The State and China's Productivity Deceleration: Firm-level Evidence

Jorge, Alvarez, Tuo Chen, Grace Li

November 2017

Abstract

This paper documents the total factor productivity (TFP) growth path from 1998 to 2013 using both the aggregate data and the firm-level data of China. We find that the TFP growth is positive from 1998 to 2011 and then turn to flat and even negative. And careful comparison between the state-owned enterprises (SOEs) and private firms reveals that the slowing down of TFP growth of the SOEs is the major contributor to the TFP growth reversal of the whole manufacturing sector. Moreover, we rejected a possible explanation that manufacturing TFP growth slowing down may due to the sectoral factor after the decomposition analysis.

1 Stylized Facts in the Time Series of TFP

1.1 Data Description

We employ two sources of data in this paper, the Chinese Statistical Yearbook and the Chinese industrial Survey. Most of the the aggregate data is downloaded from the China Industry Statistical Yearbook, including the sales income, sales cost, number of people employed, and total asset. The wage bill is from the China Labor Statistical Yearbook. The two-digit sectoral price data is downloaded from the China Statistical Yearbook[1].

The firm-level data is from Chinese Industrial Survey (1998 - 2013). This dataset is widely used to study China related questions, including a seminal work by Chang-Tai Hsieh (2009). The main difference of our data set compared to previous ones is that the it has longer historical length. Usually people use Chinese Industrial Survey (1998 - 2007) because of the homogeneity of the annual dataset: each year has almost identical variables and a unique identifier. The data after 2007 suffers from several issues. First, the 2008 year data has no identifier and 2009 miss one-third of the identifiers. This complication make

[1]The China Statistical Yearbook data can be accessed through the website of National Bureau of Statistics of China, while the China Industry Statistical Yearbook and the China Labor Statistical Yearbook can be accessed through the website of China Knowledge Resource Integrated Database (http://www.cnki.net/).

it is impossible to include 2008 and 2009 data in the panel dataset. Therefore, we prepare two versions of data set. One version includes 2008 and 2009 data which we treat as cross-sectional dataset for the calculation of the yearly cross-sectional aggregate statistics. The other version excludes the data of these two years and is a panel dataset. Second, the variables to calculate the firm-level TFP is missing in the later years. A key variable value added, which serves as the output in the production function, is missing after year 2007. So we have to calculate our own measure of the TFP. We will show in the next subsection that is will not affect the measure of TFP growth much. Third, the 2010 year data has data quality issue so that we exclude this year from the dataset.

1.2 TFP Measurement Methodology

In this section, we show the different measures of TFP, and all the measures point to one fact that Chinese industrial TFP growth has slowed down in the later years of our dataset.

We start with the estimation method with the aggregate data. We assume the production function is Cobb-Douglas and constant return to scale. And we use the value added as the output, therefore the production function has two factor input: capital and labor. The labor share is calculated as the total wage bill over the value added and the capital share is 1 minus the labor share. The TFP is calculated as the Solow residual. The whole procedure can be presented in equations 1.

$$\begin{aligned}
&\mathrm{VA}_t = \text{Sales Income}_t - \text{Sales Cost}_t \\
&\text{Real VA}_t = \tfrac{\mathrm{VA}_t}{P_t} \\
&\alpha_t^L = \tfrac{w_t L_t}{\mathrm{VA}_t} \\
&\alpha_t^K = 1 - \alpha_t^L \\
&K_t = \tfrac{\text{Total Asset}_t}{P_t^K} \\
&A_t = \tfrac{\text{Real VA}_t}{L_t^{\alpha_t^L} K_t^{\alpha_t^K}}
\end{aligned} \tag{1}$$

where "VA" stands for the valued added, and "Real VA" is "VA" deflated by the producer price index P_t, capital measure K_t is "Total Asset" deflated by the price index of investment in fixed assets P_t^K, $w_t L_t$ is the total wage bill, α_t^L and α_t^K are respectively the labor share and capital share. Variables "Sales Income", "Sales Cost", P, P^K, L_t, $w_t L_t$ and "Total Asset" are all directly observed.

We can apply the aggregate TFP measure methodology to the firm-level data as well, with some modifications. We still assume the Cobb-Douglas production function and constant return to scale. The value added, the assets, labor and wage bill are all at the firm level. Ideally, the producer price index and investment price index should be at firm-level as well. However, the best we can do here is to use the two-digit sectoral producer price index and the same industrial level investment price index. The labor share and capital share here are measured at the two-digit sectoral level. The estimation can be summarized in equations 2

$$
\begin{aligned}
&\text{Real } va_{ist} = \tfrac{va_{ist}}{P_{st}} \\
&\alpha^L_{st} = \frac{\sum\limits_{i \in s} w_{ist} L_{ist}}{\sum\limits_{i \in s} \mathrm{va}_{ist}} \\
&\alpha^K_{st} = 1 - \alpha^L_{st} \\
&K_{ist} = \tfrac{\text{Fixed Asset}_{ist}}{P^K_t} \\
&A_{ist} = \tfrac{\text{Real } va_{ist}}{L_{ist}^{\alpha^L_{st}} K_{ist}^{\alpha^K_{st}}}
\end{aligned}
\tag{2}
$$

where va_{ist} stands for the valued added of firm i in the two-digit sector at time t, and "Real va" is "va" deflated by the two -digit sectoral producer price index P_{st}, capital measure K_{ist} is "Fixed Asset" deflated by the price index of investment in fixed assets P^K_t[2], $\sum\limits_{i \in s} w_{ist} L_{ist}$ is the total wage bill of the firms within two-digit sector s, α^L_{st} and α^K_{st} are respectively the labor share and capital share in sector s. Variables va_{is}, P_s, P^K, L_{is}, $w_{is}L_{is}$ and Total Asset$_{is}$ are all directly observed.

While the estimation of the TFP in equations 2 does not use much feature of the firm-level data, we can measure the firm-level TFP with regressions. In this paper, we will employ the methodologies in Olley and Pakes (1996), Levinsohn and Petrin (2003) and De Loecker (2011).

Both Olley and Pakes (1996) and Olley and Pakes (1996) try to solve the endogeneity of input choice (or simultaneity bias) in the sense that the the firm-level productivity will be correlated with the input choice such as capital. Although the productivity is unobservable to the econometricians, a specific firm may have private information about its own productivity and makes the decision for the scale of input accordingly. This correlation between the input and the productivity makes the OLS estimates biased.

The key assumption made in Olley and Pakes (1996) is that the firm-level productivity is a function of investment and this function is invertible. Olley and Pakes (1996) also assumes the Cobb-Douglas production function, but it no longer assumes the constant return to scale. More specifically, we want to estimate the coefficients of equation 3

[2] We try the total asset as well, and the result does not change for all the analysis.

$$y_{ist} = \beta^l l_{ist} + \beta^k k_{ist} + \Sigma_s \delta_s Ind_s + \Sigma_t \delta_t year_t + \beta^{\mathrm{SOE}} SOE_{ist} + \beta^{exp} exp_{ist}$$
$$+\omega_{ist}(age_{ist}, k_{ist}, inv_{ist}) + \varepsilon_{ist} \tag{3}$$

where the lower-case letters are the log values of the variables: $y = \ln(\text{Real } va)$, $l = \ln L$, $k = \ln K$, and ω is the residual TFP that is not captured by the observables. age captures firm's the length of existence. inv is the log value of firm's investment. Ind is the variable for two-digit sector. $year$ captures the time fixed effect. SOE is a dummy indicate whether a firm is a state-owned enterprise. And exp is a dummy indicate whether a firm exports or not.

The goal of the regression is to estimate β_l and β_k and back out the production function. Then the TFP can be measured as the Solow residual.

However, if we run the regression in 3 using OLS, the coefficient β_k could be biased since the residual ω is dependent on k as well[3]. So we want to run a two-step estimation.

In the first step we run the following regression:

$$y_{ist} = \beta^l l_{ist} + \Sigma_s \delta_s Ind_s + \Sigma_t \delta_t year_t + \beta^{\mathrm{SOE}} SOE_{ist} + \beta^{exp} exp_{ist} \phi(age_{ist}, k_{ist}, inv_{ist}) + \varepsilon_{ist}, \tag{4}$$

where $\phi(age_{ist}, k_{ist}, inv_{ist}) = \beta^k k_{ist} + \omega(age_{ist}, k_{ist}, inv_{ist})$. Since we do not know the functional form of $\phi(age_{ist}, k_{ist}, inv_{ist})$, it is approximated by a higher-order polynomial of the input variables.

Now we have an estimate of function ϕ, noted as $\widehat{\phi}$. Moreover, we assume that ω_{ist} follows a Markov-chain process: $\omega_{ist+1} = g(\omega_{ist}) + \mu_{ist+1}$, and μ_{ist} is iid. Then we can write the second-step regression in equation 5:

$$\widehat{\phi}_{ist+1} = \beta^k k_{ist+1} + g(\omega(age_{ist}, k_{ist}, inv_{ist})) + \eta_{ist} \tag{5}$$

Function $g(\omega(.,.,.))$ can also be approximated by a high-order polynomial.

The methodology in Levinsohn and Petrin (2003) is very similar to Olley and Pakes (1996). But instead of using investment as the proxy for the unobserved productivity, they use the intermediate input.

$$y_{ist} = \beta^l l_{ist} + \beta^k k_{ist} + \Sigma_s \delta_s Ind_s + \Sigma_t \delta year_t + \beta^{\mathrm{SOE}} SOE_{ist} + \beta_{exp} exp^{ist}$$
$$+\omega_{ist}(age_{ist}, k_{ist}, m_{ist}) + \varepsilon_{ist} \tag{6}$$

where m_{iit} is the measured as total operation input deflated by the intermediate input price index.

[3]There are ways to deal with the bias of β_l if ω is a function of l too. In the estimation we employ the methodology in Ackerberg, Caves and Frazer (2015) to correct for the estimate of β_l. But for the purpose of illustration, we do not talk about it here.

After have the unbiased estimates $\widehat{\beta}_k$ and $\widehat{\beta}_l$, total factor productivity is measured as in equation 9 in both Olley and Pakes (1996) and Levinsohn and Petrin (2003) is

$$a_{ist} = y_{ist} - \widehat{\beta}^k k - \widehat{\beta}^l l \tag{7}$$

There is an implicit assumption in either Olley and Pakes (1996) and Levinsohn and Petrin (2003) that the difference between the firm-level price and two-digit sectoral level price is not correlated with the input choice. However, they could be correlated. For example, a monopolistic firm could charger higher price than the average sectoral price. Therefore, it gains more profit and enlarges its scale. In this scenario, $P_{ist} - P_{st}$ is positively correlated with capital and labor. The methodology in De Loecker (2011) take the potential bias caused by the price difference seriously and incorporate a CES demand system into the estimation. The regression to be run is as follows:

$$\begin{aligned} y_{ist} &= \beta^{l*} l_{ist} + \beta^{k*} k_{ist} + \beta^s y_{st} + \Sigma_s \delta_s Ind_s + \Sigma_t \delta year_t + \beta^{\text{SOE}} SOE_{ist} \\ &+ \beta^{exp} exp_{ist} + \omega_{ist}(age_{ist}, k_{ist}, m_{ist}) + \varepsilon_{ist} \end{aligned} \tag{8}$$

where y_{st} is the log value of the real term of two-digit sectoral value added, defined as $y_{st} \equiv \ln(\Sigma_{i \in s} va_{ist}) - \ln(P_{st})$, and β_s is interpreted as the inverse of the elasticity of substitution of sector s.

And TFP is measure as follows:

$$a_{ist} = (y_{ist} - \widehat{\beta}^{k*} k - \widehat{\beta}^{l*} l - \widehat{\beta}^s y_{st}) \frac{1}{1 + \widehat{\beta}^s} \tag{9}$$

1.3 TFP Growth Deceleration

In this subsection, we show that the TFP measures from different methodologies highly correlated with each other, which implies the precision of the firm-level TFP measures. Also we demonstrate that the aggregate TFP growth decelerats in later years of our data.

Table 1 and table 2 shows how closely the four measures of firm-level TFP growth correlated with each other. "DL" stands for the methodology of TFP measure using De Loecker (2011) described in equation 8 , "LP" for Levinsohn and Petrin (2003) in equation 6, "OP" for Olley and Pakes (1996) in equation 3 and "CD" for the Cobb-Douglas methodology in equation 2. The reason for having two tables with high similarity is that the variable value added is only observable between years 1999 and 2007. After year 2007, we have to construct our own measure of value added. So the "Pseudo VA" used in table 2 is calculated as the difference between the firm-level "sales income" and "sales cost". So the correlation in table 1 is calculated only for the period from 1998 to 2007, while the correlation in table 2 is calculated for the period from 1998 to 2013.

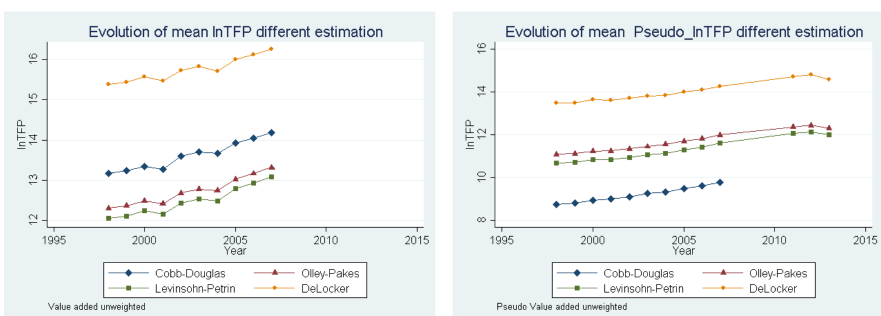
Table 1: DlnTFP Correlation measured by VA

Variables	DL	LP	OP	CD
DL	1.000			
LP	0.998	1.000		
OP	0.995	0.997	1.000	
CD	0.942	0.945	0.966	1.000

Table 2: DlnTFP Correlation measured by Pseudo VA

Variables	DL	LP	OP	CD
DL	1.000			
LP	0.997	1.000		
OP	0.991	0.997	1.000	
CD	0.969	0.977	0.982	1.000

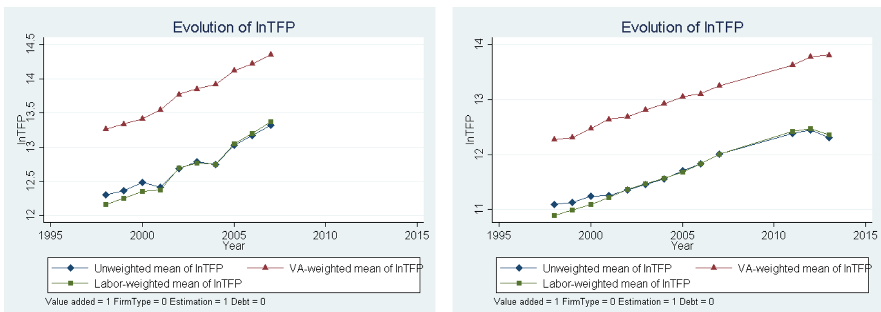
Figure 1 shows the unweighted average TFP growth path measured by different methodologies. The left panel uses the officially calculated value added while the right panel uses the "Pseudo value added" calculated as the difference between "sales income" and "sales cost". The level difference by different measures of TFP is caused by the normalization, therefore does not reveal any information. The informative pattern in this figure is the trend. We observe that the growth path of TFP is very similar across different measures in both sub-figures. This observation is consistent with the high correlation in table 1 and table 2. The right panel tells us that starting around year 2011, the TFP growth rate starts to decline, and it even becomes negative from year 2012 to year 2012.



(a) Directly Observed VA (b) Pseudo VA = Sales Income - Sales Cost

Figure 1: Unweighted Mean of lnTFP by different measures

Figure 2 plots the mean lnTFP paths by different weights using the firm-level TFP measured by methodology in De Loecker (2011). Similarly, the left panel uses officially calculated value added and the right panel uses the "pseudo value added". The slowing down of TFP growth starting from year 2011 can be observed in any type of weighted mean lnTFP on the right panel, be it the unweighted, value added weighted or labor weighted mean lnTFP. Another interesting point of figure 2 is that the labor weighted mean is about the same as the unweighted mean, while the value-added weighted mean is much higher in both sub-figures. The reason for that is that the firms with higher TFP tend to have a higher value added. However there is no clear correlation between the labor employment and TFP.



(a) Directly Observed VA

(b) Pseudo VA = Sales Income - Sales Cost

Figure 2: Mean of lnTFP by different weights

Figure 3 compares the firm-level TFP measure to the aggregate TFP measure, and shows that both measures have the same trend. The red curve plots the unweighted mean of value-added lnTFP measure, and the blue curve unweighted mean of pseudo-value-added lnTFP measure using the methodology in Olley and Pakes (1996) with the firm-level data. And we find that the two have the same trend between the period from year 1998 to 2007. The green curve plots the aggregate lnTFP measure using equation 1. It can be seen that after 2011, we observe the same reversal of TFP growth path that we observe in the firm-level data. Moreover, the aggregate TFP series has a longer historical data, indicating that after 2013, the TFP continues to drop.

2 The Role of SOEs in the Manufacturing Sector

In this section, we demonstrate that the state-owned enterprises (SOEs) are different from the non-state-owned-enterpreises (Non-SOEs)in many aspects. In particular, the timing of the privitization and the slowing down of the privitization coincides with the TFP growth and decelaraton of the manufacturing as a whole. Moreover, we show evidence that the TFP decelaration cannot be due to any sectoral difference since the growth pattern of TFP of almost all sectors

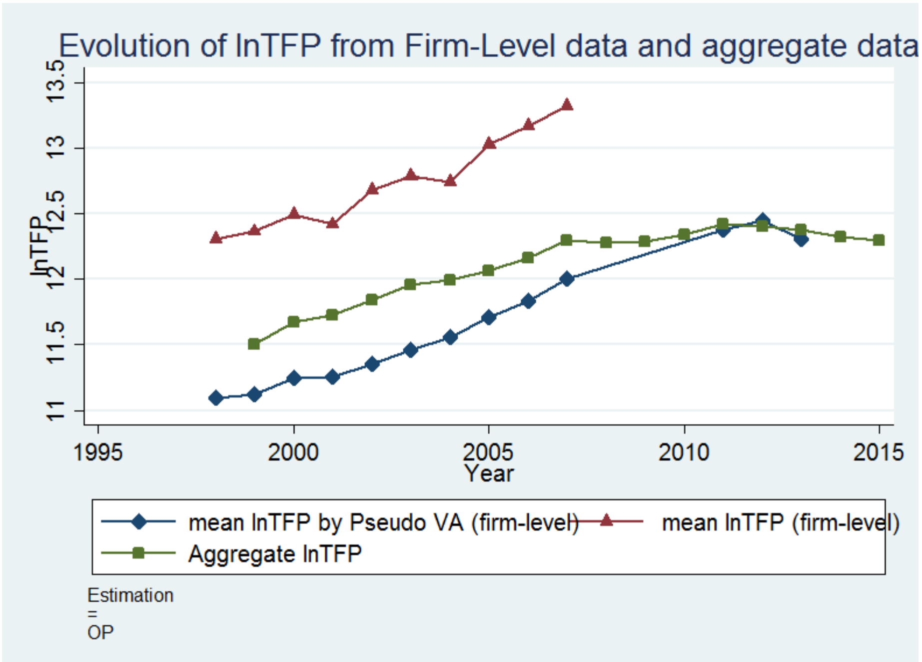


Figure 3: Firm-level Estimation and Aggregate Estimation

are the similar to that of the whole manufacturing sector.

2.1 Productivity of SOEs vs Non-SOEs

The empirical framework in this subsection is as follows:

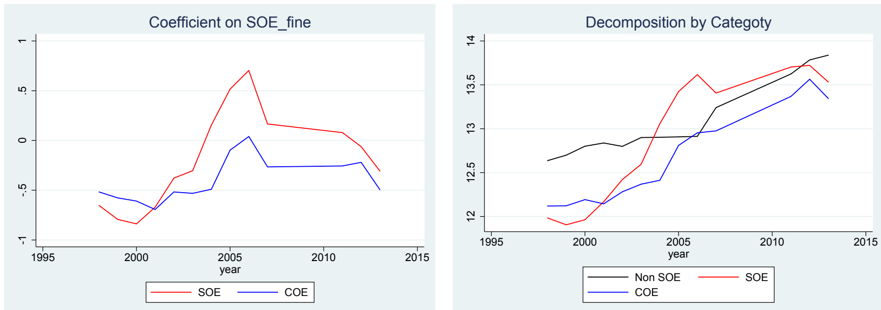
$$lnTFP_{it} = \beta_{0t} + \Sigma_g^{N_g} \beta_{gt} D_{it}^g + \varepsilon_{it} \tag{10}$$

where i is the firm identifier, t is the year subscript, g is the characteristic group subscript. N_g stands for the number of groups. D_{it}^g is a dummy variable, defined as $D_{it}^g = 0$ if $i \notin g$ and $D_{it}^g = 1$ if $i \in g$.

When we divide firms into three categories: SOEs, collectvity-owned enterprises (COEs)[4] and non-SOEs, and run the regression in equation 10 year by year, we can plot the graph in 4. Here we set the baseline group to be the group of non-SOE firms. That is why in the left panel, there are only two series of coefficients standing for the SOEs and COEs while on the right panel there are three lines including the one for the baseline group. Why we cannot have clear interpretation of the coefficnets time series, the predicted average of lnTFP in has different patten. From 1998 to 2004, SOEs are relatively less productive compared to the Non-SOEs. Then SOEs catches up with the Non-SOEs in

[4]COEs can be considered as one type of SOEs, which are owned by local governments. We will show later that COE are less important in terms of size compared to the other two types.

terms of productivity and even surpass the Non-SOEs after 2005. However, Starting from year 2011, we see a decelaraton of TFP growth in SOEs but not in non-SOE firms.



(a) Coefficients (b) Predicted average lnTFP

Figure 4: Decomposition by SOE category

Figure 5 plots the dispersion of the marginal revenue product of capital (MRPK). The calculation of MRPK follows Chang-Tai Hsieh (2009). And the dispersion is measured as the standard deviation of the log value of MRPK. The definition is shown in equation 11.

$$\begin{aligned} &MRPK_{it} \propto \frac{va_{it}}{K_{it}} \\ &\text{Dispersion}_t = std(\ln MRPK_{it}) \end{aligned} \tag{11}$$

Now we divide firms by sectors and run the same year-by-year regression in equation 10. We can get the predicted average lnTFP by sectors as shown in Figure 6. Here the sectors are grouped by their main characteristics for the purpose of clear presentation. Although it seems a bit messy in early years, the average lnTFP of all sectors displays the dip that is similar to that of the average lnTFP of the total industry as a whole. This means that the deceleration of TFP growth exists in every sectors, which contrasts to the observation that it does not in every SOE category.

2.2 Privatization Slowing Down

In this subsection, we show more difference between the SOEs and non-SOEs. More specifically speaking, we are going to show that the privatization process has slowed down and even reversed. And a possible reason for that is that the borrowing cost for non-SOEs have gone up too much.

Figure 7 presents a mirroring pattern by construction. The blue curve presents the share of real capital of non-SOEs and the red curve that of SOEs. From year 1998 to 2011, the share of real capital of non-SOEs is always increasing except for year 2004. This is mainly because of the privatization process

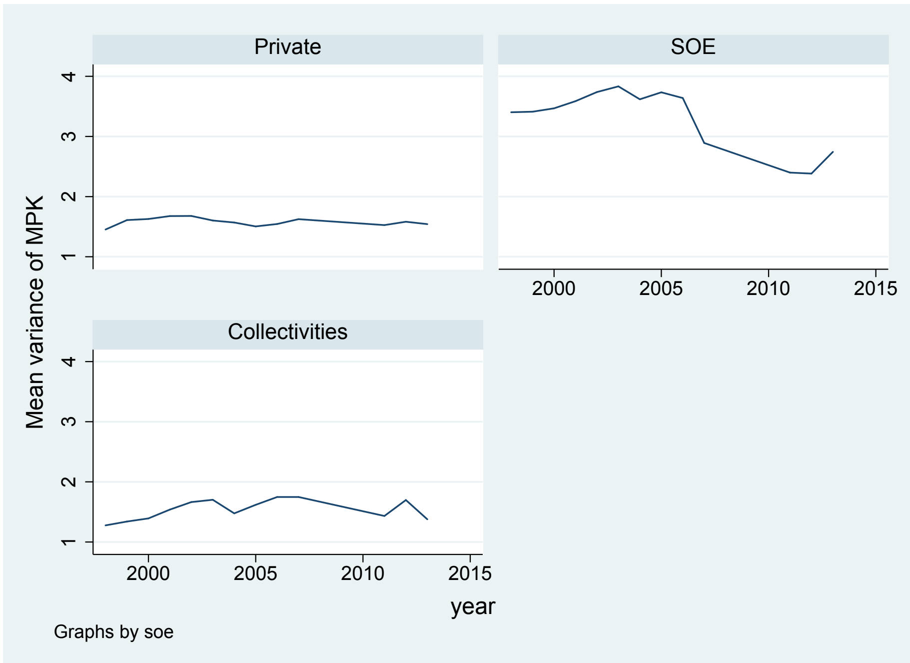


Figure 5: Dispersion of MRPK

in China. However, this privatization process seems to stop or even reverse starting from year 2011. More detailed break down can be found in table 3. The SOE category in Figure 7 contains the SOE and COE in table 3, while the non-SOE category contains the other two, POE stands for private owned firms, and FOE for foreign investor owned firms.

In table 4 and table 5 show respectively the share of labor by different ownership and the share of value added by different ownership. Both tables show a similar pattern to table 3. All three tables reveal the same signal: by any type size measure (share of capital, labor employment or value added), the SOEs have experienced a significant share drop in the economy due to the privatization process. However, the process slows down starting from 2011, which coincides with the timing of TFP drop.

In Figure 8, we show the average interest rate by SOE categories. The average interest rate is measured as the interest expense over the debt. And in Figure 8 we only include the interest rate between 0 and 1. We can see the borrowing cost of SOEs are much lower than POEs. In all the years of the data set, the average interest rate of SOEs are below 4%, while most of the years, the average interest rate of POEs are above 4% and but below 6%. After 2004, the difference between the average interest rate is even widening between the SOEs and POEs.

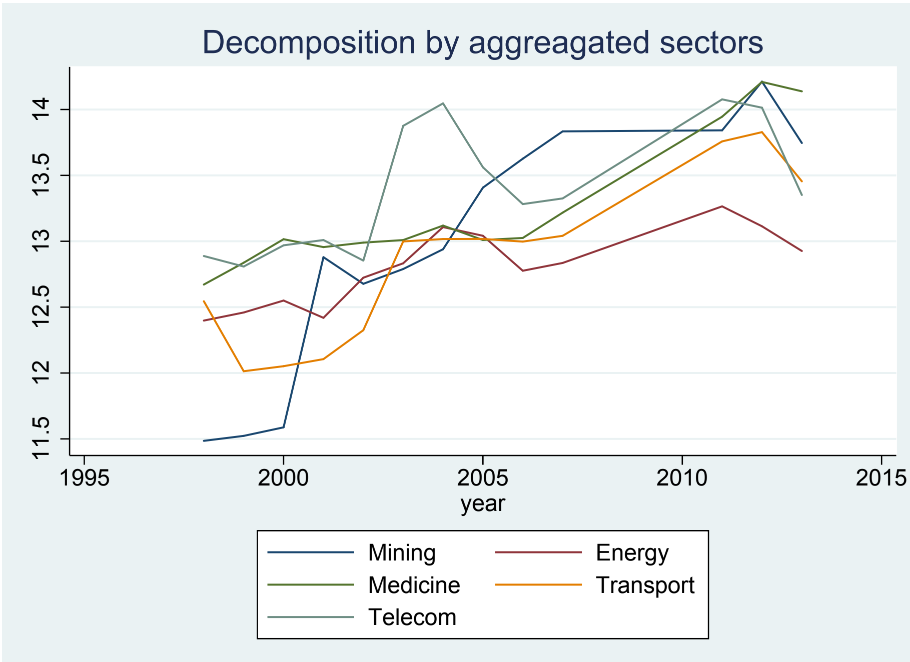


Figure 6: Decomposition by Sectors

Table 3: Summary Statistics Capital by Ownership

	Ownership			
year	**POE** %	**SOE** %	**COE** %	**FOE** %
1998	12.1	57.7	9.4	20.7
1999	13.8	55.2	8.8	22.2
2000	22.9	48.0	7.8	21.2
2001	38.6	33.7	5.8	21.9
2002	43.3	29.5	4.8	22.4
2003	44.6	28.2	4.0	23.1
2004	40.5	32.7	2.6	24.1
2005	41.8	30.0	2.0	26.1
2006	42.0	30.1	1.6	26.3
2007	50.8	21.1	1.4	26.6
2011	61.2	15.0	0.8	23.0
2012	58.1	16.2	1.0	24.7
2013	59.7	18.2	0.4	21.6
Total	45.8	40.1	5.9	24.0

Figure 7: Shares of by SOE category

Table 4: Summary Statistics Labor by Ownership

	Ownership			
year	**POE**	**SOE**	**COE**	**FOE**
	%	%	%	%
1998	16.2	52.8	19.8	11.2
1999	19.3	48.5	18.8	13.4
2000	24.5	42.5	17.1	15.9
2001	37.1	31.8	13.8	17.4
2002	41.2	28.0	11.9	18.9
2003	45.0	23.1	9.7	22.2
2004	49.2	19.4	5.5	25.9
2005	50.5	17.2	5.0	27.4
2006	52.1	15.6	4.0	28.3
2007	55.6	12.0	3.4	29.0
2011	61.4	6.6	1.3	30.7
2012	61.4	6.5	1.3	30.8
2013	64.2	6.3	1.2	28.2
Total	50.5	33.4	13.0	24.6

Table 5: Summary Statistics Value Added by Ownership

	Ownership			
year	**POE** %	**SOE** %	**COE** %	**FOE** %
1998	18.0	43.1	16.7	22.2
1999	21.1	37.5	14.8	26.7
2000	28.8	29.6	11.9	29.7
2001	44.3	20.9	7.8	27.0
2002	45.9	20.4	6.9	26.9
2003	45.6	19.1	5.5	29.8
2004	37.9	26.7	3.3	32.2
2005	40.1	27.1	3.4	29.5
2006	41.9	26.0	2.9	29.2
2007	55.3	13.1	2.4	29.3
2011	62.5	8.8	1.1	27.6
2012	64.0	7.9	1.1	27.0
2013	65.1	7.9	0.6	26.4
Total	47.9	28.0	9.3	28.4

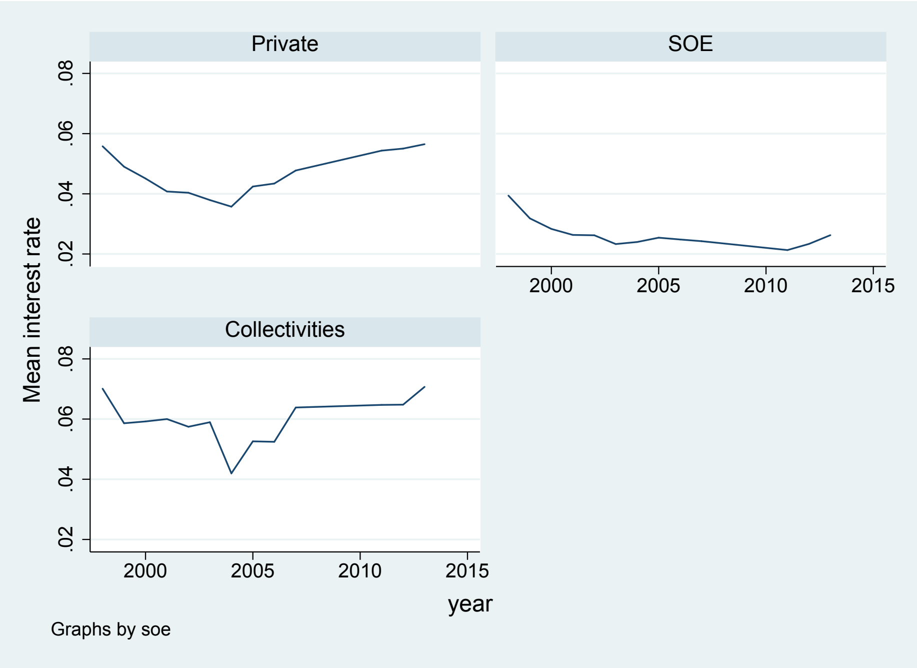


Figure 8: Interests by SOE category

3 Compositional Transition Between SOEs and Non-SOEs

In this section, we present evidence showing that in the early period (1998 - 2011), the TFP increase is mainly due to the within Non-SOEs TFP growth; while in the late period (2011 - 2013), the TFP decrease is mainly related to the within SOE TFP decline. The reallocation between SOEs and non-SOEs also plays a role in explaining the change of the TFP growth path of the two periods. With more detailed decomposition within SOEs and non-SOEs, there is a within firm TFP growth flip in SOEs, but not in non-SOEs.

Let us begin by defined the aggregate TFP measure.

We assume the total real value added can be expressed as the sum of the real value added by each firm:

$$Y_t = \Sigma A_{it} k_{it}^{\alpha} l_{it}^{\beta} \tag{12}$$

Moreover, we assume the production function of aggregate economy is also a Cobb-Douglas function:

$$Y_t = A K^{\alpha} L^{\beta} \tag{13}$$

Therefore the aggregate TFP can be expressed as follows:

$$\begin{aligned}
\mathrm{TFP}_t \equiv A_t = \tfrac{Y_t}{K^{\alpha} L^{\beta}} = \Sigma_i A_{it} \tfrac{k_{it}^{\alpha} l_{it}^{\beta}}{K_t^{\alpha} L_t^{\beta}} \\
= \Sigma_s \tfrac{\Sigma_{i \in s} k_{it}^{\alpha} l_{it}^{\beta}}{K_t^{\alpha} L_t^{\beta}} \underbrace{\Sigma_{i \in s} \frac{k_{it}^{\alpha} l_{it}^{\beta}}{\Sigma_{i \in s} k_{it}^{\alpha} l_{it}^{\beta}} A_{it}}_{\equiv TFP_t^s} \\
= \Sigma_s \tfrac{\Sigma_{i \in s} k_{it}^{\alpha} l_{it}^{\beta}}{K_t^{\alpha} L_t^{\beta}} TFP_t^s
\end{aligned} \tag{14}$$

where s is the subscript for the SOE/non-SOE categories. And $K_{st}^{\alpha} L_{st}^{\beta} \equiv$. The advantage to do this two-step weighted average is that the weight of the first-step weighted average $\Sigma_i \frac{k_{it}^{\alpha} l_{it}^{\beta}}{\Sigma_{i \in s} k_{it}^{\alpha} l_{it}^{\beta}} A_{it}$ sums up to one.

The change of TFP between two years then can be decomposed to three parts as in equation 15: changes within SOE categories, changes between SOE/non-SOE and a covariance term.

$$\Delta TFP_t^* \equiv \underbrace{\Sigma_s \frac{\Sigma_{i\in s} k_{ir}^{\alpha} l_{ir}^{\beta}}{K_r^{\alpha} L_r^{\beta}} \Sigma_{i\in s} \frac{k_{ir}^{\alpha} l_{ir}^{\beta}}{\Sigma_{i\in s} k_{ir}^{\alpha} l_{ir}^{\beta}} A_{ir}}_{\text{TFP}_r} - \underbrace{\Sigma_s \frac{\Sigma_{i\in s} k_{it}^{\alpha} l_{it}^{\beta}}{K_t^{\alpha} L_t^{\beta}} \Sigma_{i\in s} \frac{k_{it}^{\alpha} l_{it}^{\beta}}{\Sigma_{i\in s} k_{it}^{\alpha} l_{it}^{\beta}} A_{it}}_{\text{TFP}_t}$$

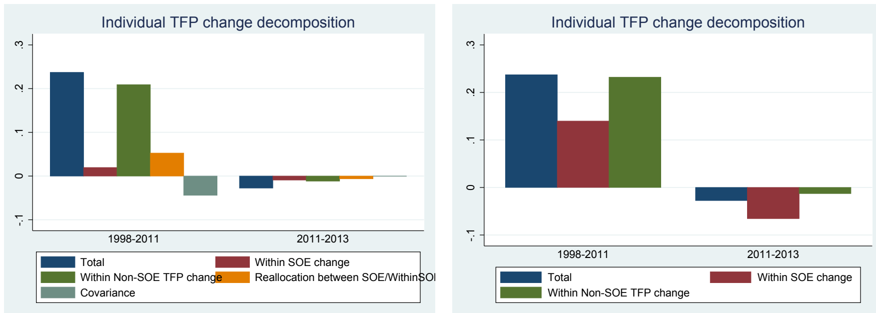
$$= \underbrace{\Sigma_s \frac{\Sigma_{i\in s} k_{ir}^{\alpha} l_{ir}^{\beta}}{K_r^{\alpha} L_r^{\beta}} [TFP_r^S - TFP_t^S]}_{\text{Within SOE/non-SOE changes}} + \underbrace{\Sigma_s [\frac{\Sigma_{i\in s} k_{ir}^{\alpha} l_{ir}^{\beta}}{K_r^{\alpha} L_r^{\beta}} - \frac{\Sigma_{i\in s} k_{it}^{\alpha} l_{it}^{\beta}}{K_t^{\alpha} L_t^{\beta}}] TFP_r^S}_{\text{Between SOE/non-SOE changes}}$$

$$\underbrace{-\Sigma_s (TFP_r^S - TFP_t^S)(\frac{\Sigma_{i\in s} k_{ir}^{\alpha} l_{ir}^{\beta}}{K_r^{\alpha} L_r^{\beta}} - \frac{\Sigma_{i\in s} k_{it}^{\alpha} l_{it}^{\beta}}{K_t^{\alpha} L_t^{\beta}})}_{\text{Covariance term}} \tag{15}$$

where r is the reference year, which we set to 2011, and t is the year in question. We cut the period from 1998 to 2013 into two sub-periods, 1998-2011 and 2011-2013. In the first sub-period, we set $t = 1998$; while in the second sub-period, $t = 2013$.

The result of the decomposition in equation 15 is presented in 9. The left panel is the weighed TFP changes, so the sum of the maroon bar, green bar yellow bar and gray bar equal to the height of blue bar. So from this graph, the most significant change is the green bar. It means that the within non-SOE TFP change contributes to the most of the TFP increase in the first sub-period. And the reallocation between SOEs and non-SOEs comes in the second place in terms of contribution to the TFP increase. During the second sub-period, it seems that the within SOE and within Non-SOE contributes equally to the TFP decline. However, we have to realized that we are using the 2011 as the reference year, the weight we use is the year when SOEs only account for less than 20% in any size measure. Therefore, if we look at the within SOE and within non-SOE change without weights in the right panel of Figure 9. The flip of the growth rate of TFP in SOEs is as significant as in non-SOEs.

Now we want to zoom into the within within SOE/non-SOE categories. We can decompose the within SOE/non-SOE change into five parts: within firm change, between firm change, entry, exit and covariance. Graphically speaking, we now decompose the maroon bar and green bar in the right panel of Figure 9 into five bars. The decomposition method is expressed in equation 16.



(a) Weighted TFP Changes (b) Unweighted TFP Changes

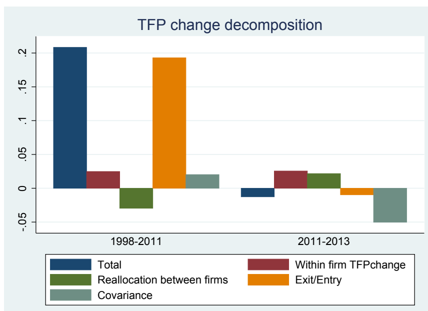
Figure 9: TFP Changes by SOE Category

$$
\begin{aligned}
&\Delta TFP_t^S \equiv TFP_r^s - TFP_t^s \\
&= \Sigma_{i\in s}\frac{k_{ir}^{\alpha}l_{ir}^{\beta}}{\Sigma_{i\in s}k_{ir}^{\alpha}l_{ir}^{\beta}}A_{ir} - \Sigma_{i\in s}\frac{k_{it}^{\alpha}l_{it}^{\beta}}{\Sigma_{i\in s}k_{it}^{\alpha}l_{it}^{\beta}}A_{it} \\
&= \underbrace{Weight_{stay^s,r}\Sigma_{i\in stay^s}\frac{k_{ir}^{\alpha}l_{ir}^{\beta}}{\Sigma_{i\in stay^s}k_{ir}^{\alpha}l_{ir}^{\beta}}[A_{ir} - A_{it}]}_{\text{Within Firm}} \\
&\underbrace{+Weight_{stay^s,r}\Sigma_{i\in stay^s}[\frac{k_{ir}^{\alpha}l_{ir}^{\beta}}{\Sigma_{i\in stay^s}k_{ir}^{\alpha}l_{ir}^{\beta}} - \frac{Weight_{stay^s,t}}{Weight_{stay^s,r}}\frac{k_{it}^{\alpha}l_{it}^{\beta}}{\Sigma_{i\in stay^s}k_{it}^{\alpha}l_{it}^{\beta}}]A_{it}}_{\text{Between Firm}} \\
&\underbrace{+Weight_{enter^s,r}\Sigma_{i\in enter^s}\frac{k_{ir}^{\alpha}l_{ir}^{\beta}}{\Sigma_{i\in enter^s}k_{ir}^{\alpha}l_{ir}^{\beta}}A_{ir}}_{\text{Enry}} \underbrace{- Weight_{exit^s,t}\Sigma_{i\in exit^s}\frac{k_{it}^{\alpha}l_{it}^{\beta}}{\Sigma_{i\in exit^s}k_{it}^{\alpha}l_{it}^{\beta}}A_{it}}_{\text{Exit}} \\
&\underbrace{-\Sigma_{i\in stay^s}(A_{ir} - A_{it})(\frac{k_{ir}^{\alpha}l_{ir}^{\beta}}{\Sigma_{i\in s}k_{ir}^{\alpha}l_{ir}^{\beta}} - \frac{k_{it}^{\alpha}l_{it}^{\beta}}{\Sigma_{i\in s}k_{it}^{\alpha}l_{it}^{\beta}})}_{\text{Covariance}}
\end{aligned} \tag{16}
$$

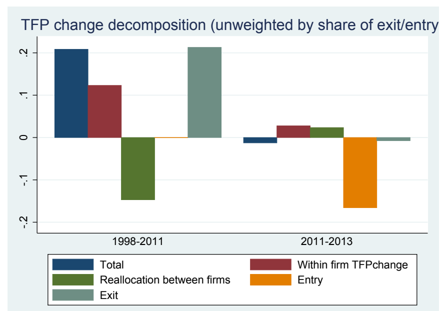
where $stay^s$ is the subset of s, standing for firms that exist both in year t and in reference year r in the category s; $enter^s$ is the subset of s, standing for firms that newly enter into the market in reference year r and do not exist in year t in the category s; $exit^s$ is the subset of s, standing for firms that exit in year t but do not exit in reference year r in the category s. Moreover,

$$
\begin{aligned}
Weight_{stay^s,r} &\equiv \frac{\Sigma_{i \in stay^s} k_{ir}^{\alpha} l_{ir}^{\beta}}{\Sigma_{i \in s} k_{ir}^{\alpha} l_{ir}^{\beta}} \\
Weight_{stay^s,t} &\equiv \frac{\Sigma_{i \in stay^s} k_{it}^{\alpha} l_{it}^{\beta}}{\Sigma_{i \in s} k_{it}^{\alpha} l_{it}^{\beta}} \\
Weight_{enter^s,r} &\equiv \frac{\Sigma_{i \in enter^s} k_{ir}^{\alpha} l_{ir}^{\beta}}{\Sigma_{i \in s} k_{ir}^{\alpha} l_{ir}^{\beta}} \\
Weight_{exit^s,t} &\equiv \frac{\Sigma_{i \in exit^s} k_{it}^{\alpha} l_{it}^{\beta}}{\Sigma_{i \in s} k_{it}^{\alpha} l_{it}^{\beta}}
\end{aligned}
\tag{17}
$$

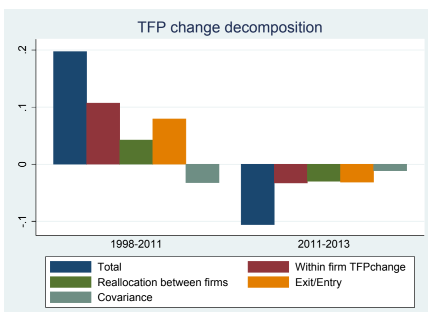
The result of the decomposition in equation 16 is shown in Figure 10. The first row is the weighted and unweighted within non-SOE TFP change while the second row is the counterpart for the SOE TFP change. The first column is the weighted TFP change while the second column is the unweighted TFP change. One interesting pattern is that for SOEs, the within firm TFP change (maroon bar) flips from positive in the first sub-period to negative in the second sub-period, while we do not observe the same change in non-SOEs. This means that on average the measured TFP experience a decline for the existing SOEs but not for the existing non-SOEs. Anther point is that the reallocation between firms (green bar) causes a TFP growth flip in SOEs but not in non-SOEs, indicating that SOEs may suffer more misallocation in the second sub-period compared to the non-SOEs.



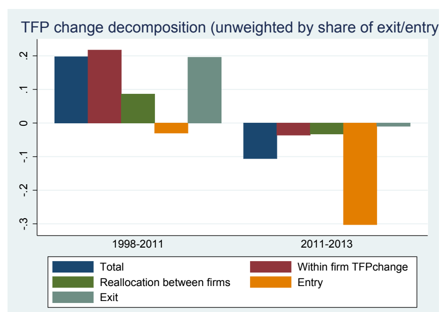
(a) Weighted within non-SOE TFP Changes



(b) Unweighted within non-SOE TFP Changes



(c) Weighted within SOE TFP Changes



(d) Unweighted within SOE TFP Changes

Figure 10: TFP Changes by within SOE/Non-SOE

References

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

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

De Loecker, Jan. 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity." *Econometrica*, 79(5): 1407–1451.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies*, 70(2): 317–341.

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