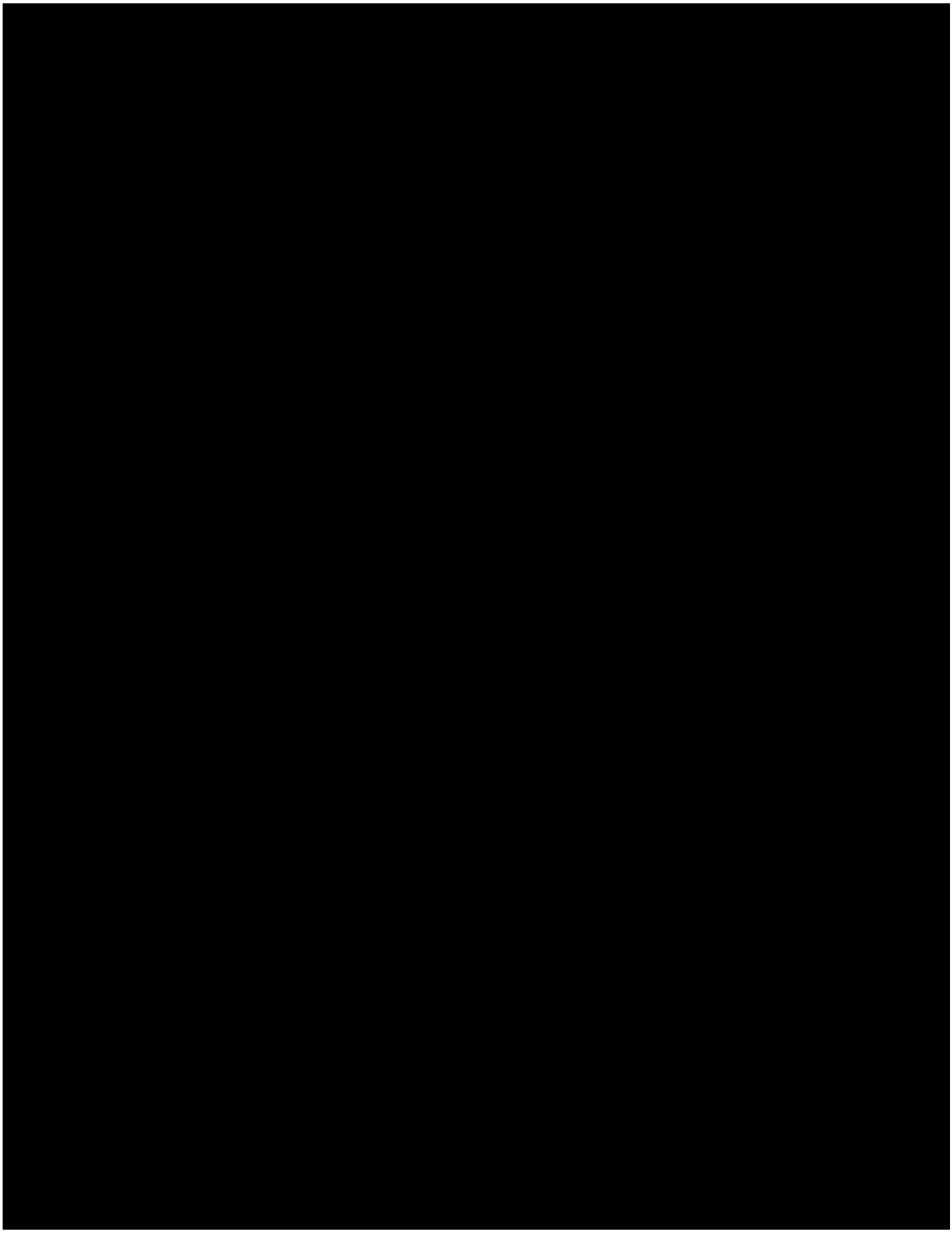


© 2021 International Monetary Fund WP/21/75

IMF Working Paper

Asia-Pacific Department

Resource Misallocation Among Listed Firms in China: The Evolving Role of State-Owned Enterprises

Prepared by Emilia Jurzyk and Cian Ruane

Authorized for distribution by Helge Berger

March 2021

***IMF Working Papers* describe research in progress by the author(s) and are published to elicit comments and to encourage debate.** The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

Abstract

We document that publicly listed Chinese state-owned enterprises (SOEs) are less productive and profitable than publicly listed firms in which the state has no ownership stake. In particular, Chinese listed SOEs are more capital intensive and have a lower average product of capital than non-SOEs. These productivity differences increased between 2002 and 2009, and remain sizeable in 2019. Using a heterogeneous firm model of resource misallocation, we find that there are large potential productivity gains from reforms which could equalize the marginal products of listed SOEs and listed non-SOEs.

JEL Classification Numbers: L10, L33, O11, O41

Keywords: State-Owned Enterprises, Misallocation, Productivity

Author's E-Mail Address: ejurzyk@imf.org; cruane@imf.org

1. Introduction

China has experienced a sustained period of growth over the last three decades. During this time, Chinese authorities have pursued major reforms, including a transition away from state management of most sectors of the economy. Indeed, between 1998 and 2005 the share of state-owned enterprises (SOEs) in industrial output fell from 50 to 30 percent. This transition coincided with rapid aggregate productivity growth, which came in part from the growth of the private sector at the expense of less productive SOEs (Hsieh and Song, 2015; Brandt, Van Biesebroeck and Zhang, 2012).

However, productivity growth and progress on SOE reforms slowed in recent years. The slogan of "Grasp the Large, Let Go of the Small" reflected the Chinese government's desire to maintain control of the largest and most profitable SOEs, many of which were publicly listed. Some of these SOEs became national champions, growing into industry leaders and now counting among the world's largest firms (IMF Fiscal Monitor, April 2020). In China itself, large publicly listed SOEs account now for 35 percent of aggregate SOE revenues. Though a new wave of SOE reforms, particularly focused on larger firms, was announced at the Third Plenum in 2013, reform progress appears to have been slow, with the government reluctant to reduce its role (Rosen et al., 2018). The question remains, therefore, whether there is scope for further reforms of the state-owned sector which could provide a boost to productivity growth. This is especially relevant given that aggregate productivity growth in China has slowed over the past decade, averaging a mere 0.6 percent between 2012 and 2017. Our paper contributes to this debate by measuring productivity gaps between listed SOEs and private firms from 2002 to 2019, and quantifying the potential gains from reforms.

We use data on publicly listed Chinese firms on the Shanghai and Shenzhen stock exchanges between 2002 and 2019 to compare the productivity of SOEs relative to 'privately-owned' firms. We define SOEs as firms where the major shareholder or ultimate controller is a central or local government agency. Private firms are those in which neither central nor local Chinese government entities have an ownership stake. A benefit of focusing on listed firms is that data is available until 2019, while other surveys (such as the commonly used Annual Survey of Industries) is only available until 2013. In addition, listed firms cover all sectors of the economy, rather than just manufacturing. While our sample only covers 3,700 firms, they are a large and growing share of the Chinese economy, accounting for 6 percent of GDP in 2019 and 10 percent of industrial value-added.

Our baseline productivity measures are labor productivity, capital productivity and a total factor productivity (TFP) constructed as value-added divided by a geometric average of capital and labor (measured as employment), corresponding to a standard Cobb-

Douglas production function. As we don't have firm-specific prices, our measure of firm-level TFP is a measure of 'revenue productivity' , which we refer to as TFPR. TFPR reflects differences in revenues that cannot be explained by variations in inputs. Thus, differences in TFPR between firms can correspond to differences in factor prices (e.g. interest rates), markups, or other taxes and subsidies but they do not necessarily reflect differences in the technology of firms, their production efficiency or the quality of their products.

Even among publicly listed firms, we document large and statistically significant differences in TFPR between SOEs and private firms. On average, SOEs have 30 percent lower TFPR than private firms in the same sector, with differences largely driven by a high number of low productivity SOEs: 15 percent of SOEs are in the bottom decile of the TFPR distribution while only 7 percent of private firms are. This TFPR gap increased between 2002 and 2009 with the listing of many new high-TFPR private firms. It has remained broadly stable until 2016, and closed slightly between 2016 and 2019. When differentiating between capital and labor productivity, we find that the SOE-private firm TFPR gap is accounted for by a very large difference in capital productivity (over 50 percent), while on average labor productivity is only 6 percent lower.

Our findings are robust to various measures of productivity, including using a gross output production function rather than a value-added production function. We also find large and significant differences between SOEs and private firms over time in accounting measures of firm profitability, such as net profits over revenues and return on assets. Despite the differences in governance and incentives between central-government owned enterprises and local-government owned enterprises (Lin and Chang, 2019), we do not find any difference between them in terms of TFPR. This could stem from the fact that only the best local SOEs (i.e. those that operate as efficiently as central SOEs) can be publicly listed.

One potential explanation for the high capital-intensity and low capital productivity of SOEs is that they have cheaper costs of financing than private firms. The existing literature provides ample evidence that state-owned enterprises enjoy preferential access to credit relative to private firms (Poncet, Steingress and Vandebussche, 2010; Li, Meng, Wang and Zhou, 2008; Boyreau-Debray and Wei, 2005; Wu, Firth and Rui, 2014; Bai, Lu and Tian, 2018; Harrison, Meyer, Wang, Zhao and Zhao, 2019). Lower interest rates would incentivize SOEs to take on more debt and accumulate more assets than private firms. We also explore this in our data by constructing the effective interest rates faced by listed firms in 2018 and 2019, measured as a ratio of interest expenses over total liabilities.[1] We find that SOEs have higher leverage than private firms, and that their ef-

[1]These are the only years for which the interest expense variable is available.

fective interest rates are lower than that for private firms, in particular when controlling for firm leverage. These differences are not small in magnitude. The effective interest rate is roughly a quarter smaller for SOEs compared to private firms, which could explain half the capital productivity gap.

To quantify the aggregate importance of the average firm-level productivity differences, we interpret our findings through the lens of a model of heterogeneous firms following Hsieh and Klenow (2009). The critical assumption is that the within-sector differences in measured SOE and private firm *average* products can be interpreted as differences in true *marginal* products which distort the optimal allocation of resources.[2] Using sector-by-sector estimates of TFPR gaps, we evaluate the gains from closing the TFPR gap in each sector. To the extent that these differences are due to SOEs using a more capital-intensive technology than private firms, having larger overhead costs (as a share of their inputs) than private firms, or understating their revenues (overstating their costs) more than private firms, our estimates provide an upper bound for the gains from closing TFPR gaps.

We find average gains of 8.2 percent over our whole sample period, which decline to 5.8 percent in 2019. We consider also a second counterfactual reform in which we close both the TFPR gap and equalize the average capital intensity of SOEs and private firms, thus removing both 'scale' distortions and 'factor' distortions (following terminology from David and Venkateswaran (2019)). In this case, we estimate larger average gains of 10.7 percent over our whole sample, and 7.5 percent as of 2019. Our findings suggest that there remain potentially large gains from SOE reform. In particular, reforms which would recognize and remove the implicit government guarantees that allow SOEs to access financing from banks and financial markets at lower rates could be particularly effective.

1.1. Related Literature

To document the existence of large productivity gaps between SOEs and private firms, many firm-level studies have used the Chinese Annual Survey of Industries, which covers all above-scale industrial firms up to 2013 (Hsieh and Klenow, 2009; Berkowitz et al., 2017; Bai et al., 2018).[3] Like us, most of this literature documents that capital productivity gaps between SOEs and private firms were larger than labor productivity gaps.

[2]We don't rely on the all TFPR dispersion to reflect dispersion in marginal products. A large part of this could simply be due to measurement error (Bils et al., 2020) or misspecification (Haltiwanger et al., 2018).

[3]Chinese Annual Survey of Industries includes all above scale state-owned firms as well as non-state firms. These are firms with sales exceeding 5 million RMB. In 2011, the designated size increased from 5 million to 20 million RMB.

Naturally, the listed firms in our sample account for a smaller share of GDP than above-scale industrial firms, though they still account for 35 percent of aggregate SOE revenues in 2019 and have broader sectoral coverage. Importantly, we have data through 2019, allowing us to examine whether the major SOE reforms announced at the Third Plenum in 2013 had any effect. Our paper is one of the first to both document the existence of productivity gaps over this extended time period and provide a quantification of the resulting misallocation losses.

Our paper also relates to an extensive literature examining the the role of resource reallocation between SOEs and private firms in China's recent economic transformation. The contribution of privatization and exit of SOEs to aggregate productivity growth remains a point of discussion. Brandt et al. (2012) find that within-firm productivity growth was an important source of growth from 1998 to 2007, though net entry contributed two thirds of TFP growth. Using a different decomposition, Hsieh and Song (2015) find that the privatization of SOEs and new entry of SOEs contributed 30 percent to aggregate productivity growth during the 2000s, with productivity growth among private firms explaining most of the rest. Given that listed firms accounted for a small share of GDP during the early 2000s, we don't assess their contribution to aggregate productivity growth over this period, but rather we consider how the losses due to resource misallocation *among* listed firms changed over time.

A more recent literature has evaluated the gains from SOE privatization between 1998 and 2013. Harrison et al. (2019) and Chen et al. (2020) both find that privatizing SOEs saw productivity improvements relative to SOEs that didn't privatize, though there remains a revenue productivity gap between ex-SOEs and private firms, with ex-SOES still benefiting from implicit subsidies and cheaper access to credit. Given the focus of "Grasp the Large, Let Go of the Small" policies was to keep large SOEs within government control, we don't see frequent privatization of SOEs among listed firms. We, therefore, focus on cross-sectional differences in revenue productivity between SOEs and private firms. For the smaller sample of firms that did switch ownership, however, we do indeed find evidence of revenue productivity gains.

Our paper also relates to an extensive literature on resource misallocation in China.[4] David and Venkateswaran (2019) decompose capital productivity dispersion among large-scale industrial firms in China, finding that permanent differences in capital productivity across firms account for most of it. Song and Wu (2015) conduct a similar exercise, also allowing for production functions to vary across firms, and still find substantial aggregate productivity losses due to misallocation. Wu (2018) and Bai et al. (2018) both

[4]The approach used in this literature began with Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). See Hopenhayn (2014) for a review of the literature.

find an important role for financial frictions in generating capital misallocation. Such financial frictions are less likely to explain our findings, however, given that we focus on large listed firms for which these tend to be less binding. Finally, Brandt et al. (2013) take a more sectoral approach, quantifying cross-sector and cross-region misallocation due to SOEs. Such cross-sector misallocation would imply larger potential gains from SOE reform than what we find in this paper.

Our paper also relates to the literature examining differences between the behavior of SOEs and private listed firms in China. Boeing et al. (2016) document that listed firms benefit more from RD than SOEs, suggesting the potential for TFP losses due to misallocation in innovation inputs. Chen et al. (2017) document that the within-business group allocation of capital is better for private firms than SOEs, and document that chairmen at SOEs tend to reallocate capital to prop up units within the same business group that operate in high unemployment areas. Such misaligned incentives could be an important explanation of our findings of persistently lower capital productivity for SOEs.

The rest of the paper is organized as follows. In Section 2. we discuss the dataset of listed firms in China and present some stylized facts. In Section 3. we estimate differences in productivity between state owned and private firms. In Section 4. we present our model of misallocation, and in Section 5. we estimate the gains from reforms.

2. Listed Firms in China

2.1. Wind Database of Listed Firms

We use data on listed firms in the Shanghai and Shenzhen stock exchanges between 2002 and 2019 from the Wind database.[5] The main firm-level variables we use are employment, operating costs, fixed assets and total revenues. We aggregate all quarterly variables to an annual frequency. We construct a measure of intermediate inputs as total operating costs minus labor costs, and construct firm-level value-added as total revenues minus intermediate inputs.[6] We use the available industry classification of 28 industries (see Data Appendix).

Hsieh and Song (2015) discuss the difficulties in measuring whether a firm is controlled by the Chinese state. Many large firms are officially registered as private when they are de facto state-owned, with the State-Owned Assets Supervision and Administration Commission (SASAC) having controlling ownership shares. We therefore follow

[5] Information available at https://www.wind.com.cn/en/data.html.

[6] We impute labor costs using firm-level employment and average urban wages from the NBS. This approach follows closely how intermediate inputs and value-added are constructed in David and Venkateswaran (2019). See Data Appendix for more details.

our database's classification of firms as state-owned or not state-owned, which uses local and central government ownership and control rather firm registration. We define SOEs as firms where the major shareholder belongs to a central or local government SASAC, a central or local state agency, or another central or local state-owned enterprise. We also use this classification to separately define local SOEs and central SOEs. We define privately owned enterprises as firms in which neither central nor local Chinese government entities have an ownership stake. All other firm types are classified as Other.[7]

While our database is smaller and less nationally representative than the commonly used Annual Industrial Survey, there are a few advantages to focusing on listed firms. Firstly, our data is available annually between 2002 and 2019, while the industrial survey is only available (with gaps) until 2013. Secondly, Chen, Chen, Hsieh and Song (2019) find evidence of deteriorating data quality in the Annual Industrial Survey from 2008 on. Listed firms have a stricter accounting framework to follow which are subject to more rigorous auditing. Finally, our data includes firms in all sectors of the economy, rather than just focusing on industrial firms.

2.2. Stylized Facts about Listed Firms

After, some basic data cleaning we are left with 3,703 firms in our final database.[8] Figure 1 shows that the number of listed firms in China increased dramatically between 2002 and 2015, from around 1,000 to over 3,000. This entire increase is due to entry of newly listed firms, as there is close to zero exit (de-listing) of incumbents. In addition, from 2006 to 2015 the vast majority of newly listed firms (over 90% on average) are privately owned. We do observe some firms switching from state-ownership to private ownership, however this is not particularly frequent among listed firms. At it's peak, 3% of state-owned firms in 2005 switched to private ownership the following year.

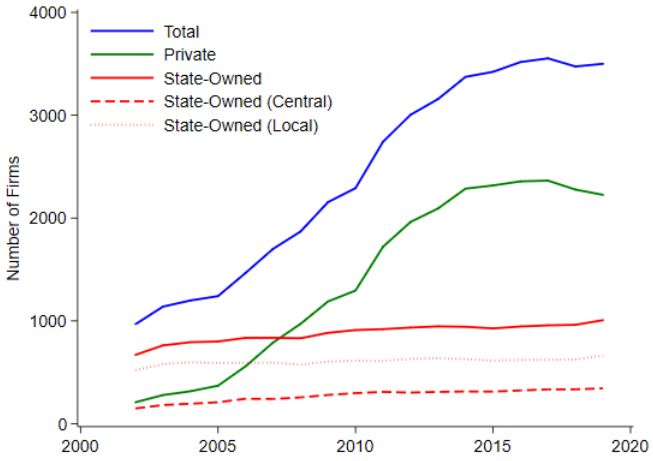
While listed firms only account for a tiny fraction of all Chinese firms, they account for a large and rising share of aggregate economic activity.[9] Figure 15 in Appendix A plots the GDP share of listed firms over time which rose to 6 percent in 2019, with listed manufacturing firms accounting for 10 percent of aggregate manufacturing value-added. In

[7] This category includes Collectively-Owned Enterprises, Foreign Companies, Joint Ventures and Public Enterprises.

[8] We drop firms with missing or negative observations of key variables. We also drop firms operating in the Financial Services sector, as well as two large petrochemical companies who account for over 20% of aggregate value-added. Finally we trim the 1% tails of TFPR deviations from the sector-year average, where TFPR is as defined in Section 3.

[9] Listed industrial firms in 2019 accounted for 0.6% of above-scale industrial firms (with over 20 million RMB in revenues).

Figure 1: Number of Listed Firms in China



Notes: This figure plots the number of firms in our cleaned database of listed firms from Wind. The number of private firms and state-owned firms are shown separately.

addition, listed SOEs account for a large and rising share of aggregate SOE output and inputs. Figure 16 in Appendix A shows that the revenue share of listed SOEs rose to 35% by 2019, coinciding with a wave of reforms in the 2000s in which many medium-sized SOEs were either closed down or privatized (Hsieh and Song, 2015).

We present summary statistics for our dataset in Table 1. We observe the average firm for 11.8 years. While there are listed firms in every industry, close to 60% of listed firms are in manufacturing (Panel B. of Table 1 and Figure 17 of Appendix A). Listed firms are very large, with average employment of 5,427 workers in 2019. The employment distributions of SOEs relative to POEs is shown in Figure 2. Though there is a lot dispersion, SOEs are considerably larger than POEs on average, having 2.5 times as many workers and a 7 times larger average capital stock. However, the SOE share of employment and value-added among listed firms has been falling since the early 2000s, from 80% of employment in 2002 to 49% by 2019 (Figure 13 in Appendix A). This coincides with a general downwards trend in the SOE share of revenues for all industrial firms, as shown in Figure 14.

3. SOE vs. POE Revenue Productivity Gaps

Table 1: Sample Summary Statistics

Panel A. *Summary Statistics*

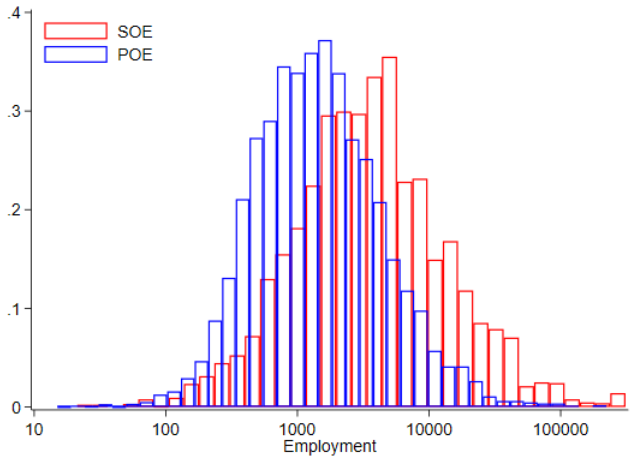
	All	Private	State-Owned
Number of firms	3,703	2,608	1,268
Number of observations	43,770	25,579	15,866
Average in 2019 (in 000s RMB):			
Revenues	10,249,220	4,680,487	22,119,446
Value-Added	1,716,727	927,411	3,402,235
Fixed Assets	3,386,439	1,216,702	8,447,579
Employment	5,472	3,572	9,403

Panel B. *# Firms per Broad Industry*

	All	Private	State-Owned
Agriculture	88	60	34
Construction	130	89	45
Hotels & Catering	34	17	23
IT	507	391	126
Manufacturing	2,146	1,653	555
Mining	185	106	92
Other	105	51	62
Real Estate	132	73	89
Transport, Storage & Post	117	40	82
Utilities	162	74	100
Wholesale & Retail Trade	97	54	60

Notes: This table contains summary statistics for the sample of firms in Wind used in the analysis in the paper. Average revenues, value-added and fixed assets are measured in 1000s of current RMB, while employment is measured in number of workers. The total number private and state-owned firms do not necessarily add up to the total as some firms change ownership structure, and therefore appear in the dataset both as state-owned and private in different years.

Figure 2: Employment Distribution for POEs and SOEs



Notes: This figure plots distribution of employment for SOEs and POEs separately, pooling data from 2015 to 2019. Data on listed firm employment is from Wind.

3.1. Productivity Measurement

As a baseline, we construct productivity measures based on a constant returns to scale Cobb-Douglas production function, where real value-added is a function of capital and labor. For firm i in sector s we have that:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s} \tag{1}$$

We construct total factor revenue productivity (TFPR) as:

$$\text{TFPR}_{si} \equiv \frac{P_{si} Y_{si}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}} = P_{si} A_{si} \tag{2}$$

where P_{si} is the price index for firm i. Note that TFPR is different from the firm's technical efficiency A_{si} (often referred to as physical productivity or TFPQ following Foster et al. (2008)), as it includes the firm's price. Like the vast majority of papers which estimate productivity differences across firms, we do not have data on firm-specific prices and therefore cannot construct TFPQ without further assumptions.[10] TFPR does not simply capture technical efficiency or product quality differences across firms, but also variation in firm-level prices. Such variation can be due to firms charging different markups

[10]These can be timing assumptions about when inputs can adjust to productivity shocks, as in Olley and Pakes (1996) and Ackerberg, Caves and Frazier (2015), or assumptions about the shape of the demand curve as in Hsieh and Klenow (2009).

or responding to idiosyncratic demand or supply shocks. However, it could also be due to firms facing different input costs (e.g. SOEs benefiting from implicit or explicit government subsidies) or facing restrictions on their input use (e.g. SOEs being required to keep on surplus workers) which shift their positions along the demand curve. In this section, we present our empirical results without imposing any assumptions on what is driving differences in TFPR between SOEs and private firms. Such assumptions are necessary however to make inferences about allocative efficiency, and we discuss this further in Section 4.

We use fixed assets as our measure of capital, and employment as our measure of labor (we do not observe firm-specific labor costs). We impose sector-specific and time invariant factor shares α_s. We construct α_s as the average cost-share of capital in each sector.[11] We define labor productivity as value-added per worker, while we define capital productivity as value-added per RMB of fixed assets. We also refer to these as the average revenue products of labor and capital; ARPL and ARPK respectively. All the results in this section are presented net of sector-year fixed effects, hence we do not deflate any of the variables.

3.2. Main Findings

Figure 3 plots the distribution of revenue productivity separately for SOEs and POEs, pooling all years between 2002 and 2019. There is huge dispersion in TFPR, even within sectors; the revenue productivity of firms at the 90th percentile is more than four times larger than that of firms at the 10th percentile. However, it is clear that the SOE revenue productivity distribution is to the left of that of private firms. Figure 4 reinforces this point by showing the share of SOEs and private firms that are in each decile of the TFPR distribution. 15% of SOEs are in the bottom decile of TFPR distribution, though only 7% of private firms are. This difference is not symmetric however; SOEs are more strongly over-represented in the left tail of the distribution than they are under-represented in the right tail.[12] A feature of the data is therefore a fat left tail of low-TFPR state-owned

[11] We first construct each sector's total labor costs by multiplying sectoral employment by the average wage in that sector (obtained from CEIC). We then construct the capital cost share in each year as $\frac{RK_s}{RK_s+wL_s}$ imposing a rental rate of 20% following Bils et al. (2020). We then average this share over time to obtain α_s. This yields an average sectoral capital share of 47%. Given these are large listed firms, we also check robustness to lower values for the rental rate on capital.

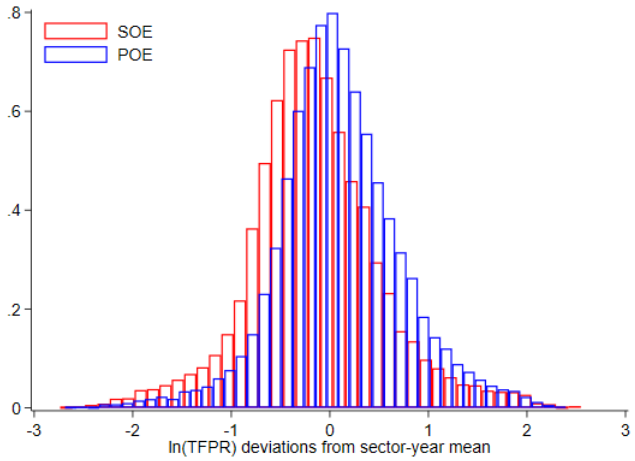
[12] The share of SOEs in the top decile of the TFPR distribution is 7.3% relative to 11.6% for private firms. However, both the share of SOEs and private firms is 1% in the top centile of the distribution, while it is 1.5% and 0.7% in the bottom centile for SOEs and private firms respectively.

firms. Column (1) of Table 2 formally reports the results from the following regression:

$$\ln(\text{TFPR})_{sit} = \alpha + \beta \text{SOE}_{sit} + \delta X_{sit} + \gamma_{st} + \varepsilon_{sit} \tag{3}$$

where SOE is an SOE dummy, X_{sit} is a vector of potential controls, and γ_{st} is a vector of industry-year fixed effects. On average across our sample period, SOE revenue productivity is a 31% lower than for private firms within the same industry and year, with a standard error of 4.6%.

Figure 3: TFPR Distributions for SOEs and Private Firms

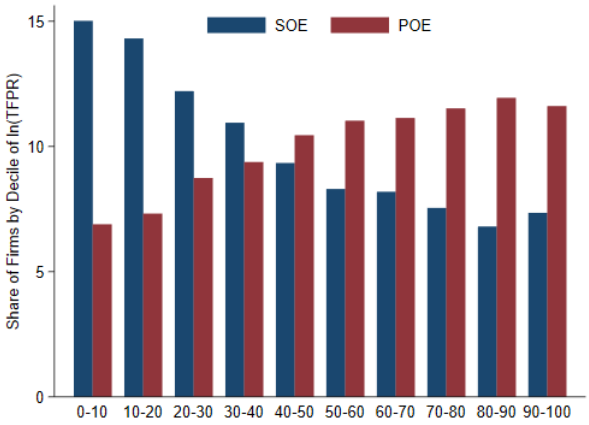


Notes: This figure plots the distribution of ln(TFPR) deviations from the sector-year average for SOEs and POEs separately, pooling data from 2002 to 2019. Data on listed firms is from Wind.

The revenue productivity gap between SOEs and private firms is almost entirely accounted for by differences in capital productivity (ARPK). Columns (2) and (3) of Table 2 show that the capital productivity gap between SOEs and private firms is 51% on average, while it is only 6% for labor productivity (ARPL). The distributions of SOE and private firms capital and labor productivity are shown in Figures 18 and 19 in the Appendix, which show a lot of overlap between the distributions but with a leftward shift for SOEs. It follows from the above findings that an important difference between Chinese SOEs and private firms is their capital intensity. Indeed, column (4) of Table 2 shows that the capital labor ratio of listed SOEs is 45% higher than for private firms.

We explore how the SOE revenue productivity gap relative to private firms evolved over time in Figure 5. In 2002, there was close to no difference in the TFPR of listed SOEs and private firms, however this gap increased dramatically between 2002 and 2008, coinciding with the listing of high-TFPR private firms. Indeed, we find that newly-listed private firms have 28% higher TFPR than existing private firms throughout our sample.

Figure 4: Share of SOEs and Private Firms by Decile of TFPR



Notes: This figure plots the share of SOEs and POEs that are in each decile of the ln(TFPR) deviations from the sector-year average. Data is pooled data from 2003 to 2019. Data on listed firms is from Wind.

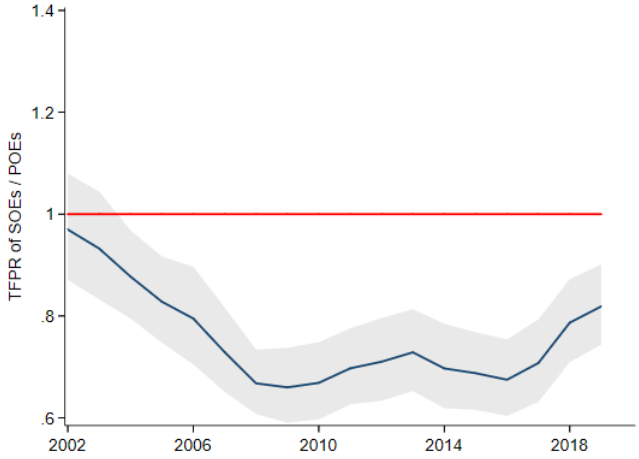
Table 2: Estimates of Average Product Gaps Between SOEs and Private Firms

	(1) ln(TFPR)	(2) ln(ARPK)	(3) ln(ARPL)	(4) ln(K/L)
SOE	-0.31***	-0.51***	-0.06	0.45***
	(0.046)	(0.06)	(0.05)	(0.08)
Other	-0.10***	-0.17***	-0.03	0.14**
	(0.03)	(0.043)	(0.039)	(0.059)
Observations	43,770	43,770	43,770	43,770
R^2	0.92	0.33	0.24	0.13

Notes: The dependent variable in all columns is ln(TFPR), and the omitted group are private firms. The Other category includes Collectively-Owned Enterprises, Foreign Companies, Joint Ventures and Public Enterprises. All regressions include industry x year fixed effects. Standard errors are clustered at industry level.

The SOE productivity gap then plateaued until 2016, before narrowing slightly between 2017 and 2019. Appendix Figures 21 and 22 document similar time-series patterns for both capital and labor productivity, while Figure 6 shows how this is also reflected in the relative capital-intensity of SOEs and private firms, with SOEs only slightly more capital intensive than listed POEs in 2002 but with a large gap materializing by 2008. Our findings of large capital productivity gaps between firms between SOEs and private firms line up with those in Berkowitz et al. (2017) and Hsieh and Song (2015) for all industrial firms prior to 2013. While our analysis is restricted to listed firms, we document that these capital productivity gaps have persisted until 2019.

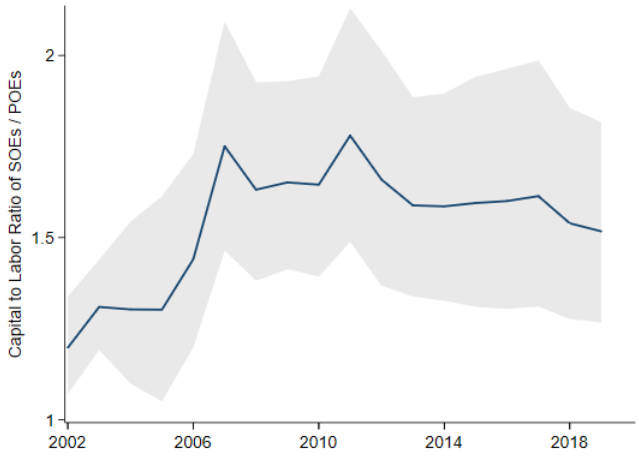
Figure 5: SOE TFPR Gaps Over Time



Notes: This figure plots the TFPR gap between SOEs and POEs between 2002 and 2019. These are obtained by regressing ln(TFPR) on an SOE dummy, controling for sector x year fixed effects. The coefficient on the dummy is then exponentiated. Data on listed firms is from Wind.

Next, we examine industry heterogeneity in the magnitude of the TFPR gaps between SOEs and private firms. We estimate Equation 3 separately for each broad industry. We report these industry productivity gaps pooling 2016 to 2019 in Figure 7. While there is a lot of heterogeneity across sectors, the SOE productivity gaps are present in all industries except for Food and Drink Manufacturing, and Real Estate. SOEs have particularly low TFPR in sectors such as Utilities, Transportation and Steel Manufacturing, with smaller differences in more high-tech sectors such as IT Services, Manufacturing of Pharmaceuticals, and Medical Equipment. Table 9 in the Appendix reports these sectoral productivity gaps by 5-year windows. Though there is some variation over time, the gaps are generally quite persistent; sectors with the initially largest productivity gaps continuing to have the largest gaps by 2019. Figure 8 also shows that the TFPR gaps are negatively correlated with the SOE employment share in the sector – sectors where

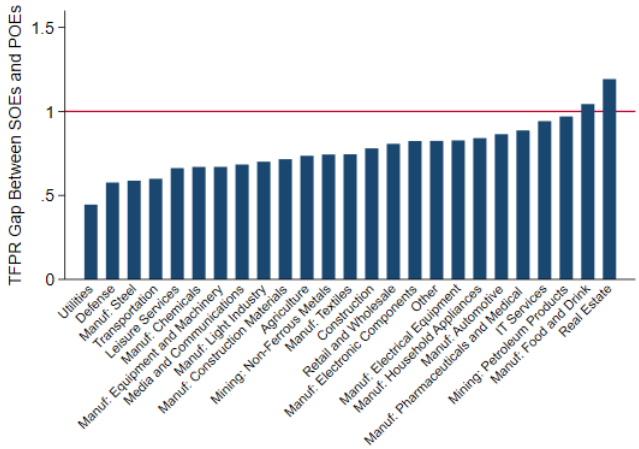
Figure 6: SOE Capital Intensity Relative to Private Firms Over Time



Notes: This figure plots the capital to labor ratio of SOEs relative to POEs between 2002 and 2019. These are obtained by regressing ln(capital to labor ratio) on an SOE dummy, controling for sector x year fixed effects, and then exponentiating the coefficient and confidence intervals. The coefficient on the dummy is then exponentiated. Data on listed firms is from Wind.

SOEs are have lower TFPR tend also to be those where SOEs account for a larger share of inputs and output. This suggests the potential for large productivity losses in certain sectors.

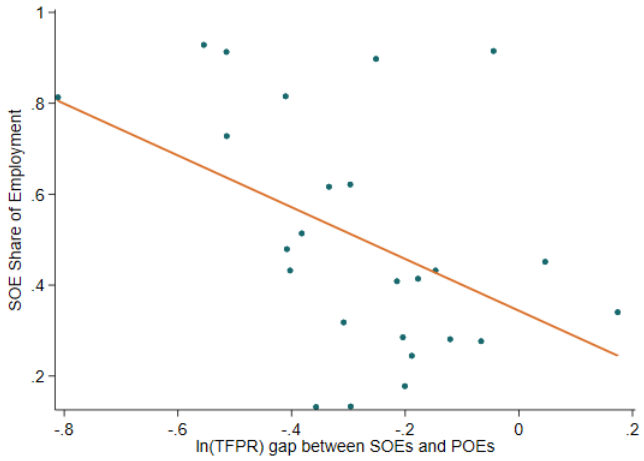
Figure 7: SOE TFPR Gaps by Sector (2016-2019)



Notes: This figure plots the TFPR gap between SOEs and POEs by sector (1 means no gap). TFPR gaps are estimated from 2016 to 2019. Data on listed firms is from Wind.

We also explore other dimensions along which our results could be heterogeneous.

Figure 8: SOE TFPR Gap by Sector vs. SOE Share of Sectoral Employment



Notes: This figure plots the TFPR gap between SOEs and POEs between 2016 and 2019 against the SOE share of total employment in that sector (averaged between 2016 and 2019). Data on listed firms is from Wind.

As previously described, SOEs can be owned or controlled by either central governments or local goverments. Central and local SOEs may therefore face different incentives, governance structures, or differentially benefit from government incentives that could affect their measured productivity. We therefore explore whether central and local SOEs exhibit different TFPR gaps relative to private firms in column (2) of Table 3. We regress ln(TFPR) on a separate dummy for central and local SOEs, again controlling for sector-year fixed effects. However, we don't find any significant difference between central and local SOEs, with both of them exhibiting large TFPR differences relative to private firms of around 30%.

Some of the differences in revenue productivity across firms could be explained by factors such as age and size. As previously noted, newly-listed firms tend to have higher TFPR than incumbents, and SOEs tend to be much larger than private firms. Younger firms may inherently have higher average products as they are still in the process of growing to their optimal size. In turn, firm size could be correlated with many factors that could affect TFPR; the share of overhead costs in total costs, markups and market power, the share of surplus inputs, the amount of collateral available for financing, implicit guarantees because of systemic importance, etc... In column (3) of Table 3 we therefore add both firm size and age as controls in Equation 3. We find that younger firms have higher TFPR (as expected) and that larger firms have lower TFPR. This second result contrasts with the finding in Hsieh and Klenow (2009) that TFPR increases with firm size in India and China. An important difference could be that we are focusing

on listed firms only however. Overhead costs and markups don't seem to be reasonable explanations of this pattern, as they would imply that markups are declining in firm size (and market share) or that overhead costs are rising as a share of inputs with firm size. Other explanations, such as better access to financing because of implicit guarantees or better collateral seem a better fit. However, even controlling for firm size and age, we document that SOEs have 17% lower TFPR than private firms.

Finally, we explore productivity differences between SOEs and private firms by looking at switchers - firms that change ownership status. However, as described in Section 2., there are not many changes between SOE and private ownership status among listed firms. In total, we identify 307 firms that at some point changed ownership status between SOE and private ownership (not exclusively in that direction). While changes in ownership status are not exogenous events (Chen et al. (2020) document that non-performing SOEs tended to be those that were privatized in the 2000s), they do help control for some time invariant unobservable characteristics that could be relevant in explaining SOE and private firm differences, such as a role of providing social goods. In column (4) of Table 3 we therefore add a firm fixed effect to Equation 3. We still find that SOEs have 17% lower revenue productivity than POEs, qualitatively similar to results in Harrison et al. (2019).

3.3. Alternative Measures of Firm Performance

In this sub-section, we show that the results presented so far hold for a wide array of measures of productivity and firm performance. While our baseline specification considers a value-added production function in capital and labor, we could alternatively consider a gross output production, where output is produced using capital, labor and intermediate inputs:

$$Q_{si} = A_{si}^{Q} \left(K_{si}^{\alpha_s} L_{si}^{1-\alpha_s} \right)^{\gamma_s} X_{si}^{1-\gamma_s} \tag{4}$$

X_{si} is the intermediate input bundle, while γ_s is the value-added share of gross output. In Table 4 we run Equation 3 but substituting in gross output TFPR on the left hand side. Similarly, we redefine capital and labor productivity as being firm revenues over capital and labor. We still find a significant 7.3% difference in revenue productivity between SOEs and private firms.[13] We also find that SOEs still have 23% lower capital productivity, and a 6% lower intermediate input productivity. Interestingly however, SOEs have a higher average product of labor than private firms, though this is likely due to the higher

[13]Note that the productivity differences based on a gross ouptput production function and a value-added production function are not directly comparable in magnitudes.

Table 3: Robustness: Estimates of SOE TFPR Gaps

	(1)	(2)	(3)	(4)
SOE	-0.31***		-0.17***	-0.17***
	(0.046)		(0.047)	(0.029)
Other	-0.10***	-0.10***	-0.027	-0.08**
	(0.03)	(0.003)	(0.0289)	(0.04)
Central SOE		-0.29***		
		(0.055)		
Local SOE		-0.32***		
		(0.050)		
Age			-0.012***	
			(0.002)	
ln(Employment)			-0.13***	
			(0.018)	
Sector x Year FEs	✓	✓	✓	✓
Firm FEs				✓
Observations	43,770	43,770	43,734	43,764
R^2	0.92	0.92	0.93	0.96

Notes: The dependent variable in all columns is ln(TFPR), and the omitted group are POEs. Standard errors clustered at sector level.

capital and intermediate input intensity of SOEs. The time-series of gross output based TFPR gaps between SOEs and private firms is similar to that for value-added, as shown in Panel (A) of Figure 9.

Finally, we consider other standard measures of accounting profitability in Table 5. SOEs have lower accounting profitability than private firms, as measured by net profits over revenues, EBITDA over revenues and return on assets (ROA, defined as net profits over total assets). ROA differences between SOEs and private firms also exhibit a similar time series pattern as shown in Panel (B) of Figure 9. All in all, our findings suggest unambiguously lower productivity and profitability for SOEs than private firms, with such differences persisting through the recent period.

Table 4: Robustness: TFPR Gaps with Gross Output Production Function

	(1) ln(TFPR)	(2) ln(ARPK)	(3) ln(ARPL)	(4) ln(ARPX)
SOE	-0.073***	-0.225**	0.227***	-0.063***
	(0.010)	(0.09)	(0.043)	(0.019)
Other	-0.025**	-0.062	-0.03	-0.024
	(0.010)	(0.066)	(0.039)	(0.015)
Observations	43,770	43,770	43,770	43,770
R^2	0.80	0.20	0.24	0.11

Notes: the omitted group in all regressions are POEs. All regressions include sector x year fixed effects. Standard errors clustered at sector level.

3.4. SOEs and Interest Rates

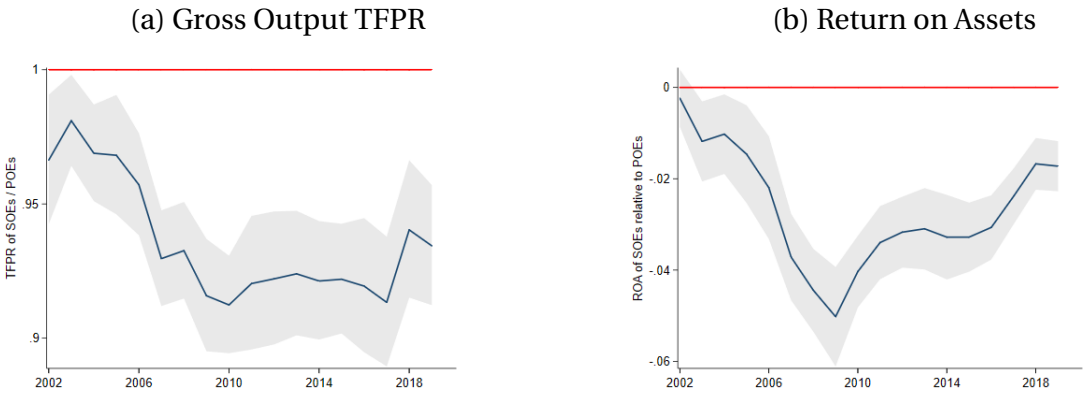
We documented in this section that SOEs are considerably more capital intensive than private firms in the same industry, and have a significantly lower capital productivity. One possible explanation for this is simply that SOEs happen to use a more capital intensive technology than private firms. However, it isn't clear why there would be such large systematic technological differences across firms with different ownership types. An alternative explanation is that SOEs have better access to low-interest loans and bank credit, which enables them to invest massively in fixed assets. Such favorable access to credit and loans for SOEs has been well documented and discussed in the literature (Li, Meng, Wang and Zhou, 2008; Boyreau-Debray and Wei, 2005; Zheng, Storesletten and Zilibotti, 2011; Wu, Firth and Rui, 2014), while more recently Bai et al. (2018) and Har-

Table 5: Robustness: Other Measures of Firm Performance

	(1) Net Profits / Revenues	(2) EBITDA / Revenues	(3) Return on Assets
SOE	-0.036***	-0.036***	-0.029***
	(0.007)	(0.010)	(0.003)
Other	-0.016**	-0.015**	-0.013***
	(0.006)	(0.006)	(0.003)
Observations	43,066	43,221	42,896
R^2	0.1	0.13	0.15

Notes: the omitted group in all regressions are POEs. All regressions include sector x year fixed effects. Standard errors clustered at sector level.

Figure 9: Robustness: Productivity and Profitability Gaps Over Time



Notes: sub-figure (a) plots the SOE-POE gross output based TFPR gap over time. Sub-figure (b) plots SOE-POE difference in return on assets (measured as net profits over total assets. Data for both sub-figures is from Wind.

rison et al. (2019) document that industrial SOEs tend to have lower effective interest rates than private firms.

We follow Harrison et al. (2019) to explore the interest rates faced by listed firms, by constructing effective interest rates as total interest expenses over total liabilities.[14] We only have data on interest expenses for 2018 and 2019, so we restrict our analysis to these years. First of all, we document in column (1) of Table 6 that firms with more leverage (as measured by total liabilities over total assets) tend to have higher effective interest rates; a 1 percentage point increase in leverage is associated with 2.7 percentage point higher effective interest rates.[15] This should be unsurprising given that higher leverage is likely to be associated with higher repayment risk. In addition, as has been widely documented already, SOEs tend to have considerably higher leverage than private firms (8.2 percentage points higher), shown in column (2). In column (3) we run a simple regression of effective interest rates on an SOE dummy and sector-year fixed effects. We find that SOEs have 0.2 percentage point lower interest rates than private firms, though the relationship is not particularly strong with a standard error of 0.1. However, in column (4) we additionally control for leverage, and find larger and more statistically significant differences between SOEs and private firms of 0.43 percentage points. This is a large difference, given that average effective interest rates are 1.94 percentage points. This implies that SOEs could be facing an effective cost of capital almost 25% lower than private firms.

As we will see in the next section, such differences in factor prices can distort the optimal allocation of capital and explain the average product of capital differences across firms that we measure in this section.

4. Macroeconomic Implications

4.1. A Model of Misallocation

Our model follows the setup from Hsieh and Klenow (2009). The economy has S sectors, L workers and an exogenous capital stock K. There are an exogenous number of heterogeneous firms N_s operating in each sector producing different varieties of goods. Rather than specifying preferences for each variety, we use the standard trick of aggregating varieties within a sector into a sectoral good, and aggregating sectoral goods into a final aggregate consumption good. A representative firm produces the final good Y in a perfectly competitive final output market. This final good is produced using value-added

[14] Results are similar if we consider current liabilities or total debt.
[15] We trim the 1% tails of leverage.

Table 6: SOEs and Effective Interest Rates

	(1) Interest Rate	(2) Leverage	(3) Interest Rate	(4) Interest Rate
SOEs		0.085***	-0.002*	-0.0043***
		(0.012)	(0.001)	(0.001)
Leverage	0.0272***			0.0295***
	(0.0024)			(0.0024)
Observations	5,724	5,724	5,724	5,724
R^2	0.19	0.21	0.088	0.21

Notes: All regressions include sector x year fixed effects. Data is for 2018 and 2019 only, as these are the only years for which we observe interest expenses. The Leverage variable is constructed as total liabilities over total assets, while the interest rate variable is constructed as interest expenses over total liabilities. Standard errors clustered at sector level.

Y_{st} from each sector s with a Cobb-Douglas production technology. We normalize the price of the final good P to 1.

$$Y = \prod_{s=1}^{S} Y_s^{\theta_s} \quad \text{where} \quad \sum_{s=1}^{S} \theta_s = 1 \tag{5}$$

Sectoral output Y_s is a CES aggregate of the value-added of firms in each sector:

$$Y_s = \left(\sum_{i=1}^{N_s} Y_{si}^{1-\frac{1}{\sigma}} \right)^{\frac{1}{1-\frac{1}{\sigma}}} \tag{6}$$

We denote by P_s the price index of output from sector s. Firms have idiosyncratic productivity draws A_{si}, and produce output Y_{si} using a Cobb-Douglas technology in capital and labor:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s} \quad \text{where} \quad 0 < \alpha_s < 1 \tag{7}$$

The output elasticities α_s are sector-specific, but common across firms within a sector. Firms are monopolistically competitive, and we invert Equation 6 to derive their downward sloping demand curve, given by $Y_{si} = Y_s \left(\frac{P_{si}}{P_s} \right)^{-\sigma}$. Firms take P_s and Y_s as exogenous. Firms also face idiosyncratic labor distortions τ_{si}^L and capital distortions τ_{si}^K.

They maximize profits Π_{si} taking input prices as given.

$$\Pi_{si} = P_{si}Y_{si} - (1+\tau_{si}^{L})wL_{si} - (1+\tau_{si}^{K})rK_{si} \tag{8}$$

We can derive the following expression for aggregate value-added:

$$Y = AK^{\widetilde{\alpha}}L^{1-\widetilde{\alpha}} \quad \text{where} \quad \widetilde{\alpha} \equiv \sum_{s=1}^{S} \alpha_s \theta_s \tag{9}$$

where A is a standard development accounting measure of aggregate TFP. A is given by:

$$A = \overline{R} \times \prod_{s=1}^{S} A_s^{\theta_s}. \tag{10}$$

where $\overline{R}$ captures the effect of the *sectoral* distortions τ_s^L and τ_s^K.[16] These sectoral distortions lead to cross-sector misallocation which lowers aggregate TFP, however these are not the focus of this paper.[17] A_s is sectoral TFP and is defined as:

$$A_s \equiv \frac{Y_s}{K_s^{\alpha_s} L_s^{1-\alpha_s}} \tag{11}$$

A_s is a function of firm-level productivities and distortions:

$$A_s = \left[\sum_{i=1}^{N_s} A_{si}^{\sigma-1} \left(\frac{\left(1+\tau_{si}^{L}\right)^{1-\alpha_s} \left(1+\tau_{si}^{K}\right)^{\alpha_s}}{\left(1+\tau_{s}^{L}\right)^{1-\alpha_s} \left(1+\tau_{s}^{K}\right)^{\alpha_s}} \right)^{1-\sigma} \right]^{\frac{1}{\sigma-1}} \tag{12}$$

Going one step further, we can decompose A_s into an allocative efficiency term and a residual term which is not a function of idiosyncratic distortions.

[16] These are defined as the value-added weighted harmonic mean of the idiosyncratic distortions; $(1+\tau_s^L) \equiv \frac{1}{\sum_i \frac{P_{si}Y_{si}}{P_sY_s} \frac{1}{1+\tau_{si}^L}}$ and similarly for τ_s^K.

[17] See Brandt et al. (2013) for a quantitative evaluation of sectoral distortions due to SOEs in China.

$$A_s = \underbrace{\left[\frac{1}{N_s}\sum_{i}^{N_s}\left(\frac{A_{si}}{\widetilde{A}_s}\right)^{\sigma-1}\left(\frac{\left(1+\tau_{si}^{L}\right)^{1-\alpha_s}\left(1+\tau_{si}^{K}\right)^{\alpha_s}}{\left(1+\tau_{s}^{L}\right)^{1-\alpha_s}\left(1+\tau_{s}^{K}\right)^{\alpha_s}}\right)^{1-\sigma}\right]^{\frac{1}{\sigma-1}}}_{AE_s=\text{Allocative Efficiency}} \times \underbrace{\widetilde{A}_s}_{\text{Residual Productivity}}.$$

This is a 'distance from the frontier' concept of allocative efficiency. AE_s is maximized and equal to 1 when there is no variation in the distortions across firms ($\tau_{si}^{L} = \tau_{s}^{L}$ and $\tau_{si}^{K} = \tau_{s}^{K}\ \forall i$). It is therefore well-suited to measuring the level effect of distortions on TFP, in contrast to alternative decompositions which capture the contribution of allocative efficiency growth to TFP growth (Baqaee and Farhi, 2019). Residual productivity captures all other aspects of sectoral productivity. It is increasing in the number of firms (due to love of variety), the dispersion in productivity A_{si} (because resources get reallocated to more productive firms with $\sigma > 1$) and in average firm productivity.

4.2. Inferring Allocative Efficiency

In this subsection we show how to infer allocative efficiency in the absence of measurement error.[18] In the previous section we defined TFPR, ARPK and ARPL as the average revenue products of the firm. Given the model assumptions of CES demand and constant returns to scale, these average products are proportional to the marginal products of the firm, and thereby proportional to the idiosyncratic distortions:

$$\text{ARPK}_{si} \equiv \frac{Y_{si}}{K_{si}} \propto (1+\tau_{si}^{K}) \tag{13}$$

$$\text{ARPL}_{si} \equiv \frac{Y_{si}}{L_{si}} \propto (1+\tau_{si}^{L}) \tag{14}$$

$$\text{TFPR}_{si} \equiv \frac{Y_{si}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}} \propto (1+\tau_{si}^{K})^{\alpha_s}(1+\tau_{si}^{L})^{1-\alpha_s} \tag{15}$$

Note that TFPR, ARPL and ARPK are not proportional to the firm's underlying 'physical productivity' A_{si}. Rather we can construct TFPQ which is proportional to A_{si} but independent of the distortions.

$$\text{TFPQ}_{si} \equiv \frac{(Y_{sit})^{\frac{\sigma}{\sigma-1}}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}} \propto A_{si}$$

[18] Bils et al. (2020) consider the case with additive measurement error. We will to this issue in Section 5.

We can then infer sectoral allocative efficiency using the following expression:

$$AE_s = \left[\sum_{i=1}^{N_s} \left(\frac{\text{TFPQ}_{si}}{\text{TFPQ}_s}\right)^{\sigma-1} \left(\frac{\text{TFPR}_{si}}{\text{TFPR}_s}\right)^{1-\sigma}\right]^{\frac{1}{\sigma-1}} \tag{16}$$

$$\text{where} \quad \text{TFPQ}_s = \left[\sum_{i=1}^{N_s} \text{TFPQ}_{si}^{\sigma-1}\right]^{\frac{1}{\sigma-1}} \tag{17}$$

$$\text{and} \quad \text{TFPR}_s = \text{ARPL}_s^{1-\alpha_s} \text{ARPK}_s^{\alpha_s} \tag{18}$$

$$\text{and} \quad \text{ARPK}_s = \left[\sum_{i=1}^{N_s} \frac{P_{si}Y_{si}}{P_sY_s} \frac{1}{\text{ARPK}_{si}}\right]^{-1} \tag{19}$$

Aggregating across sectors we obtain aggregate allocative efficiency:

$$AE_t = \prod_{s=1}^{S} AE_{st}^{\theta_{st}} \tag{20}$$

Following Bils et al. (2020) and based on evidence from Redding and Weinstein (2019), we pick a value of $\sigma = 4$ for the elasticity of substitution across plants. Misallocation is amplified under higher values of this elasticity. As previously described, we infer α_s from sectoral cost shares. We allow the sectoral shares of aggregate value-added θ_{st} to vary across years, which we recover from the data. In addition to trimming the 1% tails of TFPR, we trim the 1% tails of TFPQ for the allocative efficiency calculations.

So what determines the losses due to misallocation? Heterogeneity across firms in marginal revenue products implies the potential for a productive reallocation of inputs from low marginal product firms to high marginal product firms. A critical feature of mapping this model to the data is the assumption of constant returns to scale and CES demand, which implies that *average* revenue products are proportional to *marginal revenue products*. Dispersion in TFPR across firms is therefore associated with lower allocative efficiency. To see this more clearly, it helps to make the additional assumptions that the firm distortions are the same across inputs $\tau^K = \tau^L$, and that both the idiosyncratic distortions and productivities A are jointly lognormally distributed. We then obtain the standard formula for allocative efficiency from Hsieh and Klenow (2009):

$$\ln(\text{AE}_s) = -\frac{\sigma}{2}\text{var}_s(\ln(\text{TFPR}_{si})) \tag{21}$$

In this simple case, dispersion in TFPR is a sufficient statistic for misallocation. A given dispersion in TFPR leads to worse allocative efficiency for higher values of σ. This is because, with more substitutability between the varieties of goods produced by firms, a given distortion will lead to a larger change in inputs, and hence worse allocative efficiency. Moving back to a slightly more general case, allowing for $\tau^K \neq \tau^L$ but maintaining the lognormality assumption, we obtain the following expression for ln(AE):

$$\begin{aligned}\ln(\text{AE}_s) = &-(1-\alpha_s)[1+(\sigma-1)(1-\alpha_s)]\text{var}_s(\ln(\text{ARPL}_{si})) \\ &-\alpha_s[1+(\sigma-1)\alpha_s]\text{var}_s(\ln(\text{ARPK}_{si})) \\ &-(\sigma-1)(1-\alpha_s)\alpha_s\text{cov}_s(\ln(\text{ARPL}_{si}),\ln(\text{ARPK}_{si}))\end{aligned} \tag{22}$$

It is clear therefore, that not only does dispersion in the average product of capital and labor matter individually, but so does their covariance. The higher the covariance between the capital and labor distortions, the more extreme the tails of the distortions will be, amplifying misallocation.

In the next section, we will use the allocative efficiency formulas derived here to consider the impact of counterfactual reforms which close the TFPR gap between SOEs and private firms. However, with the assumption of lognormality, the covariance of distortions and firm productivities does not matter. We therefore also use the most general formula in Equation 16 for our main counterfactuals.

5. Gains from SOE Reforms

We showed in Section 3. that SOEs have lower TFPR than private firms, especially lower capital productivity. How should we interpret these differences in average products? The model in Section 4. interprets such differences as reflecting underlying distortions τ^K and τ^L which lower allocative efficiency. However, there are alternative possible explanations which don't imply any true resource misallocation. These include SOEs using a more capital-intensive technology to private firms, SOEs having larger overhead costs (as a share of their inputs), measurement error leading SOEs understating their revenues or overstating their costs or private firms being younger and slowly adjusting their inputs to reach their optimal size.

However, it isn't clear that non-policy explanations can rationalize our findings. Firstly, there is no a priori reason to believe that SOEs are using an inherently more capital-intensive technology than private firms. Secondly, given the high accounting standards facing all listed firms, there does not seem to be a reason to think that measurement er-

ror should be worse for SOEs than private firms.[19] Thirdly, overhead costs also seem like an unlikely explanation for our findings. De Loecker et al. (2020) measure overheads as Selling, General Administrative Expenses, which includes selling expenses (salaries of sales personnel, advertising, rent,...), general operating expenses, and administration (executive salaries, general support related to the overall administration). These should be included either in our measure of labor or our measure of intermediates, which would imply weakly larger gaps in labor productivity than capital productivity. Finally, in Table 3 we showed that there are still large TFPR gaps between SOEs and private firms even after controling for firm age. In addition, TFPR gaps remained large even in the late 2010s, at which point the vast majority of private firms in our database had been listed for more than five years, suggesting that post-listing transitions can't explain our findings.

While we cannot categorically rule out the above explanations, there is ample evidence of large policy distortions favoring SOEs in China. For example, in Section 3. we documented that SOEs may be facing 25% lower effective interest rates than private firms. This alone could account for half the 50% gap in capital productivity we document in 2019. In contrast to capital productivity, we don't find large gaps in labor productivity between SOEs and private firms. This aligns with findings in Hsieh and Song (2015) who document that ARPL gaps between SOEs and private industrial firms closed throughout the 2000s. This is consistent with the interpretation that pressure on SOEs to maintain surplus labor declined over this period.

In our counterfactual exercises, we treat the entirety of the TFPR gap between SOEs and POEs as reflecting distortions. We consider the gains from a hypothetical reform which equalizes the average TFPR of SOEs and private firms within each sector. Such reforms could include requiring banks to carry higher risk weights on SOE loans, or more generally pushing for competitive neutrality. In our first exercise, we first estimate the SOE TFPR gap in each sector and year, and then shift the whole distribution of SOE TFPR to close that gap.[20] In order to convey the intuition of our results, we use the log-normal allocative efficiency formula. In Table 7 we report across different windows the size of the TFPR gap, pre- and post-reform TFPR dispersion, pre- and post-reform allocative efficiency, and finally the gains from reform. Closing the TFPR gap between SOEs and POEs reduces overall TFPR dispersion, which as previously explained implies improvements in allocative efficiency. Larger TFPR gaps are associated with larger declines in dispersion, and larger gains from reform. These gains peak between 2007 and 2016, averaging 9.7%, before then declining to 6.0% between 2017-2019.

[19] SOEs would need to be systematically overreporting the value of their fixed assets. Underreporting of revenues would imply similar differences in capital and labor productivity.

[20] We only do this for sectors in which the SOE-POE TFPR gap is negative.

Table 7: AE Gains from Closing SOE-POE TFPR Gap with Lognormal Assumptions

	(1) 2002-2006	(2) 2007-2011	(3) 2012-2016	(4) 2017-2019
SOE-POE TFPR Gap	-0.131	-0.386	-0.350	-0.262
Baseline σ^2_{TFPR}	0.459	0.420	0.362	0.334
Counterfactual σ^2_{TFPR}	0.475	0.461	0.395	0.356
Baseline AE	0.398	0.342	0.369	0.400
Counterfactual AE	0.411	0.375	0.404	0.424
Counterfactual TFP Gains	3.18%	9.87%	9.57%	6.01%

Notes: The first row of this table plots our estimates of the average SOE-POE TFPR gap in the time-period given by the corresponding column. The second and third rows show TFPR dispersion in the data and in our counterfactual. The fourth and fifth columns show baseline allocative efficiency as well as allocative efficiency following our counterfactual reform. The last row shows the counterfactual TFP gains from this reform.

In our second exercise, we evaluate the gains from reforms using the full model rather than the lognormal formula. In this case, we once again start by estimating the SOE TFPR gap in each sector and year. However, we only shift the SOE TFPR distribution for firms with below median TFPR (relative to the sector-year average). This corresponds more closely to potential reforms which would remove distortions from SOEs with particularly low TFPR, and which therefore seem to be benefiting low effective interest rates or other implicit subsidies. It also avoids that we increase the TFPR of SOEs in the right tail of the TFPR distribution, where measurement error tends to be more of a concern (Bils et al., 2020).[21] Note that this implies that our counterfactual reform does not fully close the TFPR gap between SOEs and private firms.

In Figure 10 Panel (a), we report the level of allocative efficiency for listed firms as implied by the model. Though the model suggests that aggregate TFP could be close to doubled if all distortions were to be removed, these numbers should not be treated as the potential gains from large-scale reforms. A lot of TFPR dispersion could be due to measurement error, model misspecification, adjustment costs, overheads or markup

[21] Allocative efficiency is quite sensitive to the right tail of TFPR dispersion, as shown in White and Rotemberg (2020). As discussed in the previous section, SOEs are present in the far right tail of the TFPR dispersion almost to the same degree as private firms. We don't want changes in the TFPR of these firms to drive our results, especially given that measurement error tends to be particularly large in the right tail of the TFPR distribution (Bils et al., 2020).

Figure 10: Gains from Closing SOE TFPR Gaps

(a) Allocative Efficiency



(b) TFP Gains



Notes: sub-figure (a) plots allocative efficiency over time, both in the baseline and in the counterfactual reform scenario where TFPR gaps are closed. The counterfactual scenario shifts SOEs with below median TFPR by the size of the sectoral productivity gap (conditional on the gap being negative), but holds fixed the TFPR of SOEs with above median TFPR. Sub-figure (b) plots the implied TFP gains under the counterfactual scenario, and can be calculated as $100x(\frac{\text{AE reform}}{\text{AE baseline}} - 1)$. Data for both sub-figures is from Wind.

variation. However, the gains from closing the SOE-POE TFPR gaps shown in Figure 10 Panel (b) don't rely on the exact level of allocative efficiency being accurate, but rather on our estimated TFPR gaps in every sector. Indeed, even if a substantial share of TFPR dispersion was purely measurement error, we would find similar gains provided our estimated TFPR gaps are accurate. As previously discussed, these gaps are unlikely to be driven by measurement error or transitory factors such as adjustment costs. We therefore avoid many of the issues that typically are problematic in the misallocation literature.[22] We find large gains on average over our sample period of 8.2%. These gains decrease to 5.8% in 2019, given the slightly declining SOE TFPR gap.[23]

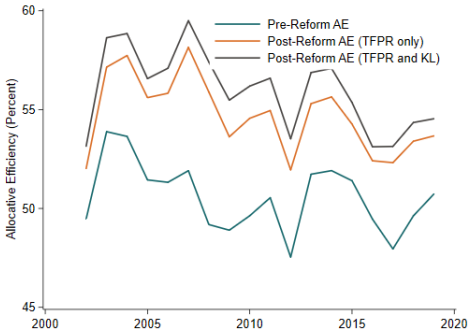
Our empirical results showed that the SOE capital productivity gap is considerably larger than the SOE labor productivity gap, implying that SOEs are considerably more capital intensive than private firms. In our final counterfactual, we consider a reform which not only closes the average TFPR gap within each sector, but also equalize the

[22] See Haltiwanger, Kulick and Syverson (2018) for a broad discussion of these issues, including the role of model misspecification. One type of misspecification is that production might be CES rather than Cobb-Douglas. Berkowitz et al. (2017) estimate elasticities between capital and labor that are greater than 1. For a given distortion, a higher elasticity of substitution implies larger losses due to misallocation (as inputs shift more across firms), and therefore our estimated gains from reform would be a lower bound. Deviations from constant returns to scale have ambiguous effects on the losses from misallocation, as discussed in Ho and Ruzic (2020).

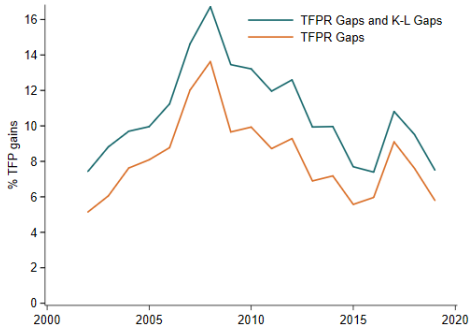
[23] If we only shift the TFPR of SOEs in the bottom quartile of the TFPR distribution, we find smaller gains of around 5% in 2019. If we shift the TFPR of SOEs outside the top quartile of the TFPR distribution, we find slightly larger gains of 7% in 2019).

Figure 11: Gains from Closing SOE TFPR Gaps and Equalizing Capital-Intensity

(a) Allocative Efficiency



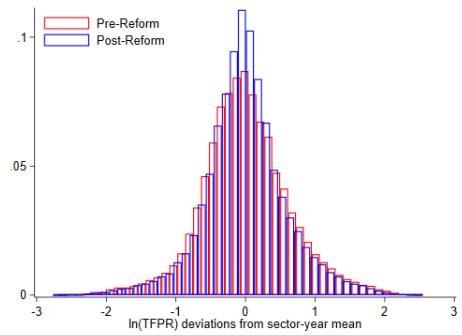
(b) TFP Gains



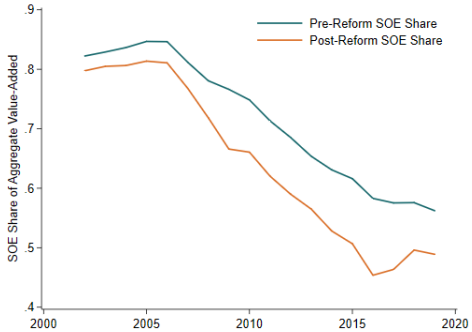
Notes: sub-figure (a) plots allocative efficiency over time, both in the baseline and in two counterfactual reform scenarios. In the first reform scenario, average TFPR gaps are closed while in the second scenario both TFPR gaps and K/L gaps are closed. The counterfactual scenario shifts SOEs with below median TFPR by the size of the sectoral productivity gap (conditional on the gap being negative), but holds fixed the TFPR of SOEs with above median TFPR (similarly with the KL gaps). Sub-figure (b) plots the implied TFP gains under the counterfactual scenarios, and can be calculated as $100\text{x}(\frac{\text{AE reform}}{\text{AE baseline}} - 1)$. Data for both sub-figures is from Wind.

Figure 12: Pre- and Post-Reform Distribution of TFPR and SOE Value-Added Share

(a) TFPR Distribution



(b) SOE Share of Aggregate Value-Added



Notes: sub-figure (a) plots the distribution of TFPR (residualized on sector-year fixed effects) pre- and post-reform. Sub-figure (b) plots the SOE share of aggregate value-added in the pre-reform scenario and post-reform scenario. Data for both sub-figures is from Wind.

average capital intensity of SOEs with that of POEs.[24] The change in allocative efficiency and corresponding gains from reform are shown in Figure 11. We find larger average gains relative to a reform which only equalizes TFPR, averaging 10.7% over the whole sample period, and 7.5% in 2019. Panel (a) in Figure 12 shows the distribution of TFPR pre- and post-reform. It is clear that the distribution is more compressed overall, though there remains significant dispersion overall. Panel (b) shows that the SOE share of aggregate value-added would fall considerably following the reform, from 58% to around 50% in 2019, as inputs reallocate from low TFPR SOEs to high TFPR private firms.

6. Conclusion

We document large and persistent productivity gaps between SOEs and private firms in China, between 2002 and 2019. These gaps are accounted for to a large extent by a low capital productivity of SOEs. We provide evidence that listed SOEs benefit from lower effective interest rates than private firms, contributing to a distorted capital-labor ratio. Our results suggest that there remains significant scope for gains from SOE reform among listed firms in China. In particular, equalizing credit access for both SOEs and private firms could contribute to productivity-enhancing capital reallocation. Though we do not have any data on firm-level productivity outside of listed firms for more recent years, industry-level data still indicates that return on assets in 2019 were 3.5% for SOEs and 6.3% for private firms, suggesting that our findings may also be extrapolated more broadly for Chinese firms. However, as Chinese SOEs are also used to provide many social and non-economic functions, complementary reforms would also be important to ensure that the gains are inclusively distributed, and workers do not lose out.

[24] Again, we only shift the capital intensity of SOEs with a greater than median capital intensity, and only for sectors in which SOEs are on average more capital intensive than private firms.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazier, "Identification Properties of Recent Production Function Estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Bai, Yan, Dan Lu, and Xu Tian, "Do Financial Frictions Explain Chinese Firms' Saving and Misallocation?," 2018. NBER Working Paper No. 24436.

Baqaee, David and Emmanuel Farhi, "Productivity and Misallocation in General Equilibrium," 2019.

Berkowitz, Daniel, Hong Ma, and Shuichiro Nishioka, "Recasting the Iron Rice Bowl: The Reform of China's State-Owned Enterprises," *The Review of Economics and Statistics*, 2017, *96* (4), 735.

Bils, Mark, Peter J. Klenow, and Cian Ruane, "Misallocation or Mismeasurement?," 2020.

Boeing, Philipp, Elisabeth Mueller, and Philipp Sandner, "China's RD explosion—Analyzing productivity effects across ownership types and over time," *Research Policy*, 2016, *45* (1), 159–176.

Boyreau-Debray, Genevieve and Shang-Jin Wei, "Pitfalls of a State-Dominated Financial System: The Case of China.," 2005. NBER Working Paper No. 11214.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 2012, *97*, 339–351.

_ , Trevor Tombe, and Xiadong Zhu, ""Factor market distortions across time, space and sectors in China," *Review of Economic Dynamics*, 2013, *16* (1), 39–58.

Chen, Donghua, Dequan Jiang, Alexander Ljungqvist, and Mingming Zhou, "State Capitalism vs. Private Enterprise," 2017.

Chen, Wei, Xilu Chen, Chang-Tai Hsieh, and Zheng Song, "A Forensic Examination of China's National Accounts," *Brookings Papers on Economic Activity*, Spring 2019.

Chen, Yuyu, Mitsuru Igami, Masayuki Sawada, and Mo Xiao, "Privatization and Productivity in China," 2020.

David, Joel M. and Venky Venkateswaran, "The Sources of Capital Misallocation," *American Economic Review*, 2019, *109* (7), 2531–2567.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, March 2008, *98* (1), 394–425.

Haltiwanger, John, Robert Kulick, and Chad Syverson, "Misallocation Measures: The Distortion That Ate the Residual," 2018.

Harrison, Ann, Marshal Meyer, Peichun Wang, Linda Zhao, and Minyuan Zhao, "Can a Tiger Change Its Stripes? Reform of Chinese State-Owned Enterprises in the Penumbra of the State," 2019. NBER Working Paper 25475.

Ho, Sui-Jade and Dimitrije Ruzic, "Returns to Scale, Productivity Measurement, and Trends in U.S. Manufacturing Misallocation," 2020.

Hopenhayn, Hugo A, "Firms, misallocation, and aggregate productivity: A review," *Annu. Rev. Econ.*, 2014, *6* (1), 735–770.

Hsieh, Chang-Tai and Michael Song, "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China," *Brookings Papers on Economic Activity*, Spring 2015, pp. 295–346.

__ and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Li, Hongbin, Lingsheng Meng, Qiang Wang, and Li-An Zhou, "Political Connections, Financing and Firm performance: Evidence from Chinese Private Firms," *Journal of Development Economics*, 2008, *87*, 283–299.

Lin, Lauren Yu-Hsin and Yun-Chien Chang, "Do State-Owned Enterprises Have Worse Corporate Governance? An Empirical Study of Corporate Practices in China," 2019.

Loecker, Jan De, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

Olley, G Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Poncet, Sandra, Walter Steingress, and Hylke Vandebussche, "Financial constraints in China: Firm-level evidence," *China Economic Review*, 2010, *21*, 411–422.

Redding, Stephen J. and David E. Weinstein, "Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences," *Quarterly Journal of Economics*, 2019, *135* (1), 503–560.

Restuccia, Diego and Richard Rogerson, "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, October 2008, *11*, 707–720.

Rosen, Daniel H., Wendy Leutert, and Shan Guo, "Missing Link: Corporate Governance in China's State Sector," 2018.

Song, Zheng and Guiying Laura Wu, "Identifying Capital Misallocation," 2015.

White, T Kirk and Martin Rotemberg, "Measuring Cross-Country Differences in Misallocation," 2020.

Wu, Guiying Laura, "Capital Misallocation in China: Financial Frictions or Policy Distortions?," *Journal of Development Economics*, 2018, *130*, 203–223.

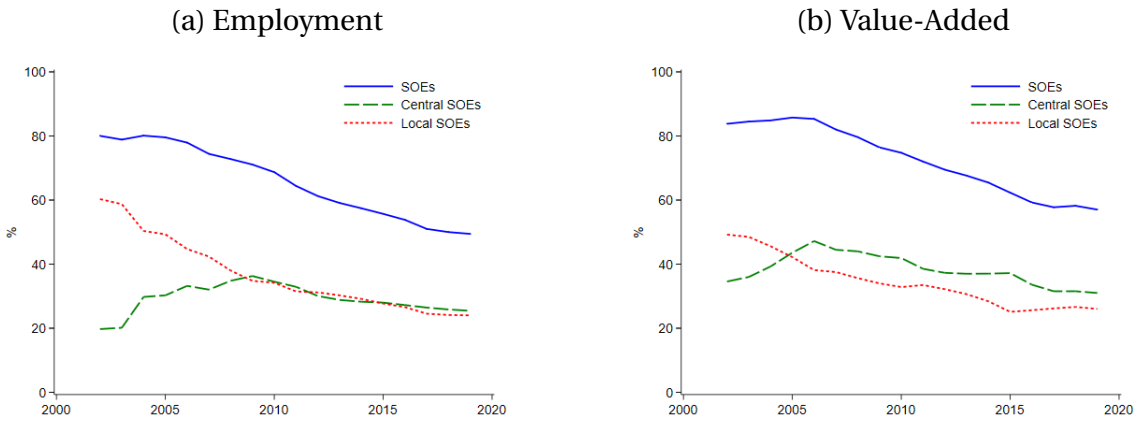
Wu, Wenfeng, Michael Firth, and Oliver M. Rui, "Trust and the Provision of Trade Credit," *Journal of Banking Finance*, 2014, *39*, 146–159.

Zheng, Song, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing Like China," *American Economic Review*, February 2011, *101*, 196–233.

Appendix
For Online Publication

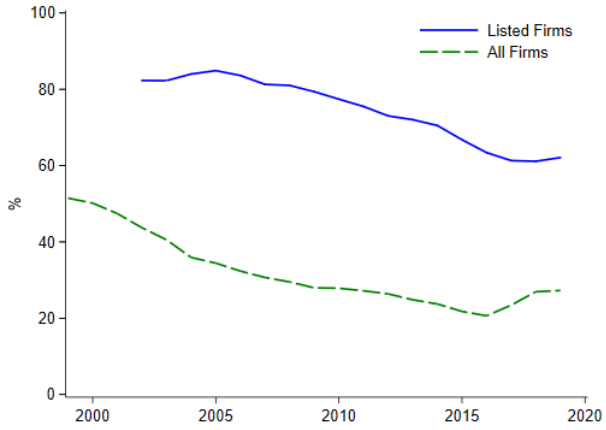
A Figures

Figure 13: SOE Share of Employment and Value-Added



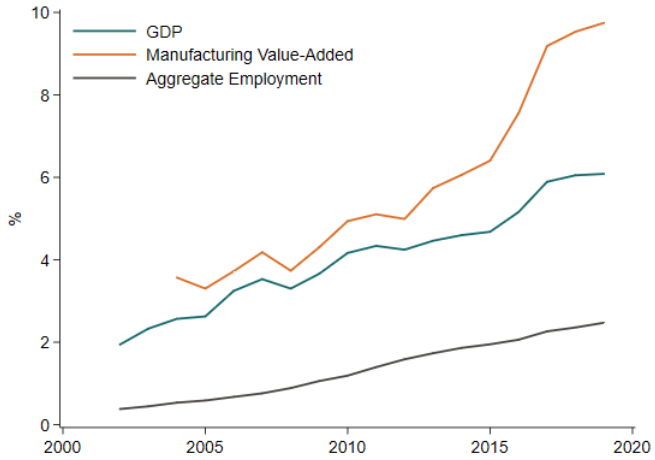
Notes: sub-figure (a) plots the employment share of all SOEs, local SOEs and central SOEs in total listed firm employment. Sub-figure (b) plots the value-added share of all SOEs, local SOEs and central SOEs in total listed firm value-added. Data for both sub-figures is from Wind.

Figure 14: SOE Share of Revenues, Listed and Aggregate



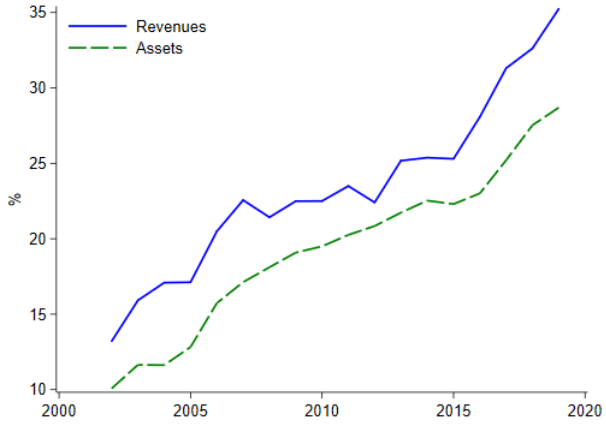
Notes: This figure plots over time the SOE share of revenues for listed firms, and for all industrial firms. Data on listed firms is from Wind, and for all industrial firms is for CEIC.

Figure 15: Listed Firm Share of GDP and Employment



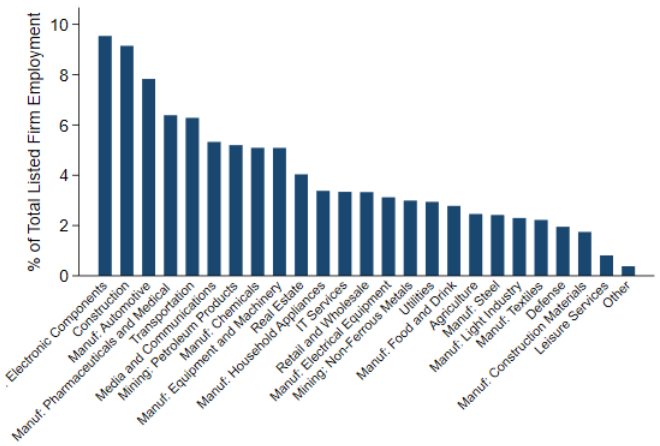
Notes: This figure plots over time the share of listed firm value-added in GDP separately for manufacturing and for all sectors, as well as the share of listed firm employment in aggregate employment. Data on listed firms is from Wind while data on aggregate GDP and employment is from the NBS.

Figure 16: Listed SOE Share of Aggregate SOE Revenues and Assets



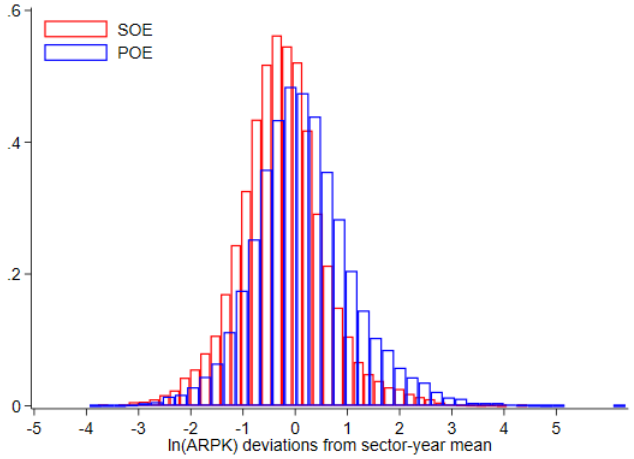
Notes: This figure plots over time the share of listed SOE revenues and assets in aggregate SOE revenues and assets. Data on listed firms is from Wind while data on aggregate SOE revenues and assets is from CEIC.

Figure 17: Sectoral Breakdown of Employment



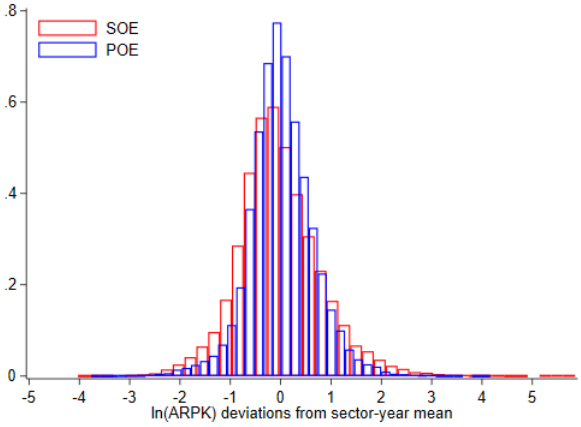
Notes: This figure plots the share of aggregate listed firm employment by sector in 2019. Data on listed firms is from Wind.

Figure 18: ARPK Distributions for SOEs and POEs



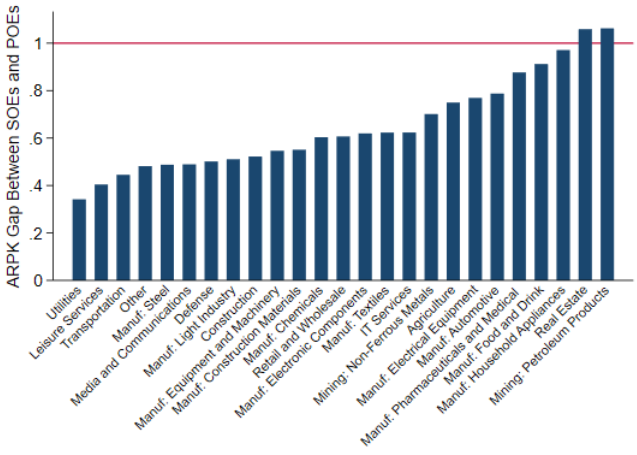
Notes: This figure plots the distribution of ln(ARPK) deviations from the sector-year average for SOEs and POEs separately, pooling data from 2002 to 2019. Data on listed firms is from Wind.

Figure 19: ARPL Distributions for SOEs and POEs



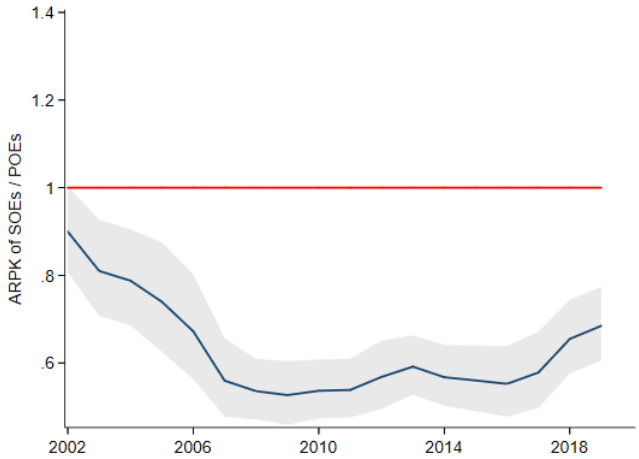
Notes: This figure plots the distribution of ln(ARPl) deviations from the sector-year average for SOEs and POEs separately, pooling data from 2002 to 2019. Data on listed firms is from Wind.

Figure 20: SOE-POE ARPK Gaps by Sector (2016-2019)



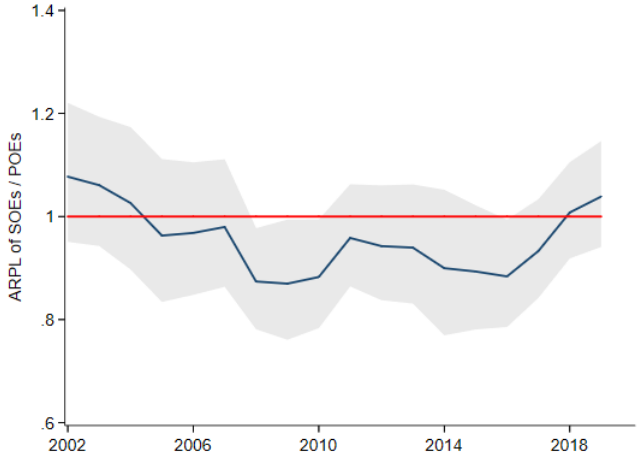
Notes: This figure plots the ARPK gap between SOEs and POEs by sector (1 means no gap). ARPK gaps are estimated from 2016 to 2019. Data on listed firms is from Wind.

Figure 21: SOE-POE Capital Productivity Gaps Over Time



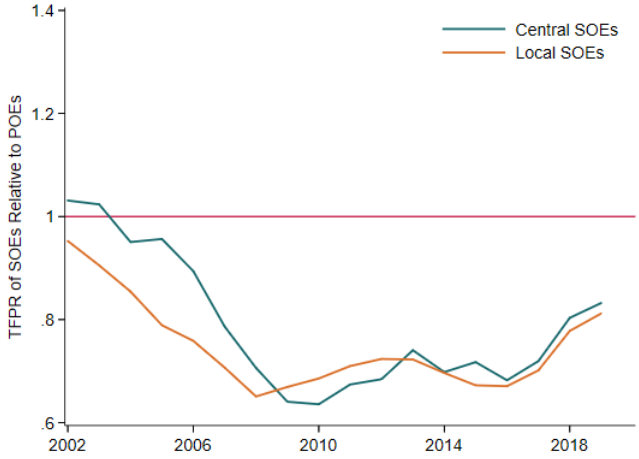
Notes: This figure plots the ARPK gap between SOEs and POEs between 2002 and 2019. These are obtained by regressing ln(ARPK) on an SOE dummy, controling for sector x year fixed effects. The coefficient on the dummy is then exponentiated. Data on listed firms is from Wind.

Figure 22: SOE-POE Labor Productivity Gaps Over Time



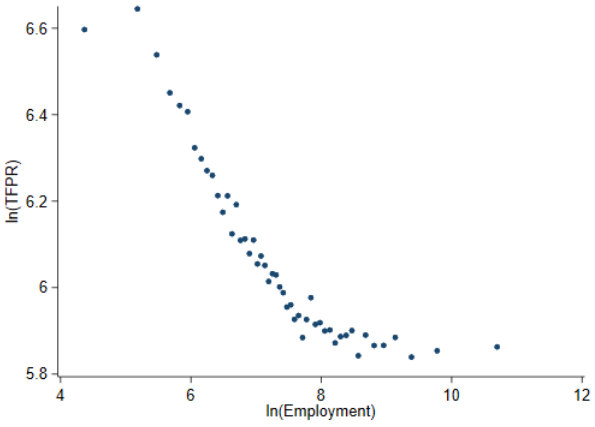
Notes: This figure plots the ARPL gap between SOEs and POEs between 2002 and 2019. These are obtained by regressing ln(ARPL) on an SOE dummy, controling for sector x year fixed effects. The coefficient on the dummy is then exponentiated. Data on listed firms is from Wind.

Figure 23: Local and Central SOE TFPR Gaps Over Time



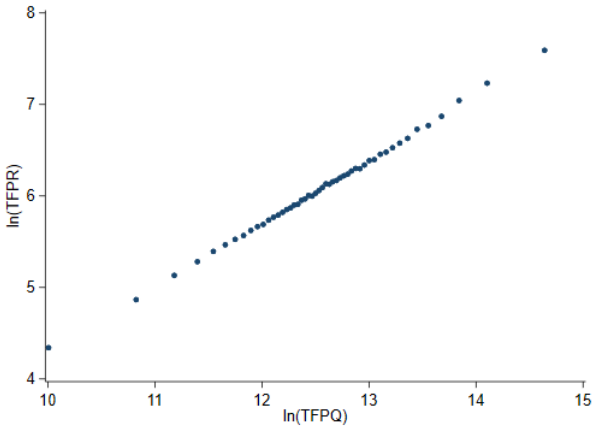
Notes: This figure plots the TFPR gap between SOEs and POEs between 2002 and 2019, separately for central and local SOEs. These are obtained by regressing ln(TFPR) on both a central SOE dummy and a local SOE dummy, controling for sector x year fixed effects. The coefficients are then exponentiated. Data on listed firms is from Wind.

Figure 24: TFPR vs. Employment



Notes: This figure plots a binned scatter plot of ln(TFPR) on ln(Employment). Both ln(TFPR) and ln(Employment) are first residualized on sector x year fixed effects. Data on listed firms is from Wind.

Figure 25: TFPR vs. TFPQ



Notes: This figure plots a binned scatter plot of ln(TFPR) on ln(TFPQ). Both ln(TFPR) and ln(TFPQ) are first residualized on sector x year fixed effects. Data on listed firms is from Wind.

B Tables

Table 8: Estimates of ARPK Gaps Between SOEs and POEs

	(1)	(2)	(3)	(4)
SOE	-0.51***		-0.41***	-0.20***
	(0.058)		(0.047)	(0.036)
Other	-0.17***	-0.10***	-0.10	-0.07
	(0.04)	(0.003)	(0.042)	(0.06)
Central SOE		-0.47***		
		(0.07)		
Local SOE		-0.53***		
		(0.059)		
Age			-0.02***	
			(0.003)	
ln(Employment)			-0.06***	
			(0.021)	
Sector x Year FEs	✓	✓	✓	✓
Firm FEs				✓
Observations	43,770	43,770	43,734	43,764
R^2	0.24	0.24	0.26	0.68

Notes: The dependent variable in all columns is ln(TFPR), and the omitted group are POEs. Standard errors clustered at sector level.

Table 9: TFPR Gaps by Broad Sector and Over Time

	2005-2009	2010-2014	2015-2019
Agriculture	-0.36	-0.46	-0.36
	(0.09)	(0.08)	(0.08)
Construction	-0.46	-0.37	-0.25
	(0.10)	(0.07)	(0.05)
Hotels & Catering	0.064	-0.29	-0.35
	(0.18)	(0.15)	(0.14)
IT	-0.43	-0.26	-0.27
	(0.05)	(0.03)	(0.03)
Manufacturing	-0.37	-0.39	-0.29
	(0.02)	(0.02)	(0.02)
Mining	0.01	-0.21	-0.25
	(0.06)	(0.05)	(0.04)
Other	-0.14	-0.37	-0.44
	(0.10)	(0.07)	(0.07)
Real Estate	0.07	0.28	0.21
	(0.11)	(0.09)	(0.07)
Transport, Storage & Post	-0.54	-0.52	-0.53
	(0.09)	(0.07)	(0.06)
Utilities	-0.37	-0.83	-0.77
	(0.08)	(0.06)	(0.05)
Wholesale & Retail Trade	-0.12	-0.08	-0.21
	(0.09)	(0.07)	(0.07)

Notes: This table reports the results of regressing ln(TFPR) on an SOE dummy controling for sector x year fixed effects. The sectors in the fixed effects are more disaggregated than the broad sectors reported in the table. Robust standard errors in brackets. All firm-level data is from Wind.