected owners). The resulting estimates are $\frac{\theta}{1-\alpha-\beta} = 10.3$ for the elasticity of the owner's size with respect to TFP and $\delta = 2.5$ for the concavity of the connection cost with respect to the number of connections. The latter implies that the average cost of a downward connection increases 11-fold when the number of connections increases from 1 to 5. Without the hierarchy of connections, special deals would only benefit a very small group of directly connected owners.[22]

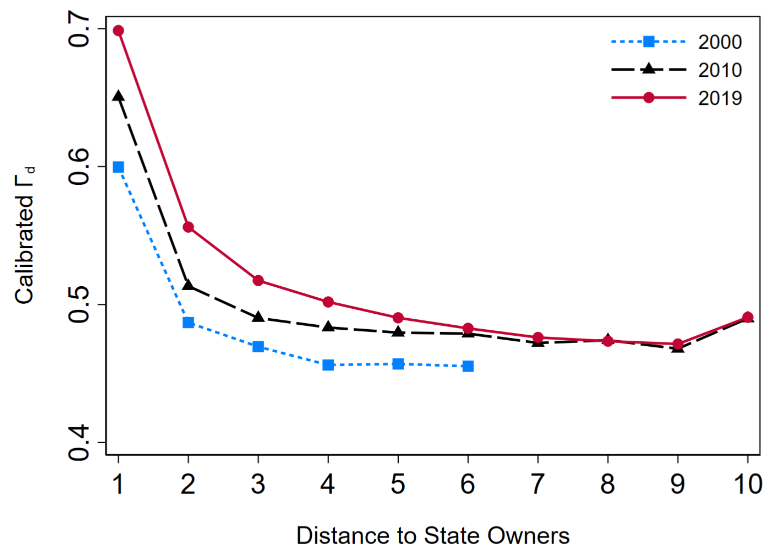
The last step is to estimate Γ_d by exactly matching the number of owners at each distance d in each year (conditional on the estimates of $\frac{\theta}{1-\alpha-\beta}$, $\lambda^{\frac{1-\alpha-\beta}{\theta}}$, and δ). The resulting estimates of the connection benefits Γ_d are shown in Figure 8. In the cross-section, the gain from becoming connected Γ_d is generally decreasing in distance to the state. This is what the model infers to match the fact that the number of connections of an owner falls with distance to the state. In the time series, connection benefits increase from 2000 to 2019 to match the increase in the span of connected owners over this time

[21] Following Hsieh and Klenow (2009), we measure firm TFP as the residual of firm value-added after controlling for the average product of labor and capital of the firm. Since this can only be done for firms in the industrial survey (remember we only observe registered capital in the registration records), we estimate firm TFP from the 2013 industrial survey. This yields a Pareto shape parameter of 255, with the scale parameter set to 1. See Appendix I for more details. While the calibrated TFP dispersion is small, we can add measurement errors or temporary TFP shocks that are irrelevant for connection choices to fit the dispersion in the data, without affecting the estimation and welfare implications below.

[22] We also estimate $\lambda^{\frac{1-\alpha-\beta}{\theta}} = 0.47$ with a standard error of 0.01. The standard errors of the estimates of $\frac{\theta}{1-\alpha-\beta}$ and δ are 0.90 and 0.08, respectively. See Appendix I for details.

period.

Figure 8: Estimated Benefit of Connection by Distance to State, 2010-2019



Note: Figure plots the connection benefit Γ_d in 2019, 2010, and 2000. Γ_d for $d \in [7, 10]$ in 2000 are omitted as less than 1,000 private owners are connected at these distances in 2000.

What exactly are the institutional reasons behind the increased benefits from connecting to state owners we infer from the data? We do not know for sure, but the growing financial resources controlled by the state is one possibility. The ratio of bank deposits to GDP increased from 1.2 in 2000 to 2 in 2019. Total credit to the real economy, measured by "social financing," grew even faster. More importantly, the financial sector is predominated by state-owned banks and more than two thirds of the credit is allocated to state-owned firms, which account for less than one third of the output. What our results suggest is that the ultimate recipient of some of the credit nominally allocated to state-owned firms may have been private owners connected to the state owners. This is consistent with the fact that private owners who are more distant to state owners are associated with higher output per unit of capital (see Figure A.7 in the appendix).

Another possibility is the large expansion of local government budgets from the emergence of local government financial vehicles and increasingly lucrative land sales. Local government debt, which was negligible in the early 2000s, rose to 40 trillion Yuan in 2015 or 58% of GDP in that year (Bai et al. (2016)). Land sales revenue, which all goes to local government budget and was less than 1% of GDP in 2000, increased to 7% of GDP in 2019. Our results suggest that much of the additional resources controlled by local governments may not have been kept in the state sector but instead were allocated to private firms connected to the state sector.

6.2. Effects of Connected Investors

From (7) there are two sources of growth. The first is the growth in the number of private owners and the shift in the TFP distribution of these owners, holding the distribution of Γ_d fixed. The second is the rightward shift in the distribution of Γ_d, which as discussed in the previous section, has a direct effect on aggregate output and also an indirect effect through increasing the number of connected owners.

We now measure the importance of the second source: to what extent the rightward shift in Γ_d shown in Figure 8 "explains" the aggregate growth in the overall economy and among private owners observed in the data, while controlling for the first source. Specifically, we change Γ_d while holding constant the number of private owners and the TFP distribution of these owners.[23] Finally, we assume fixed labor supply and small open economy so that labor cost w is endogenous but capital cost r is exogenous.

Table 8 shows the results of this exercise from 2000-2019. The top panel shows the growth rate of aggregate output (row 1), private output (row 2), and the change of output share of the private sector (row 3) in the data. The bottom panel shows the share of each of the three outcomes that can be attributed to the increase in connection benefits shown in Figure 8. The first row of the bottom panel shows that the change in the connection benefits explains 9.0% of the aggregate growth in output from 2000 to 2019, with a larger share (12.1%) in the 2010-2019 period. The aggregate output growth comes entirely from resource reallocation, implying that higher Γ_d alleviates misallocation in the estimated model. In the data, aggregate output of the private sector grew by an average 10.9% per year from 2000 to 2019. The bottom panel shows that the improvement in connection benefits accounts for 20 to 27 percent of this growth. Fi-

[23] We use the standard approach of chaining. For example, we compute growth between 2000 and 2010 allowing Γ_d to change from 2000 to 2010 values but holding the other forcing variables at their 2000 values. Then we compute growth between 2000 and 2010 by holding the other forcing variables at their 2010 values and allowing Γ_d to change from 2010 to 2000 values. We take the average of these two estimates of growth from changing Γ_d. We do the same for 2010-2019 and cumulate the growth to arrive at an estimate from 2000-2019.

nally, the third row shows that connection benefits account for almost all of the growth of the private sector's share of aggregate output over this period.

Table 8: Contribution of Δ Connection Benefits to Aggregate Growth and Δ Private Sector Share, 2000-2019

	2000-2019	2000-2010	2010-2019
Data			
Aggregate Growth Rate (per year)	9.0%	10.6%	7.3%
Private Growth Rate (per year)	10.9%	12.4%	9.3%
Δ Private Sector Output Share	22.0%	9.9%	12.1%
Contribution of Δ Connection Benefits to:			
Aggregate Growth	9.0%	7.0%	12.1%
Private Growth	23.1%	20.3%	27.2%
Δ Private Sector Output Share	95.0%	106.1%	86.0%

Note: The top panel shows the average annual growth rate of aggregate output, private sector output, and percentage points change of the private sector output share in the data. These numbers are calculated from the share of private owners in Table 2 and real GDP growth from 2000-2010 and 2010-2019 from the China Statistical Yearbook (2019) and China's Statistical Communique on the 2019 National Economic and Social Development. The bottom panel shows the contribution of changes in Γ_d we infer between 2000 and 2019 on aggregate growth and private sector share. For example, changes in Γ_d raise aggregate output by 0.8% annually, thus the contribution of $\Delta\Gamma_d$ is 9.0% (0.8% / 9.0%).

Table 9 further decomposes the contribution of the change in benefits of connections. The top panel decomposes the contribution of the change in Γ into the change that is only due to the increase in Γ for owners directly connected to the state (distance 1) vs the contribution from the increase in Γ for owners indirectly connected to the state (for distances greater than 1).[24] The top panel shows that the change in Γ for directly connected owners "accounts" for about two-thirds of the total contribution of the increase in Γ. This may seem surprising given the evidence that the largest increase in connected owners are for those that are indirectly connected. However, our calculation takes into account the effect of Γ on the number of connected owners,

[24] Again, the results shown in Table 9 are based on chaining. For example, for the results in row 1, we first change Γ_1 in 2000 to its value in 2010, while holding constant Γ_d for all other distances constant at its 2000 level. Then we change Γ_1 in 2010 back to its value in 2000, while holding constant Γ_d at all other distances constant at its 2010 values. We then take the average of these two numbers.

and we have seen that a change in Γ can have large effects on the number of connected owners at other distances.

Table 9: Decomposition of Contribution of Connection Benefits, 2000-2019

	2000-2019	2000-2010	2010-2019
By Distance of Connection Benefit:			
Δ Direct Connection Benefit	62%	74%	49%
Δ Indirect Connection Benefit	43%	31%	56%
By Directly vs. Indirectly Connected:			
Directly Connected Owners	53%	64%	42%
Indirectly Connected Owners	47%	36%	58%
By Direct vs. Indirect Effects:			
Direct Effect	36%	39%	33%
Indirect Effect	64%	61%	67%

Note: The first panel decomposes the effect of changes of Γ_d on private sector output growth to changes of Γ_d, $d = 1$ and changes of Γ_d, $d \geq 2$. The second panel decomposes the growth effect to directly and indirectly connected private owners. The third panel decomposes into the direct effect of the change in connection benefits vs. the indirect effect from the increase in the number of connected owners.

The middle panel in Table 9 decomposes the growth of connected owners into the growth of owners that are directly connected to the state vs. owners that are only indirectly connected to the state. This decomposition is, of course, an accounting decomposition and should not be interpreted as causal. Part of the growth of directly connected owners is due to the increase in benefits of connections for the indirectly connected owners, and some of growth of the indirectly connected owners comes from the increase in the benefits of connections for the directly connected owners. About 53% of the contribution of the connected sector to growth from 2000 to 2019 was "due" to the growth of directly connected owners. The relative contribution of the directly connected owners has been smaller in the last ten years, at 42%.

Finally, the bottom panel in Table 9 decomposes output growth of the private owners from the increase in Γ into the direct effect of Γ on output vs. its "indirect" effect by inducing more owners to become connected. The indirect effect accounts for almost two-thirds of the entire effect. In the absence of the network the change in the benefits from being connected would be much smaller.

We have so far assumed that connections reduces an output friction so the rise in connections shows up as reduced misallocation as in Song et al. (2011) and Hsieh and Song (2015). However, connections can also show up as TFP growth in the private sector (Brandt et al., 2012) or as a reduction in internal trade barriers and migration costs as in Tombe and Zhu (2019).

To allow for these possibilities, we entertain two alternative interpretations of Γ_d.[25] First, instead of an output tax, we model Γ_d as isomorphic to TFP. In this case, the allocation of resources across owners stays the same but their output will be different. Specifically, total profit-maximizing output of owner i at distance d is now

$$\bar{Y}_i = (1-\alpha-\beta)^{\theta-1} \left[\left(\frac{\alpha}{r}\right)^{\alpha} \left(\frac{\beta}{w}\right)^{\beta} \right]^{\frac{\theta}{1-\alpha-\beta}} \Gamma_d^{\frac{\theta}{1-\alpha-\beta}} A_i^{\frac{\theta}{1-\alpha-\beta}}.$$

The elasticity of output with respect to Γ_d is now $\frac{\theta}{1-\alpha-\beta}$ instead of $\frac{\theta}{1-\alpha-\beta} - 1$ (8) in the baseline model. After we re-estimate the model under this new formulation, we obtain the contribution of the change in Γ_d on aggregate output and the share of private owners. Table A.10 in the Appendix shows that when Γ_d is interpreted as equivalent to A_i, the effects of the estimated changes in Γ_d on aggregate TFP are smaller, while the effects on the private-sector output growth and share are quantitatively similar.

The second alternative is to interpret Γ_d as capital wedge so the cost of capital is now r/Γ_d instead of r. Profit maximizing output is now

$$\bar{Y}_i = (1-\alpha-\beta)^{\theta-1} \left[\left(\frac{\alpha}{r}\right)^{\alpha} \left(\frac{\beta}{w}\right)^{\beta} \right]^{\frac{\theta}{1-\alpha-\beta}} \Gamma_d^{\frac{\alpha\theta}{1-\alpha-\beta}} A_i^{\frac{\theta}{1-\alpha-\beta}}.$$

The elasticity of output with respect to Γ_d is now $\frac{\alpha\theta}{1-\alpha-\beta}$. We also assume a fixed supply of capital and endogenize r to equate aggregate demand for capital with the supply. After we re-estimate the parameters and forcing variables with this new formulation, Table A.11 in the Appendix shows the aggregate gains from the change in Γ_d. Here, the results are overall similar to the gains when Γ_d is interpreted as equivalent to firm TFP.

[25] See Appendix J for more details of these two models.

7. Conclusion

In this paper, we use detailed administrative data from the Chinese registration records to document the importance of "connected" investors. We report two key findings. First, in recent years, there has been a hierarchy of private owners with state-owned firms at the very top of the chain. These state owners hold equity in companies owned by a large number of private owners. In 2019 there were a hundred thousand private owners operating joint ventures with state owners. These private owners are the largest in China and also hold equity in companies owned by other, smaller, private owners, who hold equity in other companies owned by even smaller private owners, and so on. In 2019, this hierarchical chain extended to owners that were more than ten steps away from the state owners.

Our second fact is that the magnitude of this hierarchy of connected private owners is a recent phenomenon. In 2019, connected private owners accounted for 33.5% of all registered capital in China. In 2000, connected private owners only accounted for at most 14% of registered capital. The increase in the share of connected private owners between 2000 and 2019 period explains almost all the increase in the share of all private owners in China over the same period. This rise of the connected private sector is driven by two related facts. First, state-owned firms increasingly invest in more private owners. Second, the typical connected private owner itself also invests in more private owners.

We leave several important questions for future research. First, it is important to understand what exactly the benefits received by connected owners are, and the institutional forces behind these benefits. We have suggested that state control over the financial system, as well as the increased resources available to local governments may be part of the story. But this is obviously speculative, and much more work is needed to understand the phenomena we document in this paper.

Second, our estimate of the change in the share of connected owners over time assumes that a given firm does not undergo ownership changes. This assumption is obviously not true, and our estimates using the contemporary registration data in 2013 suggest that our estimates are likely to understate the the number and share of connected private owners in the past. It may be possible to unearth contemporaneous registration data in earlier years. When such data is available, we can update our estimates correspondingly.

Third, we assume that connected owners are those that are tied to state-owned firms and that private owners become "connected" only when they become linked to state-owned firms, directly or indirectly, through equity ties. Of course some of the private individuals themselves could be the connected investors, and perhaps even more connected than the state-owned firm that they are linked with. We think this is likely to be true in many cases. It also raises the question of why connected individuals would choose to "share" their equity with official state owners when the latter are less politically connected. One explanation is that the equity ties with state-owned firms could give the connected owners cover to provide favors to these firms. We do not currently have a way to identify such individuals but this is also something that future research can address.

Finally, we use a very simple model to show that the expansion in the "span" of connected investors may have increased growth by 2.5% a year between 2000 and 2019. This number is obviously tentative, and an important agenda for future research is to examine the effect of these networks on aggregate productivity with richer models.

References

Acemoglu, Daron and Pablo D. Azar, "Endogenous Production Networks," *Econometrica*, 2020, *88* (1), 33–82.

Allen, Franklin, Junhui Cai, Xian Gu, Jun "QJ" Qian, Linda H. Zhao, and Wu Zhu, "Ownership Network and Firm Growth: What Do Five Million Companies Tell About Chinese Economy," 2019.

Aminadav, Gur and Elias Papaioannou, "Corporate Control around the World," *Journal of Finance*, 2020, *75* (3), 1191–1246.

Bai, Chong-En, Chang-Tai Hsieh, and Zheng Michael Song, "The Long Shadow of China's Fiscal Expansion," *Brookings Papers on Economic Activity*, 2016.

— , — , and — , "Special Deals with Chinese Characteristics," *NBER Macroeconomics Annual*, 2019.

Baquee, David and Emmanuel Farhi, "Productivity and Misallocation in General Equilibrium," *Quarterly Journal of Economics*, 2020, *135* (1), 105–163.

Barberis, Nicholas, Maxim Boycko, Andrei Shleifer, and Natalia Tsukanova, "How does privatization work? Evidence from the Russian shops," *Journal of Political Economy*, 1996, *104*, 764–790.

Brandt, Loren, Gueorgui Kambourov, Kjetil Storesletten et al., "Barriers to Entry and Regional Economic Growth in China," 2020.

— , Johannes Van Biesebroeck, and Yifan Zhang, "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing," *Journal of Development Economics*, 2012, *97* (2), 339–351.

— , Ruochen Dai, Gueorgui Kambourov, Kjetil Storesletten, and Xiaobo Zhang, "Serial Entrepreneurship in China," 2019.

Chen, Ting and James Kai sing Kung, "Busting the "Princelings": The campaign against corruption in China's primary land market," *The Quarterly Journal of Economics*, 2019, *134* (1), 185–226.

Dai, Ruochen, Dilip Mookherjee, Kaivan Munshi, and Xiaobo Zhang, "The Community Origins of Private Enterprise in China," 2019.

Dunne, Michael, *American Wheels, Chinese Roads: The Story of General Motors in China*, John P. Wiley, 2011.

Frydman, Roman, Cheryl Gray, Mark Hessel, and Andrzej Rapaczynaki, “Private ownership and corporate performance: Evidence from the transition economies,” *Quarterly Journal of Economics*, 1999, *114*, 1153–1192.

Hsieh, Chang-Tai and Peter J Klenow, “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

— and Zheng Michael Song, “Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China,” *Brookings Papers on Economic Activity*, 2015.

Huang, Yasheng, *Capitalism with Chinese Characteristics*, Cambridge University Press, 2008.

Lardy, Nicholas, *Markets over Mao: The Rise of Private Business in China*, Peterson Institute for International Economics, 2014.

Liu, Ernest, “Industrial policies in production networks,” *Quarterly Journal of Economics*, 2019, *134* (4), 1883–1948.

Porta, Rafael La and Florencio Lopez de Silanes, “The benefits of privatization: Evidence from Mexico,” *Quarterly Journal of Economics*, 1999, *114*, 1193–1242.

— , — , and Andrei Shleifer, “Corporate Ownership around the World,” *Journal of Finance*, 1999, *54* (2), 471–517.

Shi, Xiangyu, Tianyang Xi, Xiaobo Zhang, and Yifan Zhang, ““Moving Umbrella”: Identifying Collusion Through Bureaucratic Transfers And Investment Flows,” 2020.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, “Growing like china,” *American economic review*, 2011, *101* (1), 196–233.

Taschereau-Dumouchel, Mathieu, “Cascades and Fluctuations in An Economy with An Endogenous Production Network,” 2020.

Tombe, Trevor and Xiaodong Zhu, “Trade, Migration and Productivity: A Auantitative Analysis of China,” *American Economic Review*, 2019, *109* (5), 1843–1872.

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 account. 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, suppose some one starts a company and states in the registration document 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.1 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 (in columns 2 and 4), the estimated coefficient is almost 1.

Shareholders are allowed to change (mostly increase) the registered capital of their company. To do this, they need 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 registration records is the firm's most up-to-date registered capital by the end of 2013. We do not 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.1 show the results when we merge the 2013 registration records with earlier industrial surveys in 2007,

Table A.1: Relationship between registered capital, total assets, and sales for industrial firms

	Dependent Variable: Registered Capital in 2013 Registration Records			
	2013 Survey			
	Assets (All Firms)	Assets (Drop Q1)	Sales (All Firms)	Sales (Drop Q1)
Total Assets	0.926 (0.000)	0.966 (0.000)	–	–
Sales	–	–	0.648 (0.000)	0.747 (0.000)
# of Obs.	298,551	223,912	298,589	223,941
Adjusted R^2	0.48	0.41	0.16	0.16
	2007 Survey			
	Assets (All Firms)	Assets (Drop Q1)	Sales (All Firms)	Sales (Drop Q1)
Total Assets	0.966 (0.000)	0.998 (0.000)	–	–
Sales	–	–	0.859 (0.000)	0.920 (0.000)
# of Obs.	181,940	136,449	181,967	136,468
Adjusted R^2	0.54	0.47	0.27	0.26
	2002 Survey			
	Assets (All Firms)	Assets (Drop Q1)	Sales (All Firms)	Sales (Drop Q1)
Total Assets	0.948 (0.000)	0.980 (0.000)	–	–
Sales	–	–	0.895 (0.000)	0.964 (0.000)
# of Obs.	80,911	60,682	80,926	60,692
Adjusted R^2	0.54	0.48	0.31	0.30
	1998 Survey			
	Assets (All Firms)	Assets (Drop Q1)	Sales (All Firms)	Sales (Drop Q1)
Total Assets	0.906 (0.000)	0.941 (0.000)	–	–
Sales	–	–	0.874 (0.000)	0.948 (0.000)
# of Obs.	53,776	40,331	53,803	40,352
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 surveys 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. Column 1 and 3 include all firms and Columns 2 and 4 drop the observations in the bottom quartile of assets or sales.

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.

We also investigate the potential bias caused by holding shells. Specifically, we check this by measuring the total registered capital of industrial firms in the Annual Industrial Survey that belong to a given owner and comparing this number to the total registered capital in the registration data including the registered capital of all the intermediate owners. A regression of log registered capital of firms in the industrial data on the log registered capital of all firms that belong to the same owner, including all the intermediate firms in the registration data, yields a coefficient of 0.92 and an R^2 of 0.91, which confirms that registered capital is a reasonable proxy to measure an owner's size.

B Inferred Historical Data

The registration records include both active firms and firms which have been closed. For firms that have been closed, the records contain the year in which the firms exited. By combining the information of firm's registration and exit year, we're able to infer historical data. Table A.2 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.2, 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.2: Method for Inferring Historical Data

Firm	Registration	Exit	Active Firm in		
	Year	Year	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 their immediate shareholders in 2019 (or 2013). The implicit assumption is that the immediate shareholders of a firm are constant over time. Of course this assumption is not true for some firms, as a firm can have new shareholders, some shareholders sell their equity, or the equity shares change. We do not have comprehensive data on the firm's historical shareholders in all years.[26] However, we do have contemporaneous registration data in 2013 and 2019. We can gauge the bias due to the assumption we make that the most recent shareholder information is the same as in the past using these two data-sets. Specifically, we compare the key results using data from 2013 inferred from 2019 data with those from the real 2013 data.

Table A.3: Inferred vs Real 2013 Data

	Real	Inferred
# Active Firms	14,121,908	14,125,941
# Ultimate Owners	29,082,604	30,158,962
State	93,348	64,998
Private	28,989,256	30,093,964
% Registered Capital		
State	32.6%	33.2%
Private	67.4%	66.8%
# Connected State Owners	12,009	8,847
# Downward Connections per State Owner	8.0	8.9
% Registered Capital		
Connected State Owners	30.6%	31.6%
Directly Connected Private Owners	14.6%	12.3%
Indirectly Connected Private Owners	14.5%	10.4%

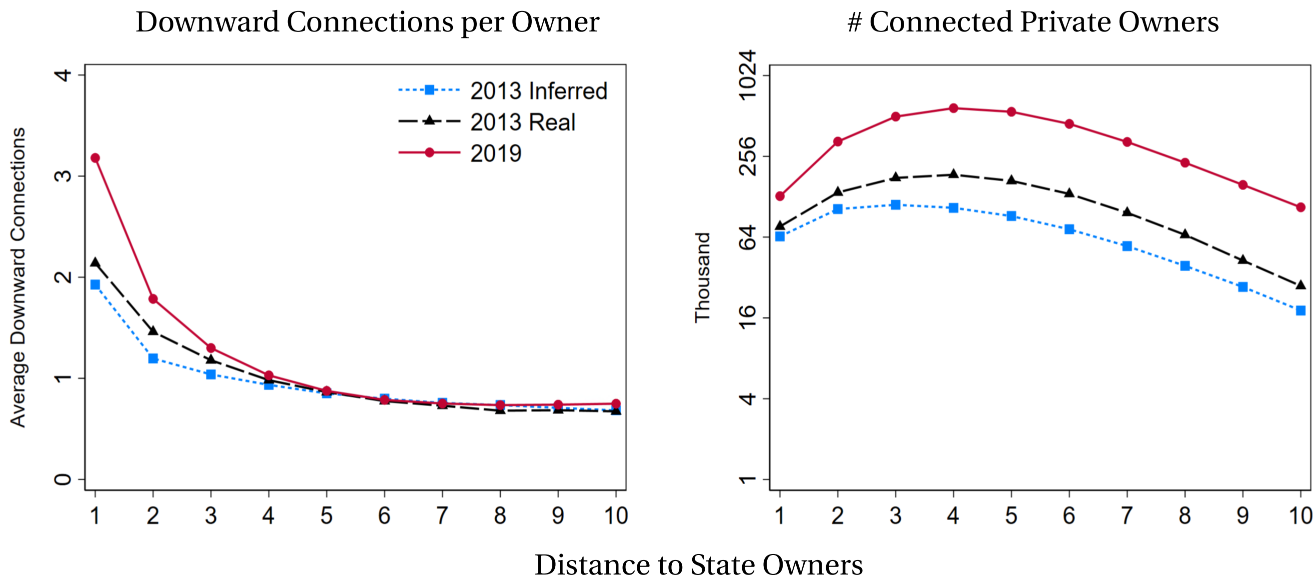
Note: "Real" uses the 2013 data. "Inferred" uses the 2019 data to calculate the statistics of firms and owners in 2013.

Table A.3 summarizes the results of the comparison. The numbers of active firms are very

[26]For some provinces and several years, there do exist some text records showing changes of firms' shareholders, but not for all provinces. For example, East Hope Aluminum was created as a joint venture between East Hope and Huanghe Aluminum and Electricity and later Huanghe sold its share to East Hope. However, we cannot find the change of immediate owners of East Hope Aluminum in the registration records. Another problem is that there is no encrypted personal ID for individual shareholders in these records. Because of this we chose not to recover the firms' real historical shareholders from these text records.

close in the inferred and the 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 understates the number of state owners and connected state owners in 2013. However, the share of registered capital of state owners and connected state owners are very close between the inferred and real data. Our inference using the 2019 data also understates the share of connected private owners in 2013 by about 6.4 percentage points. On the other hand, the number of average downward connections per state owner is slightly larger in the inferred data.

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

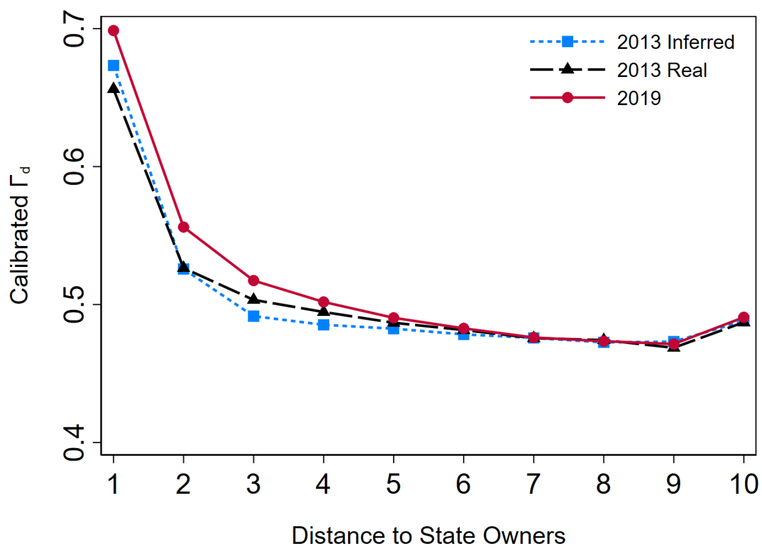


Note: Left panel shows the average number of downward connections per private owner by distance to the state using inferred 2013 data, real 2013 data and real 2019 data. Right panel shows the total number of connected private owners by distance to the state.

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 in the inferred data are just slightly smaller than that in 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 smaller number of connections per private owner (see left panel in Figure A.1).

Slightly overstating the number of average downward connections per state owner implies that we may underestimate the expansion and growth effect of directly connected private owners. On the other hand, underestimating the historical number and share of indirectly connected private owners implies that we may overestimate the expansion and growth effect of

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



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

Table A.4: $\Delta\Gamma_d$ on Private Sector Output Growth, 2013-2019 (Inferred vs Real 2013 Data)

	Inferred 2013 Data	Real 2013 Data
Private Growth	2.3%	2.7%
By Distance of Connection Benefit:		
Δ Direct Connection Benefit	0.8%	1.4%
Δ Indirect Connection Benefit	1.5%	1.3%

Note: Private sector output growth in model is average annual growth rate of total output of all private owners in the model due to the change in Γ_d we infer between 2013 (inferred and real data) and 2019.

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 using the real historical data and keeping other parameters unchanged. Figure A.2 plots these numbers, showing both Γ_d calibrated from the real 2013 data and from the 2013 data inferred from the 2019 data. As we can see, Γ_1 calibrated using inferred data is higher, and Γ_2 calibrated using inferred and real data are very close. But for $d = 3, 4, 5$, the real data give higher Γ_d. Then, as in Table 8 and 9, 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.4. When using the real historical data, the output growth caused by the increase of Γ is 2.7%, 0.4 percentage points higher than that calculated using inferred data. But the growth contributed by increase of benefit of indirect connection is smaller (1.3%) compared with that from the inferred data (1.5%).

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. To make the list complete, we include all possible combinations of division (e.g., Beijing) and department (e.g., Bureau of Finance). Then we match the name of the shareholder in the registration data with our list of state owners.[27]

We treat all the departments that belong to the same level of government as one 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 is that if a state firm is directly and 100% owned by a government, we classify it as a separate state owner. For example, although SAIC is owned by Shanghai's SASAC, we assume SAIC is a separate state owner.[28]

Table A.5 applies our identification of state owners to the Industrial Survey and compare the results with those in the literature (see, e.g., Hsieh and Song, 2015). We focus on firms in the 2013 Industrial Survey. First, following Hsieh and Song (2015), we identify a firm as state-owned if the share of its paid-in capital held directly by the state exceeds or equals 50% or when the state is reported as the controlling shareholder. The first row of Table A.5 shows the share

[27] We also identify state owners by searching several keywords in shareholder's name as supplementary to the list.

[28] We also checked around 7 thousand local government financing vehicles (LGFV). About a quarter of them are classified as state owners by our definition. For example, Shanghai Guosheng Group Co. is identified as a state owner because it is directly and 100% owned by SASAC of Shanghai. Another LGFV, Shanghai State-owned Assets Operation Co., is not identified as a state owner because it is owned by Shanghai International Group Co., which is a state owner.

Table A.5: Share of State Sector, Firms in 2013 Industrial Survey Data

	Share of State Sector			
	Sales	Assets	Paid-in Capital (Industrial Survey)	Registered Capital (Registration Data)
Hsieh and Song (2015)	22.2%	36.6%	33.1%	–
by state equity share	16.7%	27.6%	24.4%	25.2%
Threshold:				
$\geq 50\%$	16.2%	27.9%	25.0%	25.6%
$\geq 25\%$	20.7%	33.5%	30.5%	31.6%
$\geq 10\%$	23.2%	36.8%	33.8%	35.3%

Note: The first row shows the share of sales (assets and paid-in capital from 2013 industrial survey data) of firms in the 2013 Industrial Survey which are identified as state-owned following the method in Hsieh and Song (2015). More specifically, a firm is identified as state-owned when the share of paid-in capital held directly by the state exceeds or equals 50% or when the state is reported as the controlling shareholder. The second row shows the same shares using the state share in a firm by its state equity share in the 2013 registration data). The third to fifth rows shows the same shares of firms whose share of registered capital held by state owners are larger than 50%, 25% and 10% respectively.

of sales (assets and paid-in capital) of these identified state-owned firms of all firms.[29] The second row shows the same shares as in the first row, but uses the state shares in each industrial firm by its state equity share in the registration data. The implied state sales share is 16.7%, around 1/4 lower than that in the first row. However, notice that the number in the first row also includes sales of state controlled firms owned by non-state-owners of these firms. To make a more direct comparison, we set some ad hoc thresholds of state equity share for state-controlled firms and recalculate the state shares. The fourth row, for example, sets the threshold to 25%, implying that those firms with more than 25% equity ultimately owned by state owners be state-controlled. The state shares of sales and total assets are much closer to those in the first row. To conclude, although it is still likely that we miss some state-owned firms, the above numbers, together with the evidence discussed in the text, suggest that the bias cannot be large.

D Cash Flow Rights and Ultimate Owners

This appendix provides the details of calculating cash flow rights, which identify ultimate owners for each firm and determine their size. The next appendix proposes a simple way of calcu-

[29]We drop some observations in the 2013 Industrial Survey that don't have legal-person identities (e.g. branch companies) and cannot be matched with the registration data.

lating control rights, which are used to check the sensitivity of owners' size conditional on their distance to the state.

We begin with two matrices, X and Y. X_{ij} denotes the proportion of equity shares of firm i owned by another firm j. Y_{ik} denotes the proportion of equity shares of firm i owned by owner k from the four types of owners specified in Appendix C. Let M and N be the number of firms and owners, respectively. X is a $M \times M$ matrix and Y ix a $M \times N$ matrix. For each firm i, we have $\sum_{j=1}^{M} X_{ij} + \sum_{k=1}^{N} Y_{ik} = 1$.

We then derive matrix Z, where element Z_{ij} denotes the proportion of equity shares of firm i ultimately owned by owner j. Z is a $M \times N$ matrix. Z_{ij} measures cash flow rights.

$$Z = Y + \sum_{k=1}^{\infty} X^k Y.$$

If $\lim_{k\to\infty} X^k = \mathbf{0}$ and $\sum_{j=1}^{N} Z_{ij} = 1$ for any i, we can identify the ultimate owners for all firms. However, the two conditions are not always satisfied for the following three reasons in the data.

First, the records of immediate shareholders are incomplete for some firms. If $\sum_{j=1}^{M} X_{ij} + \sum_{k=1}^{N} Y_{ik} < 1$ for firm i, we might miss some of its ultimate owners. Second, cross holding is likely to fail $\lim_{k\to\infty} X^k = \mathbf{0}$. Take an extreme case for example. If two firms, u and v, hold all shares in each other, we will end up with $\sum_{j=1}^{N} Z_{ij} = 0$ for both firms. X_{uu}^k, X_{uv}^k, X_{vu}^k and X_{vv}^k in X^k will not converge to zero as k goes to infinity. Third, to facilitate computation, we set the equity share of a firm owned by another firm to zero when it is less than 0.1%. We also drop the ultimate owners who own less than 1% equity shares of a firm.

E Control Rights

We define a firm's controlling shareholder as the largest equity shareholder of the firm. When the firm's immediate shareholder is another firm, we look for its controlling shareholder until reaching the ultimate owners. If there are multiple largest shareholders, the control rights will be evenly divided.

For each firm i, denote $Y_i^{\text{1st}} \equiv \max\{Y_{ik}\}$ and denote by Y_i^{2nd} the second largest number of $\{Y_{ik}\}$. If

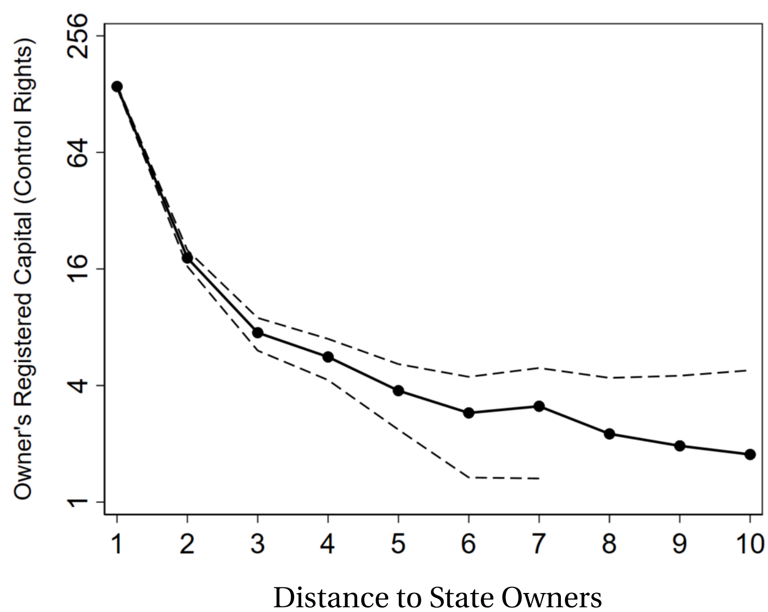
$$Y_i^{\text{1st}} > Y_i^{\text{2nd}} + \sum_{j=1}^{M} X_{ij},$$

which guarantees that no other owners can own more than Y_i^{1st} by cash flow rights. The owner who owns Y_i^{1st} is, therefore, the firm's controlling shareholder.

If the above condition is not satisfied for firm i, we need to check the condition for the firm's immediate shareholders. Incomplete information and crossing holding may also fail the identification of controlling shareholder.[30] In the 2019 registration data, the firms of which controlling shareholders can be identified account for 96% of the total registered capital.

We next calculate owner's size by assigning all registered capital of a firm to its controlling shareholder, and reproduce the main statistics in the upper panel of Table 7. The results are shown in the lower panel of the same table. We then reproduce the left panel in Figure 3 in Figure A.3 by control rights. They are very similar. Figure A.4 uses control rights to replicate Figure 4. The control right of the connected investor is typically larger than their share of the cash-flow, so this could be a reason why the ultimate owner's equity share is hidden behind several holding shells. However, it is still the case that the downward owner is typically the controlling shareholder of their businesses.

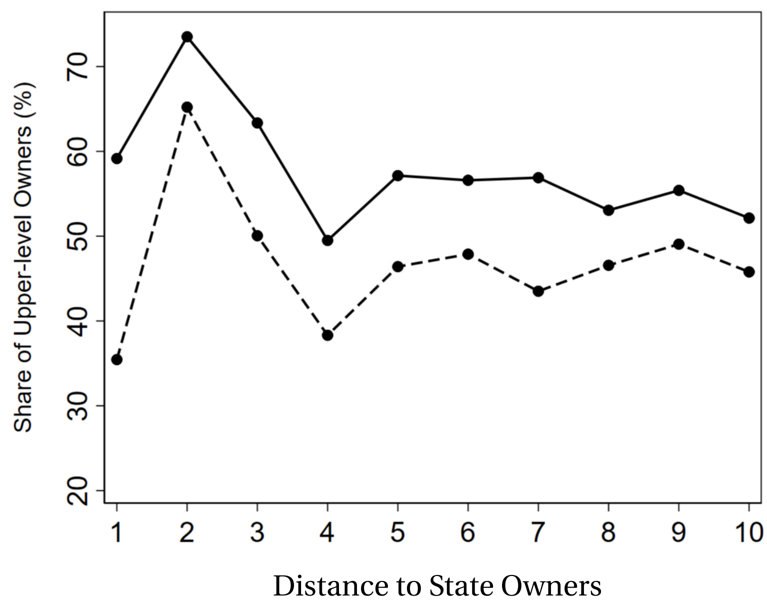
Figure A.3: Owner's Registered Capital by Control Rights, 2019



Note: This figure shows the ratio of average registered capital, by control rights, of connected private owners to the average registered capital of unconnected private owners by distance to the state (dashed lines represent 95% confidence intervals).

[30]Suppose that, for example, firm A does not satisfy the above condition and its immediate shareholder is firm B. If firm B does not satisfy the condition either, and one of firm B's immediate shareholders is firm A, we won't be able to identify controlling shareholder for the two firms.

Figure A.4: Share of Connected Investors in Private Owner's Capital by Control Rights, 2019



Note: Private owners' distances to state owners and joint ventures between connected investors and downward owners are determined by cash flow rights. Owner's registered capital is calculated by control rights. The solid line plots the share of joint ventures controlled by the connected investors weighted by joint ventures registered capital, which is the counterpart of the solid line in Figure 4. The dash line plots the ratio of the total registered capital in the joint ventures controlled by the connected investors to the total registered capital in the firms controlled by the downward owners.

F Robustness of Share of Connected Sector

In Table 7, we show the share of connected sector measured by owner's registered capital, which is the only variable we have in firm registration data to measure a firm's size. Appendix A further shows that firm's registered capital is highly correlated with its sales and total assets. Yet it is still useful to show the robustness of the results in Table 7 by measuring the share of connected sector using other variables or methods.

We first use sales and total assets provided by industrial firm survey data. To do this, we only keep owners which own at least one firm in the 2013 industrial survey. These owners' sales or total assets are defined as the sum of sales or total assets of firms in the 2013 industrial survey owned by them, weighted by their equity shares in respective firms. The results are shown in the first column of Table A.6. Table A.3 shows that connected state owners own 30.6% registered capital, while directly and indirectly connected private owners own 14.6% and 14.5% registered capital respectively. When measured by sales (upper panel of Table A.6), the share of connected

Table A.6: Share of Connected Sector Measured by Predicted Sales and Total Assets

	Real Data	Predicted Sales			
		Model (1)	Model (2)	Model (3)	Model (4)
State Owner	17.1%	22.3%	22.2%	20.6%	17.6%
Connected	16.4%	20.4%	20.2%	18.1%	15.2%
Private Owner	82.9%	77.7%	77.8%	79.4%	82.4%
Directly Connected	15.7%	12.8%	12.7%	10.9%	9.9%
Indirectly Connected	16.6%	16.2%	15.8%	14.5%	14.6%
	Real Data	Predicted Total Assets			
		Model (1)	Model (2)	Model (3)	Model (4)
State Owner	28.6%	27.1%	27.3%	26.0%	23.7%
Connected	27.9%	25.2%	25.2%	23.6%	21.4%
Private Owner	71.4%	72.9%	72.7%	74.0%	76.3%
Directly Connected	20.0%	13.8%	13.9%	12.9%	12.3%
Indirectly Connected	16.0%	15.5%	15.1%	14.9%	15.1%

Note: The first column of the upper (lower) panel shows the share of connected sector's sales (total assets) using the real data. Note that only owners which owns at least one firm in the 2013 industrial survey are included here and their sales (total assets) are defined as the sum of sales (total assets) of firms in the 2013 industrial survey which are owned by them, weighted by their equity shares in respective firms. The second to fifth columns of the upper (lower) panel show the share of connected sector's predicted sales (total assets). All owners are included here. We first predict sales (total assets) of all firms in the 2013 registration data using the models estimated in Table A.7 and then calculate each owners' predicted sales (total assets). In Model (1), we regress sales (total assets) in log term on firm's log registered capital. In Model (2), we add firm age as explanatory variable. In Model (3), we control city fixed effects, while in Model (4) we control both city and 2-digit industry fixed effects. When predicting sales or total assets using Model (4), industrial fixed effects are excluded as the registration data include service firms that are not covered by the industrial survey. The predicted sales or total assets using Model (4) only capture within-industry variations.

state owners is significantly lower (17.1%), but the share of directly and indirectly connected private owners are slightly larger (15.7% and 16.6%). When measured by total assets (lower panel of Table A.6), the share of connected state owners is very close to that in Table A.3 (27.9%), while the share of directly and indirectly connected private owners are larger (20.0% and 16.0%).

We also try another method to measure the share of connected sector. By merging the 2013 industrial firm survey data and 2013 registration data, we're able to estimate the relationship between sales or total assets and registered capital, firm age and location and industry fixed effects. Then, we can calculate predicted sales and total assets using the estimated model for all firms in the firm registration data, which in turn can be used to calculate share of predicted sales or total assets of connected sector. The regression results are shown in Table A.7, while the share of connected sector based on predicted sales or total assets are shown in the second to fifth columns in Table A.6. Generally speaking, the results based on predicted sales or assets might be slightly lower than the results based on registered capital, yet the difference is not big.

Table A.7: Sales and Total Assets vs. Registered Capital for Industrial Firms

	Dependent Variable: Sales 2013			
	(1)	(2)	(3)	(4)
Registered Capital	0.884 (0.002)	0.844 (0.002)	0.787 (0.002)	0.739 (0.002)
Firm Age	–	0.033 (0.001)	0.037 (0.001)	0.039 (0.001)
City Fixed Effects	N	N	Y	Y
2-Digit Industry Fixed Effects	N	N	N	Y
# of Obs.	267,548	267,548	267,548	264,665
	Dependent Variable: Total Assets 2013			
	(1)	(2)	(3)	(4)
Registered Capital	0.943 (0.001)	0.904 (0.001)	0.860 (0.001)	0.826 (0.001)
Firm Age	–	0.035 (0.000)	0.037 (0.000)	0.037 (0.000)
City Fixed Effects	N	N	Y	Y
2-Digit Industry Fixed Effects	N	N	N	Y
# of Obs.	270,118	270,118	270,118	267,174

Note: Observations are firms in the 2013 industrial survey data matched with firms in the 2013 registration data. The dependent variable in the upper (lower) panel is log sales (total assets) from the industrial survey. The independent variables are log registered capital, firm age, city fixed effects and industry fixed effects from the registration data. Since the industrial survey only covers firms with annual sales above 20 million Yuan, truncated regressions are used to get the unbiased estimates of the coefficients. All the regressions are weighted by firm employment from the industrial survey.

G Model Proofs

For the proofs, we will first assume that the equilibrium has an interior solution. That is, we assume that not all private owners are connected to the state in equilibrium, as what the data shows. We will show in the proof of Proposition 4 that the condition will be satisfied if the number of state owners relative to that of private owners is sufficiently small, Γ_d is sufficiently small and λ is sufficiently large.

Proof of Propositions 1 and 2

Denote by $\bar{d}$ the distance where that $\Gamma_{\bar{d}} - \underline{\Gamma} > 0$ and $\Gamma_{\bar{d}+1} - \underline{\Gamma} = 0$. Given this definition we prove no owners are connected at distance $d > \bar{d}$.

Suppose there exist a set of private owners who choose $d > \bar{d}$ in equilibrium. Their net profits from being connected compared with the profits being unconnected are proportional to

$$\sum_{d>\bar{d}} \left\{ N_d \left[\lambda(\delta-1)n_d^{\delta} - \lambda\delta n_{d-1}^{\delta-1} \right] \right\} = -N_{\bar{d}+1}\lambda\delta n_{\bar{d}}^{\delta-1} - \sum_{d>\bar{d}} N_d \lambda n_d^{\delta}.$$

Note that $\Gamma_d = \underline{\Gamma}$ if $d > \bar{d}$. Therefore owners get no extra profits from production by choosing $d > \bar{d}$ compared with being unconnected. Since the right-hand side of the equation is negative, at least one of these owners choosing $d > \bar{d}$ will be better off if she is unconnected. This contradicts the choice of $d > \bar{d}$ that maximizes her total profits given by (4). Therefore, when $\Gamma_d = \underline{\Gamma}$, no one will choose d.

We next prove $n_{d-1} > 0$ and $N_d > 0$ if $\Gamma_d > \underline{\Gamma}$. Suppose there is $d \leq \bar{d}$ such that $\Gamma_d > \underline{\Gamma}$ but $n_{d-1} = 0$. Then,

$$\pi \left[A_i \,|\text{choose } d \right] \propto \left(\Gamma_d \, A_i \right)^{\frac{1}{1-\beta}} + \lambda(\delta-1)n_d^{\delta} > \left(\underline{\Gamma} \, A_i \right)^{\frac{1}{1-\beta}} ,$$

which means that for any private owner, choosing d would make them strictly better off compared to being unconnected. Thus, there should not be any unconnected private owners in the economy, which contradicts the assumption that only a subset of private owners are connected. We establish that for any d, $n_{d-1} > 0$ if $\Gamma_d > \underline{\Gamma}$.

Since $n_d > 0$ for any $d \in [0, \bar{d}-1]$ and the number of connected private owners at distance d, N_d, equals $N_0 \prod_{i=0}^{d-1} n_i$. Therefore, $N_d > 0$ for any $d \in [1, \bar{d}]$ (recall that N_0 denotes the total number of state owners).

The next step is to prove that private owners sort into distance to the state by their TFP – i.e., the second part of Propositions 1 and Proposition 2. Suppose there exist owner i and j, with $A_i > A_j$, but they choose to be connected with the state at distance $d_i > d_j$, respectively. Owner j's choice indicates that $\pi \left[A_j \,|\text{choose } d_j \right] \geq \pi \left[A_j \,|\text{choose } d_i \right]$. Because $\Gamma_{d_j} > \Gamma_{d_i}$, $\pi \left[A \,|\text{choose } d_j \right] - \pi \left[A \,|\text{choose } d_i \right]$ is a strictly increasing function of A. We should have $\pi \left[A_i \,|\text{choose } d_j \right] > \pi \left[A_i \,|\text{choose } d_i \right]$. This contradicts the condition that owner i chooses d_i by maximizing $\pi \left[A_i \,|\text{choose } d \right]$.

Proof of Proposition 3

Note that the set of cutoff TFP $\{A_d\}_{d=1}^{\bar{d}}$ are the solutions of following indifference conditions:

$$\begin{aligned}\left(\Gamma_d^{\frac{1}{1-\beta}}-\Gamma_{d+1}^{\frac{1}{1-\beta}}\right)(A_d)^{\frac{1}{1-\beta}} &= \left[\lambda\delta n_{d-1}^{\delta-1}-\lambda(\delta-1)n_d^{\delta}\right]-\left[\lambda\delta n_d^{\delta-1}-\lambda(\delta-1)n_{d+1}^{\delta}\right],\\ \left(\Gamma_{\bar{d}-1}^{\frac{1}{1-\beta}}-\Gamma_{\bar{d}}^{\frac{1}{1-\beta}}\right)(A_{\bar{d}-1})^{\frac{1}{1-\beta}} &= \left[\lambda\delta n_{\bar{d}-2}^{\delta-1}-\lambda(\delta-1)n_{\bar{d}-1}^{\delta}\right]-\lambda\delta n_{\bar{d}-1}^{\delta-1},\\ \left(\Gamma_{\bar{d}}^{\frac{1}{1-\beta}}-\underline{\Gamma}^{\frac{1}{1-\beta}}\right)(A_{\bar{d}})^{\frac{1}{1-\beta}} &= \lambda\delta n_{\bar{d}-1}^{\delta-1}.\end{aligned} \tag{9}$$

Because $\Gamma_d > \underline{\Gamma}$ is strictly decreasing in d if $d \in [1, \bar{d}]$, the left-hand side of the three equations in (9) are positive. Thus we have:

$$\begin{aligned}\lambda\delta n_{\bar{d}-1}^{\delta-1} &> 0,\\ \lambda\delta\left(n_{\bar{d}-2}^{\delta-1}-n_{\bar{d}-1}^{\delta-1}\right) &> \lambda(\delta-1)n_{\bar{d}-1}^{\delta},\\ \lambda\delta\left(n_{d-1}^{\delta-1}-n_d^{\delta-1}\right) &> \lambda(\delta-1)\left(n_d^{\delta}-n_{d+1}^{\delta}\right).\end{aligned}$$

Recursively, we can prove that $n_0 > n_1 > \cdots > n_{\bar{d}-1} > 0$.

Lemma 1

Lemma 1. *Assume that the distribution of private owners' TFP is sufficiently compact –* $\lim_{s\to\infty} F_s(x) = 1$ *for* $x > \min\{A_i\}$, $\lim_{s\to\infty} F_s(x) = 0$ *for* $x < \min\{A_i\}$, $\min\{A_i\} > 0$, *and* $F'' < 0$. *Then, we have*

$$\frac{\partial A_j}{\partial \Gamma_d} = 0 \text{ for any } j, d \in [1, \bar{d}].$$

Proof of Lemma 1

We first differentiate (9) with respect to $\{\Gamma_d\}_{d=1}^{\bar{d}}$. If $d = j$, we have:

$$\begin{aligned}&\lambda\delta(\delta-1)\left[n_{j-1}^{\delta-2}\frac{\partial n_{j-1}}{\partial\Gamma_d}-(1+n_j)\,n_j^{\delta-2}\frac{\partial n_j}{\partial\Gamma_d}+n_{j+1}n_{j+1}^{\delta-2}\frac{\partial n_{j+1}}{\Gamma_d}\right]\\ &=\frac{1}{1-\beta}\Gamma_j^{\frac{1}{1-\beta}-1}(A_j)^{\frac{1}{1-\beta}}+\frac{1}{1-\beta}\left(\Gamma_j^{\frac{1}{1-\beta}}-\Gamma_{j+1}^{\frac{1}{1-\beta}}\right)(A_j)^{\frac{1}{1-\beta}-1}\frac{\partial A_j}{\partial\Gamma_d}.\end{aligned}$$

If $d = j + 1$, we have

$$\begin{aligned}&\lambda\delta(\delta-1)\left[n_{j-1}^{\delta-2}\frac{\partial n_{j-1}}{\partial\Gamma_d}-(1+n_j)\,n_j^{\delta-2}\frac{\partial n_j}{\partial\Gamma_d}+n_{j+1}n_{j+1}^{\delta-2}\frac{\partial n_{j+1}}{\Gamma_d}\right]\\ &=-\frac{1}{1-\beta}\Gamma_{j+1}^{\frac{1}{1-\beta}-1}(A_j)^{\frac{1}{1-\beta}}+\frac{1}{1-\beta}\left(\Gamma_j^{\frac{1}{1-\beta}}-\Gamma_{j+1}^{\frac{1}{1-\beta}}\right)(A_j)^{\frac{1}{1-\beta}-1}\frac{\partial A_j}{\partial\Gamma_d}.\end{aligned}$$

If $d \neq j$ and $d \neq j+1$, we have

$$\begin{aligned}&\lambda\delta(\delta-1)\left[n_{j-1}^{\delta-2}\frac{\partial n_{j-1}}{\partial\Gamma_d}-(1+n_j)\,n_j^{\delta-2}\frac{\partial n_j}{\partial\Gamma_d}+n_{j+1}n_{j+1}^{\delta-2}\frac{\partial n_{j+1}}{\Gamma_d}\right]\\&=\frac{1}{1-\beta}\left(\Gamma_j^{\frac{1}{1-\beta}}-\Gamma_{j+1}^{\frac{1}{1-\beta}}\right)(A_j)^{\frac{1}{1-\beta}-1}\frac{\partial A_j}{\partial\Gamma_d}.\end{aligned}$$

Denote U a $\bar{d}\times\bar{d}$ matrix with $U_{jd}=\frac{\partial n_{j-1}}{\partial\Gamma_d}$ and V a $\bar{d}\times\bar{d}$ matrix with $V_{jd}=\frac{\partial A_j}{\partial\Gamma_d}$. We can rewrite the above equations in matrix form:

$$XAU = BV + Y. \tag{10}$$

Here, X is a $\bar{d}\times\bar{d}$ upper triangular matrix $\begin{bmatrix}1 & -(1+n_1) & n_2 & 0 & \cdots & 0\\ 0 & 1 & -(1+n_2) & n_3 & \cdots & 0\\ 0 & 0 & 1 & -(1+n_3) & \cdots & 0\\ 0 & 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & 1\end{bmatrix}$, with $X_{ij}=1$ if $i=j$, $X_{ij}=-(1+n_{j-1})$ if $i+1=j$, and $X_{ij}=n_{j-1}$ if $i+2=j$, $X_{ij}=0$ otherwise. A is a $\bar{d}\times\bar{d}$ diagonal matrix with $A_{ii}=\lambda\delta(\delta-1)n_{i-1}^{\delta-2}$. B is also a $\bar{d}\times\bar{d}$ diagonal matrix with $B_{ii}=\frac{1}{1-\beta}\left(\Gamma_i^{\frac{1}{1-\beta}}-\Gamma_{i+1}^{\frac{1}{1-\beta}}\right)(A_i)^{\frac{1}{1-\beta}-1}$. Y is a $\bar{d}\times\bar{d}$ upper triangular matrix $\begin{bmatrix}Y_{11} & Y_{12} & 0 & 0 & \cdots & 0\\ 0 & Y_{22} & Y_{23} & 0 & \cdots & 0\\ 0 & 0 & Y_{33} & Y_{34} & \cdots & 0\\ 0 & 0 & 0 & Y_{44} & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & Y_{\bar{d}\bar{d}}\end{bmatrix}$, with $Y_{ij}=\frac{1}{1-\beta}\Gamma_i^{\frac{1}{1-\beta}-1}(A_i)^{\frac{1}{1-\beta}}$ if $i=j$, and $Y_{ij}=-\frac{1}{1-\beta}\Gamma_{i+1}^{\frac{1}{1-\beta}-1}(A_i)^{\frac{1}{1-\beta}}$ if $i+1=j$, $Y_{ij}=0$ otherwise.

We can further explore the relationship between U and V. As $n_{j-1}\equiv\frac{N_j}{N_{j-1}}=\frac{F(A_{j-1})-F(A_j)}{F(A_{j-2})-F(A_{j-1})}$, we have

$$\begin{aligned}\frac{\partial n_{j-1}}{\partial\Gamma_d}&=\frac{\partial n_{j-1}}{\partial A_j}\frac{\partial A_j}{\partial\Gamma_d}+\frac{\partial n_{j-1}}{\partial A_{j-1}}\frac{\partial A_{j-1}}{\partial\Gamma_d}+\frac{\partial n_{j-1}}{\partial A_{j-2}}\frac{\partial A_{j-2}}{\partial\Gamma_d}\\&=\frac{N}{N_{j-1}}\left[-f(A_j)\frac{\partial A_j}{\partial\Gamma_d}+(1+n_{j-1})f(A_{j-1})\frac{\partial A_{j-1}}{\partial\Gamma_d}-n_{j-1}f(A_{j-2})\frac{\partial A_{j-2}}{\partial\Gamma_d}\right].\end{aligned}$$

The above equation can be rewritten in matrix form.

$$U = -ZX'\mathcal{F}V, \tag{11}$$

where Z is a $\bar{d}\times\bar{d}$ diagonal matrix with $Z_{ii}=\frac{N}{N_{i-1}}$ and $\mathcal{F}$ is a $\bar{d}\times\bar{d}$ diagonal matrix with $\mathcal{F}_{ii}=f(A_i)$.

Combining (10) and (11), we have:

$$V = -\mathcal{F}^{-1}(XAZX'+B\mathcal{F}^{-1})^{-1}Y. \tag{12}$$

We are now ready to prove Lemma 1. The last equation of (9) implies

$$\lambda\delta n_{\bar{d}-1}^{\delta-1} = \left(\Gamma_{\bar{d}}^{\frac{1}{1-\beta}} - \underline{\Gamma}^{\frac{1}{1-\beta}}\right)(A_{\bar{d}})^{\frac{1}{1-\beta}} > \left(\Gamma_{\bar{d}}^{\frac{1}{1-\beta}} - \underline{\Gamma}^{\frac{1}{1-\beta}}\right)(\min\{A_i\})^{\frac{1}{1-\beta}}.$$

Because $\min\{A_i\} > 0$, $n_{\bar{d}-1}$ has a lower bound which is strictly positive. Proposition 3 ensures that $n_0 > n_1 > \cdots > n_{\bar{d}-2} > n_{\bar{d}-1}$. Since N_0 is exogenous, the number of connected private owners at each distance, $N_d = N_0 \prod_{i=0}^{d-1} n_i$, has a lower bound, denoted by $\underline{N}_d$. The rest of the proof takes three steps.

First, we prove $\lim\limits_{s\to\infty} A_{d,s} = \min\{A_i\}$, where $A_{d,s}$ is the cutoff TFP at distance d associated with $F_s(A)$. If for some d, $\lim\limits_{s\to\infty} A_{d,s} \neq \min\{A_i\}$, then there exist $\varepsilon > 0$ and $s > S$, for any positive S, such that $A_{d,s} - \min\{A_i\} > \varepsilon$. The assumption about the TFP distribution implies $\lim_{s\to\infty} F_s(x) = 1$ for $x \geq \min\{A_i\}$ and $\lim_{s\to\infty} F_s(x) = 0$ for $x < \min\{A_i\}$. So, there exists s such that $1 - F_s(A_{d,s}) < 1 - F_s(\min\{A_i\} + \varepsilon) < \frac{\sum_{i=1}^{d} \underline{N}_i}{N}$, which cannot be true.

Second, we prove $\lim\limits_{s\to\infty} f_s(A_{d,s}) = \infty$. If for some d, $\lim\limits_{s\to\infty} f(A_{d,s}) < \infty$, there exist $M > 0$ and $s > S$, for any positive S, such that $f_s(A_{d,s}) < M$. The TFP distribution implies $\forall M' > 0$, $\lim\limits_{s\to\infty} \int_{x\in\{x|f_s(x)\leq M'\}} f_s(x)dx = 0$. Thus, there exists s such that $1-F_s(A_{d,s}) < \int_{x\in\{x|f_s(x)\leq M\}} f_s(x)dx < \frac{\sum_{i=1}^{d} \underline{N}_i}{N_p}$ since $F'' < 0$, which cannot be true.

Since $\lim\limits_{s\to\infty} f_s(A_{d,s}) = \infty$, $\lim\limits_{s\to\infty} \mathcal{F}_s^{-1} = \text{diag}(0, 0, \cdots, 0)$, where $\mathcal{F}_s$ is the matrix $\mathcal{F}$ in (11) associated with $F_s(A)$. Since $A_{d,s}$, $n_{d,s}$ and $N_{d,s}$ are bounded, all the elements of $\lim\limits_{s\to\infty} (X_s A_s Z_s X_s' + B_s \mathcal{F}_s^{-1})^{-1} Y_s$ are also bounded. So, $\lim\limits_{s\to\infty} V_s = \mathbf{0}$ by (12). Lemma 1 is proved.

Proof of Proposition 4

Lemma 1 shows that the effects of changes of Γ_d on TFP cutoffs are negligible, i.e. $V = \mathbf{0}$. Combining with (10), we have:

$$U = A^{-1}X^{-1}Y.$$

More specifically,

$$A^{-1} = \frac{1}{\lambda\delta(\delta-1)} * \begin{bmatrix} \frac{1}{n_0^{\delta-2}} & 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{n_1^{\delta-2}} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{n_2^{\delta-2}} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \frac{1}{n_3 4^{\delta-2}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \frac{1}{n_{\bar{d}-1}^{\delta-2}} \end{bmatrix},$$

$$X^{-1}=\begin{bmatrix} 1 & 1+n_1 & 1+n_1+n_1n_2 & 1+n_1+n_1n_2+n_1n_2n_3 & \cdots & \frac{\sum_{d=0}^{\bar{d}-1}\prod_{i=0}^{d}n_i}{n_0} \\ 0 & 1 & 1+n_2 & 1+n_2+n_2n_3 & \cdots & \frac{\sum_{d=1}^{\bar{d}-1}\prod_{i=1}^{d}n_i}{n_1} \\ 0 & 0 & 1 & 1+n_3 & \cdots & \frac{\sum_{d=2}^{\bar{d}-1}\prod_{i=2}^{d}n_i}{n_2} \\ 0 & 0 & 0 & 1 & \cdots & \frac{\sum_{d=3}^{\bar{d}-1}\prod_{i=3}^{d}n_i}{n_3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix},$$

and

$$Y=\frac{1}{1-\beta}\left(\min\{A_i\}\right)^{\frac{1}{1-\beta}} * \begin{bmatrix} \Gamma_1^{\frac{1}{1-\beta}-1} & -\Gamma_2^{\frac{1}{1-\beta}-1} & 0 & 0 & \cdots & 0 \\ 0 & \Gamma_2^{\frac{1}{1-\beta}-1} & -\Gamma_3^{\frac{1}{1-\beta}-1} & 0 & \cdots & 0 \\ 0 & 0 & \Gamma_3^{\frac{1}{1-\beta}-1} & -\Gamma_4^{\frac{1}{1-\beta}-1} & \cdots & 0 \\ 0 & 0 & 0 & \Gamma_4^{\frac{1}{1-\beta}-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \Gamma_{\bar{d}}^{\frac{1}{1-\beta}-1} \end{bmatrix}.$$

This gives (6). As $N_j = N_0\prod_{i=0}^{j-1} n_i$, we obtain (5).

Proposition 4 also tells us that (1) n_j, $j \leq d-1$, goes to infinity when Γ_d goes to infinity; (2) n_j, for $j \in [1,\bar{d}]$, goes to infinity when λ goes to zero. Also note that N_j increases proportionally with N_0. To ensure that not all private owners are connected to the state, $\sum_{d=1}^{\bar{d}} N_d < N$, the number of state owners relative to that of private owners should be sufficiently small, Γ_d should be sufficiently small and λ should be sufficiently large.

Proof of Proposition 5

Denote by $\mathcal{Y}_j$ the total output of private owners connected at distance j.

$$\mathcal{Y}_j = N_j\left(\frac{\beta^\beta\,\Gamma_j^\beta\,\min\{A_i\}}{w^\beta}\right)^{\frac{1}{1-\beta}},$$

$$\frac{\partial\ln\mathcal{Y}_j}{\partial\ln\Gamma_d}=\begin{cases}\frac{\partial\ln N_j}{\partial\ln\Gamma_d}+\Phi_d & \text{if } j\neq d\\ \frac{\partial\ln N_j}{\partial\ln\Gamma_d}+\frac{\beta}{1-\beta}+\Phi_d & \text{if } j=d\end{cases}$$

where $\Phi_d \equiv -\frac{\beta}{1-\beta}\frac{\partial\ln w}{\partial\ln\Gamma_d}$, which captures the effect of changes of factor price caused by changes of Γ_d. We call it the general equilibrium effect.

Denoted by $\mathcal{Y}^c$ total output of all connected private owners.

$$\frac{\partial \ln \mathcal{Y}^c}{\partial \ln \Gamma_d} = \underbrace{\frac{\mathcal{Y}_d}{\mathcal{Y}^c} \frac{\beta}{1-\beta}}_{\text{Direct Effect}} + \underbrace{\sum_{j=1}^{\bar{d}} \frac{\mathcal{Y}_j}{\mathcal{Y}^c} \frac{\partial \ln N_j}{\partial \ln \Gamma_d}}_{\text{Indirect Effect}} + \Phi_d.$$

The first term, what we call *direct* effect, represents the increase of output for the owners connected at distance d caused by higher Γ_d. The second term, what we call *indirect* effect, shows the increase of output of connected private owners as higher Γ_d increases the number of connected owners.

Denote by N^c the total number of connected private owners and $N^{nc} \equiv N - N^c$ the total number of unconnected private owners. Denoted by $\mathcal{Y}^{nc} \equiv N^{nc} \left(\frac{\beta^\beta \underline{\Gamma}^\beta \min\{A_i\}}{w^\beta} \right)^{\frac{1}{1-\beta}}$ and by $\mathcal{Y}^s \equiv N_0 \left(\frac{\beta^\beta \tilde{A}_0}{w^\beta} \right)^{\frac{1}{1-\beta}}$.

$$\frac{\partial \ln \mathcal{Y}^{nc}}{\partial \ln \Gamma_d} = -\frac{N^c}{N - N^c} \sum_{j=1}^{\bar{d}} \frac{N_j}{N^c} \frac{\partial \ln N_j}{\partial \ln \Gamma_d} + \Phi_d,$$

$$\frac{\partial \ln \mathcal{Y}^s}{\partial \ln \Gamma_d} = \Phi_d.$$

It is immediate that

$$\frac{\partial \ln \frac{\mathcal{Y}^c}{\mathcal{Y}^c + \mathcal{Y}^{nc} + \mathcal{Y}^s}}{\partial \ln \Gamma_d} = \frac{\mathcal{Y}^{nc}}{\mathcal{Y}^c + \mathcal{Y}^{nc} + \mathcal{Y}^s} \left(\frac{\partial \ln \mathcal{Y}^c}{\partial \ln \Gamma_d} - \frac{\partial \ln \mathcal{Y}^{nc}}{\partial \ln \Gamma_d} \right) + \frac{\mathcal{Y}^s}{\mathcal{Y}^c + \mathcal{Y}^{nc} + \mathcal{Y}^s} \left(\frac{\partial \ln \mathcal{Y}^c}{\partial \ln \Gamma_d} - \frac{\partial \ln \mathcal{Y}^s}{\partial \ln \Gamma_d} \right) > 0.$$

Similarly, we can prove that $\frac{\mathcal{Y}^c}{\mathcal{Y}^c + \mathcal{Y}^{nc}}$ increases in Γ_d for any d.

H Effects of Connection Benefits on Aggregate Output

We first derive the elasticity of total output of private owners, denoted by $\mathcal{Y}^p$, and aggregate output, denoted by $\mathcal{Y}$, with respect to Γ_d.

$$\frac{\partial \ln \mathcal{Y}^p}{\partial \ln \Gamma_d} = \frac{\mathcal{Y}_d}{\mathcal{Y}^p} \frac{\beta}{1-\beta} + \sum_{j=1}^{\bar{d}} \left[\frac{\mathcal{Y}_j}{\mathcal{Y}^p} \frac{\partial \ln N_j}{\partial \ln \Gamma_d} \right] - \frac{\mathcal{Y}^{nc}}{\mathcal{Y}^p} \frac{N^c}{N - N^c} \sum_{j=1}^{\bar{d}} \frac{N_j}{N^c} \frac{\partial \ln N_j}{\partial \ln \Gamma_d} + \Phi_d$$

$$\frac{\partial \ln \mathcal{Y}}{\partial \ln \Gamma_d} = \frac{\mathcal{Y}_d}{\mathcal{Y}} \frac{\beta}{1-\beta} + \sum_{j=1}^{\bar{d}} \left[\frac{\mathcal{Y}_j}{\mathcal{Y}} \frac{\partial \ln N_j}{\partial \ln \Gamma_d} \right] - \frac{\mathcal{Y}^{nc}}{\mathcal{Y}} \frac{N^c}{N - N^c} \sum_{j=1}^{\bar{d}} \frac{N_j}{N^c} \frac{\partial \ln N_j}{\partial \ln \Gamma_d} + \Phi_d,$$

Denote by $\mathcal{L}_j$ the resources employed by private owners connected at distance j.

$$\mathcal{L}_j = N_j \left(\frac{\beta \Gamma_j \min\{A_i\}}{w} \right)^{\frac{1}{1-\beta}}.$$

The market clearing condition is:

$$\sum_{j=1}^{\bar{d}} \left[N_j \left(\frac{\beta\, \Gamma_j \min\{A_i\}}{w} \right)^{\frac{1}{1-\beta}} \right] + N^{nc} \left(\frac{\beta\, \underline{\Gamma} \min\{A_i\}}{w} \right)^{\frac{1}{1-\beta}} + N_0 \left(\frac{\beta \tilde{A}_0}{w} \right)^{\frac{1}{1-\beta}} = \mathcal{L}.$$

The first term on the left-hand side represents resources employed by connected private owners, denoted by $\mathcal{L}^c \equiv \sum_{j=1}^{\bar{d}} \mathcal{L}_j$. The second and third terms represent resources employed by unconnected private owners and state owners, denoted by $\mathcal{L}^{nc}$ and $\mathcal{L}^s$, respectively. The general equilibrium effect, Φ_d, follows

$$\Phi_d = -\beta \left\{ \frac{\mathcal{L}_d}{\mathcal{L}} \frac{1}{1-\beta} + \sum_{j=1}^{\bar{d}} \left[\frac{\mathcal{L}_j}{\mathcal{L}} \frac{\partial \ln N_j}{\partial \ln \Gamma_d} \right] - \frac{\mathcal{L}^{nc}}{\mathcal{L}} \frac{N^c}{N - N^c} \sum_{j=1}^{\bar{d}} \frac{N_j}{N^c} \frac{\partial \ln N_j}{\partial \ln \Gamma_d} \right\}.$$

Because $\Gamma_j > \underline{\Gamma}$, any connected private owner at distance j employs more L than an unconnected private owner – i.e. $\mathcal{L}_j / N_j > \mathcal{L}^{nc} / (N - N^c)$. Therefore, $\Phi_d < 0$.

If Γ_d is interpreted as TFP, rather than output wedge, there will be no resource misallocation. In this case, $Y_i \propto L_i$, which also means $\mathcal{Y}_j / \mathcal{Y} = \mathcal{L}_j / \mathcal{L}$ for any $j \in [1, \bar{d}]$ and $\mathcal{Y}^{nc} / \mathcal{Y} = \mathcal{L}^{nc} / \mathcal{L}$. It is easy to show that $\frac{\partial \ln \mathcal{Y}}{\partial \ln \Gamma_d} > 0$.[31] Since $\frac{\partial \ln \mathcal{Y}^p}{\partial \ln \Gamma_d} > \frac{\partial \ln \mathcal{Y}}{\partial \ln \Gamma_d}$, $\frac{\partial \ln \mathcal{Y}^p}{\partial \ln \Gamma_d} > 0$. Total output of private owners and aggregate output increase in Γ_d for any d.

I Structural Estimation

We assume that private owner's productivity follows a Pareto distribution with the scale parameter set equal to 1. We first calibrate k, the shape parameter of the Pareto distribution. Denote $\bar{A}_d^{Data}$ the average productivity of private owners at distance $d \in [1, \bar{d}]$ in the data. The idea is to back out k by matching $\bar{A}_d^{Data}$. $\bar{d}$ is set to 10. The private owners with $d \leq 10$ account for more than 98% of the registered capital in the connected private sector. Given $\{N_d/N\}$, $d = 1, 2, \cdots, \bar{d}$, $\bar{A}_d$ will only depend on k. As will be clear below, the estimated model is constructed to exactly match $\{N_d/N\}$ in the data. Therefore, we can directly back out $\bar{A}_d\left(k, \{N_d/N\}\right)$ without knowing the other parameter values. Specifically, k is calibrated by:

[31] The direct effect of changes in Γ_d on $\mathcal{Y}^p$ is $\frac{\mathcal{Y}_d}{\mathcal{Y}^p} \frac{1}{1-\beta}$, rather than $\frac{\mathcal{Y}_d}{\mathcal{Y}^p} \frac{\beta}{1-\beta}$. The direct effect of changes in Γ_d on $\mathcal{Y}$ is $\frac{\mathcal{Y}_d}{\mathcal{Y}} \frac{1}{1-\beta}$, rather than $\frac{\mathcal{Y}_d}{\mathcal{Y}} \frac{\beta}{1-\beta}$.

$$\hat{k} = \arg\min_{k} \sum_{d=1}^{\bar{d}} \left(\ln \left(\bar{A}_d^{Data} \right) - \ln \left(\bar{A}_d \left(k, \{ N_d / N \} \right) \right) \right)^2 .$$

We use the average TFP of firms owned by an owner in the 2013 industrial survey to proxy her TFP. The estimated k is 255.

Denote $\phi \equiv [\frac{\theta}{1-\alpha-\beta}, \lambda^{\frac{1-\alpha-\beta}{\theta}}, \delta]$. Given $\hat{k}$ and ϕ, we can use (9) to back out Γ_d^t for $d = 1, 2, \cdots, \bar{d}$ and $t \in \{2000, 2010, 2019\}$ by matching N_d/N 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

$$\hat{\phi} = \arg\min_{\phi} \sum_{t \in \{2000,2010,2019\}} \sum_{d=1}^{\bar{d}} \left[\left(\ln \left(Y_{d,t}^{Data} \right) - \ln \left(Y_{d,t} \left(\hat{k}, \phi, \Gamma_d^t(\hat{k}, \phi) \right) \right) \right)^2 + \left(\ln \left(\bar{Y}_{d,t}^{Data} \right) - \ln \left(\bar{Y}_{d,t} \left(\hat{k}, \phi, \Gamma_d^t(\hat{k}, \phi) \right) \right) \right)^2 \right].$$

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

Table A.8 shows the estimates of model parameters, while Figure A.5 shows the fit of the model in terms of average firm size and owner size by distance to the state.

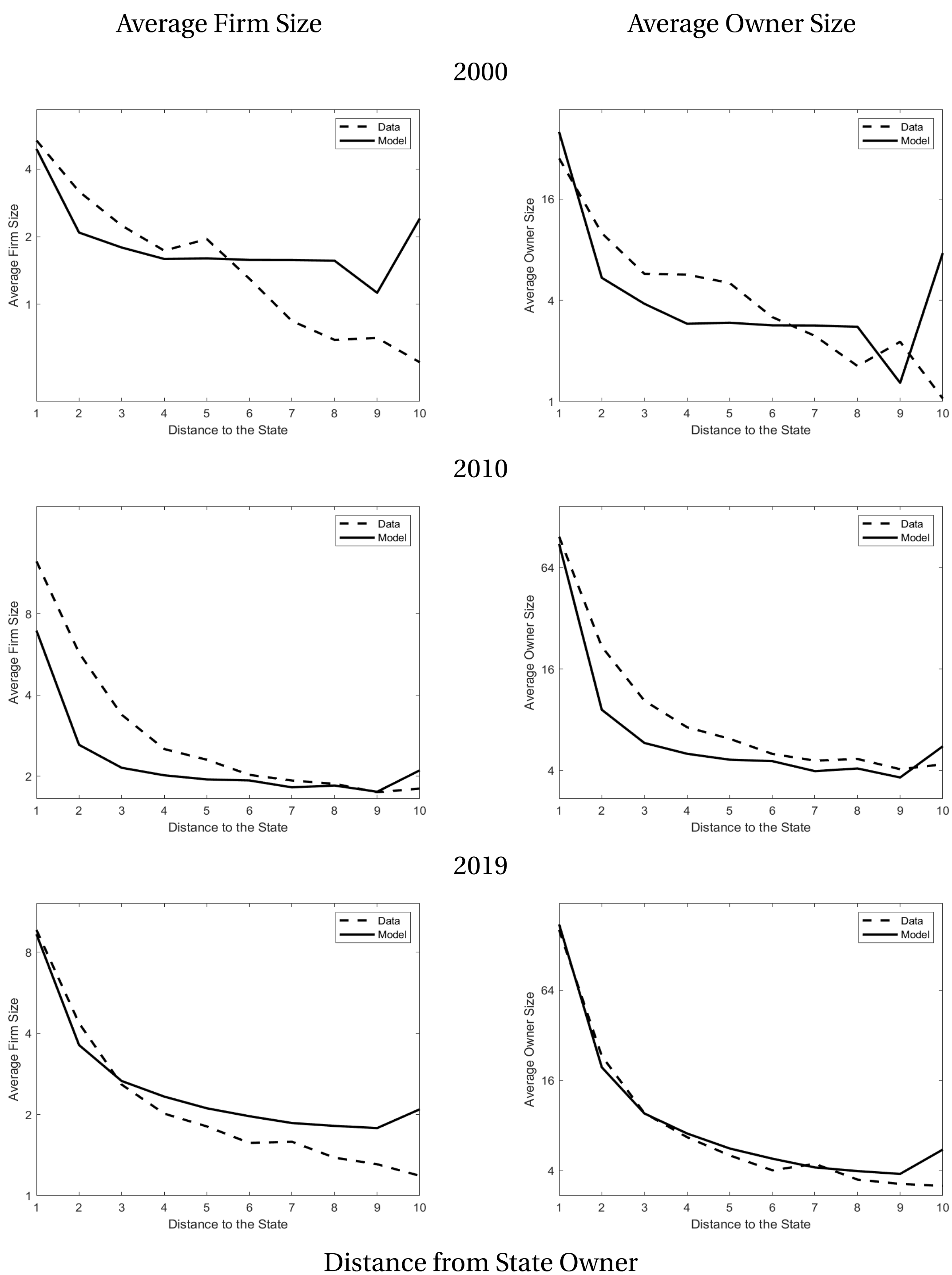
Table A.8: Model Parameters

Parameter	Description	Value	S.E.
$\frac{\theta}{1-\alpha-\beta}$	Elasticity of an owner's size to her TFP	10.32	0.90
$\lambda^{\frac{1-\alpha-\beta}{\theta}}$	Level of cost for span	0.47	0.01
δ	Convexity cost for span	2.54	0.08

J Robustness of the Effects of Changes in Connection Benefits

Our benchmark case models Γ_d as output wedge in a small open economy, where w is endogenous but r is exogenous. We now conduct two robustness checks. The first one is to model Γ_d as

Figure A.5: Target and Estimated Moments



part of owner's TFP: $Y_i = \Gamma_i\, A_i\, K_i^{\alpha}\, L_i^{\beta}$. Although the allocation of resources remains the same, firm and owner size will be different. In the benchmark case, firm's profit maximizing output is given by (2). When Γ_d is modeled as part of owner's TFP, firm and owner size become:

$$Y_i = \left[\left(\frac{\alpha}{r}\right)^{\alpha} \left(\frac{\beta}{w}\right)^{\beta} \right]^{\frac{1}{1-\alpha-\beta}} \Gamma_i^{\frac{1}{1-\alpha-\beta}}\, A_i^{\frac{1}{1-\alpha-\beta}},$$

$$\bar{Y}_i = (1-\alpha-\beta)^{\theta-1} \left[\left(\frac{\alpha}{r}\right)^{\alpha} \left(\frac{\beta}{w}\right)^{\beta} \right]^{\frac{\theta}{1-\alpha-\beta}} \Gamma_i^{\frac{\theta}{1-\alpha-\beta}}\, A_i^{\frac{\theta}{1-\alpha-\beta}}.$$

So, we have to re-estimate the model. Table A.9 reports the estimates of $\frac{\theta}{1-\alpha-\beta}$, $\lambda^{\frac{1-\alpha-\beta}{\theta}}$, and δ. Figure A.6 plots the implied connection benefits, Γ_d. Note that all the estimates are quantitatively close to those in the benchmark case. We then calculate the aggregate implications of the estimated changes in Γ_d, which are summarized in Table A.10.

Table A.9: Model Parameters (Γ_d as TFP)

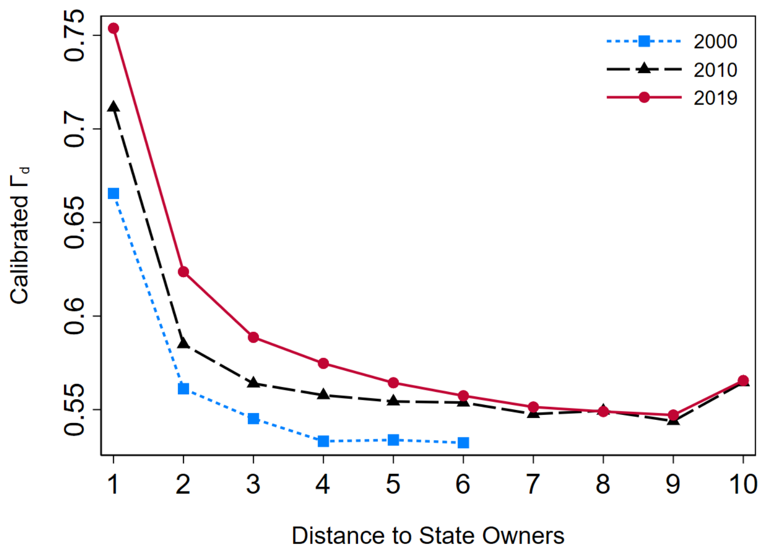
Parameter	Description	Value	S.E.
$\frac{\theta}{1-\alpha-\beta}$	Elasticity of an owner's size to her TFP	11.63	1.11
$\lambda^{\frac{1-\alpha-\beta}{\theta}}$	Level of cost for span	0.55	0.01
δ	Convexity cost for span	2.46	0.07

Note: Γ_d is part of owner's TFP.

The second robustness check is to consider capital wedge, rather than output wedge, for private owners in a closed economy, where the aggregate capital supply is inelastic and r is endogenous. Moreover, we assume that a connection to the state helps the private owner to reduce her capital wedge. This is motivated by Figure A.7, which plots the average capital productivity of firms owned by a private owner against her distance to the state, using data from the 2013 Chinese Annual Industrial Survey matched with the 2013 firm registration records. The capital productivity of firms owned by directly connected private owners are on average 40% lower than those firms owned by unconnected private owners. The capital productivity gap narrows as distance to the state increases.

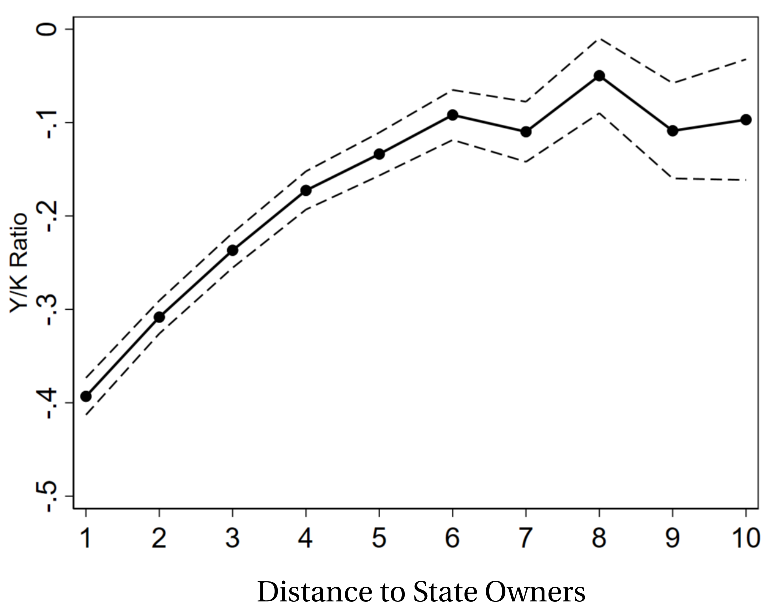
We then model Γ_d as a gross tax on the cost of capital. Private owners at distance d face a capital cost of r/Γ_d, where r is the capital cost for state owners and $\bar{\Gamma} \leq \Gamma_d \leq 1$ captures capital wedge for private owners. Γ_d decreases in d and $r/\bar{\Gamma}$ is the capital cost for unconnected private owners. Note that Γ_d has an ambiguous effect on capital misallocation. Higher Γ_d

Figure A.6: Estimated Benefit of Connection (TFP) by Distance to State, 2010-2019



Note: Figure plots the connection benefit Γ_d, which is modeled as part of owner's TFP, in 2019, 2010, and 2000. Γ_d for $d \in [7, 10]$ in 2000 are omitted as less than 1,000 private owners are connected at these distances in 2000.

Figure A.7: Capital Productivity of Private Owners, 2013



Note: Figure shows the average of log Y/K of firms owned by a connected private owner compared with firms owned by unconnected private owners by distance to the state (dashed lines represent 95% confidence intervals).

(smaller capital wedge) may alleviate capital misallocation by reducing the gap of marginal product of capital between state and connected private owners. It can also exacerbate capital misallocation by widening the gap between connected and unconnected private owners.

Firm and owner size follow:

$$Y_i = \left[\left(\frac{\alpha}{r}\right)^{\alpha}\left(\frac{\beta}{w}\right)^{\beta}\right]^{\frac{1}{1-\alpha-\beta}} \Gamma_i^{\frac{\alpha}{1-\alpha-\beta}} A_i^{\frac{1}{1-\alpha-\beta}}$$

$$\bar{Y}_i = (1-\alpha-\beta)^{\theta-1}\left[\left(\frac{\alpha}{r}\right)^{\alpha}\left(\frac{\beta}{w}\right)^{\beta}\right]^{\frac{\theta}{1-\alpha-\beta}} \Gamma_i^{\frac{\alpha\,\theta}{1-\alpha-\beta}} A_i^{\frac{\theta}{1-\alpha-\beta}}.$$

The estimation of the key parameters is identical to that with Γ_d as TFP. The implied Γ_d is different. Table A.11 reports the aggregate implications of the estimated changes in Γ_d as capital wedge.

Table A.10: Contribution of Δ Connection Benefits (TFP) to Aggregate Growth and Δ Private Sector Share

	2000-2019	2000-2010	2010-2019
% of Aggregate Growth	6.9%	4.8%	10.2%
% of Private Growth	21.4%	18.4%	25.7%
% of Δ Private Sector Output Share	95.6%	107.4%	86.0%

Note: Γ_d is part of owner's TFP.

Table A.11: Contribution of Δ Connection Benefits (Capital Wedge) on Aggregate Growth and Δ Private Sector Share (Endogenous r)

	2000-2019	2000-2010	2010-2019
% of Aggregate Growth	6.9%	5.7%	9.0%
% of Private Growth	21.5%	19.3%	24.7%
% of Δ Private Sector Output Share	95.7%	107.4%	86.0%

Note: Capital cost r is endogenous and Γ_d is modeled as capital wedge.