

NBER WORKING PAPER SERIES

CREATIVE ACCOUNTING OR CREATIVE DESTRUCTION? FIRM-LEVEL PRODUCTIVITY
GROWTH IN CHINESE MANUFACTURING

Loren Brandt
Johannes Van Biesebroeck
Yifan Zhang

Working Paper 15152
<http://www.nber.org/papers/w15152>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2009

Funding by SSHRC and CFI/OIT is gratefully acknowledged. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2009 by Loren Brandt, Johannes Van Biesebroeck, and Yifan Zhang. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing

Loren Brandt, Johannes Van Biesebroeck, and Yifan Zhang

NBER Working Paper No. 15152

July 2009

JEL No. D24,O14

ABSTRACT

We present the first comprehensive set of firm-level total factor productivity estimates for China's manufacturing sector that spans her entry into WTO. We find that productivity growth is among the highest compared to other countries. For our preferred estimate, the weighted average annual productivity growth for incumbents is 2.7% for a gross output production function and 7.7% for a value added production function over the period 1998-2006. Of the various sensitivity checks we carry out, controlling for the increase in labor quality and labor hours, as proxied by the rising real wage, has the largest (downward) effect on the productivity estimates. We further document that new entrants are a particularly dynamic force and that firms experience large productivity declines before exiting from the sample. Overall, net entry contributes roughly half to total TFP growth. Aggregate productivity growth, however, is tempered by a much lower effect of reallocation of inputs towards higher productivity firms, compared to the U.S. benchmark.

Loren Brandt
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON M5S 1A1 CANADA
brandt@chass.utoronto.ca

Yifan Zhang
Lingnan University
Department of Economics
8 Castle Peak Road Tuen Mun
New Territories Hong Kong SAR
yifan.zhang@ln.edu.hk

Johannes Van Biesebroeck
Centre for Economic Studies
K.U.Leuven
Naamsestraat 69
3000 Leuven
Belgium
and NBER
jo.vanbiesebroeck@econ.kuleuven.be

1. Introduction

China has enjoyed impressive growth in labor productivity averaging nearly 8 percent for a period now spanning three decades. Initially propelled by reform of the agricultural sector, since the mid-to-late 1980s the non-agricultural sector has emerged as the engine of growth. By 2006, non-agriculture was the source of nearly 90 percent of GDP, and employed in upwards of two-thirds of the labor force.

Considerable debate persists over the sources of this growth and the relative contributions of growth in total factor productivity (TFP) versus the mobilization of additional resources, notably physical and human capital. Much of this analysis has been at a highly aggregate level that combines agriculture and non-agriculture (Bosworth and Collins, 2008; Perkins and Rawski, 2008). These studies typically find TFP contributing approximately half of total growth. In a widely cited study focusing solely on the non-agriculture sector, Young (2003) uses officially published data to paint a much less impressive picture. Correcting for potential biases in official deflators and the measurement of human capital, Young reduces productivity growth between 1978 and 1998 from a very respectable 3 percent to a more pedestrian 1.4 percent.

Young's estimates have been revised upwards (Brandt and Zhu, 2009). There is also an issue of aggregation. China's non-agriculture sector combines secondary industry, which represents nearly half of GDP and is the source of 90 percent of her exports, with the service sector, to which reform and market liberalization has only come with a long lag. An earlier literature largely using aggregate data for industry, supplemented with small panel data sets, typically found only modest effects of reform on TFP in the state sector, but more rapid growth of TFP in the expanding non-state sector (See, for example, Jefferson and Rawski, (1994); Groves et. al., (1998)). Overall, capabilities and productivity appeared to be rising in industry, a product of the growing impact of market forces, expanded entry, and intense competition.

An analysis of behavior in Chinese manufacturing on par with that carried out for other countries has been handicapped by the lack of firm-level data sets covering the sector. This has made it difficult, for example, to examine the benefits of the massive entry of new firms, as well as exodus of incumbents. The increasing availability of firm-level data sets is helping to relax this constraint, and allow more in-depth analysis at the micro level of key aspects of behavior in manufacturing that are missed at the macro level (See, for example, Bai, Hsieh, and Qian (2006), Dougherty, Herd and He (2007), Hsieh and Klenow (2007), Park, *et al.* (2006)). This paper builds on this work.

Drawing on an unbalanced panel of firms between 1998 and 2006 that represents nearly 90 percent of the gross output in manufacturing, we present the first comprehensive set of firm-level productivity estimates for Chinese manufacturing that spans China's entry into WTO. The absolute size of China's manufacturing sector and its exports make this important in its own right.

In our analysis, we pay particular attention to the dynamics of firm entry and exit, and to the contributions that these processes are making to aggregate productivity growth in the tradition of work by Baily, Hulten and Campbell (1992), Haltiwanger (1997) and Petrin and Levinsohn (2006). Recent

qualitative work of Brandt, Rawski and Sutton (2008), as well analysis by Jefferson, Rawski and Zhang (2008) using cross-sectional data for 1998 and 2005, identifies entry and exit as potentially important drivers of the dynamism in the manufacturing sector. We provide estimates directly comparable to those of Baily et. al for China's manufacturing sector.

Over the period we examine, we find TFP growth in manufacturing firms averaging 2.7 percent for a gross output production function and 7.7 percent for a value added production function. Our results also identify net entry as the source of as much as half of productivity growth however aggregate growth in Chinese manufacturing remains constrained by limited efficiency-enhancing input allocations between active firms.

Working with firm-level data for China has its difficulties. Thus, one of the additional contributions of this paper is to carefully describe and document these data. We also make publically available complementary data we have constructed, including deflators, industry concordances, etc. that are required to make full use of the data.¹ Furthermore, in light of important concerns of Young and others, we examine the robustness of our results to a host of measurement issues. We show how alternative treatment of key variables often reduces productivity growth, but does not alter the basic picture.

The remainder of the paper is organized as follows. In the next Section we describe our methodology for measuring productivity. Section 3 describes the data set and the construction of the key variables. The Appendix and online Appendix provide more detailed documentation. In Section 4 we describe the Chinese results at the firm level, the performance of entrants and exiting firms, and the aggregate productivity growth experience. Section 5 concludes.

2. Productivity measurement

The most widely used measure of productivity is labor productivity, which is obtained by simply dividing value added by the numbers of hours worked or the number of workers. In China's national accounts the share of labor costs in total GDP is approximately one half for the full economy and even lower in manufacturing. As a result, not accounting for capital intensity is likely to paint a misleading picture. Any multi-factor productivity measure is only defined relative to a production technology, which we can characterize by the production function:

$$Q_{it} = A_{it}F_{it}(X_{it}). \tag{1}$$

Productivity is inherently a relative concept, and we can write it in general as

¹ In an online appendix at <http://www.econ.kuleuven.be/public/n07057/China/> we make available capital, intermediary input and output deflators, industry concordances over time and over different aggregation levels, and programs to match the firms over time and construct a revised capital stock series.

$$\ln \left(\frac{A_{it}}{A_{j\tau}} \right)_k = \ln \frac{Q_{it}}{Q_{j\tau}} - \ln \frac{F_k(X_{it})}{F_k(X_{j\tau})}. \quad (2)$$

For productivity growth comparisons, the same firm enters the numerator and denominator ($i=j$) and for productivity level comparisons we fix time instead ($t=\tau$). Even though we can allow the production function in (1) to differ between firms and over time, note the it subscripts on the input aggregator $F()$, which imply we have to use the same technology (k) to perform the productivity comparison in (2).

In order to accurately measure productivity, we need to accurately measure inputs and outputs and we need to estimate the input substitution possibilities that the technology allows. The first task is described at length in the next section. The second task is discussed here.

In Van Biesebroeck (2007, 2008), alternative methodologies to estimate productivity are compared. In general, the estimates tend to exhibit very high correlations however in one instance results are sensitive to the assumption of a homogenous production technology across firms.² Therefore, we implement two estimation procedures here.

For the first productivity measure, a particular functional form for the production function is assumed and its parameters are estimated using the methodology pioneered by Olley and Pakes (1996), inverting the investment equation nonparametrically to proxy for unobserved productivity. An intermediary estimation step controls for the non-random sample selection induced by the differing probability of exit for small and large low-productivity firms. Using the estimates of the production function coefficients, the productivity level for firm i at time t is calculated as

$$a_{it}^{OP} = q_{it} - \hat{\alpha}_L^s l_{it} - \hat{\alpha}_K^s k_{it}. \quad (3)$$

s superscripts on the coefficient estimates indicate that we estimate the production function separately by industry. Small cap variables represent logarithms and the three variables q , l , and k are for output (value added), labor, and capital.³ Measurement issues are discussed in the next section.

To normalize the above measure, recalling that productivity is only a relative benchmark, one can simply subtract the average productivity for all firms at time t in the same industry. An alternative is to include industry-year dummies in any regression that characterizes the patterns in the evolution or distribution of a_{it}^{OP} as the coefficients on the inputs are the same for any firm in the same industry. Productivity growth boils down to $a_{it}^{OP} - a_{it-1}^{OP}$.

² The estimation method mattered for the evaluation of learning-by-exporting effects. If the same (parametric) production technology was imposed on all observations, future exporters were found to always be more productive than nonexporters, and the difference grew larger with export market entry. Importantly, this finding disappeared if production technology was allowed to vary across firms.

³ The discussion in this section is for a value added production function. We will also present results for a gross output production function. In that case, material input always enters the equations similarly as labor input and the dependent variable q is then gross output instead of value added.

The second measure is a straightforward Törnqvist index number, as used by Solow (1957). Caves, Christensen and Diewert (1982a) illustrate how this approach can be interpreted as an exact productivity measure without estimating any parameters. The intuition is that a cost-minimizing firm will equalize the relative factor price to the local elasticity of substitution that the production technology allows. As a result, factor shares can be used to control for input substitutability.

Productivity growth is calculated in the usual way as

$$TFPG_{it}^{IN} = (q_{it} - q_{it-1}) - \overline{s_{it}}(l_{it} - l_{it-1}) - (1 - \overline{s_{it}})(k_{it} - k_{it-1}), \quad (4)$$

where $\overline{s_{it}} = (s_{it} + s_{it-1})/2$ is the average wage bill in value added. A major benefit of this approach is that it allows for technology heterogeneity in the input elasticities across firms (See Van Biesebroeck (2007)).

To compare the productivity level across firms within the same industry, Caves, Christensen, and Diewert (1982b) propose the following multilateral productivity measure:

$$\ln TFP_{it}^{IN} = (q_{it} - \bar{q}_t) - \widetilde{s_{it}}(l_{it} - \bar{l}_t) - (1 - \widetilde{s_{it}})(k_{it} - \bar{k}_t). \quad (5)$$

It amounts to a comparison with the hypothetical average firm in the industry. The weight on the labor input difference is $\widetilde{s_{it}} = (s_{it} + \bar{s}_t)/2$ and one minus this value for capital. While this measure is not transitive—the input weights differ across observations—it does allow for a comparison with the same benchmark while maintaining technology heterogeneity.⁴

3. Data

3.1 Firm-level census

We have annual firm-level data for the period 1998-2006 on all industrial firms that are identified as being either state-owned, or are non-state firms with sales above 5 million RMB⁵ (hereafter referred to as the "above-scale" industrial firms). These data come from annual surveys conducted by the National Bureau of Statistics (NBS).⁶ Industry is defined here to include mining, manufacturing and public utilities.

Statistics in Table 1A compare the coverage of our firm-level data (in the top panel) with the corresponding information from the Chinese Statistical Yearbook (in the bottom panel). In principle, the coverage should be identical. For the majority of variables reported, the differences between the first

⁴ The productivity level and growth measures obtained with the Olley-Pakes approach simply replace the varying input weights in equations (4) and (5) with the estimated input elasticity parameters.

⁵ This amounts to approximately \$US 600,000 over this period, a time when manufacturing prices were relatively stable.

⁶ Important changes in the firm identifiers used by the NBS in 1998 make it impossible to link up firms prior to 1998, and thus, our decision to begin the analysis in 1998.

two panels are less than 0.1%. For 1999, 2002, and 2003 all aggregates are identical, confirming that the firm-level data we are using are the basis for the numbers reported in the Statistical Yearbook. For 1998 and 2000 only the employment aggregates are inexplicably lower in the sample, by 8.9% and 3.4%, with no differences in any of the other variables. In 2001, 2004, and 2005 there are small additional discrepancies, but there is no real pattern: some variables match, others do not; sometimes the sample aggregate is higher, other times it is higher in the Yearbook.

[Table 1A & 1B approximately here]

In Table 1B, we make two additional comparisons. First, we extract from the 2004 Industrial Census all firms that are either SOEs or non-SOEs with sales larger than 5 million RMB, and compare them with our firm-level data set. The aggregates correspond extremely well. Second, for 2004 and for 1995, also a Census year, we examine the coverage of our sample (SOEs plus non-SOEs with sales larger than 5 million RMB) relative to the Census. In 2004, 80% of all industrial firms are excluded from our sample, however they account for only a small fraction of economic activity. These small firms employ 28.8% of the industrial workforce, but only produce 9.3% of output and generate 2.5% of exports.⁷ For 1995, the coverage with respect to sales, employment, and fixed assets is very similar to that in 2004. The similarity in coverage in 1995 and 2004 increases our confidence that the NBS decision rule on which firms to include in their annual sample is not introducing any systematic bias in our estimates.⁸

For the analysis in the paper, we focus only on manufacturing firms.⁹

Construction of a panel

Unique numerical firm IDs enable us to link firms over time. Firms occasionally receive a new ID if they go through restructuring, merger or acquisition. Where possible, we have aimed to track firms as their boundaries or ownership structure changes, using information on the firm's name, industry, address, etc. to link them.

In Table A.1, we report the percentage of firms that we match each year on the basis of firm ID and then using information other than firm ID. The fraction of firms in a year that can be linked to a firm in the previous year increases from 84.5% in the first two years (1998-1999) to almost 92% in the final two years (2005-2006). The extent to which we need to rely on identifying information other than firm IDs

⁷ The original numbers reported in the 2005 Chinese Statistical Yearbook for above-scale industrial firms show smaller totals. For example, output was 18.72 trillion RMB. The higher figure reported in the 2006 Chinese Statistical Yearbook reflects the upward revision following the 2004 Census. This suggests that before 2004, the coverage of our sample was probably in the vicinity of 80 percent.

⁸ Suppose, for example, that TFP in “excluded” firms systematically differed from those included. Changes over time in their contribution to industrial output could bias aggregate measures of TFP.

⁹ After 2002, firms with industry codes 1711, 1712, 1713, 1714, 2220, 3648, 3783, 4183 and 4280 were re-classified as service industries and we also exclude them from the sample in the earlier years.

declines substantially as the rate of restructuring slows down. Overall, 95.9% of all year-on-year matches are constructed using firm IDs and only 4.1% using other information.

The algorithm to perform the firm matches is described in some more detail in the Appendix. The program used to construct firm matches as well as the program to construct the revised capital stock series (see below) is made available in the online Appendix mentioned in the introduction.

The difficulty of following firms over time is related to changes in ownership type. While in 10% of the matches the “ownership type” of a firm changes, the proportion is significantly higher when other information is needed to trace a firm: the fraction is 24.5% in that case compared to 9.2% for matches based on IDs.¹⁰ Many of the ownership changes coincide with major restructuring that also results in new firm IDs. These events tend to occur only once over the time we observe a firm, but fully one-sixth of all firms in the sample that are observed for more than one year have their official ID changed at some point. Over time, the rate of ownership restructuring slows down markedly.

We end up with an unbalanced panel of firms that increases in size from 148,685 firms in 1998 to 278,753 in 2006. The unit of analysis is the firm, and not the plant. Other information in the survey suggests that more than 95 percent of all observations in our sample are single-plant firms. In Table 2, we report the total number of firms in each year and break this down by incumbents, entrants, and exiting firms. Between 2000 and 2001, for example, the total number of firms increases from 148,905 to 156,658, or an increase of 7,753. Gross entry was significantly larger, namely 37,729, but this was offset by the exit of 29,476 firms. The way in which we link firms, and the high percentage of the firms in our panel that experience ownership changes suggests that much of this entry and exit is *de novo*, as opposed to being a by-product of firm restructuring.

[Table 2 approximately here]

On average, the annual rate of attrition in our sample is slightly less than 14 percent. Out of our original sample in 1998 of 148,586 firms, 34,777 firms, or slightly less than a quarter, survive through 2006.

Noteworthy is the sharp increase in the number of firms in the sample between 2003 and 2004, a result of a lot more new entrants. The increase can be attributed to the 2004 Industrial Census, and the identification of firms, largely private in ownership, that should have been in the sample in earlier years, but had been left out because of a less than perfect business registry.

What is also important to note is that the average age of newly entered firms is declining over time. The median age for firms that showed up for the first time in the sample in 1999 and 2000 was 4 years. This declined to 3 years in 2001 and 2002 and further to 2 years in 2003. The additional firms picked up in

¹⁰ The ownership classification is discussed in more detail below. These statistics are for five ownership categories: state, hybrid/collective, private, Hong Kong/Macau/Taiwan, and foreign. Using the full ownership detail of 23 categories, changes are more common (13.7%), but again a lot more when non-ID information is used in the matches (41.5%).

the 2004 census year increases the median age of newly entered firms again to 3 years in 2004, but both in 2005 and 2006 this goes down again to 2 years.¹¹

In any given year, there are firms in the sample with sales that fall below the threshold of 5 million RMB. The percentage declines from 21.5% in 1998 to only 3.4% in 2005. Most of these firms are SOEs, and thus are included by fact of their ownership. Among non-SOEs, the percentage of firms with sales less than 5 million RMB rises slightly after 1998 up through 2000, and then falls sharply. In 2004, the census year, less than 1% of all non-SOEs record sales below 5 million RMB. By all indications, the 5 million RMB is not a "hard" rule, and firms whose sales slip below the threshold are not automatically removed from the sample.

In our analysis, we work with the sample as provided rather than strictly enforce the cutoff rule. This helps to maximize the number of panel observations we have. As reported above, these data also aggregate to those reported annually on Chinese manufacturing in the Chinese Statistical Yearbook.

3.2 Variable construction

Gross output

Our gross output measure is production, which differs from sales because of changes in inventories and corresponds more closely to the reported input use.

To construct an output deflator at the most detailed level possible, we use information from the 1998-2003 surveys, for which firms report the value of their output in both nominal and real prices. Output in real terms is constructed by firms using a set of "reference" prices provided by NBS. In any given year, the ratio of nominal to real output provides an index of prices in that year relative to the base. The change in this index between year t and $t-1$ then is a measure of the firm-specific price change between the two years, which we average to the four-digit industry level.¹² We drop as outliers those observations for which the price change differs by more than a half standard deviation from the mean, or approximately 15-25% of observations. We then recalculate the weighted average for each sector, using as weights each firm's current output.¹³ These changes are linked over time to provide an output deflator for each of the 423 sectors.¹⁴

¹¹The decline in age of newly entering firms shows up even more pronouncedly in the higher percentiles. For example, the 90th percentile age declines from 21 to 9 years over the sample period, with again a slight transitory up-tick in 2004.

¹² We utilize information on the changes rather than the levels because of the noise introduced into our measures by the entry and exit of firms into a sector.

¹³ We obtain extremely similar results using the average or the median when aggregating the firm-level information to the industry-level.

¹⁴ The output deflators, as well as those for capital and intermediary inputs, are available in the online Appendix mentioned in the introduction.

For 2004-2006, we use the two-digit (38 sectors) ex-factory price index in the 2007 China Statistical Yearbook. Output prices decreased between 1998 and 2002 following the Asian Financial Crises and began to rise after 2003. Between 1998 and 2006, output prices rose by only 9.3 percent, most of which occurred after 2004.

Value added

Real value added is constructed by separately deflating output, net of goods purchased in for resale and net of indirect taxes, and material inputs.

Input deflators are calculated using the output deflators and information from the 2002 National Input-Output table. The sectors defined in the Input-Output table are less disaggregated than our output deflator. The IO sectors are mainly at the 3-digit level, but there are also 2-digit and 4-digit industries. We construct a concordance table linking IO sectors and 4-digit industries. We obtain indices of output prices for each IO sector by taking an un-weighted average of prices at the 4-digit level. To obtain the input deflators for each IO sector, we calculate a weighted average of the output deflators, using as weights the coefficients in the 2002 IO table.

Between 1998 and 2006, input prices rose by 20.3 percent, more than twice as fast as output prices. Much of this reflects the rapid increase in raw material and energy prices that occurred after 2003. The input deflators are weighted averages of output deflators, with a much larger weight on industries producing raw materials and energy products than in total manufacturing output.

One potential source of measurement error is the input deflators for export processors. Legally, these firms are allowed to import raw materials and intermediates used directly in export processing duty-free. Between 1998 and 2006, there was a marked reduction in tariffs on intermediate goods, which we expect to be reflected in prices. In principle, intermediate goods prices for export processors should not have been affected, but we have no way to construct alternative indices. For these firms, our input deflators are likely biased downwards, which leads to an overestimate of real input use, and an underestimate of value-added.

Employment and wages

Firms report total annual employment and several components of employee compensation. The latter include wages, employee supplementary benefits, and unemployment insurance. Firms began to report retirement and health insurance in 2003 and housing benefits in 2004. We utilize information on the first three categories to generate consistent measures of compensation over the entire period, but conduct a robustness check including the other measures.

Reported compensation is likely to underestimate total payments to labor. Labor's share of value added in our sample is only 34.2 percent. By comparison, the national income accounts suggest a share of labor in all of industry that is between 55 and 60 percent. In some of our productivity estimation, we follow Hsieh and Klenow (2007) and inflate wage payments by a constant factor for all firms to work with a

wage share consistent with the national average. We are forced to rely on the assumption that the bias in the non-reporting of wage payments is the same across all firms.

Over time there is a slight decrease in the wage share in value added, which is the result of three evolutions. First, average firm size is declining and the wage share is strongly increasing in firm size. Second, the share of wages in state-owned firms, which initially was much higher than average, converged over time towards the lower average for private firms. And third, domestically-owned private firms, which always had the lowest average wage share, increase in importance.

Capital stock and investment

Firms do not report their fixed investment. Each firm reports information on the value of their fixed capital stock at original purchase prices, and their capital stock at original purchase prices less accumulated depreciation. As these book values are the sum of nominal values for different years, we have to make some assumptions to convert them into real values that are comparable across time and across firms. Failure to do so will likely introduce systematic biases into our estimates of the capital stock related to a firm's age.

Our procedure for estimating the real capital stock in each year begins with estimating the real value of the capital stock in the first year in which the firm appears in our data set. For simplicity of presentation here, we assume that it is 1998, the first year of our panel. Ideally, we need information on all of a firm's past investments (as well as capital disposal) to construct the real capital stock in 1998. This we do not have. As an alternative, we use information from the 1993 annual enterprise survey to construct estimates at the two-digit industry level by province of the average rate of growth of the nominal capital stock between 1993 and 1998.¹⁵ Combined with information on the age of each firm, we use these estimates to calculate the firm's nominal capital stock in the year in which it was established, or NK_E . The real capital stock, or RK_E in the same year is obtained by deflating the initial nominal capital stock with the investment deflator constructed in Brandt and Rawski (2008, Chapter 20).

The nominal capital stock in each year up through 1998 is then calculated by multiplying the initial nominal capital stock by $(1 + r_{ps})^n$, where n is the number of years since the firm was established, and r_{ps} is the estimated rate of growth in the nominal capital stock in province p in sector s . Annual investment is the change in nominal capital stock between years. The real capital stock for 1998 is calculated using the perpetual inventory method, assuming a depreciation rate of 9% and deflating annual investment using the Brandt-Rawski deflator. For years after 1998, we use the observed change in the firm's

¹⁵ We apply the same method for those firms that enter our sample after 1998, but recalculating the nominal rate of growth in the capital stock.

nominal capital stock at original purchase prices as our estimate of nominal fixed investment. The same rate of depreciation and deflator is used to roll the real capital stock estimates forward.¹⁶

Table A.2 in the Appendix illustrates the potential implications of alternative assumptions made in the construction of the capital stock. Irrespective of the construction method, capital per worker increases noticeably over the sample period. Average capital per firm tends to be more sensitive to the assumptions. The un-deflated 'net' capital stock (per firm) is relatively flat over time, while our preferred measure shows an initial increase, but a decline after 2001 as less capital-intensive firms enter the sample.

Capital per worker in our preferred benchmark estimate increases by 21.3% from 1998 to 2006, which amounts to 2.4% per year. The best estimate using official statistics, the far right results in the top panel of Table A.2, only shows an increase of 12.4% over the sample period. This adjustment lowers estimated productivity growth rates substantially, as the capital share in output is extremely high in China.

The adjustment has further impact on the comparison between firms. Table A.3 in the Appendix illustrates that our procedure leads to a larger upward adjustment in the capital stock of state-owned firms than that of other firms. This is largely because the former tend to be older. On average, the adjustment is 15 percent larger for state-owned firms. Correcting this relative underestimation of capital input for state-owned enterprises is one factor that reduces estimated productivity levels and growth rates in the capital stock for that group of firms.

Ownership

We utilize information on a firm's registered type (*qiye dengji zhuce leixing*) to classify ownership into five basic groups: state-owned, hybrid/collective, private, and two types of foreign firms, those from Hong-Kong, Macau, and Taiwan (HMT) and those from all other countries.¹⁷ We use additional information on a firm's registered capital to classify *limited liability corporations* and *shareholding corporations* as state, hybrid or private. Foreign categories combine both joint ventures and wholly-owned subsidiaries.

In Table 3, we report information on the number of firms and employment by ownership for the beginning and ending year of our sample. Over time, the role of state-owned firms falls markedly on both counts. The same is true for collective firms. The fastest growing component is private firms, which increase as a share of the total number of firms from 9.0% to 59.5%, and in terms of employment from only 5.2% to 38.6%. Foreign-invested firms increase only slightly as a share of the total number of

¹⁶ The Brandt-Rawski deflator is a chain-linked price deflator based on separate price indices for equipment-machinery, and buildings-structures. The weights are the share of these items in fixed investment, as reported by China's National Bureau of Statistics.

¹⁷ Hong Kong-based subsidiaries of MNCs are included in Foreign I. There is no easy way to untangle them from the rest of the firms, short of looking at them on a case-by-case basis.

firms, and then largely because of the increase in the number of non-HMT firms, but their share of employment more than doubles from 15.1 to 33%.

[Table 3 approximately here]

The changing importance of ownership types in Table 3 is the product of three forces: exit, new entry, as well as ownership changes that occur in the course of enterprise restructuring. Figure 1 shows entry and exit by ownership for the five categories. For private firms, for example, gross entry over the entire sample period is nearly 190,000, however nearly a third of all private firms exit subsequently. From 2002 onwards, private firms make up more than half of all entrants; from 2005 onwards, they also make up half of all exiting firms.

[Insert Figure 1 approximately here]

Table A.4 in the appendix reports for select “pairs” of years the transition between ownership categories for continuing firms. The percentage of continuing firms experiencing a change in ownership rises from 6.4% between 1998-1999 to 9.7% between 2002 and 2003, and then falls slightly to 7.8% between 2005 and 2006. Approximately one-half of the changes are associated with privatization. In each of the years, the collective and state categories lose more firms than they gain, while the private category is the destination for the majority of ownership switches.

Results in Table A.5 for the 2005 to 2006 matches further illustrate that ownership changes are three times more likely when firm information other than ID is used to match firms in adjacent years. It demonstrates the importance of matching on information other than ID to follow firms over restructuring episodes. The category by category transitions also highlight the preponderance of switching into the private category and the switching out of the state and even more the hybrid/collective category.

Industry Classification

Each firm is classified into an industry following the 4-digit Chinese Industry Classification (CIC) system that resembles the old U.S. SIC system. In 2003, the classification system was revised to incorporate more detail for some sectors, while some other sectors were merged. To make the industry codes comparable across the entire period, we constructed a harmonized classification that groups some industries prior to and post the revision.

3.3 Final sample

Some firms have missing observations for some of the variables necessary to calculate productivity. Statistics in Table 4 illustrate how many firms have missing, zero, or negative values for the real capital stock or value added, which are dropped from the analysis as the logarithms of the input and output variables are not defined. We treat firms with missing employment similarly, and further drop all firms with less than 8 employees as they fall under a different legal regime. As a result, 17% of the original

firms are dropped from the sample in 1998, but this fraction declines to approximately 6% in each year after 2001. 34,777 firms remain in the sample throughout.¹⁸

[Table 4 approximately here]

4. Results

4.1 Firm-level productivity growth

The benchmark estimates for average productivity growth rates by year are depicted in Figure 2. These measures use the index number formula in equation (4) for a value added production function on the full unbalanced panel of firms. Firm-level growth rates are aggregated using value added weights, averaged over the initial and end year.

[Figure 2 approximately here]

Output growth has been extremely rapid in the Chinese manufacturing sector, outstripping the well-documented economy-wide growth rate of 10%, by almost a factor of two. Here we find that firm-level total factor productivity growth has been extremely high as well, rising from an average of just over 2% in 1999 to almost 14% in 2006. The average over the full period stands at 9.5%.¹⁹

Figure 3 presents several alternative TFP growth estimates. These measure different objects, but the overarching message of extremely rapid productivity growth prevails. In each case, productivity growth is significantly lower in the first four years of the period, 1998-2001, than in the period following China's accession to the WTO.

[Figure 3 approximately here]

Results in the top-left quadrant show that productivity growth rates are approximately 2% higher if a parametric production function is assumed, which is estimated at the three-digit industry level using the Olley-Pakes methodology. This is consistent with firms expanding input use more rapidly for those inputs that they use relatively more intensively, which is what we expect if they are exploiting their comparative advantage.

Productivity growth rates using a gross output production function, in the top-right quadrant, are much lower. Abstracting from weighting issues and assuming that intermediate inputs are proportional to output, the ratio between TFP calculated using a gross output and a net output production function

¹⁸ The total number of active firms in the first column in Table 4 differs slightly from the corresponding column in Table 2, as firms that exit and re-enter the sample are considered operating throughout in Table 2, even though we do not observe them.

¹⁹ To construct the averages in Figure 2, the one percentile at the top and the bottom are dropped, but the effect on the average is minimal.

should be equal to the share of value-added in gross output. In the sample, the latter rises over time from 0.25 to 0.33 as the average share of intermediate inputs and energy in total output falls over time. Relative TFP using gross output behaves similarly: in the first few years it is just under one quarter of the value added TFP growth, and in the latter period it is a third. At 3.8% annually between 2001 and 2006, TFP growth on a gross output basis is also extremely high—explicit comparisons with other countries are below.

In the bottom-left quadrant, limiting the sample to the balanced panel of firms that are active throughout the entire sample period raises the average productivity growth initially. It does not make any difference to the estimates in the second half of the period. This is the result of two opposing forces that will be illustrated in greater detail below. On the one hand, many of the firms active in 1998 are performing poorly and will exit the industry in subsequent years. The survivors outperform this group by a factor of almost two in the initial years. On the other hand, many of the new entrants are very productive and enjoy especially rapid productivity growth in their first years. In the second half of the sample, the intermediate performance of survivors results in average productivity growth.

Finally, the unweighted averages of the firm-level productivity growth rates, reported in the bottom-right quadrant, are more than 2% lower per year than the value added weighted averages.²⁰ From this we can conclude that, somewhat unusually, large Chinese firms are increasing productivity at a higher than average rate. The positive correlation holds for both productivity measures, for all ownership categories, and for value added as well as employment weights.²¹

The restructuring of large state-owned firms is one explanation for the pattern in China. Alternatively, the inclusion rule in the sample based on annual sales means that some small firms only make the sample by virtue of an extremely high productivity level, from which further improvements might be difficult.

To put the extraordinary productivity growth performance documented in Figures 2 and 3 into perspective, it is important to keep in mind that at least four beneficial factors were jointly at play. The impact of each of these factors accelerated over time.

First, the sample period covers the cyclical upswing following the Asian financial crises. Second, China's entry into the WTO and its integration into the world economy lead to exports rising by more than 25% annually over the sample period, which especially benefitted the manufacturing sector. Third, restructuring of state-owned enterprises (SOEs) and collectively-owned firms started in earnest in the

²⁰ The weight used is the average of current and past value added. With current value added weights the average productivity growth would be higher still, and with lagged value added it would be lower. With employment weights instead of value added, the average would be lower as well.

²¹ There is an abundance of international evidence that firm growth is negatively correlated with size, but the evidence for productivity growth is a lot weaker. Griliches and Mairesse (1983) find a negative correlation between size and productivity growth for firms in the United States and France, but the plant-level evidence for the U.S. in Bailey, Hulten, and Campbell (1992) is weak or insignificant once they control for firm-level growth.

mid-1990s, and accelerated through the early part of the period we analyze. And fourth, liberalization and competitive pressures in the manufacturing sector exceeded those in much of the tertiary sector, especially finance and telecommunications.

We gain added perspective by comparing these numbers with similar statistics for other countries. Average statistics for both value added and gross output total factor productivity growth for several countries are reported in Table 5.²² If information was available for several sub-periods, we chose cyclical expansions to match the Chinese situation.

Compared to these other countries, China's performance remains exemplary, although not completely unheard of. It records the highest level of firm-level productivity growth of any country for which a comparable average could be found for a value added production function in manufacturing. Slovenia, also a transition economy, records a close second productivity growth at 8.5% versus 9.5% for China. After the recession of the early 1980s, the record of U.S. plants is also better than commonly assumed, with a total factor productivity growth of 3.1% per year.

[Table 5 approximately here]

The Chinese total factor productivity growth using a gross output production function averaged 2.6% over the full sample period, and 3.7% over the last five years. This record is bested by Colombia (4.7%) and Korea (3.5%).²³ Results for other countries indicate that a sustained TFP growth in excess of 2% is a rare event, even in cyclical booms.

4.2 Firm-level productivity growth: Robustness checks

An important message from Young (2003) is that measurement issues matter. Using aggregate statistics for the entire non-agriculture sector, he shows that with alternative price deflators and adjustments for input quality the productivity growth estimate is reduced from 3% to 1.4% per year over the period

²² To compare the two sets of statistics, one should scale up the gross output based numbers by one over the value added share in output.

²³ Results for Colombia are not entirely comparable as they refer to productivity calculated using physical output in the numerator. Firm-specific price increases are subtracted from TFP growth and it is shown that these are negatively correlated with physical productivity growth. As a result, productivity growth using a value concept (TFPR) is lower.

between 1978 and 1998.²⁴ In Figure 4 we explore how sensitive our average productivity growth statistics are to a variety of assumptions.²⁵

We consider five possible reasons why our benchmark estimate could be biased upward. These possibilities follow directly from the definition of productivity growth in equation (4), i.e. output growth minus weighted input growth: (1) Price inflation is underestimated. (2) Labor input growth is underestimated. (3) Initial capital stocks are overestimated. (4) The weight on capital is too low. (5) The sample selects on success.

[Figure 4 approximately here]

A direct input into the productivity calculations are the price deflators for nominal output and input to construct real value added. Up through 2003, the NBS relied on firm-level reports of output at nominal and at 'reference' prices to construct implicit price deflators at the industry level. Young argues that out of convenience firms increasingly tended to report the same value of output at both price levels.²⁶ Alternative price deflators, constructed from separate NBS administered price surveys, suggest higher price inflation and correspondingly lower productivity estimates.

To verify the sensitivity of our productivity estimates, we show alternative estimates in Figure 4 using two alternative official deflators that were constructed from independent price surveys, i.e. not from data collected through the manufacturing surveys that are the source of our firm data. First, the two-digit deflator that is already used for the 2004-2006 period is used wherever it is available (2002-2006). Second, an even more aggregate deflator, which tracks prices in 14 manufacturing sectors, is used throughout the entire period.

The changes with the first deflator are negligible; however we do find lower productivity growth using the most aggregate deflator. The magnitude of the adjustment is non-negligible, lowering productivity growth on average by 0.68% per year, but the change is smaller than Young (2003) finds for the period 1978-1998. The problems he highlights may be more severe for the rapidly expanding service industry.

²⁴ Brandt and Zhu (2009) resuscitate the higher TFP estimates. First, revisions to nominal GDP by the NBS in 2006, which were not available to Young, add nearly 0.6 percent to real output growth and TFP. Second, they construct an alternative service sector deflator that better reflects changes in the price of both consumer and business services. Young's alternative to the NBS implicit deflator for services is based solely on the price of consumer services. Nonetheless, Young's general concerns remain valid.

²⁵ The "benchmark" estimate in Figure 4 is slightly higher than the 9.5 percent average reported earlier. For purposes of consistency in our comparisons in Figure 4, we dropped any firm for which we were not able to construct all of the alternative TFP estimates in the figure. To further make sure the sample of firms did not change in the alternative calculations, we also did not drop the top and bottom percentile in the productivity growth distribution, except for our preferred estimates at the bottom.

²⁶ As we directly observe the reported output at the two price levels, we can verify Young's claim directly for the period of our analysis. The fraction of firms in the manufacturing sector that report the same output value using current and constant prices ranges from 16.9% to 21.4% and there is no distinct time trend. There are also many firms that report higher numbers at constant prices than at current prices, indicating price declines.

Using a more aggregate price index is not innocuous either. When Young (2003) introduced the new price deflator he adjusted value-added directly, rather than deflating separately gross output and input. Under the plan, input prices were held down and with liberalization they grew more rapidly than output prices, increasing value added in real terms. The one-digit deflator lacks the detail to properly take this into account.

Moreover, a more aggregate price index is likely to introduce further errors as more successful sub-sectors are able to lower costs and prices as they expand and make up a larger share of their sector. Applying the same price deflator to the entire (aggregate) sector will understate the productivity growth of the more successful sub-sectors.

The second important adjustment that Young (2003) proposes is to control for increases in human capital. Over the twenty-year period he studies, 1978-1998, when educational attainment increased rapidly, this is clearly an important issue. Again, this is likely to be less important for us as our time period is a lot shorter.

We do not observe any measure of human capital at the firm-level, but can utilize growth in the wage bill as an imperfect proxy for a broader labor input measure that includes human capital improvements. This measure grows a lot more rapidly than the absolute number of employees for three reasons: more hours worked per employee, rising marginal productivity of labor, and rents. In addition to TFP, the marginal product of labor will rise with increases in K/L and improvements in human capital.

The results in Figure 4 illustrate that this adjustment is very important. It lowers average productivity growth by 2.69% per year or even 3.33% if other employment benefits, such as pensions, employment insurance, and housing benefits, are included as well. It seems that some combination of human capital improvement and longer average working hours could explain in upwards of one quarter to one third of the measured productivity growth.

To the extent that wages rise because workers are able to appropriate some rents, the contribution of human capital is of course overestimated and productivity growth underestimated. There is some evidence that this is going on. For one, the difference between wage growth and employment growth increases markedly as the business cycle became more expansionary and labor markets were getting tighter. The average differential rises from 2.5% in 1999 to 10.3% in 2006; the median differential increased steadily from 4.8% to 8.5%. New entrants also see higher wage growth than incumbents, presumably in part because they have to compete harder to attract workers away from other firms.

An additional factor spurring wage increases is that more competitive labor markets and a more liberal stand towards migration increasingly allows the best firms to attract the best workers. This permits a more efficient employment of the available labor. This contribution to productivity growth is also taken out of the alternative productivity measures.

The third factor that potentially leads to a bias in the productivity growth calculation is the construction of the capital stock. The original capital stock measure in the data set is a hybrid sum of undeflated

investment flows and undepreciated past investments (“original value at purchase prices”). Especially if firms are older or highly capital-intensive, the imperfect capital stock measure will bias productivity growth estimates. In particular, if the initial stock is overestimated, capital growth will be underestimated and productivity growth estimates biased upward.

In the new capital stock measures that we construct, we tried to control as accurately as possible for depreciation and capital price inflation in the initial capital stocks and then use a perpetual inventory measure to roll the stock forward. Compared with the original capital stock measures, deflated using the same capital stock deflator, our benchmark productivity growth estimate is 1.45% higher—9.69% versus 8.24%. One important reason for the increase is that for many state-owned enterprises, which tend to be older and more capital intensive, the capital stock series was underestimated in the initial year. This inflated their initial relative productivity level and their capital input growth rates.

The same explanation is consistent with the higher average productivity growth that is obtained when all firms active in 1998 are dropped and the productivity growth calculations is limited to firms entering our data during the sample period. It avoids problems estimating the initial capital stock, but does lead to a different sample. On average, productivity growth is 0.99% higher per year.

Alternatively, we excluded all firms that reported a birth-year more than two years before the year they are first observed in the sample. This drops the majority of firms active in 1998 and about three quarters of newly observed firms in 1999. The fraction of newly observed firms that are indeed start-ups gradually increases over time and stands at 60% by the end of the period.

The productivity performance of new entrants is markedly better, with an average annual growth rate of 12.93%. The average is almost 1 percent higher if the definition of entrant is tightened to having a birth year at most one year prior to observed entry. Note that we average productivity growth over the entire post-entry period. The increase is the combined result of the selected sample and of the greater weight on observations later in the sample period, when growth was higher.

In both of these last two calculations, problems with the accurate measurement of the initial capital stock are avoided, but life cycle patterns in the productivity evolution are likely to play a role. We explore productivity dynamics over the firms’ life-cycle further in the next section.

A fourth and final ingredient in the productivity growth calculations that might matter is the relative weight on employment and capital input growth.²⁷ As mentioned earlier, we follow Hsieh and Klenow (2009) and inflated the average wage share to half of value added to match the fraction in the national accounts.

²⁷ We conducted several additional robustness checks that had minor effects and we do not report full results. Results are insensitive to the inclusion of firms with missing observations in intervening years; to the exclusion of firms with less than 25 employees (7% of the sample), which are most likely to be selected into the sample based on high productivity levels; or to the exclusion of firms that are only observed twice.

As capital input has grown more rapidly than labor input, this procedure leads to higher measured productivity growth. If we suspect that labor costs are underreported, we should trust the higher estimate for the wage share; if we suspect labor to claim a smaller share of value added in manufacturing than in the rest of the economy, the average unadjusted share is slightly above one third, we should trust the lower estimate. Results in Figure 4 illustrate that the magnitude of the adjustment is to raise productivity growth by almost 1% per year: 8.70% (unadjusted) versus 9.69% (adjusted).

A related issue is that several firms do not report non-wage employment benefits. To the extent that this is a reporting problem, the weight on their employment growth will be underestimated and that of all other firms overestimated, as the average share is fixed at one half. Excluding non-wage benefits for all firms results in a minor downward adjustment to productivity growth of 0.08% per year, suggesting that firms that report benefits see their capital-labor ratio grow more rapidly than other firms.

In light of the importance of the issues discussed, we incorporate some of them in our productivity growth calculations. In particular, we adjust the calculations for the fact that firms might underreport price changes, for human capital accumulation and longer average working hours through the business cycle, and for a lower weight of labor input in manufacturing. As the truth is likely in between our benchmark assumptions and the extreme assumptions in Figure 4, we go half way on each adjustment.

In particular, we use the geometric average of output growth using the detailed and aggregate (14 sector) deflator; the geometric average of labor input growth based on employment and on wage costs; and we only adjust the wage share half the way to one half. Each of these three corrections has a downward effect on productivity growth. The results, at the bottom of Figure 4, indicate a preferred productivity growth estimated based on a value added production function of 7.67% annually and 2.73% for a gross output production function.

4.3 Entering and exiting firms

The results in the preceding section average across the growth rates of all enterprises, including incumbents, new entrants, and firms about to exit the industry. Two important additional factors for aggregate productivity growth are (i) where in the distribution the new firms are entering and where the old ones are disappearing and (ii) how the productivity growth evolves over a firm's life cycle. These topics are particularly important for China given the high observed rates of firm turnover reported in Table 2.

The results of a simple OLS regression of firm-level productivity on a full set of industry-year dummies and five post-entry dummies are reported in Table 6. Additional controls include ownership type, dummies for coastal provinces, and a dummy for a founding year at most two years before the entry year.

We report results for a number of alternative specifications. First, in column (1), we compare entrants to all active firms and measure productivity using the Olley-Pakes methodology, which assumes a specific parametric form for the production function. Firms enter slightly below the average productivity level of

incumbents (-2.4% in log points), but they improve rapidly. In the first full year post-entry they already show a distinct advantage, +0.078, which declines only slightly in the following years.

[Table 6 approximately here]

In column (2), the same comparison is made, but using the non-parametric index number methodology to construct productivity measures. Allowing for heterogeneity in technology, entrants enter at a much higher productivity levels, 0.13 log points above the median (or 13.9%). In their second year this advantage rises even further, but it then declines rapidly as firms predominantly expand the use of inputs that they use most intensively.

Results for the same regressions but with productivity growth rates as the dependent variable are in columns (3) and (4). The pattern over time for the two productivity measures turns out to be very similar: a large above-average productivity boost after the initial entry year—5.4% or 4.6% higher growth than incumbents—and much smaller differences, albeit still positive ones, in subsequent years.

The decline in relative productivity in years two to four post-entry is consistent with above average productivity growth because of entry of even newer plants towards the higher end of the distribution and the exit of the most poorly performing firms. The implicit comparison with the superior newest entry cohorts is a drag on productivity levels of older entrants especially if the index number measures are used. An additional reason why the growth regressions show a continued growth advantage of recent entrants even though their productivity level advantage declines in columns (1) and (2) is that surviving entrants are selected more on productivity growth than productivity level.

The next set of results, in columns (5) and (6), redefine the definition of entrants by retaining only firms that report a birth-year of at most 2 years before the first year they are observed in the sample.²⁸ The main difference is that the initial productivity level in the first year is a lot lower, while the subsequent pattern of rapid, above average growth and subsequent convergence is maintained. This makes perfect sense as these firms are more likely to be operating for only part of the first year and have to fine tune their operations initially.

Finally, in the last two columns entrants are compared to a more limited control group: the set of incumbents that operate throughout the sample period. A difficulty interpreting the results in the first six columns is that entrants see very large changes in productivity over their first couple of years. This is even more pronounced for firms which are about to exit the industry. As a result, the comparison of a specific cohort to a mixed group of firms with different production histories is not always straightforward.

²⁸ Recall that our sample only includes firms that sell at least 5 million RMB per year. Some new entrants have been growing for a long time from a small base. In addition, the business registry for the universe of firms has been improving over time, and fewer 'old' firms tend to be picked up later on.

Compared with the group of firms operating throughout the entire sample period, the relative productivity of entrants is reduced throughout. Firms that manage to survive over the entire turbulent period are clearly not a random group.²⁹ The relative change of the post-entry coefficients over time, capturing how an entry cohort's relative position evolves, is remarkably similar in columns (7) and (8) and columns (5) and (6).

The results also indicate that firms that survive for four or more years have converged almost completely with those in the balanced panel. The $t+4$ coefficients in columns (7) and (8) are no longer significantly different from zero. For the Olley-Pakes measure this is already true for the $t+3$ coefficient. Surviving entrants seem to converge very rapidly to the productivity level of incumbents.

In all columns we observe an apparent overshooting in productivity level in the first few years post-entry, which disappears over time. To some extent, this is merely a feature of the changing experience of entrants over the sample period, which is explored in Figure 5. Here, coefficients like those in column (8) that are estimated year by year are plotted over time.

[Figure 5 approximately here]

Over time, the initial productivity level of new firms is declining as captured by the thick dotted line. Note that firms are only labeled "entrants" if in the year they are first observed they indicate a birth year at most two years ago. As the Chinese market liberalized, one might have expected the entry process to change. For example, in a model like Hopenhayn (1992) a lower productivity draw would still lead a firm to enter the industry if fixed entry costs had been reduced. The greater market opportunities, especially after WTO accession in 2001, could also lead to more experimentation and more opportunistic entry.

If market selection weeds out the worst performing entrants, we expect the average productivity for survivors in subsequent years to improve upon their entry-year productivity. This is indeed the case in every year since 2002 and is especially pronounced in 2005. In contrast, in the first years of the sample period the average performance of surviving entrants versus incumbents goes down over time, perhaps because market potential was only one of several reasons for market entry.

An important pattern to note is that the relative initial productivity of entrants declines over time, but this does not hold for productivity a few years after entry. Most pointedly, the relative performance of entrants that manage to survive three or even four years sees a distinct upward trend. For the cohort

²⁹ The difference between the full and balanced sample results is somewhat larger for the OP than for the IN results. The reason is that firms in the balanced panel only make up 20% of observations and the parametric production function estimates will mostly reflect the technology chosen by firms entering and exiting the sample. This has two consequences. First, firms in the balanced sample are larger on average, while estimated scale economies are decreasing. Second, those same firms tend to operate with more capital, which receives a lower weight in the production functions estimated on the full sample. Both features lead to higher measured productivity for firms in the balanced sample using the parametric estimates rather than the index numbers.

that entered in 1999, the performance three or four years onwards was below the productivity level of incumbents (three or four years onwards). The 2002 entry cohort, on the other hand, outperformed incumbents three or four years later on. The productivity advantage of the 2003 entry cohort relative to incumbents even ballooned to 10% after only three years (in 2006), while the 2004 cohort surpassed even that mark after only two years.

Productivity level and growth results for exiting firms are reported in the bottom panel of Table 6. There are fewer surprises to note here. In the final year of operation, firms that subsequently exit are a lot less productive than the average surviving firm. The orders of magnitude vary from -24.6% (-0.282 log points) for the index number measure relative to the full sample to -34.5% for the Olley-Pakes measure relative to the balanced sample firms.

These differences are very large and they open up mostly in the last couple of years of operation. Most of the measures, except for the index number comparison with the balanced panel, show very little difference four years before exit. Growth rate differentials in year $t-4$ are insignificant from zero as well. At that point, the productivity of future exiting firms starts to deteriorate very rapidly. The results also demonstrate that the index number comparisons, which adjust for differences in the importance of labor and capital across firms, evolve much more smoothly than the Olley-Pakes measures that impose a specific functional form on the production function. Especially for these struggling firms, this methodological difference seems to matter a lot.

4.4 Aggregate productivity growth

Next, we investigate the relative contribution of incumbent firms and changes at the extensive margin—entry and exit—for the evolution of aggregate productivity growth. The seminal contribution was the productivity decomposition of Baily, Hulten, and Campbell (1992), which was improved for unbalanced panels by Haltiwanger (1997). Defining the aggregate productivity level as $\ln TFP_t = \sum_i \theta_{it}^Q \ln TFP_{it}$, the aggregate growth rate can be decomposed as

$$\begin{aligned} \Delta \ln TFP_t = & \sum_i^{\text{survivor}} (\theta_{it}^Q \ln TFP_{it} - \theta_{it-1}^Q \ln TFP_{it-1}) \\ & + \sum_n^{\text{enter}} \theta_{nt}^Q (\ln TFP_{nt} - \ln TFP_{t-1}) - \sum_x^{\text{exit}} \theta_{xt-1}^Q (\ln TFP_{xt-1} - \ln TFP_{t-1}) \end{aligned} \quad (6)$$

where the last two terms represent the contribution of net entry and θ_i^Q is the gross output share of firm i .

Foster, Haltiwanger, and Krizan (2001) construct various decompositions for the U.S. manufacturing sector and note that the relative contribution of net entry was one of the most robust patterns. From the many alternatives they report, we reproduce in the top-left panel of Table 7—labeled ‘Baily-Hulten-

Campbell-Haltiwanger'—the results for 1977-1987 that Bartelsman and Doms (2000) also report in their survey article.

The above decomposition has generated some controversy. In particular, Petrin and Levinsohn (2006) propose the following alternative:

$$\begin{aligned} d\widehat{TFP}_t = & \sum_i^{\text{survivor}} (\theta_{it}^Y + \theta_{it-1}^Y)/2 \Delta \ln TFP_{it} \\ & + \sum_n^{\text{enter}} \theta_{nt}^Y/2 (\ln TFP_{nt} - \ln \widehat{TFP}_{n\tau}) - \sum_x^{\text{exit}} \theta_{xt-1}^Y/2 (\ln TFP_{xt-1} - \ln \widehat{TFP}_{x\tau}) \end{aligned} \quad (7)$$

There are three differences: (i) Value added weights are used instead of gross output weights. (ii) The contribution of survivors is limited to firm-level growth, the so-called “within term”, excluding the contribution of changes in weights. With some simple algebra, the first term in (6) for each survivor can be written as $(\theta_{it} + \theta_{it-1})/2 \Delta \ln TFP_{it} + \Delta \theta_{it} (\ln TFP_{it} + \ln TFP_{it-1})/2$, and only the first term is retained in (7). (iii) The productivity levels in the entry and exit terms are adjusted for the relative productivity of firms at the time of entry or exit (τ), not by the average productivity level in the initial period. Because $\ln \widehat{TFP}_{zt}$ ($z=n,x$) is inherently unobservable, regression analysis is used to construct an estimate.

Even though this decomposition has not been performed yet on (confidential) U.S. census data, we can use published statistics from Foster, *et al.* (2001) to construct an approximate decomposition that at least accounts for the second and third difference.³⁰ To facilitate the comparison with China for which we observe a shorter time period, all growth rates are expressed per year.

While the contribution of net entry increases for longer time horizons, aggregate productivity growth in the United States in the decade from 1977 to 1987 is clearly driven by the experience of incumbents. One reason is that new entrants enter at below average productivity and it takes time for the surviving entrants to catch up. Another reason is that re-allocation of inputs and output share towards more productive incumbents provide a strong ongoing productivity boost among continuing plants, a pattern documented below.

[Table 7 approximately here]

For comparability, we perform the exact same calculations on the Chinese data, i.e. sticking with gross output weights also in (7) and calculating TFP_{it} using the same input elasticities for all firms (average

³⁰ Comparing the two decompositions, we can write the modified ‘PL’ net entry (NE) term as [modified NE] = [usual NE]/2 - $(\ln \widehat{TFP}_{n\tau} - \ln TFP_{t-1}) \sum_n \theta_{nt}^Y/2 + (\ln \widehat{TFP}_{x\tau} - \ln TFP_{t-1}) \sum_x \theta_{xt-1}^Y/2$. The entry and exit dummies from regressions in Foster, *et al.* (2001) are used to substitute the two terms in brackets.

input shares). The cumulative aggregate productivity growth of 0.369 (4.0% annually), is slightly higher than the gross output production function results reported in Figure 3.³¹

The Chinese total growth largely exceeds the corresponding U.S. statistic, which only stands at 1.0% per year over the full period or 1.5% over the higher growth period 1982-1987. If the changes in output share among continuing firms are ignored and the productivity differences of entrants and exiting firms at time of entry or exit are accounted for, according to the PL decomposition, the difference is much smaller. The U.S. annual technical change of 0.8% is now almost one half of the Chinese 1.9% performance. These numbers can be interpreted as welfare gains associated with technical change if markets are competitive. Furthermore, the statistics are even more similar for the within term, the firm-level productivity growth for incumbents. Over the entire period, the annual growth rate is 1.1% for China and 0.7% for the US.

The most important difference between the Chinese and the U.S. results is the greater relative importance of net entry for China.³² The exact magnitude of the difference depends on the time period and decomposition chosen, but the direction is nearly always the same. Over the full sample period and for the PL decomposition, which controls for the initial productivity advantage of entrants, net entry is responsible for 0.8% of aggregate productivity growth per year; this represents 42% of the total, three times as much as in the United States. If we include the initial advantage of entrants in their contribution, as in the BHCH decomposition, their annual effect in China amounts to a staggering 2.7% aggregate productivity growth annually, ten times the U.S. equivalent (0.27%), and more than half of aggregate total TFP growth. The lower output share, but higher survival rate for U.S. entrants contributes to this difference.

Even though new entrants tend to be more productive than the exiting firms they replace, they tend to be small and will only account for a substantial fraction of output after several years. Even more than for the United States, the contribution of net entry increases with the length of the time period. This is particularly pronounced for the PL decomposition, which is intuitive as it controls for the higher productivity of Chinese entrants at the time of entry.

The difference between the two decompositions is larger for China than for the United States. Entering U.S. plants are estimated to be only 0.3% more productive than incumbents in the initial period (Foster,

³¹ These results are not directly comparable with those from Figure 3. The earlier numbers are year-on-year productivity growth estimates averaged over the entire sample, and sample period. The current aggregate is the change in the weighted average productivity level of all firms active in 1998 versus all firms active in 2006. What happens in between is irrelevant now.

³² A potential concern is that firms with missing information may artificially introduce entry and exit into the sample for China. The effect of this is likely to be marginal because we only use firms active in the first or last year of the period over which we do the calculation. Alternative estimates, where we drop from the sample all observations for a firm with any missing information in some year, are very similar.

et al. 2001, Table 10), while Chinese entrants already show a 9.4% productivity advantage.³³ Similarly, exiting U.S. plants are estimated to be 1.9% less productive than average, while the difference is 6% for China.³⁴

In addition, the impact of output share reallocations, $\Delta\theta_{it}(\ln TFP_{it} + \ln TFP_{it-1})/2$, which is added to the continuing plants contribution in the first decomposition but not in the second, is only 0.9% in the United States versus 2.3% in China. Over the sub-periods in the sample, this difference is even larger. Petrin and Levinsohn (2006) note that this “between term” only contributes to welfare if factor or product markets are not competitive, which Hsieh and Klenow (2009) argue is exactly the case in China.

A more direct way of looking at the impact of the allocation of inputs across firms of differing productivity levels is provided in Figure 6, where we replicate Figure 1 from Bartelsman and Dhrymes (1998) in the left panel. In an efficient economy we expect that resources are put to their most efficient use. Some aggregate growth will result from above-average productive plants commanding a greater share of inputs over time, which is exactly the case in the United States.

The graph shows the evolution of the unweighted productivity average and the aggregate productivity level, both normalized by their level in the initial period. Log productivity is purged from time and four-digit industry effects and exponentiated before aggregating. The plant-weight to construct the aggregate level is the following input aggregate, $L_{it}^{\alpha_L} K_{it}^{1-\alpha_L}$, but results are similar for output weights. The gap between the two lines is the extent to which the input weight is being shifted towards more productive plants.³⁵ Results in the top graphs use residuals from a Cobb-Douglas production function estimated with least squares. In the bottom graph, a Törnqvist index (Solow residual) for productivity is constructed using plant-specific factor shares as input weights.

[Figure 6 approximately here]

The U.S. pattern is remarkable. The unweighted average slumps by 0.45% annually if the Cobb-Douglas productivity measures are used and is unchanged for the corrected Solow Residual. A decline reflects lower dispersion as the average of the log-productivity measures are normalized to zero in each year and the exponential function is convex. The weighted average on the other hand, which corresponds to the evolution of aggregate productivity, increases annually by 2% in the case of the Cobb-Douglas and by 1.2% for the Solow Residual. Note that these effects are solely the result of changing input weights

³³ This advantage is larger than both differences in Table 5 because we follow Foster, *et al.* (2001) and only include year dummies, not industry-year dummies, and do not limit the comparison group to firms from the balanced panel.

³⁴ There are slight differences in China for the two periods. Between 1998 and 2002, entrants were 11.1% more productive than incumbents, while exiting firms were 5.4% less productive. For the post-2002 period, the percentages were 7.0 and -8.1%, respectively.

³⁵ The productivity decomposition in Olley and Pakes (1996) for the U.S. telecommunications equipment illustrates the same phenomenon. The gap between the two lines in Figure 4 corresponds to the importance of the covariance term in the popular Olley-Pakes decomposition.

and are in addition to 0.8% and 0.1% average annual productivity increase common across plants. The different evolution between the unweighted and weighted average can be the results of either larger plants increasing their productivity advantage or of plants with the highest productivity level increasing their weight over time. Other evidence of great persistence in relative productivity levels leads Bartelsman and Dhrymes (1998) to prefer the latter explanation.

In the right panel of Figure 6 we plot the corresponding patterns for China (in black).³⁶ The decline in the unweighted average appears again, but is slightly less pronounced than in the United States for the top measure. There is also no recovery towards the end of the period, suggesting a more persistent decline in dispersion. Regression to the mean is one explanation; a tendency for entrants to have productivity levels closer to the mean than surviving incumbents or exiting firms is a second. In addition, while the mean of log productivity is normalized to zero, the median declines over time indicating a downward shift in the mass of the distribution. While positive outliers survive, negative outliers disappear and the upward shifted mean is re-normalized every year.

A more pronounced difference between the U.S. and Chinese experience is the vastly smaller contribution to the aggregate productivity growth of inputs shifting to more productive firms. For the Cobb-Douglas results, the reallocation of resources provides some positive effect, but the cumulative effect over eight years is barely 4% (the difference between the dashed and solid black lines). For the Corrected Solow Residuals, shifts in input weights are virtually unrelated to productivity differences. The difference with the U.S. pattern is so large that we had to rescale the vertical axis by a factor of two for China.

As a robustness check, we also construct the patterns for China on the balanced panel of firms operating throughout the eight year period. The absence of entry and exit diminishes the decline in the unweighted average, consistent with the explanations above, but the contribution of input reallocations is scarcely more important.

In sum, we have shown that relative to the U.S. experience, productivity growth in China's manufacturing sector is to a much greater extent due to changes at the extensive margin, entry and exit. The absolute contribution of firm-level productivity growth is similar in both countries. Moreover, aggregate growth is constrained by extremely limited efficiency-enhancing input reallocations. These results confirm a wealth of anecdotal evidence that describes the continued support of less efficient manufacturing firms with significant ties to the state.³⁷ The comparison with the United States provides some illustration of how much potential productivity growth is thus sacrificed.

³⁶ Bartelsman and Dhrymes (1998) pool plants from three related two-digit industries and their Cobb-Douglas results use the same input coefficient estimates for all. For China, we re-estimate the production function for each two-digit industry.

³⁷ In 2005, for example, 36 percent of all industrial SOEs were loss-making compared to 20 percent of all other non-state firms with sales larger than 5 million RMB.

5. Conclusions

Drawing on an unbalanced panel of firms that covers most of China's manufacturing sector, the purpose of this paper has been to examine the dynamics of productivity growth over a period that spans China's entry into WTO. By all indications, productivity growth has been very rapid, a finding that appears to be robust to a host of measurement issues. Improvements in productivity of "continuing firms", either as a result of restructuring efforts, or investments in capability building are an important part of the picture. Equally, if not more important, are the gains to creative destruction, i.e. entry and exit, that China's decentralized reforms have increasingly allowed. However, these gains remain tempered by limitations on efficiency-enhancing input reallocations to more productive firms that are constraining their growth.

These findings have important implications. First, they affirm the important role that policy liberalizing entry as well as facilitating exit are playing in Chinese manufacturing. Both are likely to be very important moving forward. Second, the run-up in productivity growth we estimate for the years after 2002 provides a potential explanation for the acceleration in manufacturing exports, which represent between 25-30% of total manufacturing output. The rapid productivity growth also suggests that there may have been more room for China to let the exchange rate appreciate than policy makers believed.³⁸ And third, despite the dynamism we document, our results point to continued constraints on the growth of some of the most productive of firms. Problems in the allocation of credit, and biases in favor of larger firms with state-sector connections are a potential source. More work is needed to identify the exact nature of these constraints.

³⁸ Tightness in labor markets, especially in the coastal provinces, was sending similar signals. In the fall of 2007, for example, turnover in labor in manufacturing firms in Guangdong was running as high as 50 percent.

Appendix: Algorithm to match firm observations over time

Our matching process can be divided into three stages.

Stage 1: Match the samples of two consecutive years

First, we match the firms by firm ID. Table A.1 reports for each year the percentage of firms that can be matched by firm ID. The firm IDs are alpha-numeric. Most firms IDs are unique, however each year between 10 and 30 firms have duplicate IDs. Before we matched the firm IDs, we changed all letters in firm IDs to upper case. The difference in upper/lower case accounts for about 1-2% of the unmatched firms.

Second, for those firms that cannot be matched by ID, we use additional information to link them. We create new codes that use various combinations of firm name (in Chinese), name of legal person representative (in Chinese), geographic code, industry code, founding year, phone number, address, name of main product. As shown by Table A.1, this procedure enables us to match annually an additional 1%-8% of all firms.

Stage 2: Match three consecutive years

Firms may disappear from the sample and reenter later. We tried to link the re-entrants to earlier years (about 1000-3000 firms each year) using procedures similar to those described in stage 1. Consider the following case: From stage 1, Firm A in Year 1 is not matched to any firm in Year 2. But in stage 2 it's linked to Firm C in Year 3 that has already been matched to Firm B in Year 2. Then we can conclude that Firm A and Firm B are the same firm.

Stage 3: Create a ten-year panel

We merged the files created in stage 2 to construct a ten-year panel.

Online Appendix

<http://www.econ.kuleuven.be/public/n07057/China/>

The following information is available in the online appendix:

1. Industry codes: files with the concordance of pre-2003 CIC, post-2003 CIC, and more aggregate IO categories into a consistent 4-digit industry classification.
2. Deflators for outputs and inputs (3 alternatives: firm-level based 4-digit, 14-industry, 2-digit), capital goods deflator (2 alternatives: Brandt-Rawski and official series)
3. Program used to link firms over time using ID and other identifying variables
4. Program to calculate the new capital stock series based on book values of capital, firm age, and province-industry average investment growth rates

Bibliography

- Ahn, S., K. Fukao, H. U. Kwon. "The internationalization and performance of Korean and Japanese firms: An empirical analysis based on micro-data." *Seoul Journal of Economics*, 17(4), Winter 2004, 439-482.
- Aw, B. Y., X. Chen, and M. J. Roberts. "Firm-level evidence on productivity differentials and turnover in Taiwanese manufacturing." *Journal of Development Economics* 66(1), October 2001, 51-86.
- Bai, C.-E., C.-T. Hsieh, and Y. Qian. "The return to capital in China." *Brookings Papers on Economic Activity* 2, 2006, 61-88.
- Baily, M. N., C. Hulten, and D. Campbell. "Productivity dynamics in manufacturing plants." *Brookings Papers: Microeconomics* 4(1), 1992, 187-267.
- Bartelsman, E. J. and P. J. Dhrymes. "Productivity dynamics: U.S. manufacturing plants, 1972-1986." *Journal of Productivity Analysis* 9(1), January 1998, 5-34.
- Bartelsman, E. J. and M. Doms. "Understanding productivity: Lessons from longitudinal microdata." *Journal of Economic Literature* 38, September 2000, 569-594.
- Brandt, Loren, Thomas Rawski, and John Sutton, "Industrial Development in China." In Loren Brandt and Thomas G. Rawski (Eds.), *China's Great Economic Transformation*, New York: Cambridge University Press, 2008, pp. 569-632.
- Brandt, L. and X. Zhu. "Explaining China's Growth." Mimeo. University of Toronto. 2009.
- Caves, D. W., L. R. Christensen, and E. W. Diewert. "The economic theory of index numbers and the measurement of input, output, and productivity." *Econometrica* 50(6), Nov. 1982a, 1393-1414.
- Caves, D. W., L. R. Christensen, and E. W. Diewert. "Multilateral comparisons of output, input, and productivity using superlative index numbers." *Economic Journal* 92(365), March 1982b, 73-86.
- Chow, G. C. and K.-W. Li, "China's economic growth: 1952-2010." *Economic Development and Cultural Change* 51(1), 2000, 247-256.
- De Loecker, J. and J. Konings. "Job reallocation and productivity growth in an emerging economy. Evidence from Slovenian manufacturing." *European Journal of Political Economy* 22, 2006, 388-408.
- Dougherty, Sean, Richard Herd and Ping He. "Has a private sector emerged in China's industry? Evidence from a quarter of a million Chinese firms." *China Economic Review*. 18, 2007, pp. 309-334.
- Eslava, M., J. Haltiwanger, A. Kugler, and M. Kugler. "The effects of structural reforms on productivity and profitability enhancing reallocation: Evidence from Colombia." *Journal of Development Economics* 75, 2004, 333-71.
- Foster, L., J. Haltiwanger, and C. Krizan. "Aggregate productivity growth: Lessons from microeconomic evidence." In C. R. Hulten, E. R. Dean, and M. J. Harper (Eds.), *New Developments in Productivity Analysis*, Volume 63 of NBER Studies in Income and Wealth, pp. 303-63. Chicago, IL: University of Chicago Press.
- Foster, L., J. Haltiwanger, and C. Syverson. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* 98(1), March 2008, 394-425.
- Griliches, Z., and J. Mairesse. "Comparing Productivity Growth: An Exploration of French and U.S. Industrial and Firm Data." *European Economic Review* 21, March-April 1983, 89-119.

- Haltiwanger, J. C. "Measuring and analyzing aggregate fluctuations: The importance of building from microeconomic evidence." *Federal Reserve Bank St. Louis Review* 79(3), May-June 1997, 55-77.
- Hopenhayn, H. A. "Entry, exit, and firm dynamics in long run equilibrium" *Econometrica*, 60(5), September 1992, 1127-1150.
- Hsieh, C.-T. and P. J. Klenow. "Misallocation and manufacturing TFP in China and India." Working Paper, Department of Economics, Stanford University, July 2007.
- Islam, N., E. Dai, and H. Sakamoto. "Role of TFP in China's growth." *Asian Economic Journal* 20(2), 2006, 127-159.
- Jefferson, G. H., T.G. Rawski, and Y. Zhang. "Productivity growth and convergence across China's industrial economy." *Journal of Chinese Economic and Business Studies* 6(2), May 2008, 121-140.
- Naughton, B. "A political economy of china's economic transition." In L. Brandt and T. Rawski (Eds.), *China's Economic Transition: Origins, Mechanisms, and Consequences*. 2008, New York: Cambridge UP.
- Olley, G. S. and A. Pakes. "The dynamics of productivity in the telecommunications equipment industry." *Econometrica* 64(6), 1996, 1263-97.
- Park, A., D. Yang, X. Shi, and Y. Jiang. "Exporting and firm performance: Chinese exporters and the Asian financial crisis." *Review of Economics and Statistics* (forthcoming) July 2006.
- Pavcnik, N. "Trade liberalization, exit, and productivity improvement: Evidence from Chilean plants." *Review of Economic Studies* 69(1), January 2002, 245-76.
- Petrin, A. and J. Levinsohn. "Measuring aggregate productivity growth using plant-level data." Working Paper, University of Minnesota, Department of Economics, April 2006.
- Solow, R. M. "Technical change and the aggregate production function." *Review of Economics and Statistics* 39, 1957, 312-320.
- Tybout, J. R. and M. D. Westbrook. "Trade liberalization and the dimensions of efficiency change in Mexican manufacturing industries." *Journal of International Economics* 39, 1995, 53-78.
- Van Biesebroeck, J. "Firm size matters: Growth and productivity growth in African manufacturing." *Economic Development and Cultural Change* 53(3), March 2005, 545-84.
- Van Biesebroeck, J. "Robustness of productivity estimates." *Journal of Industrial Economics* 55(3), September 2007, 529-569.
- Van Biesebroeck, J. "The sensitivity of productivity estimates: Revisiting three important productivity debates." *Journal of Business and Economic Statistics* 26(3), July 2008, 321-338.
- World Bank. *Vietnam Development Report 2006: Business*. Joint Donor Report to Vietnam Consultative Group Meeting, Hanoi, December, 2007.
- Young, A. "The tyranny of numbers: Confronting the statistical realities of the East Asian growth experience." *Quarterly Journal of Economics* 110(3), August 1995, 641-680.
- Young, A. "Gold into base metals: Productivity growth in the People's Republic of China during the reform period." *Journal of Political Economy* 111(6), 2003, 1220-1261.

Table 1A: Comparison of sample coverage with China Statistical Yearbook**(a) Firm-level data set**

Year	Number of	Value added	Sales	Output	Employment	Export	Net value of fixed assets
1998	165,118	1.94	6.41	6.77	56.44	1.08	4.41
1999	162,033	2.16	6.99	7.27	58.05	1.16	4.73
2000	162,883	2.54	8.42	8.57	53.68	1.46	5.18
2001	169,030	2.79	9.24	9.41	52.97	1.61	5.45
2002	181,557	3.30	10.95	11.08	55.21	2.01	5.95
2003	196,222	4.20	14.32	14.23	57.49	2.69	6.61
2004	279,092	6.62	20.43	20.16	66.27	4.05	7.97
2005	271,835	7.22	24.69	25.16	68.96	4.77	8.95
2006	301,961	9.11	31.36	31.66	73.58	6.05	10.58

(b) China Statistical Yearbook 2006 (Table 14-6) and China Statistical Abstract 2006

Above Scale Industrial Firms

Year	Number of	Value added	Sales	Output	Employment	Export	Net value of fixed assets
1998	165,080	1.94	6.41	6.77	61.96	1.08	4.41
1999	162,033	2.16	6.99	7.27	58.05	1.15	4.73
2000	162,885	2.54	8.42	8.57	55.59	1.46	5.18
2001	171,256	2.83	9.37	9.54	54.41	1.62	5.54
2002	181,557	3.30	10.95	11.08	55.21	2.01	5.95
2003	196,222	4.20	14.32	14.23	57.49	2.69	6.61
2004	276,474	5.48	18.78	20.17	66.22	4.05	7.38
2005	271,835	7.21	24.46	25.16	67.85	4.79	8.81
2006	301,961	9.11	31.36	31.66	73.58	5.96	10.58

Table 1B: Comparison of sample coverage with 2004 Census**(a) Firm-level data set**

Year	Number of firms	Sales	Output	Employment	Export	Original value of fixed assets
2004	279,092	20.43	20.16	66.27	4.05	12.54

(b) Census

Year	Number of firms	Sales	Output	Employment	Export	Original value of fixed assets
Above Scale Industrial Firms						
1995	110,977	4.78		61.03		4.02
2004	279,040	20.42	20.17	66.22	4.05	12.58
Below Scale Industrial Firms						
1995	382,227	0.42		22.66		0.42
2004	1,098,789	1.98	2.06	26.82	0.11	1.24
Fraction below scale						
1995	77.5%	8.1%		27.1%		9.5%
2004	79.7%	8.8%	9.3%	28.8%	2.5%	9.0%

Notes: Above scale means SOEs and all other firms with sales above 5 million RMB. All values are denoted in Trillion RMB and employment in millions of workers. All industrial firms are included, while the analysis in the paper is limited to firms in the manufacturing sector.

Table 2: Evolution of the panel over time

	Active firms	Destination		Origin	
	Total	Continuing	Exiting	Incumbents	Entrants
1998	148,685	127,330	21,355		
1999	147,553	126,220	21,333	127,330	20,223
2000	148,905	119,429	29,476	126,220	22,685
2001	156,658	138,416	18,242	119,429	37,229
2002	167,193	147,977	19,216	138,416	28,777
2003	183,699	152,226	31,473	147,977	35,722
2004	259,239	220,256	38,983	152,226	107,013
2005	251,067	230,023	21,044	220,256	30,811
2006	278,753			230,023	48,730

Note: Firms that exit and later re-enter the sample are considered to be operating throughout, even though we do not observe them.

Table 3: Number of firms and employment by ownership category

	Number of Firms				Employment (millions)			
	1998		2006		1998		2006	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
State	53,396	35.9%	32,488	11.7%	25.4	55.1%	14.0	22.1%
Hybrid/collective	55,907	37.6%	20,456	7.3%	11.8	25.6%	4.0	6.3%
Private	13,333	9.0%	165,912	59.5%	2.4	5.2%	24.5	38.6%
Hong Kong, Macau, Taiwan	15,472	10.4%	28,740	10.3%	3.8	8.2%	10.2	16.1%
Foreign	10,574	7.1%	31,157	11.2%	2.7	5.9%	10.7	16.9%
Total	148,682	100%	278,753	100%	46.1	100%	63.4	63.4m

Table 4: Number of observations with nonmissing input and output variables

	Active firms	Firms with missing, zero, or negative			Remaining observations	
	Total	Capital	Value added	Employment (≤ 8)	All firms	Balanced panel
1998	148,685	5,121	12,243	13,444	123,304	34,777
1999	146,101	3,349	10,837	5,796	132,972	34,777
2000	147,253	2,542	9,666	9,393	128,721	34,777
2001	154,324	2,424	7,261	3,810	144,375	34,777
2002	165,861	2,328	8,156	3,590	155,183	34,777
2003	181,079	2,036	6,580	2,474	172,097	34,777
2004	258,945	5,009	14,539	6,709	240,805	34,777
2005	251,061	2,302	10,555	2,396	237,674	34,777
2006	278,753	2,362	13,458	3,163	262,117	34,777

Table 5: Average TFP growth in different countries

country	study	period	sector	TFP growth
<u>Firm-level: value added production function</u>				
China		1998-2006	manufacturing	0.095
Slovenia	De Loecker-Konings (2006)	1994-2000	manufacturing	0.085
Vietnam	World Bank (2007)	2001-2003	manufacturing	0.073
U.S.	Baily-Hulten-Campbell (1992)	1982-1987	selected manufacturing	0.031
Chile	Pavcnik (2002)	1979-1986	manufacturing	0.028
U.S.	Baily-Hulten-Campbell (1992)	1982-1987	selected mfct, excl. 3533, 3711	0.019
<u>Firm-level: gross output production function</u>				
China		1998-2006	manufacturing	0.026
Colombia	Eslava, <i>et al.</i> (2004) ¹	1982-1998	manufacturing	0.047
Korea	Ahn-Fukao-Kwon (2004)	1990-1998	manufacturing	0.035
Taiwan	Aw-Chen-Roberts (2001)	1986-1991	9 manufacturing sectors	0.021
Mexico	Tybout-Westbrook (1995)	1984-1990	manufacturing	0.019
U.S.	Haltiwanger (1987)	1982-1987	manufacturing	0.017
Japan	Ahn-Fukao-Kwon (2004)	1994-2001	manufacturing	0.003

Note: ¹ The numerator in the productivity growth in the Eslava *et al.* (2004) paper refers to physical output, i.e. controls for price differences. As productivity and prices are negatively correlated, more efficient firms charge lower prices, this leads to a higher productivity growth measure than using revenue in the numerator.

Table 6: Productivity evolution**(a) Entering firms**

Relative to:	full sample		full sample		full sample		balanced panel	
Productivity:	level		growth		level		level	
Entrants:	all		all		only recent		only recent	
	OP	IN	OP	IN	OP	IN	OP	IN
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
t_0 (entry)	-0.024	0.130			-0.088	0.046	-0.146	-0.004
	(0.003)	(0.003)			(0.004)	(0.004)	(0.005)	(0.006)
$t+1$	0.078	0.169	0.053	0.045	0.089	0.151	0.030	0.100
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.004)	(0.005)	(0.006)
$t+2$	0.071	0.131	0.021	0.011	0.076	0.134	0.017	0.084
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.005)	(0.006)	(0.006)
$t+3$	0.071	0.066	0.009	0.005	0.059	0.076	0.000	0.025
	(0.003)	(0.003)	(0.002)	(0.002)	(0.005)	(0.005)	(0.006)	(0.006)
$t+4$	0.062	0.034	0.015	0.012	0.052	0.046	-0.008	-0.006
	(0.004)	(0.004)	(0.003)	(0.003)	(0.006)	(0.006)	(0.007)	(0.007)

(b) Exiting firms

Relative to:	full sample		full sample		balanced panel	
Productivity:	level		growth		level	
	OP	IN	OP	IN	OP	IN
t_0 (exit)	-0.359	-0.282	-0.163	-0.173	-0.424	-0.331
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
$t-1$	-0.061	-0.129	-0.014	-0.022	-0.125	-0.177
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
$t-2$	-0.140	-0.111	-0.059	-0.062	-0.204	-0.160
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
$t-3$	-0.200	-0.063	-0.040	-0.035	-0.264	-0.112
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)
$t-4$	0.037	-0.047	0.003	0.005	-0.029	-0.094
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)

Table 7: Decomposition of annual TFP Growth for Manufacturing Establishments**(a) United States**

method:	Baily-Hulten-Campbell-Haltiwanger				Petrin-Levinsohn		
Census period	Total Growth	Continuing plants	Net entry	Technical change	Continuing plants	Adjusted net entry	
1977-87	0.010	0.008 (74%)	0.003 (26%)	0.008	0.007 (86%)	0.001 (14%)	
1977-82	0.005	0.004 (75%)	0.001 (25%)	0.003	0.003 (91%)	0.000 (9%)	
1982-87	0.015	0.013 (86%)	0.002 (14%)	0.012	0.011 (93%)	0.001 (7%)	

Source: The first set of statistics are from Bartelsman and Doms (2000), who took them from Foster, Haltiwanger, and Krizan (2001). The second set of statistics are approximate results for the Petrin and Levinsohn (2006) decomposition for the United States, calculated based on published statistics from Haltiwanger (1997) and Foster, *et al.* (2001).

Notes: In the BHCH decomposition, the 'Continuing plants' term is the sum of the within, between, and covariance terms. In the 'net entry' term the productivity level of entrants and exiters is normalized by the aggregate initial productivity level. Output weights are used. In the PL decomposition, the 'Continuing plants' term is limited to firm-level growth (the within term in BHCH). The contribution of entrants and exiters is normalized by the average productivity at the entry or exit point for those plants, which is obtained using regression analysis. Value added weights are used, which are averaged over the initial and end period, such that entrants or exiters are weighted by half of their value added share

(b) China

method:	Baily-Hulten-Campbell-Haltiwanger				Petrin-Levinsohn		
Census period	Total Growth	Continuing firms	Net entry	Technical change	Continuing firms	Adjusted net entry	
1998-2006	0.040	0.013 (33%)	0.027 (67%)	0.019	0.011 (58%)	0.008 (42%)	
1998-2002	0.032	0.017 (53%)	0.015 (47%)	0.010	0.010 (97%)	0.000 (3%)	
2002-2006	0.049	0.026 (54%)	0.022 (46%)	0.027	0.021 (78%)	0.006 (22%)	

Source: Own calculations based on Chinese firm-level data

Figure 1: Entry and Exit by Ownership Type

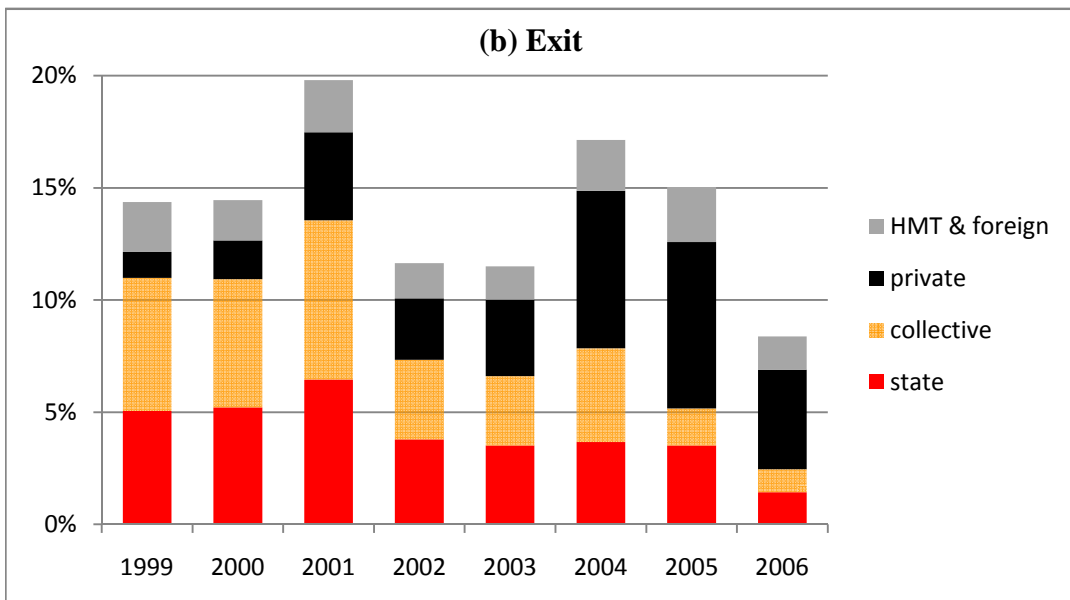
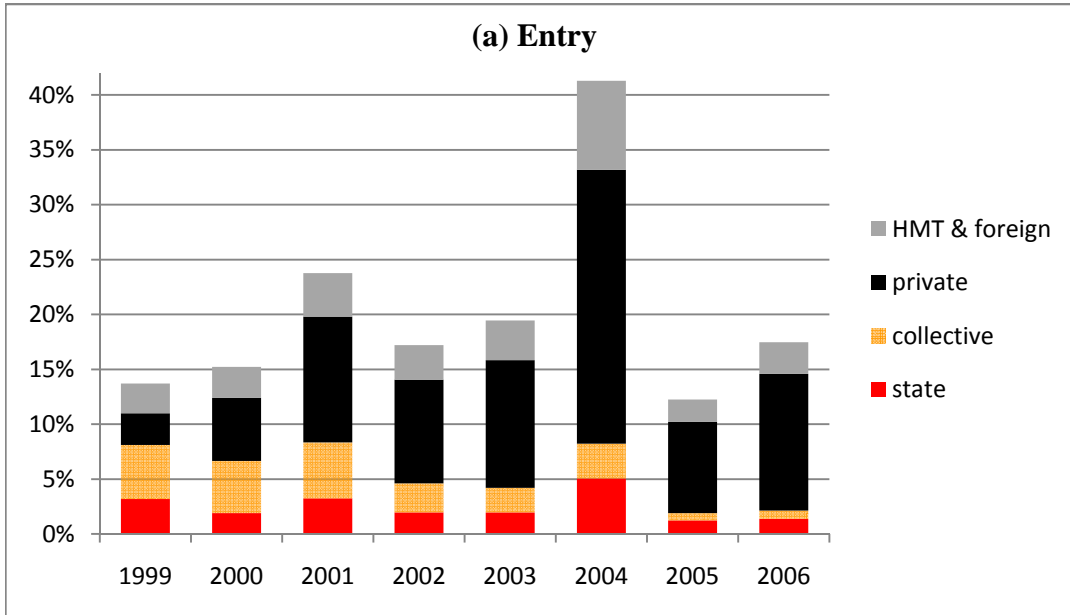


Figure 2: Total Factor Productivity growth

(unbalanced sample, value added weights, index numbers)

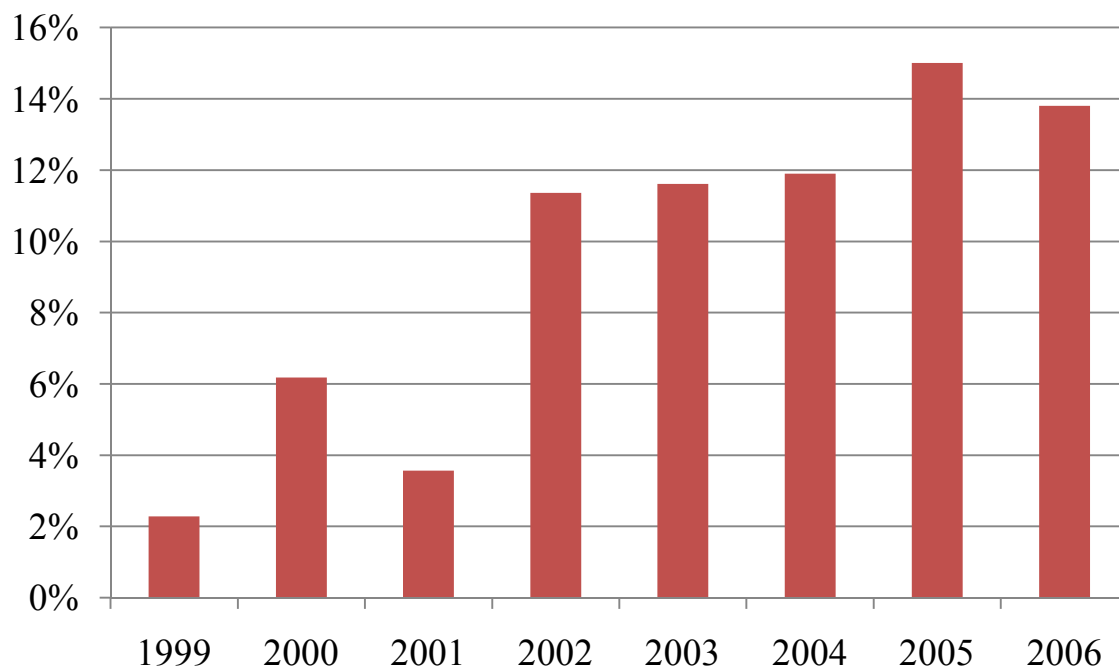
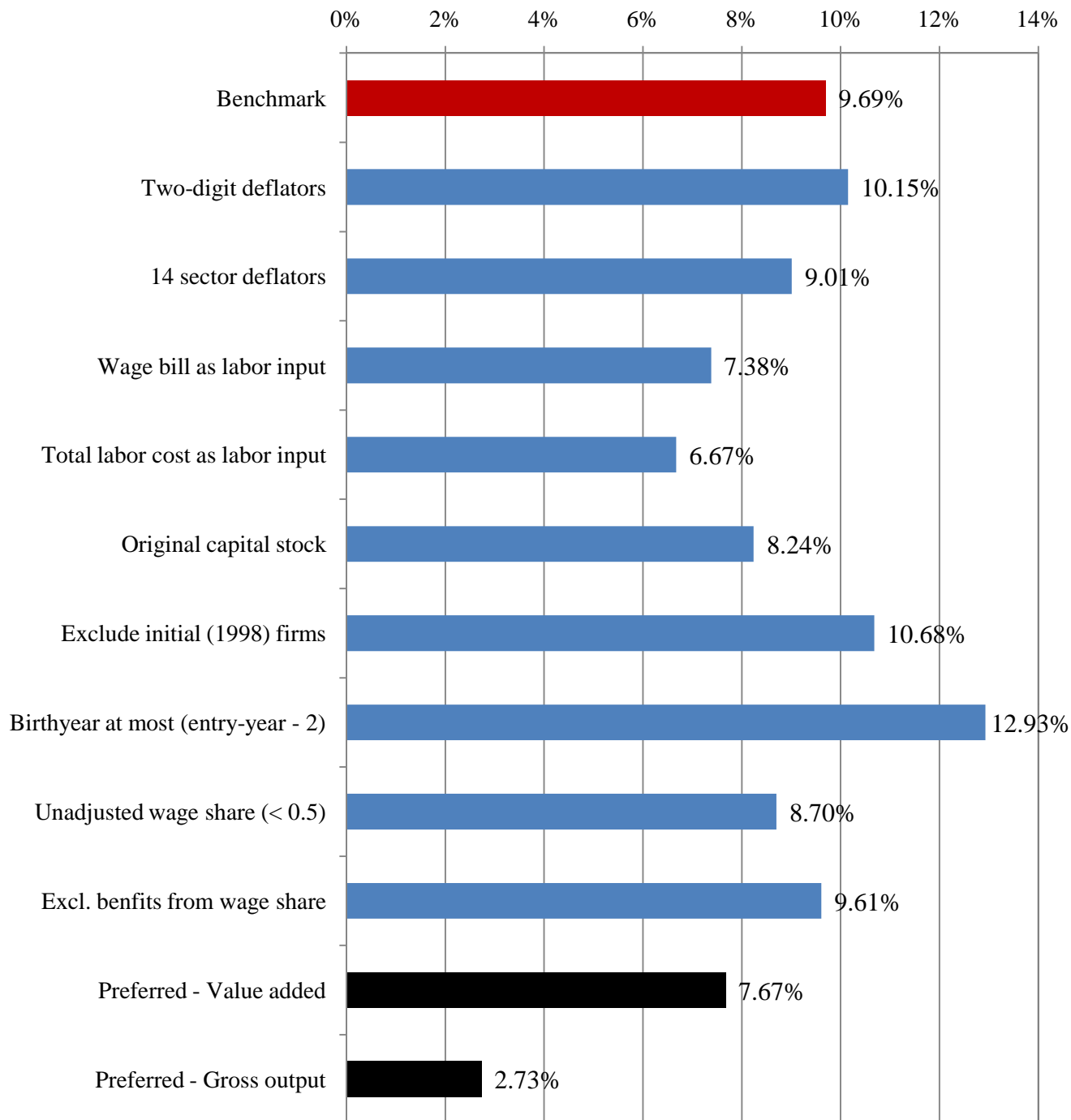


Figure 3: TFP growth estimates calculated differently



Note: The benchmark estimates always use the index numbers to calculate productivity (instead of the Olley-Pakes parametric methodology); use value added weights (instead of unweighted averages); are calculated on the full sample of firms (instead of the balanced panel); and use a value added production function (instead of gross output).

Figure 4: Robustness of TFP growth estimates to a variety of assumptions



Note: The sample for the calculations in bars 1-6 and 9-10 differs slightly from that in Figures 2 and 3 and Table 5 earlier, as firms with some missing data are dropped throughout to make the samples comparable across bars. Therefore the benchmark estimate of 9.69% is slightly higher than the 9.5% in Table 5. For the same reason, the top and bottom percentile of productivity growth observations are not dropped either. For the last two (“preferred statistics”) calculations, the sample is again identical to the sample in earlier Figures and Tables.

Figure 5: Productivity-levels for new entrants versus incumbents

(Index numbers, only recently-founded firms)

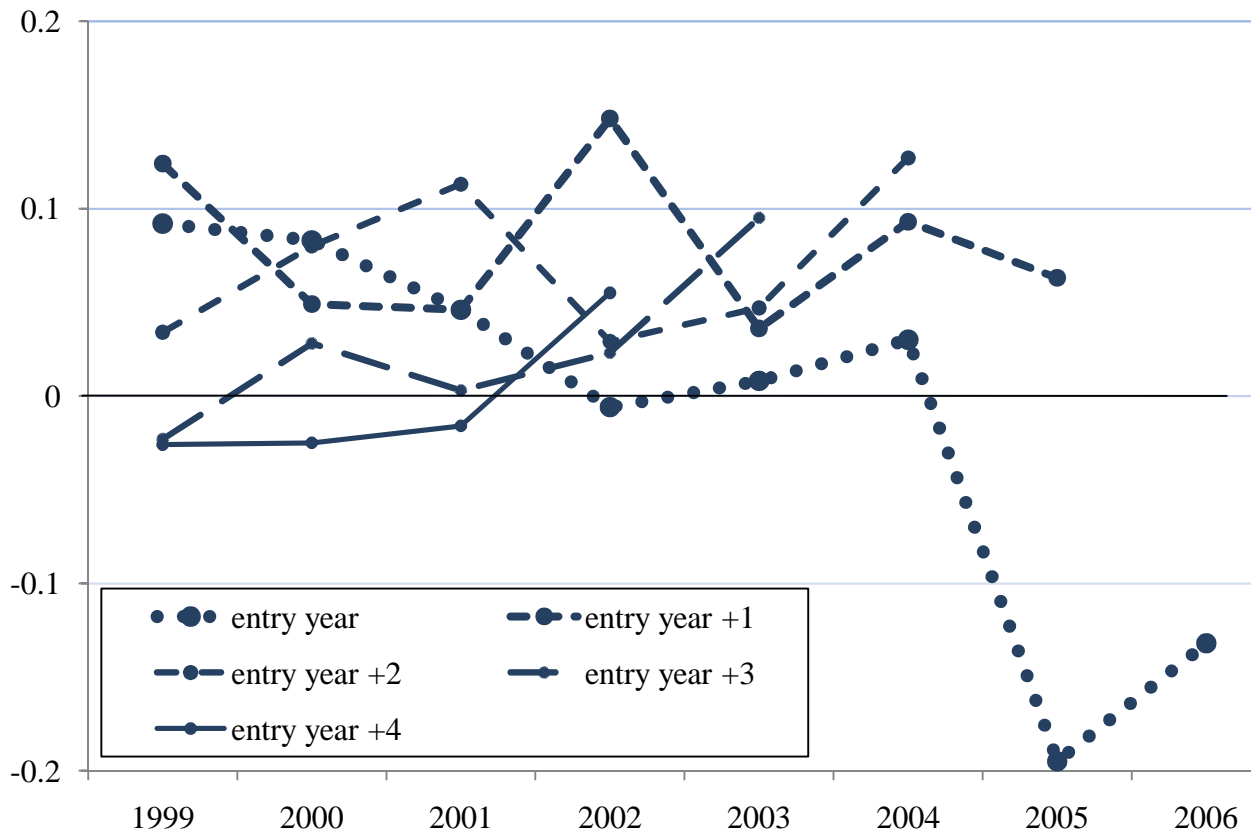
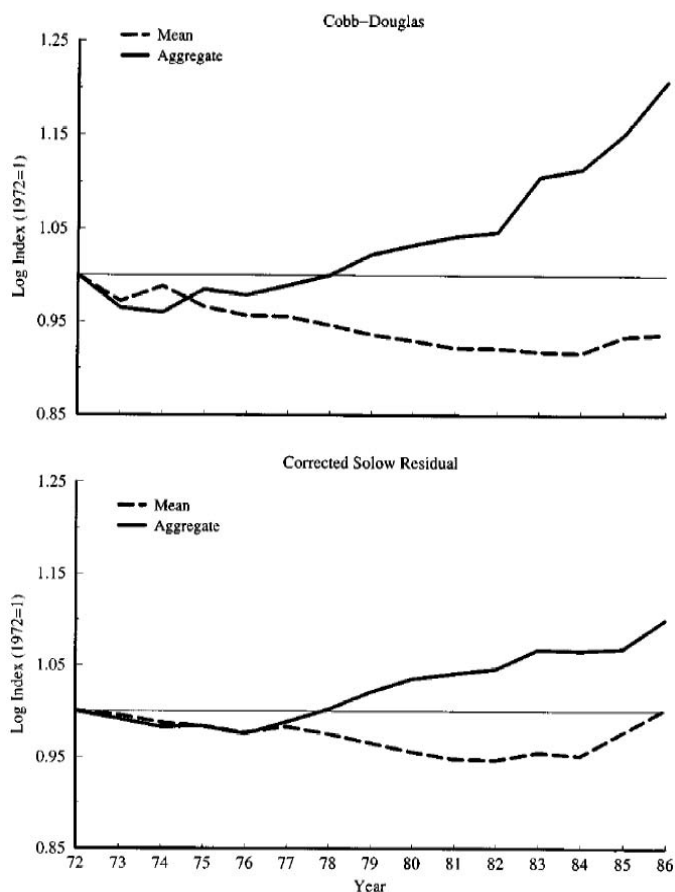


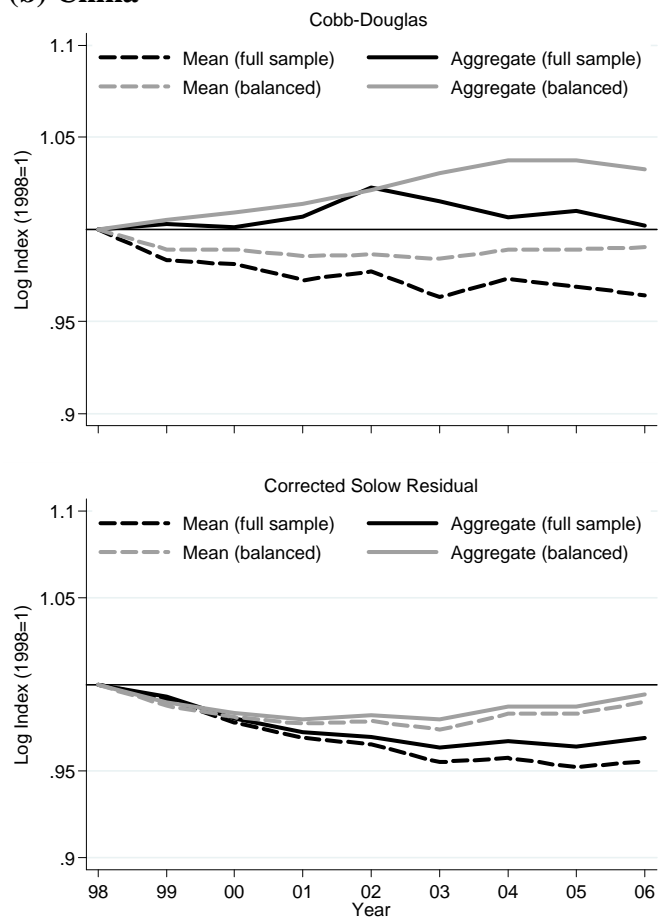
Figure 6: Productivity residuals

(a) United States



Source: Bartelsman and Dhrymes (1998), Figure 1

(b) China



Source: Own calculations based on Chinese firm-level data

Notes: Cobb-Douglas productivity measures are calculated from a production function estimated by least squares. Solow Residuals refer to a Tornqvist index number using the wage share as weight on labor input and the material share for materials; in the United States a "capital share" is observed, while in China constant returns to scale are assumed. For both measures, residuals are used, purging the productivity numbers from year effects and 4-digit industry effects. The dashed lines are unweighted averages, the solid lines weigh plants or firms by their aggregate input ($L^{\alpha_L} K^{\alpha_K} M^{\alpha_M}$) share.

Table A.1: Fraction of observations matched to previous year observations

	Matched by firm ID	Matched by other Info	Total
1999	80.6	4.1	84.7
2000	81.9	2.4	84.3
2001	73.8	4.6	78.3
2002	84.5	2.9	87.3
2003	82.5	4.0	86.6
2004	73.5	8.0	81.5
2005	82.4	1.7	84.1
2006	90.1	1.5	91.6

Table A.2 Comparison of capital stock estimates

Number of firms		fa_net			fa_net with official deflator			fa_original		
	sum	K/L	K/firm	sum	K/L	K/firm	sum	K/L	K/firm	
1998	130,819	2.82	66.3	21,548	2.84	66.8	21,704	4.16	89.4	31,770
1999	138,298	3.24	63.3	23,436	3.28	64.0	23,696	4.84	87.1	34,962
2000	133,402	3.24	71.9	24,294	3.24	71.9	24,294	4.87	100.0	36,489
2001	147,685	3.49	69.5	23,661	3.48	69.2	23,569	5.38	98.5	36,413
2002	159,506	3.75	70.9	23,592	3.73	70.4	23,449	5.87	102.4	36,880
2003	174,790	4.14	72.8	23,672	4.02	70.8	23,025	6.57	105.8	37,566
2004	249,881	4.96	75.8	19,833	4.57	69.9	18,271	7.79	107.3	31,166
2005	246,405	5.67	81.2	23,025	5.15	73.7	20,883	9.00	116.3	36,528
2006	273,746	6.66	87.7	24,319	5.70	75.0	20,808	10.50	125.6	38,381

Number of firms		Our benchmark			Our calculations with official investment deflator			Our calculations treating all years as 1998 benchmark		
	sum	K/L	K/firm	sum	K/L	K/firm	sum	K/L	K/firm	
1998	130,819	3.72	79.4	28,445	3.64	78.3	27,791	3.64	78.3	27,791
1999	138,298	4.20	77.8	30,382	4.13	76.8	29,837	4.00	72.0	28,955
2000	133,402	4.16	86.8	31,169	4.10	86.0	30,722	3.83	78.4	28,722
2001	147,685	4.43	82.5	30,011	4.38	81.7	29,625	4.03	73.5	27,271
2002	159,506	4.60	83.6	28,929	4.55	82.8	28,592	4.20	73.4	26,373
2003	174,790	5.00	84.9	28,629	4.95	84.2	28,335	4.53	73.8	25,909
2004	249,881	5.88	86.0	23,515	5.82	85.4	23,304	5.20	73.9	20,804
2005	246,405	6.56	91.6	26,626	6.51	90.9	26,407	5.82	77.8	23,604
2006	273,746	7.41	96.3	27,078	7.35	95.6	26,862	6.53	81.1	23,862

Notes: The sum of capital is denominated in trillion RMB and capital per worker or per firm in thousands of RMB.

Firms with employment below 8 workers and with negative capital stocks or employment have been excluded.

"fa_original" represents fixed assets at original purchase prices. "fa_net" is "fa_original" less depreciation.

Table A.3: Potential bias in capital stock estimate

	State-owned firms	Other firms	Relative bias
1998	0.308	0.269	1.14
1999	0.311	0.294	1.06
2000	0.308	0.296	1.04
2001	0.304	0.268	1.13
2002	0.293	0.263	1.11
2003	0.313	0.263	1.19
2004	0.297	0.249	1.19
2005	0.328	0.271	1.21
2006	0.361	0.301	1.20

Note: The table reports the average difference in $\ln K$ in our benchmark estimate and $\ln K$ in "fa_net with official deflator".

Table A.4: Transitions out of each ownership category

	1998-99		2002-03		2005-06	
	switch-out	switch-in	switch-out	switch-in	switch-out	switch-in
State	4.7%	3.8%	13.2%	11.4%	21.0%	18.2%
Hybrid/collective	8.6%	4.7%	19.1%	6.9%	21.7%	12.0%
Private	11.0%	25.8%	6.4%	14.0%	4.7%	7.1%
HMT/foreign	2.8%	2.9%	2.0%	2.5%	2.2%	2.4%
Total	6.4%	6.4%	9.7%	9.7%	7.8%	7.8%

Note: Switching-out percentages are relative to the number of firms in that category in the previous year; switching-in percentages are relative to total firms in the category at the end.

Table A.5: Transitions between ownership categories by match type

2005-06 matches based on firm ID (230,027 firms)

		Ownership type in 2006				Total	
		State	Hybrid/ collective	Private	HMT/ foreign		
ownership	State	79.5	2.8	16.9	0.8	29,105	(12.9%)
type	Hybrid/collective	4.7	79.0	16.1	0.2	20,262	(9.0%)
in	Private	3.0	1.0	95.5	0.5	125,604	(55.5%)
2005	HMT/ foreign	0.6	0.1	1.3	98.0	51,304	(22.7%)

2005-06 matches based on other information (3,752 firms)

		Ownership type in 2006				Total	
		State	Hybrid/ collective	Private	HMT/ foreign		
ownership	State	52.7	6.2	33.5	7.6	565	(15.1%)
type	Hybrid/collective	5.5	42.5	47.5	4.5	381	(10.2%)
in	Private	4.8	1.6	84.0	9.6	2,356	(62.8%)
2005	HMT/ foreign	4.9	1.1	18.4	75.6	450	(12.0%)

Note: Compared to the ownership categories in Table 3, the firms owned by Hong Kong, Macau, and Taiwan investors are grouped with foreign firms.