I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16259

The Births of New Private-Owned Enterprises in an Environment of State-Owned Enterprises

Zhong Zhao
Liang Zheng

JUNE 2023

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16259

The Births of New Private-Owned Enterprises in an Environment of State-Owned Enterprises

Zhong Zhao
Renmin University of China and IZA

Liang Zheng
Central University of Finance and Economics

JUNE 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Births of New Private-Owned Enterprises in an Environment of State-Owned Enterprises*

The impact of the incumbent state-owned enterprises (SOEs) on the births of new private-owned enterprises (POEs) in China is a central concern for the government and society. In this paper, we apply agglomeration theories to distinguish the linkages between SOEs and POEs. Using China's 2008 economic census, the 2007 Input-Output Table, and the 2005 population mini census, we measure the formation of new POEs at the city-industry level, and the agglomeration forces of distance proximity to inputs, outputs, labor, and technology. More explicitly, we measure the extent to which local SOEs provide relevant inputs, consume outputs, employ similar workers, and use similar technology. Our findings indicate that overall, incumbent SOEs hinder the formation of new POEs. For manufacturing, the entry of new POEs is significantly lower in places where more upstream SOEs are concentrated. For services, the entry of new POEs is significantly lower in places where more upstream and downstream SOEs are concentrated. However, the agglomeration effects from the incumbent POEs are either insignificant or significantly positive.

JEL Classification: L26, L60, L80, R10, R12

Keywords: new firm formation, state-owned enterprise, firm ownership, agglomeration

Corresponding author:
Liang Zheng
School of Management Science and Engineering
Central University of Finance and Economics
579M+X65
Chang Ping Qu
Bei Jing Shi
China
E-mail: lzheng@cufe.edu.cn

* We thank Bill Kerr, Janet E. Kohlhase, and participants at the First Chinese Labor Economist Forum, the 4rd International RUSE Workshop in China, the 7th IZA/CIER Annual Workshop on Research in Labor Economics, and seminars at Renmin University of China, Shanghai Jiao Tong University and University of Nottingham Ningbo Campus for very helpful comments. This research is supported by "Thematic Research Project on China's Income Distribution" (Project No. 21XNLG03) with funding from the research fund of Renmin University of China. All results have been reviewed to ensure that no confidential information is disclosed. The views expressed herein are those of the authors and do not necessarily reflect the views of the data providers. Any errors are our own.

1 Introduction

The formation of new firms is an important indicator of the dynamics of an economy. In China, after the economic reform that began in 1978 and especially since the retrenchment of the State-owned Enterprises (SOEs) in the middle 1990s, the private sector has become increasingly significant. Mixed ownership has become a hallmark of the Chinese economy; in 2019 the private sector employed more than eighty percent of China's total urban labor force (*China Statistical Yearbook* 2020).

Nonetheless, the private sector does not grow evenly, neither along time nor across regions. There are significant regional variations. Growth of the private sector and the births of new private-owned enterprises (POEs) are inevitably influenced by SOEs and other factors as well, For example, Guo et al. (2014) investigate the effect of political connections on POEs and find that POEs with political connections have enjoyed significant rent after China's 2002 Constitution amendment. In this paper, we explore a potentially important factor behind the regional variations of POE formations. We focus on the impact of the incumbent SOEs on the births of new POEs in China.[1] Intuitively, upstream and downstream SOEs may have different impacts. SOEs that employ similar workers can generate competition or spillover effects. The direction of the impact of the SOEs on the formation of POEs is not unambiguous.

In this study, we rely on agglomeration theories to distinguish the linkages between incumbent SOEs and new POEs. Marshall (1890) provides three distinct theories to explain the concentration of firms in a particular location: firms in agglomeration may gain the benefits from geographic proximity to suppliers or customers, a thick labor market, and technologies. More specifically, we analyze whether the entry of POEs is related to the extent to which local SOEs provide relevant inputs, consume outputs, employ similar workers, and use similar technology.

Although Marshall's proposition is intuitive, the empirical studies to test and distinguish the agglomeration theories face many challenges.[2] A growing body of literature attempts to disentangle the agglomeration theories by combining firm-level microdata with the information on inter-industry relations (e.g., input-output table and metrics of occupational similarity) (Ellison et al. 2010; Faggio et al. 2017; Diodato et al. 2018; O'Sullivan and Strange 2018; Faggio et al. 2020). One field of study utilizing this approach tries to explain clusters of entrepreneurship and new firm births.[3] Glaeser

[1] The POEs in this paper include private-owned firms and mixed-ownership firms which are controlled by the private sector but exclude foreign-owned or foreign-controlled firms; the SOEs in this paper do not include collective-owned firms.

[2] For selective literature reviews please refer to Rosenthal and Strange (2004), Puga (2010), and Combes and Gobillon (2015). Chatterji et al. (2014) offer an excellent review on the clusters of entrepreneurship and new firm formation.

[3] Rosenthal and Strange (2003) point out that using new firms as an indirect test of agglomeration poses both advantages and disadvantages. The positive side is that new firms are less likely to be constrained by previous decisions and alleviate the concern of reverse causality. The negative side is that there are no new firms in some cities, which causes a truncated econometric issue. In this paper, we mainly apply a Tobit model and compare the results from OLS and Tobit models in the robustness check section. Our main results are similar from both models.

and Kerr (2009) are among the first to explore which agglomeration mechanisms are at play. They find that the entry of U.S. manufacturing firms is higher in cities with more upstream and downstream firms as well as firms that employ the same sort of workers and technologies. Using Spanish data, Jofre-Monseny et al. (2011) conclude that all three Marshallian agglomeration mechanisms matter in attracting new manufacturing firms. Ghani et al. (2014) find that in India both manufacturing and service firm entry is related to the presence of suppliers, customers, and workers with adequate skills.

The agglomeration economics literature suggests that agglomeration effects may differ by the types of firms. First and most importantly, there has been a long-term debate on whether the benefits of agglomeration come from firms within in the same industries that generate localization externalities or firms in other industries that generate urbanization externalities (Rosenthal and Strange 2004; Puga 2010; Combes and Gobillon 2015). Second, the gains from agglomeration may attenuate across firm locations. For example, Rosenthal and Strange (2008) show that workers' wage premiums are more strongly generated by nearby workers within five miles relative to workers outside of the five-mile proximity radius.

The third dimension of heterogeneous agglomeration effects is firm size. Chinitz (1961) emphasizes that the presence of many small firms rather than dominant large firms is conducive to the entry of new firms since large firms both provide to and need less goods and services from other firms. Rosenthal and Strange (2010) find that small firms generate greater agglomeration externalities. Faggio et al. (2017) also support Chinitz's theory and finds that input sharing encourages the coagglomeration of industries with more small firms. This has important implications for China. After the privatization of small SOEs in the late 1990s, SOEs have tended to be large and may arrange their inputs and outputs at the national level and then have limited impact on the births of new local firms. Indeed the average employment size of the SOEs is 350 for manufacturing and 51 for services, compared with 41 and 14 for the POEs based on the data used in this paper.

Fourth, the temporal dimension of agglomeration effects is another concern. Past studies find that firms currently may be more successful in places where, historically, there are a number of relevant firms which accumulate knowledge and skills (Glaeser et al. 1992; Henderson et al. 1995). By comparing historical data, Diodato et al. (2018) find that industries tend to proximate to input providers or output customers in the past decades but to qualified workers in the recent decades. Fifth, firm affiliation status matters. Henderson (2003) distinguishes non-affiliate and corporate plants and finds that non-affiliate plants generate stronger agglomeration externalities. Sixth, the characteristics of firm owners play a role, For example, Ghani et al. (2013) show that female entrepreneurship is strongly correlated with local incumbent female firms.

However, the role of ownership remains unexplored. This paper focuses on the firm ownership dimension which is certainly important in a mixed ownership economy. To the best of our knowledge, no previous studies have tested agglomeration theories in this regard. This research brings the ownership dimension to the agglomeration economics literature, which is the first contribution of our paper. Mixed ownership is not only important to China, but also to many other countries, such as India, as well. SOEs provide a wide range of goods and services in core industries, especially utilities, real estate, finance, transportation, and capital-intensity industries. In a study of 40 countries, OECD (2017) estimates that SOEs outside of China employ over 9.2 million people and are valued at over USD 2.4 trillion in 2015. E.g., SOEs in Norway hire 9.5% of non-agricultural employees. Agarwal et al. (2022) report that as of the 2019-2020 fiscal year, SOEs in India account for roughly 22% of GDP in total assets and 12% of GDP in fixed assets. [4] More than 50% of SOEs engage in services, 40% in manufacturing, and the remaining mainly in mining and exploration.

On the substantive side, this paper is one of the first studies to analyze the agglomeration effects of SOEs on new POE formations.[5] Recent years, China's Communist Party's Central Committee and the State Council issued guidelines to deepen SOE reforms. The guidelines emphasize the ultimate importance of SOEs in the Chinese economy, and the new policies bring the study of SOEs back to center stage in China. Our results contribute to this renewed debate about the role of SOEs and are relevant to the policymakers who aim at reforming SOEs and promoting entrepreneurship. This is our second contribution. [6]Overall, we find that incumbent SOEs hinder the births of POEs. For manufacturing, the entry of new POEs is significantly lower in places where more upstream SOEs are located. For services, the entry of new POEs is significantly lower in places where more upstream SOEs and downstream SOEs are concentrated. However, the agglomeration effects from the incumbent POEs on the formation of new POEs are insignificant or significantly positive. For manufacturing, upstream POEs and POEs that use similar technology facilitate the births of new POEs. We further explore potential mechanisms through three lenses. First, the negative impact of SOEs on new POEs is not attributed to the large firm size of SOEs. Second, the presence of SOEs affects the finance and business environment faced by private firms. Third, SOEs have substantial advantages in hiring workers over POEs in the same labor market.

The remaining of the paper is organized as follows: Section 2 briefly introduces the institutional background. Section 3 describes data sets and key variables including the construction of agglomeration metrics. Sections 4 and 5 are empirical models and main results. Section 6 discusses potential mechanisms and Section 7 offers concluding remarks.

[4] These numbers exclude public sector banks and insurance companies.

[5] Existing literature mainly study an overall correlation between SOEs and new POEs, but they do not distinguish industrial linkages based on Marshallian agglomeration theories. E.g., Brandt et al. (2018) find that the size of SOEs has a causal impact on the entry of new POEs by examining city-level correlations, Zheng and Zhao (2017) prove that SOEs significantly reduce new private firm formation within the same service sectors.

[6] This paper is related to a strand of literature that assesses the employment effect of public sector on private sector. Using English data, Faggio and Overman (2014) find additional public sector job has limited effect in the short run (2003–2007) but significantly reduce total private sector jobs in the long run (1999–2007). Using Spanish city data over a long period, Jofre-Monseny et al. (2020) find public sector jobs crowd-in private employment. Other studies exploit historical incidences about public job relocation in U.K. (Faggio 2019) and in German (Becker et al. 2021).

2 Institutional Background

The development of POEs in the People's Republic of China (PRC) can be roughly divided into four phases.

The first phase is the transformation, diminution, and abolition of POEs from 1949 to 1956. After the founding of the PRC in 1949, China initially adopted a mixed economy policy, allowing for the co-existence of multiple ownerships. However, the government rapidly changed its policy, and adopted a central planning economic policy, and quickly nationalized the economy. In 1953, POEs employed 3.67 million people in China; another 8.98 million were self-employed, accounting for 46% of the urban labor force. In 1956, these numbers were drastically reduced to 0.03 million and 0.16 million, respectively, accounting for only 0.6% of China's urban labor force (Hu, 2014).

The second phase of POE development in the PRC is from 1956 to 1978. Along with the establishment of the socialist economic system, the private sector vanished in China. In the rural areas, farmers were organized into a commune system. In the urban areas, State-owned and collective enterprises controlled almost the whole economy. Economically, SOEs produced more than 90% of industrial output in 1956; legally, the private sector was outlawed during this period. This situation remained unchanged until China's economic reform began in 1978.

The third phase is from 1978 to 1992. China's economic reform originated in rural areas. The most fundamental change in the rural areas is the household responsibility system, which emerged at the end of the 1970s and eventually replaced the commune system in early 1980. This reform restored the central role of the family in production activity in rural China and returned economic freedom to the farmer. In contrast with rural areas, in urban areas, the reform was carried out piece by piece. As an exploratory project, in 1979, the PRC government created special economic zones in four of the country's coastal cities. Besides preferential policies, such as special tax laws, for the special economic zones, the government also allowed foreign direct investment and permitted firms in these zones to operate in accordance with the principles of a market economy rather than a planned economy.

Another important event, which happened after 1978, is the return of more than 17 million young people who had been sent down to the countryside for re-education during the cultural revolution. The sudden influx of these young people created serious employment problems in urban areas. The government was unable to allocate a job for all the returned young people and had to allow them to be self-employed.

However, until the middle 1980s, the Chinese urban economy was still dominated by SOEs and Collective Enterprises. The industrial structure breakthrough occurred in the 1980s with the expansion of Town and Village Enterprises (TVEs). Unlike the old-styled SOEs, the TVEs relied on the market instead of a planned economy for inputs and outputs. Though the TVEs still initially belonged to the collective sector, many later transformed into POEs. Meanwhile, small POEs started to emerge in urban areas but their numbers grew slowly because of institutional and ideological discrimination against the private sector.

The private sector did not obtain its legal status in China until 1988 when China amended its constitution. The amended constitution states that ”the state permits the private sector of the economy to exist and develop within the limits prescribed by law. The private sector of the economy is a complement to the socialist public economy. The state protects the lawful rights and interests of the private sector of the economy, and exercises guidance, supervision, and control over the private sector of the economy.” In 1992 the 14th Congress of the Communist Party of China formally adopted a socialist market economy as its long-term policy. In this year, 2.32 million people were employed by the private sector in China; 24.68 million more were self-employed, in comparison with virtually no POEs and only 0.14 million self-employed in 1978 (Quan 2008).

The last phase, from 1992 to the present, embarked on a significant corporation and privatization of SOEs. Up until 1995, economic reform in urban China failed to improve SOEs’ competitiveness and profitability. Most of the SOEs, especially the small ones, continued to lose money, which intensified the financial risks of ownership and jeopardized the country’s economic growth (Wu 2005). As a result, central and local governments were eager to get rid of these money-losing SOEs. Small and medium-sized SOEs were generally controlled by local governments (county and city governments), while larger ones were controlled by central governments. The 15th Communist Party Congress adopted an important policy that guided reform during this period: ”grasping the large, and letting the small go.” Under this guidance, SOE reform was carried out by different levels of government. A milestone was the promulgation of the Company Law in 1994, which provided a legal framework to diversify ownership SOEs, with many large SOEs gradually converting into corporations.

In contrast, small and medium-sized SOEs were transformed into POEs, often by selling these firms to their employees or outside investors. The process of privatizing small SOEs initially started from some pioneering counties and then spread to the whole nation. Evidence shows that most small SOEs were privatized by the late 2000s (Bai et al. 2009; Cao et al. 1999).

During this period, both SOEs and POEs have become inseparable components of the Chinese economy, though SOEs still enjoy preferential treatment from the government, such as market monopoly power protected by the government and concessionary loans from State-owned banks.

3 Data and Key Variables

3.1 Primary Data

To measure the number of new POEs and incumbent SOEs, our primary data are drawn from the second economic census of China carried out by the National Bureau of Statistics of China (NBS) in 2008. This economic census covers all legal units in all sectors at the end of 2008.[7] Legal units (faren in Chinese) include corporation legal units (qiye faren), nonprofit public-service legal units (shiye faren), and other types of legal units.[8] Two points are noteworthy. First, this paper studies the entry of corporation legal units, which are equivalent to the standard concept of profit-making firms. Second, since many SOEs are nonprofit public-service legal units in services, this paper measures incumbent industrial conditions using all types of legal units. To keep notation simple, we continue to use the term firms in the following sections and firms refer to legal units.[9]

For each firm, the data provide a wide range of firm characteristics, including firm location, type of industry, the status of registration, total employees, year of entry, type of shareholding, etc. Our definitions of SOEs and POEs are based on both registration status and shareholding. Based on registration status, the types of firm ownership include SOEs, collective-owned firms, POEs, foreign-owned firms, and a range of mixed-ownership firms. Our definition of SOEs and POEs includes mixed-ownership firms that are controlled by the state or private sector, respectively.

New POEs are defined as those POEs created in the last twelve months by the end of 2008. We have two measures for new POEs: one is the number of total employees employed by the new POEs, and the other is the number of new POEs formed in 2008. Our main analysis, based upon past studies (e.g., Glaeser and Kerr 2009), focuses on the first measure, and we use the second measure in the robustness check section. The results from both measures are qualitatively similar.

Incumbent SOEs are defined as the SOEs established before 2008, so the stock of SOEs is predetermined prior to the formation of the new POEs in our analysis. We drop missing and miscoded data on firm location, total employees and the year of entry. One concern is that China's central government implements particular policies to limit the entry of private firms into specific industries. To address this concern, we do not consider industries with no entry of POEs at any city. Public management and social organization are also excluded. One caveat of this firm-level data is that all employees in a multi-location firm are assigned to its headquarters location, potentially leading to measurement error. Since the number of these multi-location firms accounts for only about five percent of the total sample and our key variables are aggregated industry employment in a city, we assume that this issue does not have a significant effect on our main results.

We conduct our analysis separately for manufacturing and services. Our

[7] According to NBS, China's economy is divided into three sectors. The primary sector consists of agriculture, forestry, animal, husbandry, and fishery. The secondary sector consists of mining, manufacturing, construction, and production and supply of electricity, gas, and water. The tertiary sector, i.e., the service sector, includes all other industries, not in the primary and secondary sectors.

[8] The NBS defines legal units as "legal unit refers to an economic unit meeting the following criteria:

a) established by law with its own name, internal organization, and locations, and capable of fulfilling independently its civil obligations; b) with independent ownership or rights (or authorized with rights) of using assets and bearing liabilities, with authority to sign contracts with other units; and c) with independent financial accounting, capable of compiling assets and liability tables."

[9] This paper uses the term firm and enterprise interchangeably.

sample consists of 287 cities—283 prefecture-level cities and 4 municipalities, 160 three-digit manufacturing industries, and 163 three-digit service industries. [10] In total, there are 45,920 and 46,781 city-industry pairs for manufacturing and services, respectively. While we focus on new POEs, our analysis covers a large percentage of all new firms. For employment in all firms created in the last twelve months, POEs account for about 70 percent of employment in manufacturing firms and 60 percent in service firms. Table 1 shows summary statistics of employment, agglomeration metrics, and chinitz metrics. The average number of workers in new POEs in a city-industry are 65 for manufacturing and 52 for services.

3.2 Proximity to Input Suppliers and Output Customers

To explain the spatial variation of new POE formation, our research tests the local industrial conditions based on Marshall's original agglomeration theories. Sections 3.2 and 3.3 test whether new firms tend to locate in places where more upstream and downstream firms and firms that employ similar workers are concentrated. Section 3.4 investigates another type of agglomeration: whether the entry of new firms is higher or lower in places where more firms that use similar technology are concentrated. We construct these metrics separately for all firms, SOEs, and POEs.

Marshall (1890) pioneers the analysis of the concentration of firms in particular locations and suggests three main advantages of agglomeration. First, firms benefit from the reduction of shipping costs by locating near input suppliers and output customers. We use China's 2007 Input-Output Table to capture the strength of input-output linkages. This input-output table classifies economic activities into 135 product sectors, each of which consists of one or several three-digit industries.[11] Let $Input_{i\leftarrow j}$ labels the share of industry $i^t s$ inputs that provided by industry j, and $Output_{i\rightarrow j}$ labels the share of industry $i^t s$ outputs that consumed by industry j. These shares range from zero (no dependence on inputs or outputs) to one (complete dependence).[12] Using these input or output shares as weights, we construct the weighted sums of incumbent employment across all industries as follows:

$$Input_{ic} = \sum_{j\neq i}\left(\frac{Emp_{jc}}{Emp_j} Input_{i\leftarrow j}\right),$$

and

$$Output_{ic} = \sum_{j\neq i}\left(\frac{Emp_{jc}}{Emp_j} Output_{i\rightarrow j}\right), \quad (1)$$

where Emp_j is the total incumbent employment (in all legal units) in industry j, and Emp_{jc} is the total incumbent employment (in all legal units) for industry j in city c. $Input_{ic}$ measures the strength to which all

[10] The prefecture-level city contains a city proper and surrounding rural areas.

[11] The 135 product sectors contain 5 primary sectors, 90 secondary sectors (including 81 manufacturing sectors), and 40 service sectors.

[12] These shares are measured using all intermediate inputs and outputs (including intermediate and final use) in the 2007 Input-Output Table of China.

incumbent firms provide the main inputs for industry i in city c, and $Output_{ic}$ measures the strength to which all incumbent firms are the main customers for industry i in city c. For manufacturing and services, the mean values of these two variables (multiplied by 100) are 0.272 and 0.317 respectively, with a larger value indicating that more relevant upstream and downstream firms have agglomerated.

Applying this same methodology, we construct $Input_{ic}^{SOE}$ and $Output_{ic}^{SOE}$ to measure the extent to which incumbent upstream and downstream SOEs are concentrated in city c:

$$Input_{ic}^{SOE} = \sum_{j \neq i} (\frac{Emp_{jc}^{SOE}}{Emp_j} Input_{i \leftarrow j}),$$

and

$$Output_{ic}^{SOE} = \sum_{j \neq i} (\frac{Emp_{jc}^{SOE}}{Emp_j} Output_{i \rightarrow j}), \qquad (1')$$

where Emp_{jc}^{SOE} is the incumbent employment in SOEs (including all types of legal units) for industry j in city c. Similarly, we construct $Input_{ic}^{POE}$ and $Output_{ic}^{POE}$ to measure the extent to which incumbent upstream and downstream POEs are concentrated in city c:

$$Input_{ic}^{POE} = \sum_{j \neq i} (\frac{Emp_{jc}^{POE}}{Emp_j} Input_{i \leftarrow j}),$$

and

$$Output_{ic}^{POE} = \sum_{j \neq i} (\frac{Emp_{jc}^{POE}}{Emp_j} Output_{i \rightarrow j}), \qquad (1'')$$

where Emp_{jc}^{POE} is the incumbent employment in POEs for industry j in city c.

3.3 Proximity to Firms that Use Similar Workers

The second advantage of agglomeration is that concentrations of firms form a thick labor market, which promotes efficient matches between employees and employers, and reduces the risks of workers being unemployed because they can find jobs more easily. This argument implies that proximity to suitable labor markets may increase the productivity levels of workers and firms and influence the location choices of new firms. Of course, firms that compete for similar workers may also have a negative effect. The net effect will depend on which one dominates. Following past studies (e.g., Glaeser and Kerr 2009; Jofre-Monseny et al. 2011), we look at the occupation similarity among industries as a proxy for labor similarity. We draw data from the 2005 1% population census to construct occupation similarity.[13] In total, the 2005 census classifies workers into 73 two-digit occupations and 95 two-digit industries.[14] The variable LS_{ij} measures the occupation similarity between

[13] The full sample of the one-percent 2005 population census is not available, but this paper as well as most of past studies employs a one-fifth random subsample which contains about 2.5 million population.

[14] Persons aged 15 and above are required to report their occupations and industries where they work. We proceed as follows: (1) drop missing and miscoded data on industry and occupation; (2) For each industry, we calculate the share of employment in each occupation.

industries *i* and *j*:

$$LS_{ij} = \frac{1}{\frac{1}{2}\sum_o |L_{io} - L_{jo}|},$$

where L_{io} is the share of industry i's employment that occupation o accounts for. This index LS_{ij} is the inverse of a dissimilarity index, which aggregates the absolute deviation in occupation composition between two industries. The index LS_{ij} is greater than one, and a higher value indicates higher levels of occupation similarity between two industries.

Using occupation similarity index as weights, the variable $Labor_{ic}$ measures the degree to which all incumbent firms employ similar workers as industry *i* in city *c*:

$$Labor_{ic} = \sum_{j \neq i}\left(\frac{Emp_{jc}}{Emp_j} LS_{ij}\right). \tag{2}$$

The mean value of this variable is 0.440 for manufacturing and 0.429 for services. A higher value indicates that all incumbent firms employ more similar workers. In the same manner as above, we construct $Labor_{ic}^{SOE}$ and $Labor_{ic}^{POE}$ to measure the extent to which incumbent SOEs or POEs that employ similar workers as industry *i* are concentrated in city *c*:

$$Labor_{ic}^{SOE} = \sum_{j \neq i}\left(\frac{Emp_{jc}^{SOE}}{Emp_j} LS_{ij}\right), \tag{2'}$$

and

$$Labor_{ic}^{POE} = \sum_{j \neq i}\left(\frac{Emp_{jc}^{POE}}{Emp_j} LS_{ij}\right). \tag{2''}$$

3.4 Proximity to Firms that use similar technology

We next consider another question: whether the emergence of new POEs is affected if more firms use similar technology. Firms in agglomeration could learn quickly from other firms; thereby sparking new ideas and innovations. Marshall (1890) described that "the mysteries of the trade become no mystery, but are, as it were, in the air." Following prior studies (e.g., Glaeser and Kerr 2009; Ellison et al 2010), we apply patent data to measure similarities in technologies between industries. Data are drawn from He et al. (2018) who match patent application data from China's State Intellectual Property Office (SIPO) to the annual survey of industrial firms (ASIF). We are allowed to access over 300,000 matched patent data from 1998 to 2009 with identified firm ID in ASIF.[15] We classify patent data into 121 types of 3-digit IPC codes. For example, code H04 in international patent classification (IPC) represents telecommunication technology. Note that the patent data are available for manufacturing only. For each 3-digit manufacturing industry, we calculate the total number of each type of patent. With this data in hand, the

[15] China's State Intellectual Property Office (SIPO) provides three types of patents: invention patent, utility patent, design patent. This paper uses invention patent only which reflects innovations and improvements in products, methods, and technology.

technology similarity index is calculated as follows:

$$TS_{ij} = \frac{1}{\frac{1}{2}\sum_w |P_{iw} - P_{jw}|},$$

where P_{iw} is the share of industry i's patents that type w account for. The number of TS_{ij} as LS_{ij} is also greater than one, with the larger value indicating more similar technology usage between the two industries. Overall, the mean value of this index is 1.121.

Using technology similarity index as weights, the variable $Tech_{ic}$ measures the degree to which all incumbent firms employ similar technology as industry i in the city c

$$Tech_{ic} = \sum_{j \neq i}\left(\frac{Emp_{jc}}{Emp_j} TS_{ij}\right). \quad (3)$$

The mean value of this variable is 0.691 for manufacturing. A higher value indicates that all incumbent firms employ more similar technology. In the same manner as above, we construct $Tech_{ic}^{SOE}$ and $Tech_{ic}^{POE}$ to measure the extent to which incumbent SOEs or POEs that employ similar workers as industry i are concentrated in city c:

$$Tech_{ic}^{SOE} = \sum_{j \neq i}\left(\frac{Emp_{jc}^{SOE}}{Emp_j} TS_{ij}\right), \quad (3')$$

and

$$Tech_{ic}^{POE} = \sum_{j \neq i}\left(\frac{Emp_{jc}^{POE}}{Emp_j} TS_{ij}\right). \quad (3'')$$

4 Empirical Model

The preceding section constructs a set of agglomeration metrics that measure local industrial conditions. To explore the impact of SOEs on the entry of new POEs, we estimate the following model:

$$\begin{aligned}\ln(Entry_{ic}) = {} & \alpha_0 + \alpha_e ln(Emp_{ic}) + \alpha_I Input_{ic} + \alpha_O Output_{ic} + \alpha_L Labor_{ic} \\ & + \alpha_T Tech_{ic} + \alpha_e^s ln\left(Emp_{ic}^{SOE}\right) + \alpha_I^s Input_{ic}^{SOE} + \alpha_O^s Output_{ic}^{SOE} \\ & + \alpha_L^s Labor_{ic}^{SOE} + \alpha_T^s Tech_{ic}^{SOE} + \lambda_i + \theta_c + \epsilon_{ic}, (4)\end{aligned}$$

where $\ln(Entry_{ic})$ is the log employment in new POEs for industry i in city c, $ln(Emp_{ic})$ is the log employment in all incumbent firms for industry i in city c, $ln\left(Emp_{ic}^{SOE}\right)$ is the log employment in incumbent SOEs for industry i in city c. A set of agglomeration metrics discussed in the preceding section is added to measure the extent to which all incumbent firms and all incumbent SOEs in city c that separately: (1) provide inputs to industry i ($Input_{ic}$ and $Input_{ic}^{SOE}$); (2) purchase outputs of industry i ($Output_{ic}$ and $Output_{ic}^{SOE}$); (3) hire similar employees as those hired by industry i ($Labor_{ic}$ and $Labor_{ic}^{SOE}$); (4) use similar technologies as those used by industry i ($Tech_{ic}$ and $Tech_{ic}^{SOE}$). λ_i and θ_c are industry and city fixed effects, respectively, and ϵ_{ic} is the error term. Since there are roughly two thirds of city-industry pairs without new private employment, we estimate Equation (4) using a Tobit model to account for the censoring of entry employment at zero. Standard errors are clustered at the city level.

Identifying a causal relationship between local industrial conditions and new firm entry is difficult. In the context of studying the whole manufacturing and services, the identification issue becomes more difficult. Glaeser and Kerr (2009) use predicted industrial distribution by natural cost advantages, and Ellison et al. (2010) instrument for U.S. industrial distribution using U.K industrial distribution. In the context of studying specific industries and cities, researchers explore exogenous shock to local industrial conditions. For example, Greenstone et al. (2010) utilize the natural experiment from "million-dollar plants".

There are two potential threats to bias the estimates in this paper. First, this paper may suffer from omitted variable bias. Controlling for industry and city fixed effects could not eliminate the bias from omitted variables that vary with industries and cities. For example, the negative correlation between the high concentration of SOEs and the low entry of private firms is due to governments' policy preferences that favor state-owned enterprises and discriminate against private firms in particular industries. Notice that the effect of incumbent SOEs on private firm entry is generally negative. If the omitted variables such as natural advantages and favorable place-based policies promote the clusters of both incumbent SOEs and new private firms, the negative coefficients are underestimated.

Second, reserve causality is another source of bias. Although using new firms in year t as dependent variables and incumbent firms in year t-1 may mitigate the reserve causality, high persistence of city-industry employment pattern may bias this time-series approach (Glaeser and Kerr 2009). Existing firms may relocate to places with predicted high firm entry. For example, if a special economic zone is planned to be constructed and attract many new firms in the coming years, existing input suppliers and output consumers may relocate to that location in advance. This scenario may underestimate the negative SOE effect and overestimate the positive POE effect. As a result, our findings reported throughout this paper are best interpreted as partial correlations rather than as causal effects. Thus, when interpret findings in our paper, it is important to bear above mentioned threats in mind.

5 Empirical Results

This section reports our main findings. We first discuss the general metrics constructed for all incumbent firms (Tables 2a and 2b) and then move on to discuss the metrics constructed for incumbent SOEs (Tables 3a and 3b) and POEs (Tables 4a and 4b).

5.1 Results of Overall Agglomeration Effects

Tables 2a and 2b analyze the general metrics constructed for all incumbent firms. Table 2a presents the results for manufacturing. Column 1 includes just the incumbent city-industry employment within the same industry, which is found to be positively correlated with private manufacturing entrants. The strong impact from the own industries on entry has been confirmed as an important stylized fact by past studies (Glaeser and Kerr 2009; Ghani et al. 2014). Columns 2-5 separately add the Marshallian agglomeration metrics of proximity to input suppliers, output customers, labor, and technology. The entry of new manufacturing POEs is found to be higher in cities where more upstream firms, downstream firms, and firms that use similar workers and technologies are concentrated. The last column includes all variables and the results remain unchanged except for proximity to labor.

Table 2b shows the results for services. This table is organized the same as Table 2a. As our manufacturing sector results reveal, existing firms within the same industries are important for POE entrants in the services sector. However, the entry of private service startups is not associated with local input suppliers, output customers, and incumbents that use similar workers.

5.2 Agglomeration Effects of the Incumbent SOEs

Tables 3a and 3b analyze the metrics constructed for incumbent SOEs. Table 3a shows the results for manufacturing firms. Since a location with particular advantages may attract more SOEs and POEs in the same industries, it is necessary to take into account overall industrial conditions. Therefore, the common controls in each specification contain the existing employment from own industries and the general metrics in Table 2a. To help interpret our results, imagine that there are two cities with the same amount of employment in each industry, but one city has more workers employed in SOEs. In this scenario, we are testing whether the differences in the employment landscape will have an impact on the entry of new POEs.

Column 1 adds the incumbent SOEs within the same industries. This coefficient is insignificant, implying that existing SOEs within the same industries have no impact on private manufacturing entry. Columns 2-5 add the Marshallian metrics for SOEs one by one. Overall, we find that proximity to downstream SOEs is not correlated with private manufacturing entry, whereas proximity to upstream SOEs does have a negative impact. We find a negative and significant coefficient on proximity to labor employed in SOEs. Proximity to SOEs that use similar technology appears to be insignificant. The last column combines all metrics, and most of the coefficients remain unchanged except that the coefficient on labor proximity loses significance.

Table 3b shows the results for services. The existing SOEs within the same industries are found to significantly reduce private service entry. The agglomeration effects of upstream and downstream SOEs are all negative and significant in each specification. For example, the coefficient on proximity to upstream SOEs in the last column is -0.412 and implies that the employment of new private service firms will decrease in cities where more local SOEs provide inputs. A one-standard-deviation increase in the agglomeration of upstream SOEs reduces roughly eight percent of employment in new private service firms. In unreported results from the McDonald and Moffitt decomposition, for this metric, the marginal effect on the probability of firm births is -0.036 for an average city-industry environment, while the marginal effect on employment conditional on firm births is -0.173. We also found that in services, private service entry is not associated with incumbent SOEs that hire similar workers. These SOE agglomeration metrics are correlated. To address multicollinearity, we report a *P* value of the F test on SOE agglomeration metrics. For this table, the *P* value is less than 1% and indicates that the SOE agglomeration metrics are jointly significant.[16]

5.3 Agglomeration Effects of the Incumbent POEs

Tables 4a and 4b highlight the incumbent POEs. Table 4a shows the results for manufacturing. Applying our earlier methodology, we first take into account the overall industrial conditions. With regard to this concern, the common controls in each specification contain the existing employment in all firms from own industries and the general metrics in Table 2a.

Column 1 adds the incumbent POEs within the same industries. This coefficient is positive and significant, which implies a positive agglomeration effect. Columns 2-5 separately add the Marshallian metrics for POEs and find that proximity to upstream POEs appears to facilitate private manufacturing entry, but proximity to the downstream POEs appears to have no effect. In contrast with SOEs, agglomeration externalities do not exist if the new and incumbent POEs employ similar workers. Interestingly, proximity to POEs that use similar technology significantly promote new private manufacturing firm entry. These results are unchanged if all metrics are added in the last column.

Table 4b shows the results for services. The existing POEs within the same industries are found to significantly increase private service entry. Results from the agglomeration metrics of POEs show that new private service entry is not correlated with the concentration of incumbent POEs that provide inputs, consume outputs, and use similar workers. In other words, agglomeration of POEs from other industries appears to not affect on the entry of new private service firm.

5.4 Robust Analysis

Tables 5 and 6 list our robustness checks. Table 5 focuses on the links between

[16] To assess the potential collinearity between these agglomeration metrics, we compute a set of collinearity diagnostic measures including VIF, tolerance, and R-squared separately for manufacturing and services samples. Overall, the worrisome variables appear to be the metrics of measuring the proximity to firms that use similar labor and technology. The corresponding VIF for these metrics is generally larger than 10. After dropping the labor metric of SOEs, the mean VIF of all agglomeration metrics is less than 10.

new POEs and incumbent SOEs. To save space, we suppress the estimates for overall industrial conditions. Panels A and B consider manufacturing and services, respectively. Column 1 gives our baseline estimates, which are taken from the last columns in tables 3a and 3b. One concern is that local governments may promote or protect particular industries. In columns 2 and 3, we use several strategies to deal with this issue. In column 2, we drop cities in minority regions (Xinjiang, Xizang, and Qinghai provinces) and Hainan province (a special economic zone). These places are cities in minority regions or special economic zones, which may implement particular policies toward local industries. In column 3, we drop the bottom ten percent of the manufacturing or service industries with the smallest employment ratio of new POEs to all incumbent firms. The entry of these less entrepreneurial industries may be more likely to be driven by local policies. Overall, our main findings are more or less robust and remain qualitatively similar.

Past studies have recognized that nonlinear models with fixed effects lead to biased and inconsistent estimates due to the incidental parameter problems when the length of the panel is fixed (Chamberlain 1984; Maddala 1987). Two alternative models are estimated to address this issue. First, in column 4, we estimate random effect Tobit models where unobserved industry effects are assumed to be random. The results for manufacturing are consistent with our main arguments that the agglomeration of SOEs decreases the entry of manufacturing POEs. However, the results for services are not robust: the coefficients on the agglomeration of upstream and downstream SOEs lose significance. Second, in Column 5, we present OLS estimates instead of Tobit estimates. The OLS estimates using censored data are also biased. Therefore, the magnitudes of the OLS and Tobit estimates may not be comparable. Nevertheless, using OLS model produces more significant coefficients, which reinforce the negative impact of SOE agglomeration. In the last column, we use the numbers of new POEs as an alternative dependent variable and find that six out of seven coefficients on SOE agglomeration metrics are negative and significant. Overall, none of the robust analyses show a positive and significant impact from SOE agglomeration.

Table 6 focuses on the links between new POEs and incumbent POEs. The table is organized as the same as Table 5. As before, the pattern of baseline results changes sightly. The results, using firms counts as a dependent variable, are encouraging and show the positive impact of incumbent POEs: the number of new service POEs is significantly higher in places where more upstream POEs, downstream POEs, and POEs that use similar workers are concentrated.

One concern is that the findings may just reflect a random distribution pattern of new and incumbent firms. To address this concern, we conduct a series of placebo permutation tests. More specifically, the dependent variable (i.e., new POE employment in a city-industry pair) is randomly distributed across cities and the independent variables remain unchanged. Using the randomly permuted dataset, the baseline Tobit model is re-estimated. Figures 1a-2b report the empirical density distribution of placebo estimates from permuting dependent variable 500 times. The vertical line in each graph indicates the true estimate from the original data, and the p-value is the fraction of placebo estimates which are equal to or larger in absolute value than the corresponding true estimates from the baseline Tobit model. Overall, regarding the coefficients found significant in preceding analyses, most of them are at the margin of the empirical density distribution. E.g. , figure 1b shows that the p-value of the coefficient on proximity to upstream SOEs is 0.032. One caveat is that roughly 60% percent of city-industry pairs are zero, which may reduce the variability of permuted datasets. E.g., proximity to POEs that use similar technology is found to have a highly significant impact , while figure 2a shows that this coefficient is at the margin but not at the extreme, with a p-value of 0.108.

6 Mechanism

This section further explores the underlying mechanism behind the negative linkage between new POEs and incumbent SOEs. Instead of discussing a single coefficient, we pay attention to explain why the environment of incumbent SOEs has an overall negative impact on the entry of POEs. We discuss three possible channels: firm size, finance and business environment, and the employee welfare differences between SOEs and POEs.

6.1 Does Firm Size Matter?

The negative (statistically significant or insignificant) correlation between new local POEs' entry and existing upstream and downstream SOEs is not too surprising. One possible explanation for this phenomenon is that large firms, like SOEs, may hinder new firm formation and local growth. Almost 50 years ago, Chinitz (1961) compared the industrial structure of New York and Pittsburgh and emphasized that new firms tend to emerge in places with many small diversified firms (like New York) relative to places with dominant large firms (like Pittsburgh), because large firms may have limited input-output linkage with local firms and then may not help cater to newcomers. SOEs tend to be large firms and may organize their production activities across cities. As a result, SOEs may not be willing to buy from and sell goods and services to local firms.

To test Chinitz's hypothesis, we construct the average size of upstream SOEs and POEs:

$$Chinitz_{ic}^{SOE} = \sum_{j \neq i} \left(\frac{Emp_{jc}^{SOE}}{Firm_{jc}^{SOE}} Input_{i \leftarrow j}\right),$$

and

$$Chinitz_{ic}^{POE} = \sum_{j \neq i} \left(\frac{Emp_{jc}^{POE}}{Firm_{jc}^{POE}} Input_{i \leftarrow j}\right),$$

where $Firm_{jc}^{SOE}$ and $Firm_{jc}^{POE}$ are the number of incumbent SOEs and POEs in industry j in city c, respectively; other notations follow our previous interpretations. These two Chinitz metrics calculate the weighted average firm size of upstream SOEs and POEs. A higher positive value indicates that the size of relevant firms is larger on average. Table 1 summarizes the chinitz metrics. For example, the weighted average size of upstream SOEs divided by 1,000 are 0.262 and 0.172 for manufacturing and services, respectively.

We re-estimate the baseline models by adding the chinitz metrics, and we examine whether the agglomeration of upstream SOEs becomes weaker after controlling for the size of SOEs. Table 7 presents the results for manufacturing in columns 1-4 and services in columns 5-8. The results in the odd columns are taken from the baseline results, while the results in the even columns add the chinitz metrics. We find that the coefficients on proximity to upstream SOEs change slightly when controlling for the size of upstream SOEs, indicating that the negative impact of upstream SOEs does not result from the chinitz effect. Overall, the size of SOEs has limited power to explain the negative agglomeration of upstream SOEs. Another interesting finding in column 4 shows that the coefficient on the average size of upstream POEs is negative and significant, indicating that smaller suppliers of POEs significantly improve the entry of POEs. This finding supports the existence of chinitz effect for POEs in manufacturing.

6.2 Does Finance and Business Environment Matter?

Another distinct difference between SOEs and POEs is that SOEs are controlled by the State. Such state ownership may affect the interaction between SOEs and POEs. First, SOEs may receive orders from the State to buy from or sell goods and services to particular markets. Moreover, SOEs may continue to have a soft budget and do not perform as profit-maximizing firms, Therefore, their motivation to connect with local firms may be weak.

Second, governments and government-owned banks, in particular, may implement favorable policies toward SOEs and discriminative policies toward POEs. For example, Brandt and Li (2003) show that in rural China local banks tend to discriminate against POEs, which leads to POEs being less likely to obtain loans from formal financial sectors. Haggard and Huang (2008) document that POEs suffer from unfair competition with SOEs for financial support and investment opportunities. Li et al. (2008) conduct interesting work and find that among POEs political connections with governments (e.g., the Communist Party membership of firm owners) seems to positively impact firm performance. Overall, these past studies suggest that POEs located in cities with more competitive SOEs may be less successful.[17]

We use the 2012 enterprise survey conducted by the World Bank to explore the effects of SOEs on the POEs' finance and business environment. The data

[17] In a broad sense, existing studies explore many factors related to the institutional environment such as local protectionism, local investment climate, economic liberalization, and opening to trade. In a study of China, Lu and Tao (2009) document that from 1998 to 2005, levels of manufacturing agglomeration are much lower than that in developed countries such as Finance and the United States. They argue that local protectionism obstructs the process of manufacturing agglomeration. In a study of India and Indonesia, Deichmann et al. (2008) emphasize that effective law and regulation appear to be more powerful policies to attract firms than direct fiscal incentives. In a study of interstate firm relocation within the U.S., Conroy et al. (2016) find that Business climate play a limited role in attracting firm relocation.

are carried out in China between December 2011 and February 2013 and are collected from 2,700 privately-owned and 148 state-owned firms.[18] The data contain a variety of questions on firms' views about finance and business environment. Appendix table A1 presents the summary statistics by ownership for key variables, city characteristics, and firm characteristics. The first panel reports five questions regarding finance. The first two questions are dummy variables for whether the firm has a loan and whether collateral is required, respectively. The third and fourth questions provide the value of the collateral and recent loans, and the last question is also a dummy variable with 1 indicating whether a recent loan application is approved. The values of finance variables are quite similar between POEs and SOEs. E.g., the share of firms with a loan is 31.3% for POEs and 30.6% for SOEs, respectively.

The following questions in table A1 are about the business environment. The respondents describe to what degree Political Instability or Corruption or Court is an obstacle to the current operations of this establishment. The enterprise data provide five response options from no obstacle to very severe obstacle. By contrast, it appears that the business environment is more favorable for SOEs than POEs. E.g., 78.24% of POEs report that political instability is not an obstacle for business operation, whereas this number is 84.62% for SOEs. Concerning corruption, the share of firms reporting no obstacle is 75.12% for POEs and 80.42% for SOEs, respectively. The share of SOE employment in a city is measured using the 2008 economic census, and other city-level data are drawn from the 2008 *China City Statistical Yearbook.*

To conduct our analysis, the following model is estimated:

$$y_{ijc} = \beta_0 + \beta_1 Share_c + \beta_2 City_c + \beta_3 Firm_i + \varphi_i + v_{ijc},$$

where y_{ijc} represents the outcomes (e.g., have a loan) for firm i in industry j and city c, $Share_c$ is the share of city c's employment that SOEs account for, $City_c$ and $Firm_i$ represents city and firm characteristics, respectively, φ_i controls for industry fixed effect, and v_{ijc} is the error term. β_1 is the coefficient of interest, indicating whether the finance and business environment of POEs and SOEs are worse in cities with more SOEs. Table 8 reports the results for finance outcome. Columns (1)-(5) are estimated using logit, except that columns (3) are estimated using OLS. In cities with more SOEs, POEs are less likely to have a loan and more likely to require collateral. Besides, POEs' loan is less likely to be approved. In contrast, the likelihood of having a loan is unaffected for SOEs. Table 9 reports the results for business outcome using ordered logit. Overall, political instability, corruption, and court justice are more likely to be an obstacle to the operation of POEs in cities with more SOEs. In contrast. SOEs suffer less from the business environment. These results in this section show that the presence of more SOEs appears to affect the operation of POEs through the finance and business environment.

[18] World Bank. China Enterprise Survey (ES) 2012, Ref. CHN 2012 ES v01 M. Dataset downloaded from https://microdata.worldbank.org/index.php/catalog/1559 on 07/07/2019.

6.3 Does Employee Welfare Matter?

Another channel is that SOEs may compete with POEs for workers. SOEs may have several competitive advantages over POEs. SOEs may offer a stable job with economic and social benefits such as local *hukou*. Meanwhile, workers employed in SOEs have almost no risk of being unemployed. Compared with POEs, SOEs may pay higher wages for the same skilled workers in some industries. The wage differential between SOEs and POEs is demonstrated by previous studies (e.g., Chi et al. 2012). The disadvantage of POEs in the same labor market may hinder the formation of new POEs for two reasons. First, job-hopping between POEs and SOEs may be less frequent, which reduces the effects of a thick labor market. Second, workers in SOEs may have a weak motivation to exchange and learn cutting-edge knowledge, thereby reducing the technological spillovers between POEs and SOEs.

We estimate a seemingly unrelated regression model to examine the welfare differences between workers employed in SOEs and POEs. The data are drawn from the 2005 1% population census. We select full-time workers aged 18 to 55 employed in SOEs or POEs. Full-time workers are those who are not in school, worked at least 20 hours in the last week, and earned at least 500 renminbi in the last month. We focus on many welfare outcomes including monthly wage, job contract, *hukou* type, and insurance participation. The job contract is a dummy variable that equals 1 if the worker has a fixed-term contract and 0 if otherwise; *hukou* type is a dummy variable that equals 1 if the worker has a local *hukou* and 0 if otherwise.[19] Insurance participation includes unemployment, pension, and medical insurance. All are dummy variables that equal 1 if the worker participates in this kind of insurance and 0 if otherwise.

Welfare outcomes are regressed on age and dummy variables for firm ownership type, high school degree, college degree or higher, male, marital status, Han nationality, occupations, industries, and cities. Firm ownership type is a dummy variable that equals 1 for SOEs and 0 for POEs. Table 10 gives the results, while table A2 summarizes these variables. The coefficients on SOEs are highly significant and show that employees in SOEs tend to earn higher wages, obtain a fixed-term job contract, have a local *hukou*, and have a higher probability of participating in all kinds of insurances. Notice that the results for the local *hukou* may reflect employment discrimination because SOEs may tend to hire people with a local *hukou*.

7 Conclusion

This paper examines the impact of incumbent SOEs on the births of new POEs in China. Our findings indicate that significant overall agglomeration effects exist, and the effects vary by sector, ownership, and industrial relations. Overall, incumbent SOEs hinder the births of POEs. For manufacturing, the

[19] A worker does not have a local hukou if her or his hukou is not in their current residential county.

entry of new POEs is significantly lower in places where more upstream SOEs are concentrated. For services, the entry of new POEs is significantly lower in places where more upstream SOEs and downstream SOEs are concentrated. However, the agglomeration effects from the incumbent POEs on the formation of new POEs are insignificant or significantly positive. For manufacturing, the upstream POEs and POEs that use similar technology can facilitate the births of new POEs.

The implication of our topic to the literature on agglomeration economics is that firms in cities and industrial clusters may not contribute to the agglomeration effects on firm entries equally. In particular, the concentration of firms like SOEs may generate weak agglomeration effects but amplify the crowding out effects. Furthermore, whether geographic proximity of firms could encourage new entries depends on how firm owners make production decisions and how they exchange goods, people, and ideas. For example, firms that could make production decisions independently may lead to larger agglomeration benefits than firms that could not function independently. One possible related finding is that Henderson（2003）find that in the high-tech industry, corporate plants benefit more from single plants (or non-affiliate plants) than corporate plants. Meanwhile, fair financial and labor market institutions could enhance the agglomeration effects and mitigate the crowding out effect of incumbent firms on new firm formation.

References

Agarwal, Ruchir, Elif Arbatli-Saxegaard, Lesley Fisher and Xuehui Han. 2022. "India's State-Owned Enterprises", IMF Working Paper No. 22/165.

Bai, Chong-En, Jiangyong Lu, and Zhigang Tao. 2009. "How Does Privatization Work in China?" *Journal of Comparative Economics* 37 (3): 453–470.

Becker, Sascha O., Stephan Heblich, and Daniel M. Sturm. 2021. "The Impact of Public Employment: Evidence from Bonn." Journal of Urban Economics 122 : 103291.

Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten. 2018. "Barriers to Entry and Regional Economic Growth in China." *Working Paper*

Brandt, Loren and Hongbin Li. 2003. "Bank Discrimination In Transition Economies: Ideology, Information, or Incentives?" *Journal of comparative economics* 31 (3): 387– 413.

Cao, Yuanzheng, Yingyi Qian, and Barry R. Weingast. 1999. "From Federalism, Chinese Style to Privatization, Chinese Style." *Economics of Transition* 7 (1): 103–131.

Chamberlain, Gary. 1984. "Chapter 22 Panel data." In *Handbook of Econometrics*, vol. 2. Elsevier, 1247–1318.

Chatterji, Aaron, Edward Glaeser, and William Kerr. 2014. "Clusters of Entrepreneurship and Innovation." In *Innovation Policy and the Economy*, vol. 14, edited by Josh Lerner and Scott Stern. University of Chicago Press, 129–166.

Chi, Wei, Richard B. Freeman, and Hongbin Li. 2012. "Adjusting to Really Big Changes: The Labor Market in China, 1989-2009." Working Paper 17721, National Bureau of Economic Research.

Chinitz, Benjamin. 1961. "Contrasts in Agglomeration: New York and Pittsburgh."*American Economic Review* 51 (2): 279–289.

Combes, Pierre-Philippe and Laurent Gobillon. 2015. "The Empirics of Agglomeration Economies." In *Handbook of Urban and Regional Economics*, vol. 5, edited by Gilles Duranton, Vernon Henderson, and William Strange. North Holland, 247–348.

Conroy, Tessa, Steven Deller, and Alexandra Tsvetkova. 2016. "Regional Business Climate and Interstate Manufacturing Relocation Decisions." Regional Science and Urban Economics 60: 155-168.

Deichmann, Uwe, Somik V. Lall, Stephen J. Redding, and Anthony J. Venables. 2008. "Industrial Location in Developing Countries." The World Bank Research Observer 23(2): 219-246.

Diodato, Dario, Frank Neffke, and Neave O'Clery. 2018. "Why do Industries Coagglomerate? How Marshallian Externalities Differ by Industry and Have Evolved Over Time." *Journal of Urban Economics* 106: 1-26.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns." American Economic Review 100 (3): 1195-1213.

Faggio, Giulia. 2019. "Relocation of Public Sector Workers: Evaluating a Place-based

Policy.” Journal of Urban Economics 111: 53-75.
Faggio, Giulia, and Henry Overman. 2014. “The Effect of Public Sector Employment on Local Labour Markets.” Journal of Urban Economics 79: 91-107.
Faggio, Giulia, Olmo Silva, and William C. Strange. 2017. “Heterogeneous Agglomeration.” *The Review of Economics and Statistics* 99 (1): 80-94.
Faggio, Giulia, Olmo Silva, and William C. Strange. 2020. “Tales of the City: What do Agglomeration Cases Tell Us about Agglomeration in General?” *Journal of Economic Geography* 20 (5): 1117-1143.
Ghani, Ejaz, William R. Kerr, and Stephen O’Connell. 2013. “Local Industrial Structures and Female Entrepreneurship in India.” *Journal of Economic Geography* 13 (6): 929– 964.
Ghani, Ejaz, William R. Kerr, and Stephen O’Connell. 2014. “Spatial Determinants of Entrepreneurship in India.” *Regional Studies* 48 (6): 1071–1089.
Glaeser, Edward L., Hedi D. Kallal, Jose Scheinkman, and Andrei Shleifer. 1992. “Growth in Cities.” *Journal of Political Economy* 100 (6): 1126–1152.
Glaeser, Edward L. and William R. Kerr. 2009. “Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution can We Explain?” *Journal of Economics & Management Strategy* 18 (3): 623–663.
Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. “Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings” *Journal of Political Economy* 118(3): 536-598
Guo, Di, Kun Jiang, Byung-Yeon Kim, and Chenggang Xu. 2014. “Political Economy of Private Firms in China.” *Journal of Comparative Economics* 42 (2): 286–303.
Haggard, Stephan and Yasheng Huang. 2008. “The Political Economy of Private Sector Development in China.” In *China’s Great Economic Transformation*, edited by Loren Brandt and Thomas G. Rawski. New York: Cambridge University Press.
He, Zi-Lin, Tony W. Tong, Yuchen Zhang, and Wenlong He. 2018. “A Database Linking Chinese Patents to China’s Census Firms.” *Nature Scientific Data* 5: 180042.
Henderson, Vernon. 2003. “Marshall’s Scale Economies.” *Journal of Urban Economics* 53 (1): 1–28.
Henderson, Vernon, Ari Kuncoro, and Matt Turner. 1995. “Industrial Development in Cities.” *Journal of Political Economy* 103 (5): 1067–1090.
Hu, Angang. 2014. *The Modernization of China’s National Governance*. Beijing: China Renmin University Press.
Li, Hongbin, Lingsheng Meng, Qian Wang, and Li-An Zhou. 2008. “Political connections, financing and firm performance: Evidence from Chinese private firms.” *Journal of Development Economics* 87 (2): 283–299.
Lu, Jiangyong, and Zhigang Tao. 2009. “Trends and Determinants of China's Industrial Agglomeration.” Journal of Urban Economics 65(2): 167-180.
Jofre-Monseny, Jordi, Raquel Marin-Lopez, and Elisabet Viladecans-Marsal. 2011. “The Mechanisms of Agglomeration: Evidence from the Effect of Inter-industry Relations on the Location of New Firms.” *Journal of Urban Economics* 70 (2): 61–74.
Jofre-Monseny, Jordi, José I. Silva, and Javier Vázquez-Grenno. 2020. “Local Labor

Market Effects of Public Employment." Regional Science and Urban Economics 82: 103406.
Maddala, Gangadharrao S. 1987. "Limited Dependent Variable Models Using Panel Data." *Journal of Human Resources* 22 (3): 307–338.
Marshall, Alfred. 1890. *Principles of Economics*. London: Macmillan.
OCED, 2017. "The Size and Sectoral Distribution of State-Owned Enterprises", OECD Publishing, Paris.
O'Sullivan, Arthur, and William C. Strange. 2018. "The Emergence of Coagglomeration." Journal of Economic Geography 18 (2): 293-317.
Puga, Diego. 2010. "The Magnitude and Causes of Agglomeration Economies." *Journal of Regional Science* 50 (1): 203–219.
Quan, Zhezhu. 2008. "Thirty Years of Private Sector Development: History and Future Prospect." Government report, All-China Federation of Industry and Commerce.
Rosenthal, Stuart S. and William C. Strange. 2003. "Geography, Industrial Organization, and Agglomeration." *Review of Economics and Statistics* 85 (2): 377–393.
Rosenthal, Stuart S. and William C. Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies." In *Handbook of Regional and Urban Economics,* vol. 4, edited by Vernon Henderson and Jacques-Francois Thisse. Amsterdam: North-Holland, 2119–2171.
Rosenthal, Stuart S. and William C. Strange. 2008. "The Attenuation of Human Capital Spillovers." *Journal of Urban Economics* 64 (2): 373–389.
Rosenthal, Stuart S. and William C. Strange. 2010. "Small Establishments/Big Effects: Agglomeration, Industrial Organization and Entrepreneurship." In *Agglomeration Economics*, edited by Edward L. Glaeser. The University of Chicago Press, 277–302.
Wu, Jinglian. 2005. *Understanding and Interpreting Chinese Economic Reform*. Texere.
Zheng, Liang and Zhong Zhao. 2017. "What Drives Spatial Clusters of Entrepreneurship in China? Evidence from Economic Census Data." *China Economic Review* 46:229-248

Table 1: Summary Statistics

Variables	Manufacturing Mean	Manufacturing Std. Dev.	Services Mean	Services Std. Dev
Employment at the city-industry level:				
Employment in new POEs	65	276	52	278
Employment in all incumbent firms	2,108	9,371	1,753	7,049
Employment in incumbent SOEs	208	1,666	786	3,711
Employment in incumbent POEs	1,249	5,107	608	3,215
Agglomeration metrics:				
Proximity to all upstream firms (× 100)	0.272	0.478	0.317	0.592
Proximity to all downstream firms (× 100)	0.174	0.360	0.158	0.277
Proximity to all firms that use similar workers	0.440	0.636	0.429	0.637
Proximity to all firms that use similar technology	0.691	1.076		
Proximity to upstream SOEs (× 100)	0.079	0.251	0.099	0.212
Proximity to downstream SOEs (× 100)	0.045	0.147	0.053	0.098
Proximity to SOEs that use similar workers	0.163	0.251	0.170	0.268
Proximity to SOEs that use similar technology	0.075	0.124		
Proximity to upstream POEs (× 100)	0.132	0.216	0.135	0.221
Proximity to downstream POEs (× 100)	0.084	0.167	0.068	0.121
Proximity to POEs that use similar workers	0.177	0.239	0.165	0.226
Proximity to POEs that need similar technology	0.412	0.541		
Chinitz metrics:				
Average size of upstream SOEs (/ 1000)	0.262	0.728	0.172	0.269
Average size of upstream POEs (/ 1000)	0.036	0.045	0.030	0.017

Notes: For the metrics of proximity to firms that are suppliers or customers, they are calculated from the 2007 input-output table and the 2008 economic census. For the metrics of proximity to firms that use similar types of labor, they are calculated from the 2005 1% population census and the 2008 economic census. For the metrics of proximity to firms that use similar technology, they are calculated from the He et al. (2018) and the 2008 economic census. Other variables are calculated from the 2008 economic census.

Table 2a: The Effects of All Incumbent Firms on the Births of New Manufacturing POEs

	(1)	(2)	(3)	(4)	(5)	(6)
	DV: ln(Employment in new manufacturing POEs)					
ln(Employment in all incumbent firms)	0.747***	0.729***	0.745***	0.744***	0.738***	0.723***
	(0.015)	(0.014)	(0.015)	(0.015)	(0.014)	(0.014)
Proximity to all upstream firms		0.409***				0.344***
		(0.053)				(0.050)
Proximity to all downstream firms			0.193***			0.158***
			(0.053)			(0.054)
Proximity to all firms that use similar workers				1.224***		0.362
				(0.377)		(0.299)
Proximity to all firms that use similar technology					1.312***	0.908***
					(0.199)	(0.192)
Constant	-2.477***	-3.027***	-2.710***	-9.462***	-6.357***	-7.878***
	(0.206)	(0.208)	(0.214)	(2.153)	(0.635)	(1.780)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Log likelihood	-55130.004	-55078.065	-55123.109	-55117.647	-55082.352	-55043.928
Pseudo R^2	0.238	0.239	0.239	0.239	0.239	0.239
Censored observations	27009	27009	27009	27009	27009	27009
Observations	45920	45920	45920	45920	45920	45920

Notes: For the metrics of proximity to firms that are suppliers or customers, they are calculated from the 2007 input-output table and the 2008 economic census. For the metrics of proximity to firms that use similar types of labor, they are calculated from the 2005 1% population census and the 2008 economic census. For the metrics of proximity to firms that use similar technology, they are calculated from the He et al. (2018) and the 2008 economic census. Other variables are calculated from the 2008 economic census. The dependent variable is the log employment in new manufacturing POEs at the city-industry level. Estimations use Tobit models. Standard errors clustered by city are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$.

Table 2b: The Effects of All Incumbent Firms on the Births of New Service POEs

	(1)	(2)	(3)	(4)	(5)
	DV: ln(Employment in new service POEs)				
ln(Employment in all incumbent firms)	0.540***	0.540***	0.541***	0.541***	0.541***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Proximity to all upstream firms		0.001			0.009
		(0.045)			(0.046)
Proximity to all downstream firms			-0.062		-0.037
			(0.069)		(0.062)
Proximity to all firms that use similar workers				-0.203	-0.180
				(0.126)	(0.120)
Constant	-2.834***	-2.840***	-2.730***	-1.595*	-1.712**
	(0.433)	(0.465)	(0.444)	(0.881)	(0.857)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes
Log likelihood	-45428.755	-45428.754	-45428.358	-45427.739	-45425.518
Pseudo R^2	0.335	0.335	0.335	0.335	0.335
Censored observations	29092	29092	29092	29092	29092
Observations	46781	46781	46781	46781	46781

Notes: For the metrics of proximity to firms that are suppliers or customers, they are calculated from the 2007 input-output table and the 2008 economic census. For the metrics of proximity to firms that use similar types of labor, they are calculated from the 2005 1% population census and the 2008 economic census. Other variables are calculated from the 2008 economic census. The dependent variable is the log employment in new service POEs at the city-industry level. Estimations use Tobit models. Standard errors clustered by city are reported in parentheses.

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < .01$.

Table 3a: The Effects of Incumbent SOEs on the Births of New Manufacturing POEs

	(1)	(2)	(3)	(4)	(5)	(6)
	DV: ln(Employment in new manufacturing POEs)					
ln(Employment in all incumbent firms)	0.721***	0.719***	0.721***	0.721***	0.721***	0.719***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Proximity to all upstream firms	0.344***	0.479***	0.346***	0.345***	0.347***	0.480***
	(0.050)	(0.087)	(0.050)	(0.050)	(0.050)	(0.084)
Proximity to all downstream firms	0.157***	0.175***	0.180***	0.165***	0.157***	0.213***
	(0.054)	(0.055)	(0.059)	(0.053)	(0.054)	(0.057)
Proximity to all firms that use similar workers	0.356	0.344	0.374	0.932***	0.380	0.853***
	(0.299)	(0.304)	(0.300)	(0.327)	(0.299)	(0.318)
Proximity to all firms that use similar technology	0.907***	0.825***	0.893***	0.862***	0.969***	0.790***
	(0.192)	(0.179)	(0.193)	(0.187)	(0.205)	(0.189)
ln(Employment in incumbent SOEs)	0.006	0.006	0.006	0.006	0.005	0.007
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Proximity to upstream SOEs		-0.281**				-0.272**
		(0.119)				(0.117)
Proximity to downstream SOEs			-0.074			-0.098
			(0.131)			(0.125)
Proximity to SOEs that use similar workers				-1.962**		-1.632
				(0.960)		(1.005)
Proximity to SOEs that use similar technology					-0.876	-0.245
					(0.822)	(0.868)
Constant	-7.859***	-7.502***	-7.915***	-5.429**	-7.420***	-5.442**
	(1.780)	(1.817)	(1.776)	(2.125)	(1.764)	(2.160)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
P value of F test						0.097
Log likelihood	-55043.661	-55038.727	-55043.497	-55041.750	-55043.059	-55036.920
Pseudo R^2	0.239	0.239	0.239	0.239	0.239	0.240
Censored observations	27009	27009	27009	27009	27009	27009
Observations	45920	45920	45920	45920	45920	45920

Notes: See the notes under table 2a

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < .01$.

Table 3b: The Effects of Incumbent SOEs on the Births of New Service POEs

	(1)	(2)	(3)	(4)	(5)
	DV: ln(Employment in new service POEs)				
ln(Employment in all incumbent firms)	0.610***	0.608***	0.610***	0.610***	0.608***
	(0.017)	(0.016)	(0.016)	(0.017)	(0.016)
Proximity to all upstream firms	0.011	0.131***	0.007	0.001	0.112***
	(0.043)	(0.035)	(0.042)	(0.044)	(0.040)
Proximity to all downstream firms	-0.013	-0.029	0.176*	-0.025	0.110
	(0.061)	(0.064)	(0.102)	(0.063)	(0.107)
Proximity to all firms that use similar workers	-0.172	-0.204*	-0.198	0.219	-0.080
	(0.118)	(0.116)	(0.121)	(0.365)	(0.371)
ln(Employment in incumbent SOEs)	-0.099***	-0.097***	-0.099***	-0.099***	-0.098***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Proximity to upstream SOEs		-0.455***			-0.412***
		(0.138)			(0.147)
Proximity to downstream SOEs			-0.665**		-0.499*
			(0.263)		(0.278)
Proximity to SOEs that use similar workers				-0.723	-0.259
				(0.612)	(0.626)
Constant	-1.644*	-1.041	-1.340	-1.729**	-0.900
	(0.847)	(0.888)	(0.853)	(0.851)	(0.902)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes
P value of F test					0.007
Log likelihood	-45364.543	-45357.313	-45361.375	-45364.030	-45355.501
Pseudo R^2	0.336	0.336	0.336	0.336	0.336
Censored observations	29092.000	29092.000	29092.000	29092.000	29092.000
Observations	46781	46781	46781	46781	46781

Notes: See the notes under table 2b
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < .01$.

Table 4a: The Effects of Incumbent POEs on the Births of New Manufacturing POEs

	(1)	(2)	(3)	(4)	(5)	(6)
	DV: ln(Employment in new manufacturing POEs)					
ln(Employment in all incumbent firms)	0.350***	0.350***	0.350***	0.350***	0.349***	0.349***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Proximity to all upstream firms	0.325***	0.217***	0.325***	0.324***	0.338***	0.242***
	(0.045)	(0.054)	(0.045)	(0.045)	(0.046)	(0.055)
Proximity to all downstream firms	0.141***	0.140***	0.160**	0.143***	0.140***	0.184***
	(0.052)	(0.053)	(0.067)	(0.051)	(0.052)	(0.068)
Proximity to all firms that use similar workers	0.343	0.307	0.338	-0.259	0.320	0.145
	(0.282)	(0.287)	(0.282)	(0.522)	(0.274)	(0.539)
Proximity to all firms that use similar technology	0.820***	0.791***	0.821***	0.803***	0.100	0.163
	(0.182)	(0.176)	(0.183)	(0.176)	(0.284)	(0.257)
ln(Employment in incumbent POEs)	0.421***	0.419***	0.421***	0.421***	0.421***	0.420***
	(0.024)	(0.023)	(0.024)	(0.023)	(0.023)	(0.023)
Proximity to upstream POEs		0.432***				0.377**
		(0.167)				(0.165)
Proximity to downstream POEs			-0.061			-0.143
			(0.149)			(0.147)
Proximity to POEs that use similar workers				1.629		0.370
				(1.152)		(1.144)
Proximity to POEs that use similar technology					1.564***	1.372***
					(0.497)	(0.459)
Constant	-7.562***	-7.202***	-7.542***	-6.381***	-6.933***	-6.380***
	(1.687)	(1.700)	(1.692)	(1.799)	(1.619)	(1.908)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
P value of F test						0.000
Log likelihood	-54799.601	-54793.307	-54799.532	-54798.526	-54794.731	-54789.431
Pseudo R^2	0.243	0.243	0.243	0.243	0.243	0.243
Censored observations	27009	27009	27009	27009	27009	27009
Observations	45920	45920	45920	45920	45920	45920

Notes: See the notes under table 2a

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < .01$.

Table 4b: The Effects of Incumbent POEs on the Births of New Service POEs

	(1)	(2)	(3)	(4)	(5)
	DV: ln(Employment in new service POEs)				
ln(Employment in all incumbent firms)	0.235***	0.235***	0.235***	0.235***	0.236***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Proximity to all upstream firms	0.003	-0.043	0.010	0.003	-0.035
	(0.040)	(0.060)	(0.041)	(0.041)	(0.057)
Proximity to all downstream firms	-0.004	0.003	-0.118	-0.003	-0.108
	(0.054)	(0.056)	(0.113)	(0.054)	(0.110)
Proximity to all firms that use similar workers	-0.311***	-0.311***	-0.316***	-0.286*	-0.258
	(0.101)	(0.100)	(0.101)	(0.158)	(0.158)
ln(Employment in incumbent POEs)	0.409***	0.409***	0.409***	0.409***	0.408***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Proximity to upstream POEs		0.195			0.196
		(0.215)			(0.214)
Proximity to downstream POEs			0.300		0.298
			(0.282)		(0.276)
Proximity to POEs that use similar workers				-0.110	-0.251
				(0.691)	(0.683)
Constant	-0.774	-0.813	-0.744	-0.768	-0.770
	(0.739)	(0.742)	(0.739)	(0.751)	(0.751)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes
P value of F test					0.662
Log likelihood	-44814.016	-44818.087	-44817.492	-44817.559	-44818.080
Pseudo R^2	0.344	0.344	0.344	0.344	0.344
Censored observations	29092	29092	29092	29092	29092
Observations	46781	46781	46781	46781	46781

Notes: See the notes under table 2b

* p < 0.10, ** p < 0.05, *** p < .01.

Table 5: Robust Estimation for incumbent SOEs

	Base estimation	Drop particular cities	Drop low entry industries	RE Tobit	OLS regression	Use firm counts as DV
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Manufacturing						
ln(Employment in all incumbent firms)	0.719***	0.721***	0.730***	0.908***	0.248***	0.279***
	(0.015)	(0.015)	(0.015)	(0.009)	(0.008)	(0.010)
ln(Employment in incumbent SOEs)	0.007	0.007	0.006	-0.031***	0.062***	0.015***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.006)	(0.003)
Proximity to upstream SOEs	-0.272**	-0.269**	-0.276**	-0.464***	-0.259***	-0.150***
	(0.117)	(0.119)	(0.112)	(0.090)	(0.081)	(0.054)
Proximity to downstream SOEs	-0.098	-0.093	-0.123	-0.242*	-0.397***	-0.171***
	(0.125)	(0.126)	(0.126)	(0.135)	(0.076)	(0.049)
Proximity to SOEs that use similar workers	-1.632	-1.664*	-1.964*	0.443	-1.274*	-1.305***
	(1.005)	(1.001)	(1.007)	(0.301)	(0.725)	(0.473)
Proximity to SOEs that use similar technology	-0.245	-0.334	-0.443	0.375	-0.906	-0.655**
	(0.868)	(0.853)	(0.890)	(0.235)	(0.557)	(0.320)
General Metrics Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes		Yes	Yes
City fixed effects	Yes	Yes	Yes		Yes	Yes
Log likelihood	-55036.920	-54472.223	-53014.046	-57109.231		-30652.840
Pseudo R^2	0.240	0.238	0.227			0.411
Adjusted R^2					0.535	
Censored observations	27009	26206	22873	27009		27009
Observations	45920	44960	41328	45920	45920	45920
Panel B: Services						
ln(Employment in all incumbent firms)	0.608***	0.607***	0.618***	0.851***	0.228***	0.321***
	(0.016)	(0.017)	(0.017)	(0.012)	(0.008)	(0.011)
ln(Employment in incumbent SOEs)	-0.098***	-0.099***	-0.096***	-0.093***	-0.038***	-0.030***
	(0.009)	(0.010)	(0.009)	(0.009)	(0.007)	(0.005)
Proximity to upstream SOEs	-0.412***	-0.441***	-0.382***	0.044	-0.232***	-0.070
	(0.147)	(0.151)	(0.138)	(0.117)	(0.081)	(0.085)
Proximity to downstream SOEs	-0.499*	-0.482*	-0.390	-0.210	-0.939***	-0.791***
	(0.278)	(0.281)	(0.273)	(0.272)	(0.230)	(0.177)
Proximity to SOEs that use similar workers	-0.259	-0.266	-0.415	-0.748***	-1.801***	-2.046***
	(0.626)	(0.636)	(0.632)	(0.156)	(0.600)	(0.557)
General Metrics Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes		Yes	Yes
City fixed effects	Yes	Yes	Yes		Yes	Yes
Log likelihood	-45355.501	-44433.658	-44925.024	-47548.770		-27840.517
Pseudo R^2	0.336	0.336	0.316			0.495
Adjusted R^2					0.670	
Censored observations	29092	28483	24569	29092		29092
Observations	46781	45803	42189	46781	46781	46781

Notes: Panels A and B consider manufacturing and services, respectively. For each panel, Column 1 is taken from Columns 4 in Tables 3a and 3b. Column 2 drops cities in Xinjiang, Xizang, Qinghai, and Hainan provinces. Column 3 drops city-industry pairs with the low entry of new POEs. Column 4 estimates random effect Tobit models where unobserved industry effects are assumed to be random. Column 5 reports OLS estimates. Column 6 uses firm counts as alternative dependent variables. Standard errors clustered by city are reported in parentheses except for random effect Tobit models. Data come from the 2008 economic census, the 2007 Input-Output Table, the 2005 1% population census, and He et al. (2018).

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$.

Table 6: Robust Estimation for incumbent POEs

	Base estimation	Drop particular cities	Drop low entry industries	RE Tobit	OLS regression	Use firm counts as DV
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Manufacturing						
ln(Employment in all incumbent firms)	0.349***	0.353***	0.356***	0.363***	0.056***	0.120***
	(0.024)	(0.024)	(0.024)	(0.020)	(0.008)	(0.009)
ln(Employment in incumbent POEs)	0.420***	0.416***	0.423***	0.546***	0.247***	0.185***
	(0.023)	(0.024)	(0.025)	(0.020)	(0.010)	(0.010)
Proximity to upstream POEs	0.377**	0.384**	0.356**	0.406***	0.419***	0.279***
	(0.165)	(0.167)	(0.166)	(0.124)	(0.135)	(0.082)
Proximity to downstream POEs	-0.143	-0.147	-0.108	0.086	0.247*	-0.002
	(0.147)	(0.148)	(0.149)	(0.170)	(0.126)	(0.065)
Proximity to POEs that use similar workers	0.370	0.439	0.445	0.854**	-0.808	0.099
	(1.144)	(1.146)	(1.151)	(0.366)	(0.888)	(0.601)
Proximity to POEs that use similar technology	1.372***	1.366***	1.538***	0.424***	2.389***	1.193***
	(0.459)	(0.460)	(0.485)	(0.116)	(0.405)	(0.233)
General Metrics Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes		Yes	Yes
City fixed effects	Yes	Yes	Yes		Yes	Yes
Log likelihood	-54789.431	-54234.492	-52776.591	-56510.717		-30276.667
Pseudo R^2	0.243	0.241	0.230			0.418
Adjusted R^2					0.543	
Censored observations	27009	26206	22873	27009		27009
Observations	45920	44960	41328	45920	45920	45920
Panel B: Services						
ln(Employment in all incumbent firms)	0.236***	0.233***	0.244***	0.326***	0.062***	0.143***
	(0.017)	(0.017)	(0.017)	(0.014)	(0.005)	(0.010)
ln(Employment in incumbent POEs)	0.408***	0.410***	0.406***	0.503***	0.255***	0.203***
	(0.015)	(0.015)	(0.015)	(0.012)	(0.009)	(0.009)
Proximity to upstream POEs	0.196	0.215	0.209	0.099	0.158	0.193*
	(0.214)	(0.218)	(0.200)	(0.175)	(0.153)	(0.116)
Proximity to downstream POEs	0.298	0.282	0.232	0.623**	1.161***	0.424**
	(0.276)	(0.276)	(0.277)	(0.290)	(0.183)	(0.167)
Proximity to POEs that use similar workers	-0.251	-0.285	-0.045	1.098***	2.663***	2.601***
	(0.683)	(0.693)	(0.697)	(0.227)	(0.718)	(0.763)
General Metrics Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes		Yes	Yes
City fixed effects	Yes	Yes	Yes		Yes	Yes
Log likelihood	-44816.945	-43907.805	-44406.497	-46592.999		-27251.765
Pseudo R^2	0.344	0.344	0.324			0.505
Adjusted R^2					0.690	
Censored observations	29092	28483	24569	29092		29092
Observations	46781	45803	42189	46781	46781	46781

Notes: See the notes under table 5

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < .01$.

Table 7: Firm Size and Overall Agglomeration Effects

	Manufacturing				Services			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	DV: ln(Employment in new new manufacturing POEs)				DV: ln(Employment in service POEs)			
ln(Employment in all incumbent firms)	0.719***	0.718***	0.349***	0.350***	0.608***	0.608***	0.236***	0.236***
	(0.015)	(0.015)	(0.024)	(0.024)	(0.016)	(0.016)	(0.017)	(0.017)
ln(Employment in incumbent SOEs)	0.007	0.007			-0.098***	-0.098***		
	(0.008)	(0.008)			(0.009)	(0.009)		
Proximity to upstream SOEs	-0.272**	-0.335**			-0.412***	-0.480***		
	(0.117)	(0.131)			(0.147)	(0.161)		
Proximity to downstream SOEs	-0.098	-0.098			-0.499*	-0.498*		
	(0.125)	(0.125)			(0.278)	(0.277)		
Proximity to SOEs s that use similar worker	-1.632	-1.668*			-0.259	-0.281		
	(1.005)	(1.011)			(0.626)	(0.628)		
Proximity to SOEs that use similar technology	-0.245	-0.339						
	(0.868)	(0.863)						
Average size of upstream SOEs		0.045				0.064		
		(0.039)				(0.080)		
ln(Employment in incumbent POEs)			0.420***	0.418***			0.408***	0.408***
			(0.023)	(0.023)			(0.015)	(0.015)
Proximity to upstream POEs			0.377**	0.422**			0.196	0.220
			(0.165)	(0.164)			(0.214)	(0.219)
Proximity to downstream POEs			-0.143	-0.140			0.298	0.289
			(0.147)	(0.148)			(0.276)	(0.276)
Proximity to POEs that use similar workers			0.370	0.311			-0.251	-0.259
			(1.144)	(1.144)			(0.683)	(0.683)
Proximity to POEs that use similar technology			1.372***	1.371***				
			(0.459)	(0.458)				
Average size of upstream POEs				-0.924**				-1.110
				(0.421)				(1.348)
Constant	-4.267***	-4.261***	-5.138***	-5.106***	-3.960***	-3.981***	-3.564***	-3.526***
	(0.450)	(0.448)	(0.449)	(0.448)	(0.442)	(0.442)	(0.427)	(0.427)
General Metrics Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log likelihood	-55036.920	-55035.385	-54789.431	-54787.037	-45355.501	-45355.158	-44816.945	-44816.487
Pseudo R^2	0.240	0.240	0.243	0.243	0.336	0.336	0.344	0.344
Censored observations	27009	27009	27009	27009	29092	29092	29092	29092
Observations	45920	45920	45920	45920	46781	46781	46781	46781

Notes: The construction of independent variables is described in the text. Columns 1-4 consider the impact for manufacturing, while columns 5-8 consider services. The dependent variables are the log employment in new POEs at the city-industry level. Estimations use Tobit models. Standard errors clustered by city are reported in parentheses. Data come from the 2008 economic census, the 2007 Input-Output Table, the 2005 1% population census, and He et al. (2018).

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$.

Table 8: The Impact of SOEs on the Finance of POEs

	POEs				SOEs
	Have a loan	Collateral required	Value of collateral	Loan approved	Have a loan
	(1)	(2)	(3)	(4)	(5)
Share of city	-4.928***	4.645**	0.265	-8.815***	-2.221
employment in SOEs	(0.560)	(1.955)	(0.309)	(2.303)	(2.321)
ln(Population)	0.633***	1.381**	0.088	1.744**	1.174
	(0.161)	(0.573)	(0.112)	(0.767)	(0.816)
ln(Per capita FDI)	-0.741***	-0.226	0.057	0.049	-0.400
	(0.088)	(0.193)	(0.055)	(0.385)	(0.307)
ln(Density)	-0.984***	2.289***	-0.041	-0.583	-0.303
	(0.190)	(0.495)	(0.129)	(0.953)	(0.751)
ln(Per capita GDP)	0.821***	0.938**	-0.070	-0.700	0.959
	(0.161)	(0.395)	(0.095)	(0.616)	(0.742)
Share of GDP	0.018*	0.023	-0.008	0.115*	-0.022
in services	(0.009)	(0.047)	(0.011)	(0.064)	(0.054)
ln(Loan GDP ratio)	0.863***	-1.333**	0.193	-0.671	1.160
	(0.158)	(0.565)	(0.156)	(0.796)	(0.761)
ln(Sales in 2009)	0.212***	-0.211*	0.043	0.262	0.162
	(0.047)	(0.118)	(0.053)	(0.186)	(0.138)
ln(Employees in 2009)	0.205***	0.117	0.069**	0.608**	0.135
	(0.058)	(0.160)	(0.032)	(0.272)	(0.178)
Female manager	0.116	-0.072	-0.107	-0.654	-0.025
	(0.165)	(0.365)	(0.131)	(0.694)	(0.727)
Manager's working	0.024***	0.020	0.009*	-0.029	0.031
experience	(0.006)	(0.019)	(0.005)	(0.026)	(0.019)
Value of the most		0.170	0.815***		
recent loan		(0.107)	(0.061)		
Constant	-7.615***	-33.542***	2.446	-6.902	-17.159
	(2.420)	(7.980)	(1.564)	(10.381)	(11.385)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Adjusted R^2			0.864		
Pseudo R^2	0.159	0.208		0.245	0.130
log likelihood	-1261.873	-197.399		-87.101	-74.157
Observations	2410	509	392	486	137

Notes: Columns (1)-(5) are estimated using logit, except that columns (3) are estimated using OLS. Standard errors are reported in parentheses. Data come from the 2008 economic census, 2008 China City Statistical Yearbook, and 2012 Enterprise Survey conducted by the World Bank.

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$.

Table 9: The Impact of SOEs on the Business Environment of POEs

	POEs			SOEs		
	Political instability	Corruption	Courts	Political instability	Corruption	Courts
	(1)	(2)	(3)	(4)	(5)	(6)
Share of city	1.361***	1.427***	1.644***	2.153	2.882	4.224*
employment in SOEs	(0.526)	(0.503)	(0.522)	(2.536)	(2.433)	(2.479)
ln(Population)	-0.133	-0.098	-0.115	-0.164	0.127	0.008
	(0.168)	(0.164)	(0.168)	(0.872)	(0.805)	(0.812)
ln(Per capita FDI)	-0.028	-0.018	0.073	-0.076	0.075	0.255
	(0.093)	(0.088)	(0.093)	(0.440)	(0.411)	(0.430)
ln(Density)	0.825***	0.920***	1.221***	0.369	0.843	0.921
	(0.206)	(0.196)	(0.203)	(1.086)	(0.983)	(1.055)
ln(Per capita GDP)	0.646***	0.612***	0.583***	0.685	0.623	0.519
	(0.164)	(0.158)	(0.163)	(0.775)	(0.757)	(0.747)
Share of GDP	-0.057***	-0.077***	-0.070***	-0.045	-0.042	0.019
in services	(0.012)	(0.012)	(0.012)	(0.063)	(0.056)	(0.054)
ln(Loan GDP ratio)	0.068	0.330*	0.104	-0.159	-0.157	-1.458
	(0.182)	(0.176)	(0.182)	(0.938)	(0.871)	(0.930)
ln(Sales in 2009)	-0.024	-0.023	-0.011	-0.070	-0.179	-0.160
	(0.045)	(0.043)	(0.045)	(0.187)	(0.183)	(0.190)
ln(Employees in 2009)	0.078	0.065	0.051	0.068	0.182	0.240
	(0.058)	(0.056)	(0.058)	(0.211)	(0.202)	(0.211)
Female manager	0.049	0.192	-0.005	0.396	0.612	-0.086
	(0.170)	(0.161)	(0.170)	(0.877)	(0.767)	(0.886)
Manager's working	-0.018***	-0.010	-0.017**	-0.020	-0.035	-0.007
experience	(0.007)	(0.007)	(0.007)	(0.028)	(0.026)	(0.026)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.042	0.046	0.063	0.032	0.048	0.065
log likelihood	-1564.883	-1649.755	-1486.096	-66.565	-77.477	-74.570
Observations	2454	2454	2454	136	136	136

Notes: Columns (1)-(6) are estimated using ordered logit. Standard errors are reported in parentheses. Data come from the 2008 economic census, 2008 China City Statistical Yearbook, and 2012 Enterprise Survey conducted by World Bank.

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$.

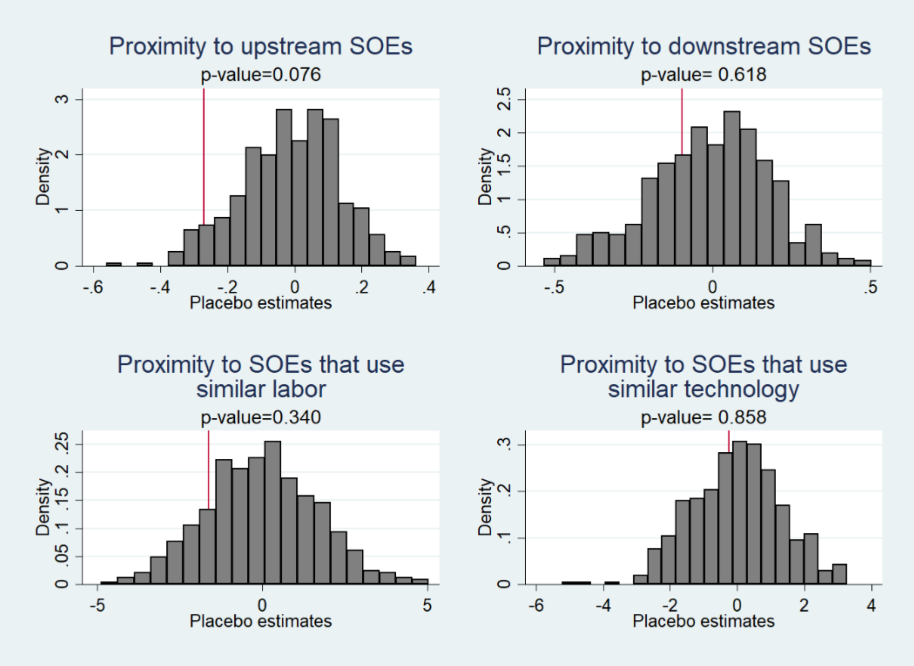
Table 10: Testing the Welfare Differences Between SOEs and POEs

	ln(Wage)	Fixed-term Job Contract	Local Hukou	Unemployment Insurance	Pension Insurance	Medical Insurance
	(1)	(2)	(3)	(4)	(5)	(6)
State-owned enterprises	0.031***	0.444***	0.076***	0.424***	0.414***	0.330***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Age	0.001***	0.001***	0.009***	0.005***	0.008***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
High school degree	0.115***	0.119***	0.058***	0.152***	0.188***	0.123***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
College degree or higher	0.378***	0.199***	0.056***	0.268***	0.282***	0.211***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Han nationality	0.025***	0.022***	0.056***	0.028***	0.043***	0.056***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)
Male	0.151***	-0.021***	-0.040***	-0.022***	-0.025***	-0.032***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Married	0.070***	0.001	-0.019***	0.000	0.018***	0.011***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
Constant	6.788***	0.256***	0.115**	-0.079*	-0.221***	-0.007
	(0.048)	(0.049)	(0.045)	(0.043)	(0.043)	(0.049)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
City dummies	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.423	0.345	0.339	0.470	0.491	0.348
Observations	199647	199647	199647	199647	199647	199647

Notes: The fixed-term job contract is a dummy variable that equals 1 if the worker has a fixed- term contract and 0 if otherwise; *hukou* type is a dummy variable that equals 1 if the worker has a local *hukou* and 0 if otherwise. Insurance participation includes unemployment, pension, and medical insurances. All are dummy variables that equal 1 if the worker participates in this kind of insurance and 0 if otherwise. Estimations use a seemingly unrelated regression model. Standard errors are reported in parentheses. Data come from the 2005 1% population census.

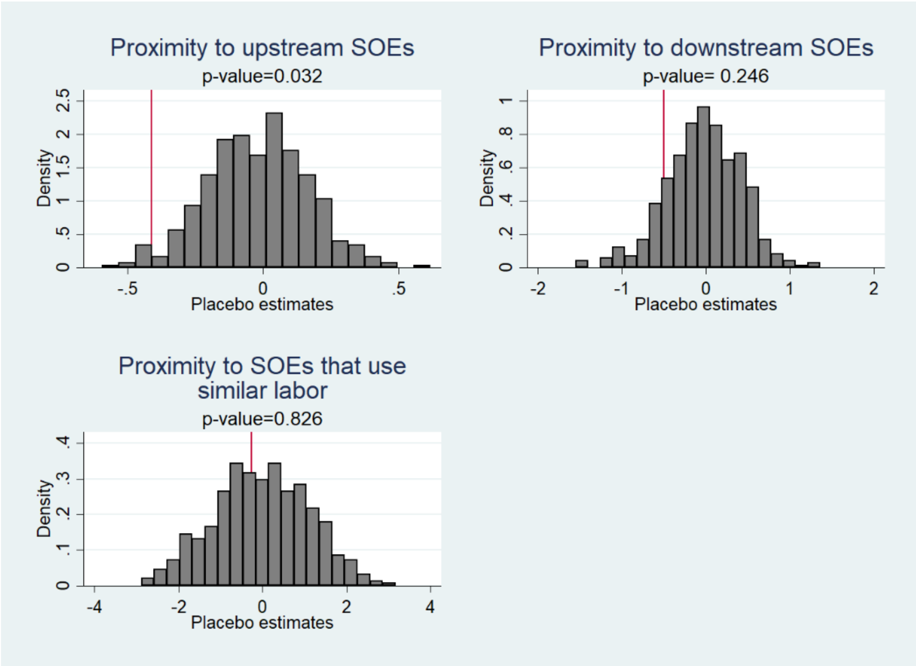
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < .01$.

Figure 1a: Placebo Permutation Test for the Impact of Incumbent SOEs on New Manufacturing POEs



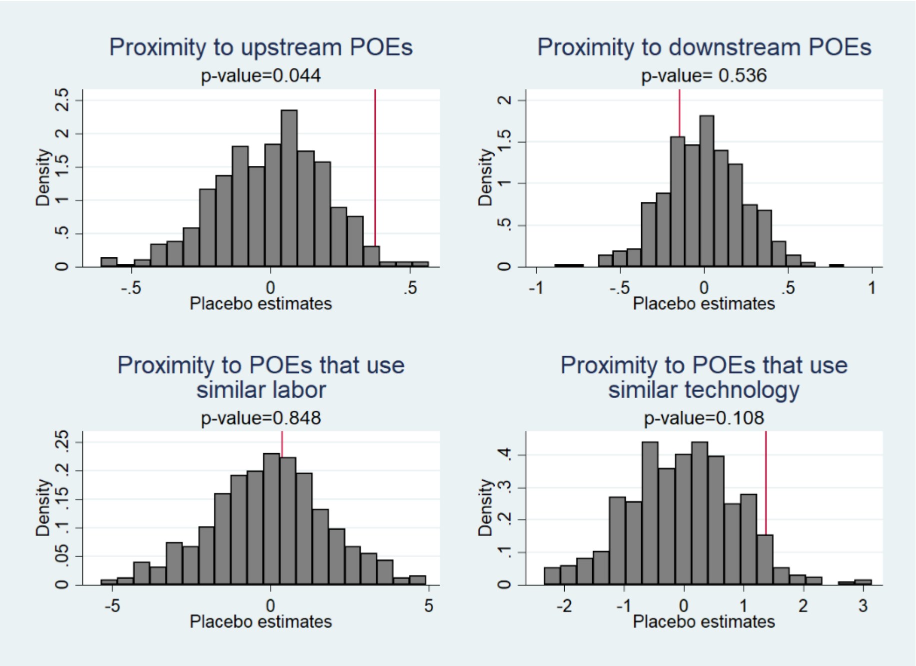
Notes: We conduct placebo permutation tests in which the city-industry pairs (e.g., the new POE employment) are randomly distributed across cities and the metrics of agglomeration remain unchanged. These histograms plot the empirical density function of placebo estimates from permuting city-industry pairs 500 times. The vertical line indicates the true estimate computed on the original data. The *p*-value of each permutation placebo test is the fraction of placebo estimates that are equal to or larger in absolute value than the corresponding estimate from the baseline Tobit model.

Figure 1b: Placebo Permutation Test for the Impact of Incumbent SOEs on New Service POEs



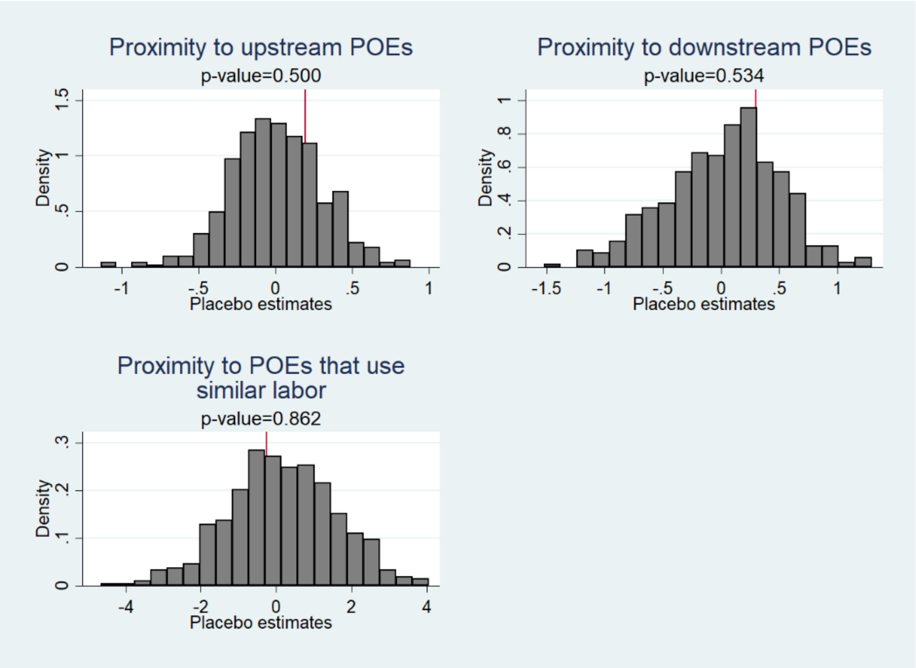
Notes: See the notes under figure 1a

Figure 2a: Placebo Permutation Test for the Impact of Incumbent POEs on New Manufacturing POEs



Notes: See the notes under figure 1a

Figure 2b: Placebo Permutation Test for the Impact of Incumbent POEs on New Service POEs



Notes: See the notes under figure 1a

Table A1: Descriptive Statistics of POEs and SOEs in Enterprise Survey Data

	POEs			SOEs		
	Mean	S.D.	Obs	Mean	S.D.	Obs
Finance						
Have a loan (%)	31.3	46.4	2588	30.6	46.2	144
Collateral required (%)	78.1	41.4	752	72.5	45.2	40
Value of collateral (million RMB)	15.771	1.484	459	16.969	1.379	23
Value of Recent Loan (million RMB)	20.850	83.190	536	47.344	104.190	29
Loan approved (%)	94.9	21.9	652	97.2	16.7	36
Political instability			2,647			143
No obstacle (%)	78.24			84.62		
Minor obstacle	17.19			12.59		
Moderate obstacle	3.51			2.8		
Major obstacle	0.76			0		
Very severe obstacle	0.3			0		
Corruption			2,649			143
No obstacle (%)	75.12			80.42		
Minor obstacle	20.27			16.78		
Moderate obstacle	3.62			2.1		
Major obstacle	0.72			0.7		
Very severe obstacle	0.26			0		
Courts			2,646			143
No obstacle (%)	76.49			81.12		
Minor obstacle	20.33			16.08		
Moderate obstacle	2.83			2.1		
Major obstacle	0.23			0.7		
Very severe obstacle	0.11			0		
City characteristics						
Share of city employment in SOEs	0.228	0.130	2700	0.248	0.126	148
Population (million)	6.910	2.491	2700	6.882	2.229	148
Per capita FDI (thousand dollar)	0.449	0.409	2700	0.381	0.371	148
Density (thousand/km^2)	0.779	0.266	2700	0.733	0.245	148
Per capita GDP (thousand RMB)	71.215	63.440	2700	62.817	57.483	148
Share of GDP in services	0.434	0.096	2700	0.423	0.089	148
Loan GDP ratio	1.093	0.467	2700	1.068	0.488	148
Firm characteristics						
Sales in 2009 (million RMB)	127.958	1098.337	2540	442.085	1777.740	142
Number of employees in 2009	196.506	976.345	2622	1193.379	6213.558	145
Female manager (%)	0.110	0.313	2696	0.101	0.303	148
Manager's experience	16.337	7.520	2640	18.993	8.429	142

Notes: Data come from the 2008 economic census, 2008 China City Statistical Yearbook, and 2012 Enterprise Survey conducted by the World Bank.

Table A2: Descriptive Statistics of Workers

	Mean	Standard Deviation
Monthly wage	1203	1136
Fixed-term job contract	0.528	0.499
Local *hukou*	0.704	0.456
Unemployment insurance	0.377	0.485
Pension insurance	0.508	0.500
Medical insurance	0.559	0.496
State-owned enterprise	0.405	0.491
Age	34.309	9.398
High school degree	0.286	0.452
College degree or higher	0.178	0.382
Han nationality	0.958	0.201
Male	0.612	0.487
Married	0.752	0.432

Notes: Data come from the 2005 1% population census. All variables are dummies excerpt for wage. Wage is in 2005 Chinese Yuan.