8-30-2017 10:00 AM

Essays on Growth and Input Misallocation in China

Wenya Wang, *The University of Western Ontario*

Supervisor: Dr. James MacGee, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Economics

Abstract

My thesis consists of three chapters that contribute to the study of input misallocation and TFP growth in China.

In Chapter 2, I compare the misallocation of intermediate goods to those of capital and labor, which have been extensively studied in the literature. To measure misallocation, I compute the dispersion of marginal products of intermediate goods across firms, and the potential output gains by eliminating this dispersion in China Industrial Enterprise Survey (CIES) data. Although the within-industry dispersion of marginal products of intermediates is smaller than that of capital and labor, gross output and value added gains from reallocating intermediate goods are 6 and 14 times those from capital and labor reallocations. If intermediate goods, capital and labor are reallocated to equalize their marginal products, the total value added gain in the CIES is 550%, much greater than the 98% obtained under the value added approach in the literature (i.e. Hsieh and Klenow, 2009). This suggests that with its 74% revenue share and input complementarity, distortions and frictions through intermediate goods could be a promising channel to account for sizable misallocation in China's data. I further find suggestive evidence of pre-order friction: intermediate goods need half a year to pre-order, which gives rise to borrowing constraints in paying for intermediates. Similar to capital, marginal products of intermediates are found to be more dispersed among potentially constrained firms with low net worth, as one would expect if borrowing constraints exist.

Motivated by the findings in Chapter 2, Chapter 3 quantifies the novel role of pre-order friction and borrowing constraints on intermediate goods in accounting for misallocation in the CIES data (Hsieh and Klenow, 2009; Brandt, Van Biesebroeck, and Zhang, 2012). With a gross output production function, I incorporate intermediate goods frictions into the firm investment model of Cooper and Haltiwanger (2006). Firms order and prepay for a fraction of intermediate goods one period in advance (pre-order), and face one borrowing constraint on capital and intermediate goods. Firms also face capital adjustment costs. I measure misallocation by the potential gross output gain as a percentage of actual gross output, if intermediate goods, capital and labor are hypothetically reallocated to equalize marginal products across firms. Over 1998-2007, gross output misallocation in the CIES data averages 140 percent. The model accounts for around 70 percent of this misallocation, when calibrated to key moments in firm-level debt, productivity and market share distribution in the CIES data. Half of the misallocation in

the model is attributed to intermediate goods frictions: 34 percent from borrowing constraints, and 16 percent from pre-order. While borrowing constraints on capital induce small misallocation, capital adjustment costs account for the other half. Larger misallocation with intermediate goods frictions than without arises from its large gross output revenue share and recurrent need of financing. This tightens the borrowing constraint and interrupts the self-financing mechanism for capital accumulation. Further, as in Chapter 2 for the data, I find that value added approach in literature also underestimates misallocation by ignoring misallocation of intermediate goods for the model. The importance of intermediate goods frictions in misallocation could be applicable to other countries with an underdeveloped financial system.

Chapter 4 decomposes China's fast aggregate manufacturing productivity growth into firm-level technological growth, intensive reallocation of inputs across existing firms and extensive reallocations through net entry. Following Baily, Hulten, and Campbell (1992)'s approach of aggregate productivity decomposition, I find that extensive reallocation accounts for 93% and 144% of 5-year aggregate productivity growth in 1998-2003 and 2002-2007 in the CIES. In contrast, intensive reallocation contributes -10% and -93% to the growth. These estimates are however biased by a left-censoring problem, since CIES does not survey non state owned firms with sales less than 5 million yuan. I propose a methodology accordingly to recover the three sources of growth in China's manufacturing sector. I find that when China's data is taken as the manufacturing universe, the role of extensive reallocation in aggregate productivity growth is overstated by a quarter to two thirds during 1998-2003 and 2002-2007. Most of the overstated magnitude in extensive reallocation is picked up by the intensive reallocation among existing firms. Compared to U.S., intensive reallocation in China's manufacturing sector is still smaller, but larger than that directly implied in the CIES. This also indicates that the left-censoring problem in other countries should be taken into account when analyzing micro-level sources of aggregate productivity growth across countries.

Keywords: Intermediate Goods, Gross Output Misallocation, Value Added Misallocation, Pre-Order, Borrowing Constraints, Left-Censoring, TFP Growth, Technological Growth, Intensive Reallocation, Extensive Reallocation

Acknowlegements

I would like to first express my sincere gratitude to my dissertation committee members. Over the years, I have enjoyed numerous Friday-afternoon conversations with my advisor James MacGee in discussing my research. These become such a memorable part of my study because of not only his intellectual influence, but also his constant encouragement and pushing me to be a better researcher. He sets an excellent role model as a researcher, mentor and advisor. I deeply thank Simona Cociuba and David Rivers for their invaluable advise and generous help in the job search process. Other than my thesis committee, I would also like to extend my thanks to Igor Livshits for his incisive feedback and help in my academic development, Rui Castro and Terry Sicular for their inspiring discussion in my research, and Guoxiong Zhang for his kind help in access to data.

I am grateful to all the support and services that I constantly ask for from our economics staff, Yvonne Adams, Sandra Augustine, Karin Feulgen, Jennifer Hope, Leslie Kostal, Debra Merrifield, Maureen O'Connell, and Sharon Phillips. I also enjoy the companion and inspiration from my fellows. While this list is too long, I would like to extend particular thanks to Yuxi Yao, Sha Wang, Jin Zhou, Youngming Park, Hiro Mori, Yifan Gong, Youjin Choi, Hyeongsuk Jin, George Orlov, Peter and Zinaida Foltin.

My last but deepest thanks are to my family. Throughout my life, my parents have dedicated to provide me and my younger sister the best education and environment to grow. Their dedication and unconditional love inspire and motivate me from time to time. My PhD endeavor would not be possible without the love and support from my husband.

Dedicated to My Parents

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Introduction

China's manufacturing sector has undergone several major institutional changes during 1998-2007. These changes include privatization of some state owned firms, trade liberalization with accession to WTO and financial market reforms (Xu, 2011; Zhu, 2012). There is a growing consensus that these reforms have facilitated reallocation of inputs and increased aggregate productivity (e.g. Yu, 2010; Song, Storesletten, and Zilibotti, 2011; Tombe and Zhu, 2015). However, there is also evidence that substantial misallocation, defined as the gap between actual and potential output when marginal products of inputs are equalized, still exists and is greater than that in the more efficient benchmark of U.S.. (Hsieh and Klenow, 2009; Brandt, Van Biesebroeck, and Zhang, 2012). Two natural questions arise given these facts. What frictions or distortions could quantitatively account for the still existing misallocation in China after these reforms? How are these reforms reflected in the reallocation across firms during this time period?

To answer the first question, Chapter 2 of this thesis studies the misallocation of the third input – intermediate goods and its quantitative magnitude in China Industrial Enterprise Survey (CIES) data. Because of its 74% gross output revenue share, intermediate goods frictions are promising to account for misallocation. Chapter 3 thus builds on the suggestive evidence of pre-order friction and borrowing constraints on intermediate goods found in Chapter 2, and quantifies their roles in accounting for the substantial misallocation in China.

For the second question, earlier studies either focus on a particular reform,[1] or use the left-censored CIES data to see the total effect of reforms (Van Biesebroeck, 2008; Brandt, Van Biesebroeck, and Zhang, 2012; Ding, Guariglia, and Harris, 2016). Since 80% of manufacturing firms with sales less than 5 million yuan are

[1]For example, there is trade liberalization in Yu (2010), privatization of state owned firms in Song, Storesletten, and Zilibotti (2011).

not included in the CIES, Chapter 4 corrects the left-censoring and recovers the true reallocations in China's manufacturing sector.

The case of China is interesting for two reasons. First, persistent misallocation exists despite of policy reforms in many developing countries (Moll, 2014; Buera and Shin, 2013). Using Chinese firm-level data to explore intermediate goods misallocation and reallocation in its manufacturing sector could help to understand misallocation in other countries, e.g. India and Mexico. Second, as the second largest economy since 2010, China's current and potential growth from improving allocation efficiency is closely related to productivity growth and welfare in other countries (Di Giovanni, Levchenko, and Zhang, 2014; Hsieh and Ossa, 2016).

My thesis takes a firm as the production unit, which uses capital, labor and intermediate goods to produce gross output under some technology. Perhaps motivated by the macro approach to study cross-country GDP differences, the misallocation literature has focused on value added output, and studied misallocation in capital and labor. The magnitude of misallocation in intermediate goods have not been explored, despite its largest gross output revenue share (over 50%) across countries (Jones, 2011). Chapter 2 of my dissertation documents intermediate goods misallocation in the CIES data. I explore (i) whether there exists misallocation of intermediate goods; (ii) what frictions potentially account for misallocated intermediate goods; (iii) bias on the magnitude of potential value added output gain (i.e. Hsieh and Klenow, 2009) by neglecting intermediate goods misallocation.

I find that although the dispersion of the marginal products of intermediates is smaller than those of capital and labor, the gross output and value added gains from intermediate goods reallocation are about 6 and 14 times those from capital and labor, respectively. I also find suggestive evidences of pre-order friction and borrowing constraints on intermediates in the CIES data. The time period from ordering intermediates to receiving sales revenue lasts about 6 months. Further, the dispersion of marginal products of intermediates is higher among firms with low net worth. This is consistent with borrowing constraints, as one would expect more dispersion of marginal products among constrained firms.

Misallocation of intermediate goods implies that the value added approach in the literature may underestimate the magnitude of misallocation. Specifically, the value added approach nets intermediates from gross output and computes value added productivity. Capital and labor are then reallocated to firms with highest value added productivities to equalize their marginal products across firms. Under this approach, the total value added gain in the CIES is 98%, much smaller than

550% when all three inputs are reallocated simultaneously to firms with highest gross output productivities. This is because first, the value added approach neglects the extra output gain from reallocating intermediates *via* input complementarity. Secondly, it also biases the misallocation since firm-level value added productivity no longer reflects its gross output productivity.

Motivated by the suggestive evidence of pre-order friction and borrowing constraints for intermediates in Chapter 2, Chapter 3 quantifies how much these two frictions can account for misallocation in China. I incorporate pre-order and borrowing constraints on intermediate goods into the standard firm investment model of Cooper and Haltiwanger (2006). The model also features borrowing constraints on capital (Midrigan and Xu, 2014; Moll, 2014; Bai, Lu, and Tian, 2016 among many others) and capital adjustment costs (Asker, Collard-Wexler, and De Loecker, 2014). In the quantitative analysis, the model is calibrated to match key moments in firm-level debt and productivity, as well as the market share distribution over firm age groups in the CIES data over 1998-2007.

I find that the model generates substantial misallocation and accounts for 69% of measured misallocation in the CIES data. Counterfactual experiments suggest that borrowing constraints on intermediates and on capital, pre-order on intermediates and capital adjustment costs account for 23%, less than 1%, 11% and 35% of misallocation in the CIES data, respectively. In the model, borrowing constraints on intermediate goods generate much larger misallocation than that on capital. This small impact of financial frictions on capital investment is consistent with Midrigan and Xu (2014) and Moll (2014), due to persistent firm-level productivities and firms' ability to save such that the stationary distribution have few constrained firms. With intermediate goods borrowing constraints, capital financing is crowded out directly intra period, and also indirectly through a lower level of net worth in the future. This slows capital accumulation and leaves more firms constrained.

While Chapter 2 and 3 focus on average misallocation during 1998-2007, Chapter 4 studies the dynamic reallocations of inputs across firms in China through the lens of aggregate productivity growth. Following the U.S. aggregate productivity decomposition literature (Baily, Hulten, and Campbell, 1992; Foster, Haltiwanger, and Krizan, 2001), I decompose 5-year growth in gross output weighted firm-level productivity in the CIES into: technological growth, intensive reallocation and extensive reallocation. Specifically, technological growth measures the average firm-level productivity growth. The intensive reallocation measures reallocations among existing firms, while the extensive one measures whether more productive entrants replace less productive exiters over time.

A key innovation of this chapter is to take the left-censoring of the CIES into account when implementing the decomposition. Unlike U.S. census data, the CIES data excludes non state owned firms with sales less than 5 million yuan, and consequently misses 80% of manufacturing firms in China. I develop a method to recover the true technological growth, intensive and extensive reallocation in China's manufacturing sector. I show this possibility under several assumptions of firm-level productivity process, closure behavior and the shape of sales distribution, using the CIES and several aggregate statistics about firms with sales less than 5 million yuan.

Aggregate manufacturing productivity grows 18% in 1998-2003 and 16% in 2002-2007 in the CIES. If one takes the CIES data as the manufacturing universe, the fraction accounted by extensive reallocation is correspondingly 93% and 144%, and that by intensive reallocation is -10% and -93%. However, the left-censoring problem biases the extensive margin upwards and the intensive margin downwards. In the decomposition that corrects for the left-censoring, extensive reallocation accounts for 74% of aggregate productivity growth in 1998-2003 and 86% in 2002-2007, while intensive reallocation accounts for 24% and -40%, respectively. Two sources account for the bias. First, manufacturing firms with sales less than 5 million cutoff are neglected in the former exercise. These firms increase aggregate productivity through intensive reallocation and technological growth. Second, incumbent firms cross the 5 million cutoff are mis-classified as entrants and exiters in the CIES, which consequently mis-attributes some of the intensive reallocation in the manufacturing sector to extensive reallocation in the former exercise. Overall, intensive reallocation in China's manufacturing sector is still less than that in the U.S. manufacturing sector, but larger than what is directly implied in the CIES data. The extensive reallocation during 1998-2007 remains to be China's main driving force of aggregate productivity growth after correcting the left-censoring problem.

My thesis offer several policy implications. First, despite its fast growth, China's financial underdevelopment may distort allocation of intermediate goods as well as capital investment across firms, and therefore causes aggregate output loss. When this intermediate goods channel is taken into account, there could be more efficiency enhancing from the ongoing financial market reforms in China, e.g. a better enforcement of bankruptcy law, construction of personal credit system and interest rate liberalization. Second, institutional reforms during 1998-2007 decrease the magnitude of gross output misallocation, consistent with earlier studies about their positive role in improving allocation. In particular, Chapter 3 suggests that the potential gross output gain decreases from 1.68 in 1998 to 1.24 in 2007. Chapter 4 finds that privatization of state owned firms increases aggregate productivity by 0.27% over 5 years. Third, according to the cross-country analysis in Asturias, Hur, Kehoe, and Ruhl (2017), the net entry margin from decreased entry costs and trade liberalization in early 2000s is likely to be exhausted when China transits into a lower growth regime. Therefore, reforms facilitating reallocation of inputs among existing firms may provide a more sustainable path for productivity growth in China. Example reforms of this type potentially include financial market reforms and lowering domestic trade costs (Young, 2000; Tombe and Zhu, 2015; Yang, 2016).

Bibliography

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic Inputs and Resource (Mis)Allocation." *Journal of Political Economy* 122(5):1013–1063.

Asturias, Jose, Sewon Hur, Timothy J. Kehoe, and Kim J. Ruhl. 2017. "Firm Entry and Exit and Aggregate Growth." *Research Department Staff Report, Federal Reserve Bank of Minneapolis* .

Bai, Yan, Dan Lu, and Xu Tian. 2016. "Do Financial Frictions Explain Chinese Firms' Saving and Misallocation?" *2016 Meeting Papers. No. 998. Society for Economic Dynamics, 2016* .

Baily, Martin Neil, Charles Hulten, and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity. Microeconomics* 1992:187–267.

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

Buera, Francisco and Yongseko Shin. 2013. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy* 121(2):221–272.

Cooper, Russell W. and John C. Haltiwanger. 2006. "On the Nature of Captial Adjustment Costs." *Review of Economic Studies* 73:611–633.

Di Giovanni, Julian, Andrei A Levchenko, and Jing Zhang. 2014. "The Global Welfare Impact of China: Trade Integration and Technological Change." *American Economic Journal: Macroeconomics* 6 (3):153–183.

Ding, Sai, Alessandra Guariglia, and Richard Harris. 2016. "The Determinants of Productivity in Chinese Large and Medium-Sized Industrial firms, 1998-2007." *Journal of Productivity Analysis* 45 (2):131–155.

Foster, Lucia, John C. Haltiwanger, and Cornell John Krizan. 2001. "Aggregate Productivity Growth: Lessons from Microeconomic Evidence." *New Developments in Productivity Analysis* :203–372.

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

Hsieh, Chang-Tai and Ralph Ossa. 2016. "A Global View of Productivity Growth in China." *Journal of International Economics* 102:209–224.

Jones, Charles I. 2011. "Intermediate Goods and Weak Links in the Theory of Economic Development." *American Economic Journal: Macroeconomics* 3(2):1–28.

Midrigan, Virgiliu and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review* 104(2):422–458.

Moll, Benjamin. 2014. "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation." *American Economic Review* 104(10):3186–3221.

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

Tombe, Trevor and Xiaodong Zhu. 2015. "Trade, Migration and Productivity: A Quantitative Analysis of China." *Manuscript, University of Toronto* .

Van Biesebroeck, Jo. 2008. "Aggregating and Decomposing Productivity." *Review of Business and Economics* 53 (2):122–146.

Xu, Chenggang. 2011. "The Fundamental Institutions of China's Reforms and Development." *Journal of Economic Literature* 49(4):1076–1151.

Yang, Yang. 2016. "Transport Infrastructure, City Productivity Growth and Sectoral Reallocation: Evidence from China." *Job Market Paper,University of California Los Angeles* .

Young, Alwyn. 2000. "The Razor's edge: Distortions and Incremental Reform in the People's Republic of China." *The Quarterly Journal of Economics* 115 (4):1091–1135.

Yu, Miaojie. 2010. "Processing Trade, Firm Productivity, and Tariff Reductions: Evidence from Chinese Products." *Economic Journal* 125(585):943–988.

Zhu, Xiaodong. 2012. "Understanding China's Growth: Past, Present, and Future." *Journal of Economic Perspectives* 26 (4):103–24.

Chapter 2

Intermediate Goods Misallocation in Chinese Firm-level Data

2.1 Introduction

Substantial measured misallocation of inputs across firms has been documented using the China Industrial Enterprise Survey (CIES) data (e.g. Hsieh and Klenow, 2009; Brandt, Van Biesebroeck, and Zhang, 2012). A number of studies follow the seminal paper Hsieh and Klenow (2009) and quantify the magnitude of output (value added) gain if capital and labor are reallocated to firms with highest value added productivities. This value added approach implicitly assumes that intermediate goods are efficiently allocated across firms, which is a potentially strong assumption. Since intermediate goods input has the largest gross output revenue share, there could be a large bias in measured misallocation when intermediate goods are misallocated.

This chapter quantifies the magnitude of misallocation of intermediate goods in the CIES data. Following Hsieh and Klenow (2009), I define misallocation as the gap between actual output and potential output when the marginal products of inputs are hypothetically equalized across firms as in a static firm model. Unlike Hsieh and Klenow (2009), I use gross output (rather than value added) as the output measure so as to quantify the gross output misallocation. I employ two measures to quantify intermediate goods misallocation in the CIES 1998-2007. First, I document the dispersion of marginal products of intermediates across firms within 2-digit China Industrial Classification code industries (3-digit NAICS equivalent). Second, I compute the potential gross output and value added gains by equalizing marginal products of intermediates across firms within 2-digit industries, *holding capital and labor at the firm-level as observed*.

These measures are compared to those for capital and labor, in order to investigate the relative importance of intermediates in accounting for misallocation. I find that although marginal products of intermediates are less dispersed, the gross output and value added gains from intermediate goods reallocation are about 6 and 14 times of those from capital and labor reallocations. This sizable output gain from intermediate goods reallocation is consistent with its highest gross output revenue share. Thus, frictions or distortions on intermediates, if there is any, could be a promising channel to account for the substantial misallocation in China's data.

The measured misallocation of intermediates across firms results in a bias in the standard value added approach in quantifying misallocation. Specifically, the value added approach in the literature nets intermediates from gross output and computes value added productivity under a value added production function. The magnitude of misallocation is to quantify the total value added gain when capital and labor are reallocated to firms with the highest *value added productivities*. This approach underestimates misallocation, since it first neglects the potential output gain from reallocating intermediates. Unless intermediate goods are perfectly substitutable or additively separable with value added bundle, input complementarity implies that output increases due to not only reallocation of intermediates, but also the reallocation of capital and labor *along with* intermediates. Secondly, it may also bias the misallocation since the value added productivity no longer reflects firm-level gross output productivity.

To quantify this bias, I compute total value added gain in the CIES by two approaches of reallocation. One is the value added approach in the literature. The second is the gross output approach, i.e. I quantify the gross output and value added gains by reallocating capital, labor and intermediates *simultaneously* to firms with the highest *gross output productivities*. Results suggest that the first approach gives 98% total value added gain in the CIES, while the second approach gives 550%. In other words, the value added approach underestimates misallocation substantially when intermediate goods are misallocated.

I further explore what intermediate goods frictions can potentially account for the intermediate goods misallocation. I find suggestive evidences of two frictions, pre-order friction and borrowing constraints on intermediates financing in the CIES data. First, unlike what is implicitly assumed in the value added approach, intermediate goods choice is not static. In practice, it takes time for firms to pre-order intermediates, organize production, sell goods and receive sales revenue. The entire process lasts about 6 months in the CIES. This also creates borrowing needs for working capital if firms have insufficient internal funds.

Therefore, I investigate whether firms with low net worth are constrained in financing intermediates. Using capital stock as the proxy of net worth, I find that the dispersion of marginal products of intermediates is 50% higher among firms in the bottom quartile of capital distribution, compared to those in the top quartile. A similar observation holds for the case of capital, which is consistent with a story that financial frictions not only hamper capital investment, but also distort intermediate goods choice at the firm-level.

This chapter is related to a growing literature on capital and labor misallocation. Hsieh and Klenow (2009) documented how reallocations of capital and labor across firms can increase total value added in Chinese and Indian firm-level data. For labor misallocation, Hopenhayn and Rogerson (1993), Boedo and Mukoyama (2012), Da-Rocha, Tavares, and Restuccia (2015) and Mukoyama and Osotimehin (2016) studied how labor firing costs hinder labor reallocation and lower the level and growth of aggregate productivity. The strand of capital misallocation is more broad and connects to the literature on financial development and cross-country GDP differences. The most popular explanations for capital misallocation are adjustment costs and borrowing constraints. These studies include Bartlesman, Haltiwanger, and Scarpetta (2013), Asker, Collard-Wexler, and De Loecker (2014), Moll (2014), Buera and Shin (2013), Midrigan and Xu (2014), Bai, Lu, and Tian (2016) etc. Instead of capital and labor misallocation, this chapter investigates misallocation of the third input, intermediate goods. I find that intermediate goods are the most important input to account for sizable misallocation in China's data, consistent with its highest gross output revenue share

This chapter is also related to the work on misallocation with intermediate goods. Dias, Marques, and Richmond (2016) and Bils, Klenow, and Ruane (2017) also used gross output production function and computed gross output misallocation by reallocating intermediate goods along with capital and labor in Portugal and Indian firm-level data. While these studies focus on the impact of gross output misallocation on TFP and its measurement problem, this chapter focuses solely on the empirical fact on intermediate goods misallocation and its potential explanations. The argument about the bias of using value added approach is also novel in the literature.

The rest of the chapter is organized as follows. Section 2 discusses the data source. Section 3 measures intermediate goods misallocation, compared to capital and labor misallocation. Section 4 quantifies the bias of value added approach in quantifying misallocation. Section 5 investigates two suggestive intermediate goods frictions and Section 6 concludes.

2.2 Data

My three chapters are all based on the China Industrial Enterprise Survey (CIES) from 1998 to 2007, which has been extensively used in the literature (see, e.g. Hsieh and Klenow, 2009; Brandt, Van Biesebrock, and Zhang, 2014; Bai, Lu, and Tian, 2016). For this reason, this section also serves as a data introduction for the following two chapters.

The CIES covers all state owned manufacturing firms, and non state owned manufacturing firms with sales more than 5 million yuan every year before 2007 (approximately 800 hundred thousand U.S. dollars). The observation of production units in the CIES data are firms, not establishments. In 2004, 95% of firms have only one establishment. According to National Bureau of Statistics, starting from 2007, the 5 million threshold sales is required for both state owned and non state owned firms to be included in the CIES data. In 2011, the threshold sales increases to 20 million yuan.

The dataset combines firm-level information on balance sheets, income and cash flow statements. Variables of interest in this thesis include gross output (sales), book value of capital, employment, wage bill, intermediate goods cost, birth years, inventory, account receivables, ownership and industries (see Table 2.1 for variable definitions). Industries in this dataset are classified according to the 4-digit China Industrial Classification code (CIC). In all, there are 544 4-digit CIC industries, and 29 2-digit CIC (3-digit NAICS equivalent) industries in the manufacturing sector.[1]

To track firms over time, I construct an unbalanced panel from 1998 to 2007 using firm information. Specifically, firm i at t and firm j at $t+1$ are matched sequentially according to registration I.D., name of the firm, name of legal representatives, phone number, industry and main products, following the method of Brandt, Van Biesebroeck, and Zhang (2012). Since ownership is not used for matching, a privatized SOE would be matched with the same firm in later years if any of the matching criteria is unchanged. For merges, unfortunately, this dataset

[1]Example 4-digit industries are manufacture of candy and confectionery, manufacture of leather shoes, manufacture of biochemistry pesticide and manufacture of metal forming machine to name a few. See Table A.1 in appendix for a full list of 2-digit CIC industries.

Table 2.1: Variable Definitions in CIES Data

Variable	Definition
Registration ID	Unique 9-digit identifying number upon a firm's registration
Ownership (SOE)	Firms with Registration Classification 110, 120, 141, 143, 145, 151 (see Holz, 2013)
Ownership (NonSOE)	Firms with Registration Classification 130, 142, 149, 159, 160, 171, 172, 173, 174, 190, 219, 220, 230, 240, 310, 320, 330, 340 (see Holz, 2013)
Birth year	First year that a firm operates. This could be earlier than the registration year. If an SOE firm switches to non state owned, the birth year is that of the old firm
Industry	4-digit China Industrial Classification code GB/T 4754-1994 before 2003, GB/T 4754-2002 between 2003 and 2011, and GB/T 4754-2011 since 2012
Gross Output	Market value of finished and semi-finished products that are produced in the calendar year
Book Value of Capital	Purchased value of fixed asset net accumulated depreciation
Employment	Average employment over the calender year; includes both full-time and part-time employment
Total Wage Payable	Total wage bill for all workers, including full-time and part-time. Benefits are not included
Intermediate Inputs	Value of raw materials, energy and semi-finished products purchased outside of a firm *and* used in production in the calendar year
Inventory	Value of raw materials and semi-finished products to use in the next production stage, and finished products to sell
Account Receivables	Value of sales (not restricted to sales in the calendar year) to be collected from buyers.

Sources: National Bureau of Statistics at Shanghai, http://www.stats-sh.gov.cn/, and Beijing http://www.bjstats.gov.cn/

cannot distinguish it from entry and exit.[2]

With this panel data, the real capital stock of firm i at time t, k_{it}, is constructed

[2]If firm i and j merge at time $t+1$ and the new company uses any of the above matching criteria as firm i at $t+1$, firm j would be viewed as an exit. If the new firm at $t+1$ cannot be matched with i or j in any criteria, firms i and j are viewed as exits and this new firm is viewed as an entry.

using a perpetual inventory method

$$k_{it} = (1-\delta)k_{it-1} + \frac{\Delta\text{Book Value Capital}_{it}}{\text{Capital Goods Deflator (Base Year=1998)}} * 100 \tag{2.1}$$

The initial real capital stock is the book value of capital at $t-1$ if the opening year of firm i is $t-1$. For firms established before 1998, I follow Brandt, Van Biesebroeck, and Zhang (2012) and set the firm level real capital stock proportional to total real capital stock of the province it is located, according to the firm's fraction of provincial nominal capital stock. Real capital stock at the province level is easier to compute because of longer time series in National Bureau of Statistics (NBS) yearbooks. In further calculating firm-level productivity, gross output and intermediate goods are deflated using industry-specific deflators. Wages are deflated using CPI from NBS.

Table 2.2 presents summary statistics of the main variables. During 1998-2007, the number of manufacturing firms in the CIES more than doubled, from 147,690 to 304,599 with an annual growth rate of 8.4%. This growth rate is much higher than the 1% annual growth of manufacturing firms in U.S. during 1977 to 1987 (U.S. Business Dynamic Statistics). Further, the composition of ownership also changed dramatically in the CIES. Fraction of non state owned firms increased from 38% in 1998 to 94% in 2007. Similar trends exist for their employment and gross output shares in the CIES data.

Compared to state owned firms, non state owned firms are on average smaller in employment. This size gap get widened during 1998-2007. An average SOE firm is about 23% larger than an average non state owned firm in 1998, and 89% larger in 2007. For gross output, an average state owned firm is 71% of an average non state owned firm in 1998, but increases to 1.89 in 2007. Similar patterns hold for real capital stock and intermediate goods. These time trends are consistent with the privatization of loss making state owned firms since 1998, which selects large ones remain to be state owned. Yet, state owned firms are less productive than private owned firms for each year, although their average labor productivity shows a catching-up trend with non state owned firms during 1998-2007.

Although the CIES data surveys all state owned firms, it only includes non state owned firms with at least 5 million yuan in sales. Compared to the aggregate information in Economic Census 2004 and 2008,[3] firms in the CIES data are roughly

[3]Aggregate Economic Census information for 2004, 2008 and 2013 are available at National Bureau of Statistics website http://www.stats.gov.cn/tjsj/pcsj/. The 2004 census information

Table 2.2: Summary Statistics of CIES Data 1998-2007, 1998 Constant Million Yuan

Year	Number	NonSOE %	NonSOE Emp. %	NonSOE Output %	Employment (Mean)	GrossOutput (Mean)	WageBill (Mean)	Real Capital (Mean)	Intermediate Input (Mean)	Labor Productivity
1998										
-all	147,960	38%	32%	46%	308.47	39.45	2.46	26.39	30.77	0.0276
-SOE	91,560	-	-	-	337.40	34.21	2.39	27.62	26.44	0.0230
-nonSOE	56,400	-	-	-	274.14	47.94	2.57	24.40	37.80	0.0370
1999										
-all	146,321	44%	39%	52%	321.31	43.18	2.56	26.96	33.32	0.0298
-SOE	82,274	-	-	-	349.23	37.06	2.51	29.09	28.32	0.0250
-nonSOE	64,047	-	-	-	304.39	51.04	2.63	24.22	39.75	0.0371
2000										
-all	147,106	52%	47%	60%	309.60	50.62	2.79	27.16	38.13	0.0390
-SOE	70,192	-	-	-	343.53	42.32	2.79	29.02	31.43	0.0317
-nonSOE	76,914	-	-	-	297.02	58.19	2.79	25.48	44.25	0.0469
2001										
-all	154,276	64%	58%	69%	283.05	53.91	2.87	26.16	40.96	0.0440
-SOE	55,658	-	-	-	325.73	47.01	3.21	30.01	35.12	0.0365
-nonSOE	98,618	-	-	-	275.22	57.81	2.67	23.99	44.24	0.0493
2002										
-all	165,583	71%	65%	73%	276.28	59.06	2.98	25.60	45.14	0.0482
-SOE	48,155	-	-	-	334.66	55.67	3.37	31.96	41.84	0.0413
-nonSOE	117,428	-	-	-	268.52	60.46	2.82	23.00	46.50	0.0520
2003										
-all	179,265	79%	73%	78%	269.48	70.24	3.14	25.46	52.45	0.0636
-SOE	37,829	-	-	-	349.39	72.39	4.07	37.03	53.75	0.0534
-nonSOE	141,436	-	-	-	260.08	69.67	2.89	22.37	52.10	0.0676
2004										
-all	250,090	88%	83%	85%	224.07	66.02	2.88	21.04	48.44	0.0777
-SOE	30,490	-	-	-	309.48	77.41	4.33	36.08	56.93	0.0662
-nonSOE	219,600	-	-	-	214.39	64.43	2.67	18.96	47.26	0.0801
2005										
-all	249,891	90%	85%	86%	235.58	80.35	3.33	24.23	56.96	0.0974
-SOE	25,008	-	-	-	351.88	105.76	5.47	47.82	76.72	0.0825
-nonSOE	224,883	-	-	-	227.28	77.53	3.09	21.63	54.76	0.1002
2006										
-all	277,468	92%	88%	88%	227.03	89.73	3.62	24.99	62.63	0.1174
-SOE	21,226	-	-	-	358.41	131.84	6.66	56.43	95.23	0.1021
-nonSOE	256,242	-	-	-	220.14	86.25	3.37	22.41	59.93	0.1195
2007										
-all	304,599	94%	90%	89%	221.25	101.71	4.07	26.14	70.98	0.1370
-SOE	17,215	-	-	-	403.50	183.73	8.78	72.44	134.74	0.1214
-nonSOE	287,384	-	-	-	213.30	96.79	3.79	23.37	67.16	0.1389

Labor productivity is defined as value added divided by employment.

Table 2.3: Aggregate Statistics of Above & Below 5 Million Sale Manufacturing Firms

	Number	Gross Output (billion)	Total Wage (billion)	Employment (10,000)
2004				
Below	1,001,587	1,867.76	196.54	2,413.28
Above	256,999	17,528.35	791.97	5,667.34
% of above size firms	20.42%	90.37%	80.12%	70.13%
2008				
Below	1,356,124	3,318.36	382.68	2,889.91
Above	396,950	44,135.83	2,678.62	7,731.57
% of above size firms	22.64%	93.01%	87.50%	72.79%

Current price. Source: The First and Second Economic Census (2004,2008), National Bureau of Statistics

is also available at China Data Online. There are earlier censuses, such as Industrial Census (1995), The First and Second Production Units Census (1996, 2001), that are either before 1998 when the 5 million cutoff is adopted, or are less aligned with the CIES in measurement.

the top 20% manufacturing firms in sales according to Table 2.3.[4] They hire around 70% of manufacturing workers, pay 80% of manufacturing wage bill and produce more than 90% of manufacturing gross output. This difference between the CIES data and the China's manufacturing sector is important and will be revisited in Chapter 3 and 4.

2.3 Misallocation of Intermediate Goods

This section presents evidence of intermediate goods misallocation using the CIES data. First, intermediate goods cost accounts for roughly 70% of gross output revenue in China, which is higher than the share of capital and labor combined. Second, there is dispersion of marginal products of intermediate goods across firms within 2-digit industries. Third, despite of a lowest dispersion in marginal products, eliminating intermediates misallocation gives the highest total gross output and value added gains in the CIES.

2.3.1 Intermediate Goods Share

Intermediate goods have a high revenue share in production in China. At the aggregate CIES data level, intermediate goods share is defined as the ratio of aggregate intermediate goods to aggregate gross output.

Intermediate goods include goods, energy and services that are purchased outside of firms *and* used in production. While CIES data does not break up costs on each component, materials are about 80% of intermediate input, according to World Bank Data Survey (2012). Table 2.4 shows that the aggregate intermediate goods is 74% of aggregate gross output in the CIES, with a similar share for state owned and non state owned firms. The share of intermediate goods is concentrated within 0.7-0.8 for most industries, while a few industries have a share close to 0.5 or 0.9 (see Figure 2.1). While a slightly lower intermediate goods share 68% is reported in Jones (2011) using input-output information for the entire economy,[5] China resembles South Korea and Japan in early 1970s and exceeds the average intermediate goods share 50% among OECD countries (Jones, 2011).

[4] I abuse the language a little here since state-owned firms with sales below 5 million yuan are also included in the CIES. This impact is, however, minor as state-owned shares are small in 2004.

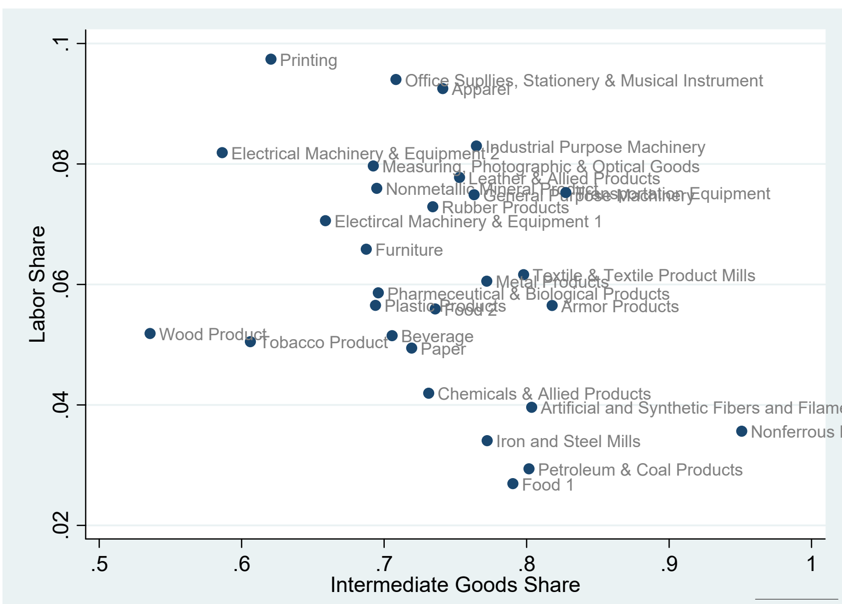
[5] Input-output table for China can be downloaded from World Input-Output Database `http://www.wiod.org/new_site/database/niots.htm`

Table 2.4: Shares of Intermediate Goods in Gross Output, CIES

	All	Top 5 Industries	State-Owned	Non-State Owned
1998	78.01%	79.31%	77.29%	78.84%
1999	77.18%	78.33%	76.43%	77.88%
2000	75.34%	76.09%	74.27%	76.04%
2001	75.96%	76.71%	74.71%	76.53%
2002	76.43%	76.95%	75.15%	76.91%
2003	74.67%	75.51%	74.25%	74.78%
2004	73.38%	74.17%	73.55%	73.35%
2005	70.88%	70.64%	72.54%	70.63%
2006	69.79%	68.27%	72.23%	69.49%
2007	69.80%	68.95%	73.33%	69.39%
1998-2007 Average	74.14%	74.49%	74.38%	74.38%

Note the share of state-owned firms is declining over time. Top 5 industries are the industries with top 5 highest total gross output in a given year.

Figure 2.1: Revenue Share of Intermediate Goods and Labor, 2-digit CIC Industry



Labor share are defined as the ratio of total wage over total gross output in industry s.

2.3.2 Productivity Measure

A key variable for measuring misallocation is firm-level productivity. Productivity for firm i in industry s, z_{is}, is defined as follows:

$$z_{is} = logy_{is} - \alpha_l^s logl_{is} - \alpha_m^s logm_{is} - \alpha_k^s logk_{is} \tag{2.2}$$

under the Cobb-Douglas production function assumption

$$y_{is} = exp(z_{is})k_{is}^{\alpha_k} l_{is}^{\alpha_l} m_{is}^{\alpha_m} \tag{2.3}$$

where y_{is} and m_{is} are real gross output and intermediate inputs. The capital measure k_{is} at firm-level is the real capital stock constructed in Equation (2.1). I use the wage bill to proxy labor inputs l_{is} as in Hsieh and Klenow (2009) to control for the human capital heterogeneity across different types of workers, which is evident in the 2004 CIES data.[6] In particular, 17.32% workers have some college degree in an average CIES firm in 2004. There is a significant dispersion of this share across firms. The 90-10 percentile difference is 43.62% and 5 times of the median fraction 8.31%.

I compute shares of labor α_l^s and intermediate goods α_m^s as the medians of firm-level revenue shares within 2-digit CIC industries. While costs on intermediate goods and labor are observable, costs on capital at firm-level is constructed by $0.85 - \alpha_l^s - \alpha_m^s$. The return to scale parameter $\alpha_l^s + \alpha_m^s + \alpha_k^s = 0.85$ is calibrated from the later numerical exercise in Section 3.3.1 of Chapter 3. Alternatively, I can impute α_k^s by imputing capital rental rate of 13% from World Bank Data Survey (2011). Section A.2 in appendix introduces the imputation method and corresponding magnitudes of misallocation under this alternative α_k^s choice. Results are similar to the magnitude of misallocation presented below.

2.3.3 Measuring Misallocation of Intermediate Goods

This subsection outlines some preliminary evidence of misallocation of intermediate goods in the CIES data. Following Hsieh and Klenow (2009), misallocation in intermediate goods is large, if its marginal products are dispersed such that reallocating the intermediate goods across firms significantly increases aggregate gross output.

The marginal products for any input x at firm i in industry s is given as:

$$MP_{ix} = \alpha_x^s \frac{y_i}{x_i}, \; x \in \{k, l, m\} \tag{2.4}$$

while the share of inputs α_k^s, α_l^s and α_m^s are the same as in Equation (2.2).

There are significant dispersions in the marginal products of all three inputs

[6] 2004 data has more information on education and qualifications of workers, which are absent in regular years.

in data. Figure 2.2 plots histograms of MP_{im}, MP_{ik} and MP_{il} across firms in the pooled 1998-2007 data, trimmed by the top and bottom 1% of marginal products for each input. Compared to capital and labor, marginal products of intermediate goods show less dispersion. Similar patterns of distribution exist among non state owned firms, which suggests that factors other than ownership drive dispersions in marginal products.

Since industries differ in the types of intermediate goods and capital they use, I examine misallocation across firms within 2-digit CIC industries. Within each industry, I first compute the coefficient of variation (CV) of marginal products for each year. These industry-level CVs are averaged using industry level output weights for a CV measure at an "average" industry level. These measures are then averaged over years.[7] Table 2.5 shows that the CV of MP_{im} for an "average" industry is 0.76, and smaller than those of capital, 12.45, and labor, 3.78. This is consistent with Figure 2.2, suggesting that intra industry variations drive patterns of marginal product distributions.

Another measure of misallocation is to compute gains from the reallocation of inputs. To see which input is the most important in driving misallocation, I assess gains by reallocating one input to equalize its marginal products across firms within 2-digit industries, *holding the other two inputs at the firm-level fixed*. The gross output gain for industry s is computed as:

$$\frac{\sum_{i\in s} exp(z_i)(k_i)^{\alpha_k^s}(l_i)^{\alpha_l^s}(m_i^1)^{\alpha_m^s} - \sum_{i\in s} y_i}{\sum_{i\in s} y_i}$$

where m_i^1 is the hypothetical intermediate goods input for firm i such that $MP_{im}^1 = \alpha_m^s \frac{y_s}{m_i^1} = constant$ and $\sum_{i\in s} m_i^1 = \sum_{i\in s} m_i$ for industry s. Consequently, the value added gain for industry s is:

$$\frac{\sum_{i\in s} exp(z_i)(k_i)^{\alpha_k^s}(l_i)^{\alpha_l^s}(m_i^1)^{\alpha_m^s} - \sum_{i\in s} y_i}{(1-\alpha_m^s)\sum_{i\in s} y_i}$$

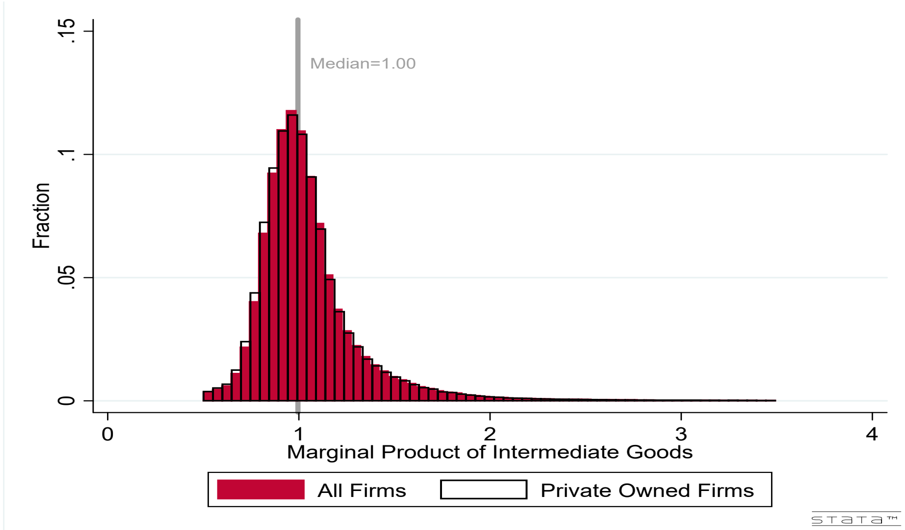
I compute output-weighted average gross output and value added gains each year and then average these gains over years.[8] In Table 2.5, on average, a 2-digit

[7] In this exercise and the following reallocation exercise, I trim the productivity distribution every year to be consistent with the model distribution in Chapter 3. Marginal products are not trimmed here. The reason is that for the case of marginal products of intermediate goods, top 1% firms have a smaller capital stock than all firms, suggesting that these firms could be financially constrained.

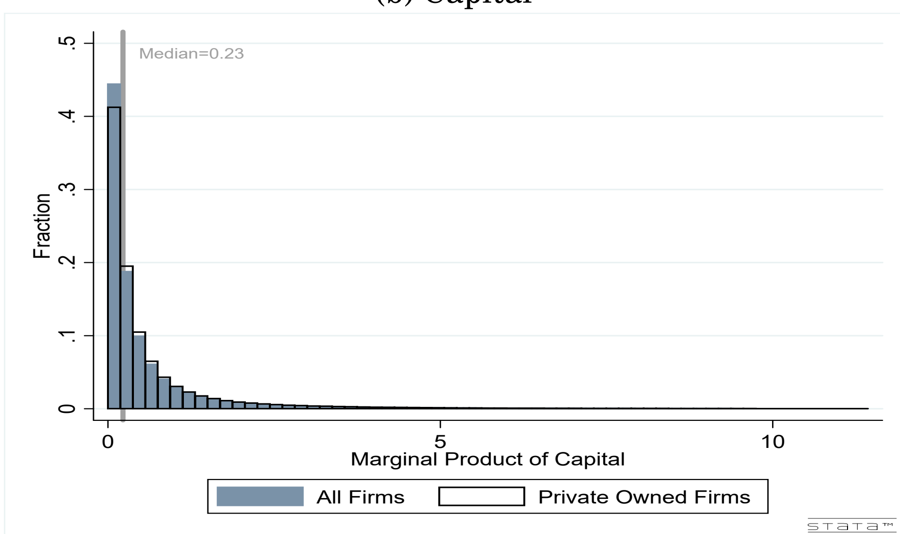
[8] The output weight is gross output for gross output gain, and value added for value added gain. For this reason, output-weighted gains for an average industry equal aggregate gains.

Figure 2.2: Marginal Products Histograms, 1998-2007 Pooled

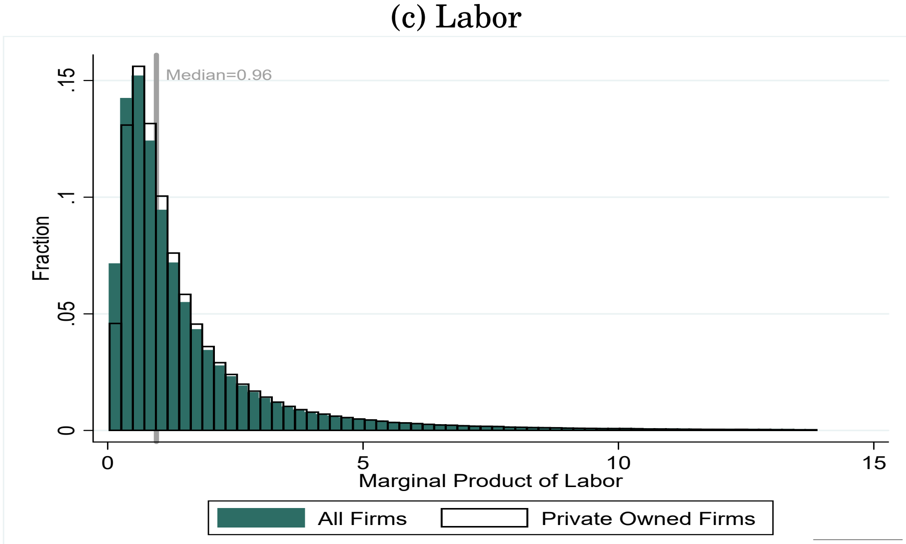
(a) Intermediate Goods



(b) Capital



(c) Labor



CIC industry increases its gross output by 35%, if marginal products of intermediate goods are equalized across firms. This implies an even larger value added gain, 135%, because of the $\frac{1}{1-\alpha_m^s}$ multiplier. These gains are much larger than those for capital and labor. In magnitude, a 135% value added gain from intermediate goods is comparable to the value added gain in Hsieh and Klenow (2009) that reallocate capital and labor after netting out intermediate goods from gross output.

Similar results hold in the private sector, suggesting that misallocation exists intra ownership type, not only across ownership types. Specifically, reallocation of intermediate goods alone across private owned firms within 2-digit industries generates 32% of gross output gain, and 123% of value added gain. These gains are again much larger than those for capital and labor.

Table 2.5: Dispersion in Marginal Products and Output Gains by Reallocating One Input within CIC 2-digit Industries, Output Weighted, 1998-2007 Average

	Intermediate Goods	Capital	Labor
All Firms			
CV	0.76	12.45	3.78
Gross output gain	34.81%	5.66%	2.46%
Value added gain	135.01%	21.83%	9.49%
Private Owned Firms			
CV	0.60	11.33	3.22
Gross output gain	31.55%	6.07%	2.43%
Value added gain	123.08%	23.86%	9.52%

This contrast between the output gain measure and the CV measure highlights the importance of input shares in generating misallocation defined as in Hsieh and Klenow (2009). Although dispersions of marginal products are indicative for misallocation, inputs with low dispersions and high revenue shares may account for more misallocation than inputs with high dispersions and low revenue shares.

Summary Intermediate goods are 74% of gross output revenue. This suggests that if it is distorted and misallocated across firms, the potential gross output and value added gains are large. This section confirms this intuition. Although marginal products of intermediates are less dispersed within 2-digit industries, the gross output and value added gains from reallocating intermediates are 6 and 14 times as large as that from capital and labor reallocation. This misallocation could not be accounted by the existence of state-owned firms, since a similar magnitude of misallocation of intermediates is found across non state owned firms.

2.4 Impact on the Magnitude of Misallocation

If intermediate goods are misallocated as described in Section 2.3, a natural starting point is to compute gross output misallocation, and then calculate total value added gains. This section investigates the implication of intermediates misallocation on the magnitude of misallocation when one uses the value added approach in the literature.

Under this value added approach, reallocation of capital and labor are based on value added productivity z_{is}^{va} for firm i in industry s. This productivity is:

$$z_{is}^{va} = log(y_{is} - m_{is}) - \frac{\alpha_l^s}{1-\alpha_m^s} log l_{is} - \frac{\alpha_k^s}{1-\alpha_m^s} log k_{is} \tag{2.5}$$

where factor shares α_k^s, α_l^s and α_m^s are the same as in Equation (2.2).

Based on z_{is}^{va}, value added gain for industry s when reallocating capital and labor is

$$\frac{\sum_{i\in s} exp(z_{is}^{va})(k_{is}^{va})^{\frac{\alpha_k^s}{1-\alpha_m^s}}(l_{is}^{va})^{\frac{\alpha_l^s}{1-\alpha_m^s}} - \sum_{i\in s}(y_{is} - m_{is})}{\sum_{i\in s}(y_{is} - m_{is})} \tag{2.6}$$

where k_{is}^{va} and l_{is}^{va} are hypothetical capital and labor, such that their value added marginal products, $(\frac{\alpha_k^s}{1-\alpha_m^s})exp(z_{is}^{va})(k_{is}^{va})^{\frac{\alpha_k^s}{1-\alpha_m^s}-1}(l_{is}^{va})^{\frac{\alpha_l^s}{1-\alpha_m^s}}$ and $(\frac{\alpha_l^s}{1-\alpha_m^s})exp(z_{is}^{va})(k_{is}^{va})^{\frac{\alpha_k^s}{1-\alpha_m^s}}(l_{is}^{va})^{\frac{\alpha_l^s}{1-\alpha_m^s}-1}$, are constant across firms within industry s.

However, the above value added approach ignores the potential misallocation of intermediate goods. There are thus two problems in using value added approach to quantify misallocation. First, value added productivity z_{is}^{va} may not reflect firms' true productivities z_{is}. If intermediate goods are efficiently allocated, $z_{is}^{va} = \frac{1}{1-\alpha_m^s} z_{is}$. This relation fails if intermediate goods allocation is distorted. In particular, if distortion is positively correlated with productivity z_{is}, value added productivity z_{is}^{va} from Equation (2.5) would be lower (higher) than the value when distortions are absent for a high (low) productivity z_{is} firm. The opposite happens if this correlation is negative (see Appendix Section A.3). This causes a bias on the magnitude of misallocation in Equation (2.6) as capital and labor reallocated to high productivity firms z_{is} might be too few or too many, depending on how distortion correlates with productivity z_{is}.

A second bias is due to the fact that Equation (2.6) does not capture the gains from input complementarity. This means that if capital-labor are reallocated *along with* intermediates, there would be more output gain than the sum of gains from reallocating capital-labor and from reallocating intermediates. Table 2.6 illustrates this point. During 1998-2007, the average gross output gain is 8.69%, if only capital and labor are reallocated to firms of the highest gross output productivities within 2-digit industries. The sum of gross output gains from this exercise and gains from reallocating intermediate goods *alone* is 43.5%, much smaller than the 138.72% gain when capital and labor are reallocated along with intermediates. In other words, the input complementarity between intermediates and capital-labor generates 94.51% total gross output gain in the CIES data.[9]

Table 2.6: Input Complementarity in Gross Output Misallocation, Industry Gross Output Weighted

Year	1. Reallocate k,l	2. Reallocate m	3. Reallocate k,l,m	Complementary Misallocation =3-(1+2)
1998	8.09%	43.52%	162.75%	111.14%
1999	8.20%	34.91%	146.51%	103.41%
2000	8.00%	39.14%	151.40%	104.27%
2001	8.50%	31.58%	132.79%	92.71%
2002	8.62%	33.05%	134.34%	92.66%
2003	8.92%	27.92%	119.80%	82.96%
2004	9.23%	35.63%	143.44%	98.58%
2005	9.25%	34.49%	130.77%	87.03%
2006	8.98%	32.79%	129.71%	87.94%
2007	9.11%	35.09%	135.71%	91.51%
Average	8.69%	34.81%	138.72%	95.22%

Given the above arguments, I compute the value added gain as in Equation 2.6 in data, in comparison to value added gain when I reallocate capital, labor and intermediate goods across firms within 2-digit industries. In the latter gross

[9]Table 2.6 also suggests some interesting time trends. During 1998-2007, capital and labor misallocation increases over time while intermediate goods misallocation generally declines. The overall gross output misallocation follows the trend of intermediate goods misallocation and also declines.

output approach, gross output gain is

$$\frac{\sum_{i\in s} exp(z_{is})(k_{is}^{go})^{\alpha_k^s}(l_{is}^{go})^{\alpha_l^s}(m_{is}^{go})^{\alpha_m^s} - \sum_{i\in s} y_{is}}{\sum_{i\in s} y_{is}} \tag{2.7}$$

and value added gain is $\frac{1}{1-\alpha_m^s}$ times this gross output gain. k_{is}^{go}, l_{is}^{go} and m_{is}^{go} are hypothetical capital, labor and intermediate goods such that MP_{ik}, MP_{il} and MP_{im} are constant within industry s.

Table 2.7: Output Gains under Gross Output vs Value Added Approaches, Output Weighted, 1998-2007 Average

	Gross output approach	Value added approach in literature
Gross output gain	138.01%	-
Value added gain	550.74%	98.12%

Table 2.7 presents misallocation results under both value added and gross output approaches. Similar to Hsieh and Klenow (2009), China's total value added in this data increases by almost 100%, if capital and labor are reallocated across firms based on value added productivity. This number is yet much smaller than the value added gain under gross output approach. If intermediate goods are reallocated along with capital and labor, gross output gains by 138% and value added gains by 551%.[10] This suggests that ignoring misallocation of intermediate goods may significantly underestimate the true value added misallocation.

2.5 Intermediate Goods Frictions

Given the substantial misallocation of intermediate goods, this section explores two potential frictions that could distort allocation of intermediate goods. In the CIES, firms need time to order (pre-order) intermediate goods to produce and finally sell to buyers. The time period from buying intermediate goods to receiving sales revenue could last 6 months. This suggests that intermediate goods may need to be financed externally, giving rise to borrowing constraints on intermediate goods purchases.

[10] My number in gross output gain is close to the gross output gain in Bils, Klenow, and Ruane (2017) in India's data.

2.5.1 Pre-Order for Intermediate Goods

One potential explanation for the dispersion in marginal products of intermediate goods across firms is real frictions. If firms adjust intermediate goods in response to productivity shocks in an undistorted competitive market, misallocation in intermediate goods should not exist. In practice, production takes time from purchasing intermediate goods to production, and from sales of output to collection of sales.

To examine the production process, I borrow the concept of operating cycle from the trade credit and working capital management literature (e.g. Jose, Lancaster, and Stevens, 1996; Petersen and Rajan, 1997). The operating cycle, OC, for firm i is defined as:

$$OC_i = \frac{\text{Inventory}_i\text{+Account Receivables}_i}{\text{Sales}_i} * 365 \tag{2.8}$$

where Inventory_i, $\text{Account Receivables}_i$ and Sales_i are the corresponding accounting entries in the CIES data. OC measures days between intermediate goods purchases[11] and collection of sales, standardized in 365 days.

The operating cycle can be further decomposed into two parts, from materials to finished products and from sales of finished products to collection of sales. The first part corresponds to a measure, Days in Inventory, DI_i, defined as:

$$DI_i = \frac{\text{Inventory}_i}{\text{Sales}_i} * 365 \tag{2.9}$$

and the second part corresponds to another measure, Days in Receivables, DR_i, defined as:

$$DR_i = \frac{\text{Account Receivables}_i}{\text{Sales}_i} * 365 \tag{2.10}$$

In the CIES data, I calculate the mean and the median OC, DI and DR across firms in each year, and average the means and medians over 1998-2007. Table 2.8 shows that there are an average of 160 days, and a median of 108 days between intermediate goods purchases and collection of sales revenue. Days on inventory are 87, and days on receivables 75, with each accounting for roughly half of the operating cycle. Total operating cycle is longer for state owned firms, and similar regardless of whether firms export or not (see Table 2.8). These evidences suggest

[11]Inventory includes materials, semi-finished products and finished products, with each takes up about a third of inventory in the World Bank Enterprise Survey (2012).

a pre-order real friction,[12] as well as a working capital need on intermediate goods.

Table 2.8: Operating Cycles, Days in Inventory and Days in Receivables, CIES 1998-2007

	All	Ownership			Exporter Status	
		Private	State	Foreign	Exporter	Non-exporter
OC						
Mean	161	134	204	155	166	149
Median	108	97	130	120	108	107
DI						
Mean	86	69	113	81	88	81
Median	47	41	57	54	45	51
DR						
Mean	75	65	91	74	77	68
Median	43	39	49	51	43	43

The mean and median *OC*, *DI* and *DR* are calculated across firms for each group in each year, and then average over 1998-2007.

2.5.2 Financial Frictions on Intermediate Goods

To finance purchases of intermediate goods and capital, firms may need to borrow. If borrowing constraints on intermediate goods are important, one should expect them to be binding for low net worth firms. In turn, this suggests that marginal products of intermediate goods and capital among low net worth firms shall be more dispersed than those among high net worth firms.

Using capital stock as a proxy for net worth, I compare CVs of marginal products of intermediate goods and capital among firms with the top quartile capital stock to those among firms with the bottom quartile of capital stock. Since more than 60% of 4-digit CIC industries have less than 200 firms, CVs are first calculated at the 2-digit CIC industry level for each year. Mean CVs and confidence intervals reported in Table 2.9 are then calculated when pooling CVs at 2-digit CIC industry level over years. Firms in the bottom quartile of capital stock distribution have a 50% higher dispersion in the marginal product of intermediate goods, and 103% higher in that of capital. The differences are statistically significant as the 95% confidence intervals for the top subsample do not contain the average CVs in the

[12] A similar concept, *time-to-ship*, exist in international trade literature. See Leibovici and Waugh (2016), Hummels and Schaur (2013)

Table 2.9: CV of Marginal Products among Top and Bottom Quartile Firms in Capital Stock

	Intermediate Goods		Capital	
	Top	Bottom	Top	Bottom
Mean	0.25	0.37	0.98	1.99
95% C.I.	[0.24, 0.26]	[0.33, 0.41]	[0.95,1.00]	[0.91, 2.07]
Nobs	290		290	

Each observation is industry-year level, with a industry defined at CIC 2-digit level. The top quartile group is defined as top 25% firms in capital stock at this industry-year level, and the bottom group as the bottom 25%.

bottom subsample, and vice versa.

Table 2.9 is consistent with a story that financial frictions not only hamper capital investment, but also distort intermediate goods choice at the firm-level. While the former has been extensively studied in the literature, the latter motivates this thesis to model financial frictions on intermediate goods in discussing misallocation in Chapter 3.

Summary This section suggests two specific distortions in intermediate goods. First, firms order intermediate goods 87 days before production and 160 days before realization of sales. Second, a smaller dispersion of its marginal products among large firms has its consistency with borrowing constraints on intermediate goods.

2.6 Conclusion

While capital and labor misallocation have been extensively studied in the literature, the empirical facts about intermediate goods misallocation have not been investigated. This is important since if intermediate goods are misallocated in the data, the value added approach in quantifying misallocation misses the extra output gains by reallocating intermediate goods. It would also invalidate the value added approach since value added productivity is a biased measure of firm-level productivity contaminated by intermediate goods distortions and frictions.

This chapter documents intermediate goods misallocation in the CIES data. I explore (i) whether there exists misallocation of intermediate goods; (ii) bias on the magnitude of value added output loss (i.e. Hsieh and Klenow, 2009) by neglecting intermediate goods misallocation; (iii) what frictions potentially account for misallocated intermediate goods;

With a 74% gross output revenue share, reallocating intermediate goods alone within 2-digit industries generates sizable gross output and value added gains. These gains are about 6 and 14 times those from reallocating capital and labor, respectively. Misallocation of intermediate goods implies that the value added approach in the literature may underestimate the magnitude of misallocation. Under this approach, the total value added gain in the CIES is 98%, much smaller than 550% when all three inputs are reallocated simultaneously to firms with highest gross output productivities.

I also find suggestive evidences of pre-order friction and borrowing constraints on intermediates in the CIES data. The time period from ordering intermediates to receiving sales revenue lasts about 6 months. This long operating cycle seems to cause borrowing constraints on purchasing intermediate goods. Specifically, the dispersion of marginal products of intermediates is higher among firms with low net worth. A similar observation is found for the case of capital. Therefore, borrowing constraints not only hamper capital investment, but also distort intermediate goods choice at the firm-level.

Bibliography

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic Inputs and Resource (Mis)Allocation." *Journal of Political Economy* 122(5):1013–1063.

Bai, Yan, Dan Lu, and Xu Tian. 2016. "Do Financial Frictions Explain Chinese Firms' Saving and Misallocation?" *2016 Meeting Papers. No. 998. Society for Economic Dynamics, 2016* .

Bartlesman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review* 103(1):305–334.

Bils, Mark, Peter J. Klenow, and Cian Ruane. 2017. "Misallocation or Mismeasurement?" *Working Paper* .

Boedo, Hernan J Moscoso and Toshihiko Mukoyama. 2012. "Evaluating the Effects of Entry Regulations and Firing Costs on International Income Differences." *Journal of Economic Growth* 17 (2):143–170.

Brandt, Loren, Johannes Van Biesebrock, and Zhang. 2014. "Challenges of Working with the Chinese NBS Firm-level data." *China Economic Review* 30:339–352.

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

Buera, Francisco and Yongseko Shin. 2013. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy* 121(2):221–272.

Da-Rocha, Jos´e-Mar´ıa, Marina Mendes Tavares, and Diego Restuccia. 2015. "Firing Costs, Misallocation, and Aggregate Productivity." *Working Paper, University of Toronto* .

Dias, Daniel A., Carlos Robalo Marques, and Christine Richmond. 2016. "Misallocation and Productivity in the Lead Up to the Eurozone Crisis." *Journal of Macroeconomics* 49:46–70.

Holz, Carsten A. 2013. "Chinese Statistics: Classification Systems and Data Source." *Eurasian Geography and Economics* 54 (5-6):532–571.

Hopenhayn, Hugo A. and Richard Rogerson. 1993. “Job Turnover and Policy Evaluations: A General Equilibrium Analysis.” *Journal of Political Economy* 101:915–938.

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

Jones, Charles I. 2011. “Intermediate Goods and Weak Links in the Theory of Economic Development.” *American Economic Journal: Macroeconomics* 3(2):1–28.

Jose, Manuel L., Carol Lancaster, and Jerry L. Stevens. 1996. “Corporate Returns and Cash Conversion Cycles.” *Journal of Economics and Finance* 20(1):33–46.

Midrigan, Virgiliu and Daniel Yi Xu. 2014. “Finance and Misallocation: Evidence from Plant-Level Data.” *American Economic Review* 104(2):422–458.

Moll, Benjamin. 2014. “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation.” *American Economic Review* 104(10):3186–3221.

Mukoyama, Toshihiko and Sophie Osotimehin. 2016. “Barriers to Reallocation and Economic Growth: the Effects of Firing Costs.” *Mimeo, University of Virginia* .

Petersen, Mitchell A. and Raghuram G. Rajan. 1997. “Trade Credit: Theories and Evidence.” *Review of Financial Studies* 10 (3):661–691.

Chapter 3

Intermediate Goods Frictions and Misallocation in China

3.1 Introduction

Substantial measured misallocation of inputs across firms has been documented in China Industrial Enterprise Survey (CIES)(e.g. Hsieh and Klenow, 2009; Brandt, Van Biesebroeck, and Zhang, 2012). According to Hsieh and Klenow (2009), output could be doubled if marginal products of capital and labor were equalized across firms. Several explanations of misallocation in a quantitative model have been proposed, focusing on firm-level distortions on labor[1] and capital, e.g. financial frictions and adjustment costs. However, the substantial magnitude of misallocation in the firm-level data found in Hsieh and Klenow (2009) remains largely unexplained.[2]

The potential impact of intermediate goods frictions on misallocation has not been investigated. Similar to capital, intermediate goods at the firm-level are subject to borrowing constraints in financing and real frictions in adjustment. According to Chapter 2 using the CIES data, intermediate goods need time to order, and are purchased about half a year before receipt of sales from buyers. This also creates a borrowing need for intermediate goods expenditure.

How much can pre-order and borrowing constraints on intermediate goods quan-

[1]See, for example, Hopenhayn and Rogerson (1993), Da-Rocha, Tavares, and Restuccia (2015), Mukoyama and Osotimehin (2016) for labor firing cost and misallocation.

[2]One strand of the literature models the distribution of firm-level productivity as a result of costly experiments or human capital investment (e.g. Da-Rocha, Tavares, and Restuccia, 2016, Gabler and Poschke, 2013 and Castro and Sevcik, 2016). The impact of distortions on misallocation can be amplified through distorting allocations of inputs as well as the productivity improvement process. However, the output loss implied in these studies deviates from the exercise in Hsieh and Klenow (2009) that assumes a fixed productivity distribution for computing misallocation.

titatively account for misallocation in China? To answer this question, I incorporate these two frictions on intermediate goods, as well as borrowing constraints on capital, into a standard firm investment model of Cooper and Haltiwanger (2006). Similar to their model, firms in this chapter maximize the discounted net present value of future dividends, facing an idiosyncratic AR(1) productivity process and capital adjustment costs.

Unlike the standard heterogeneous firm investment model, this paper models intermediate goods frictions as well as borrowing constraints on capital. Specifically, firms order and prepay for a fraction of intermediate goods one period ahead (pre-order). Firms also face fixed and convex adjustment costs when choosing their capital for next period. Payments of intermediate goods and capital investment are financed by retained earnings, and borrowings if necessary. Firm-level borrowing is subject to a constraint that endogenously depends on firms' default risk and net worth. When stochastic productivities are realized at the beginning of a period, firms choose to continue or exit under limited liability. If continue, they choose optimal labor, and intermediate good usage that cannot exceed the pre-ordered level. In other words, the intermediate goods adjustment cost is infinitely large when firms scale it up and zero when they scale it down.

To quantify how much the model can account for measured misallocation in data, I calibrate the model to match key moments in firm-level debt and productivity, as well as the market share distribution over firm age groups in the CIES data over 1998-2007. The calibration takes into account the fact that the CIES has a threshold sales of 5 million yuan, and only includes top 20% manufacturing firms in the sales distribution.[3] The purpose is to reasonably capture the fact that a large fraction of firms take years to accumulate capital before they grow above the threshold sales.

Following Hsieh and Klenow (2009), the measure of misallocation is defined as gross output gain if marginal products of intermediate goods, capital and labor were *all* equalized across firms. I find that the model generates substantial misallocation and accounts for 69% of measured misallocation in the CIES data. Specifically, gross output gain is 96% of the actual output in the simulated data, and averages 140% in the CIES data over 1998-2007. In other words, gross output could be nearly doubled in the model, and more than doubled in the CIES, if

[3] To be more specific, the CIES data has a 5 million yuan sales threshold for private-owned firms to be included. Such a threshold gives approximately the top 20% firms in sales according to the aggregate statistics in 2004. Further details would be discussed in Section 3.3.1.

marginal products of intermediate goods, capital and labor were equalized across firms.

There are four frictions that cause misallocation in the benchmark model: borrowing constraints and pre-order on intermediate goods, and borrowing constraints and adjustment costs on capital. To decompose the contribution of each friction to misallocation, I eliminate frictions one by one from the benchmark model. My first counterfactual experiment removes borrowing constraints on intermediate goods. The resulting potential output gain in this counterfactual specification is 64%. This suggests that borrowing constraints on intermediate goods induce 33% (32/96) of misallocation in the benchmark model, and accounts for 23% (32/140) of misallocation in the CIES data. Second, I further remove pre-order on intermediate goods by allowing firms to choose the static optimal amount of intermediate goods after their productivity shocks. This experiment lowers the misallocation to a 49% potential gross output gain. Therefore, pre-order on intermediate goods accounts for 16% (15/96) of misallocation in the model and 11% (15/140) of that in the CIES data. The two frictions on intermediate goods together generate 49% of misallocation in the benchmark model, and 34% of that in the CIES data.

I find that further eliminating borrowing constraints has a small impact on misallocation. The potential gross output gain in this counterfactual specification is 48%, only 1% lower than that in the second. This implies that while capital adjustment costs account for 50% (48/96) and 35% (48/140) of misallocation in the model and in the CIES data, borrowing constraints on capital induce a small amount of misallocation. This result is consistent with Midrigan and Xu (2014) and Moll (2014), and is a consequence of persistent productivities and firms' ability to save such that top productive firms own a large share of capital regardless of the constraint.

The intuition for why intermediate goods frictions have a large impact on misallocation is due to its high revenue share in production, and the recurrent need to finance because of one-period depreciation. For illustration, set the reference model as the one with adjustment costs and borrowing constraints on capital only in which firms choose the static optimal amount of intermediate goods. At time t, pre-order and down-payment for intermediate goods increase the borrowing need for firms, on top of capital investment. Capital investment is therefore crowded out, which lowers profit and net worth at time $t+1$. Additionally, at time $t+1$, firms cannot buy more intermediate goods intra period after high productivity shocks because of pre-order. The consequent profit is lower. The lower current net worth at $t+1$ caused by the above two effects further lowers capital investment and more importantly, intermediate goods for time $t+2$, and so forth. This low net worth effect is long lasting since financing intermediate goods is recurrent by one-period depreciation and puts much more stress on borrowing constraints. Consequently, capital accumulation is slowed and the stationary distribution features firms with lower capital stocks.[4]

The impact of intermediate goods frictions on measured misallocation is of general interests and not specific to China. Jones (2011) documents that intermediate goods revenue share is around 50% in most countries, while the working capital management literature about other countries (e.g. Jose, Lancaster, and Stevens, 1996) document a similar time length between intermediate goods purchases and collection of sales. Further, financial markets are underdeveloped in most developing countries, for instance, India and Mexico (e.g. Ghate and Kletzer, 2012; Pratap and Urrutia, 2012), and even in modern developed countries in a certain historical stage.[5]

This chapter is related to a growing literature on misallocation.[6] Several papers have studied productivity and misallocation in China. Hsieh and Klenow (2009) first document large firm-level distortions and substantial misallocation in manufacturing firm-level data. Brandt, Van Biesebroeck, and Zhang (2012) further document limited input reallocation across firms in China despite a high TFP growth over 1998-2007. Several explanations have been proposed in the literature: capital misallocation caused by preferred lending to state-owned firms (Brandt, Tombe, and Zhu, 2013), trade and migration costs (Tombe and Zhu, 2015), entry costs (Brandt, Kambourov, and Storesletten, 2016), and financial frictions (Bai, Lu, and Tian, 2016). Unlike these studies, this paper focuses on the novel intermediate goods frictions and provides a quantitative model to assess their roles in accounting for misallocation.

This chapter is also related to work on capital misallocation across firms. Bartlesman, Haltiwanger, and Scarpetta (2013) argue a dynamic capital investment model with other firm-level distortions important in accounting for cross-country differ-

[4]The crowding-out effect of working capital on capital investment is consistent with Fazzari and Petersen (1993). They find that working capital at the firm-level has a negative impact on fixed capital investment for financially constrained U.S. firms with zero dividend payments during 1970-1979.

[5]See, for instance, Ziebarth (2013) that finds a comparable amount of misallocation in the 19th century U.S. manufacturing sector as in modern China and India.

[6]See three recent surveys, Restuccia and Rogerson (2013), Hopenhayn (2014) and Buera, Kaboski, and Shin (2015), for a comprehensive review of the literature.

ences in size-productivity covariances. Asker, Collard-Wexler, and De Loecker (2014) find that stochastic productivities combined with capital adjustment costs are important in accounting for dispersions in marginal products of capital. A number of papers study financial frictions and capital misallocation. The argument is that firms of low net worth are financially constrained, and cannot consequently choose inputs in response to productivities. However, a counter-argument lies in the dynamic process that firms self-finance and grow out of constraints, which results in a small misallocation in the steady state. This self-financing does not undo misallocation if (1) productivity is less persistent (see Caselli and Gennaioli, 2013; Moll, 2014); (2) firms cannot save (see Amaral and Quintin, 2010); (3) transition dynamics start from a pre-misallocated economy (see Buera and Shin, 2013); (4) entrepreneurs slowly accumulate wealth before paying entry costs (see Midrigan and Xu, 2014); and (5) the borrowing constraint is endogenous (see Bai, Lu, and Tian, 2016). This paper combines the real frictions and financial frictions of capital, and extend them to the discussion of intermediate goods. The key of misallocation in the stationary distribution in this paper comes from the recurrent financing need of intermediate goods that slows down firms' capital accumulation.

Lastly, this chapter is related to the work linking intermediate goods with TFP. Jones (2011) and Bartelme and Gorodnichenko (2015) argue that vertical linkages of industries through intermediate goods can amplify distortions in resource allocation and explain cross-country TFP differences. Mendoza and Yue (2012) and Pratap and Urrutia (2012) find working capital constraints on intermediate goods in emerging markets important in explaining TFP drops post financial crisis. This paper differentiates from these studies in arguing the role of intermediate goods frictions in slowing capital accumulation and causing misallocation.

The rest of this chapter is structured as follows. Section 2 presents the model. Section 3 calibrates the model, computes the misallocation in the model and in the data, and implements decomposition exercises of misallocation contributed by each friction. Section 4 concludes.

3.2 Model

To quantify how borrowing constraints and pre-order on intermediate goods generate misallocation, this paper incorporates these frictions into a standard firm investment model of Cooper and Haltiwanger (2006). Two types of agents live in the model: firms and financial intermediaries. Firms organize production and maximize net present value of dividends, given financial and real frictions on both intermediate goods and capital. Firms endogenously exit when the net present value of dividends is smaller than exiting and liquidating its assets. Under limited liability, default happens when the liquidation value cannot cover debt repayment. Financial intermediaries consequently choose a break-even interest rate that reflects this default probability. The equilibrium in loanable funds market leads to an endogenous borrowing constraint that shapes firm dynamics.

3.2.1 Firms

The infinite horizon economy is populated with a mass M_t of heterogeneous firms at time t that grows over time. A firm is a decreasing-return-to-scale technology that produces gross output with inputs of intermediate goods, capital and labor, given an exogenous and stochastic productivity. Firms maximize their present values of future dividends and live forever until they exit. After entry, firms cannot issue new equity and only access to one-period borrowings and savings at financial intermediaries. When exit, firms pay debt up to its liquidated assets under the limited liability.

Production Function Firms produce in a gross output production:

$$y_t = exp(z_t)k_t^{\alpha_k} l_t^{\alpha_l} m_t^{\alpha_m} \tag{3.1}$$

where y_t is gross output, z_t is productivity, and k_t, l_t, and m_t are capital, labor and intermediate goods with their respective revenue shares α_k, α_l, and α_m. The production technology is assumed to satisfy decreasing return to scale, $\alpha_k + \alpha_l + \alpha_m < 1$
.

Firm-level productivity z_t is stochastic and evolves according to an AR(1) process:

$$z_{t+1} = (1 - \rho)\mu_z + \rho z_t + \epsilon_{t+1} \tag{3.2}$$

where μ_z is its unconditional mean and common across firms, ρ describes the persistence of productivity, and ϵ_{t+1} is the shock term that follows $N(0, \sigma_\epsilon^2)$. The consequent unconditional distribution for productivity z_t is $N(\mu_z, \sigma_z^2)$, where $\sigma_z = \frac{\sigma_\epsilon}{\sqrt{1-\rho^2}}$.

Given productivity z_t, capital k_t and intermediate goods m_t, firms choose labor input l_t to maximize its gross output net of labor payment, π_t:

$$\pi_t = \max_{l_t} p_y y_t(z_t, k_t, m_t, l_t) - w l_t \tag{3.3}$$

where p_y is output price and w is wage. The separability of labor inputs from other choices as above is because labor inputs can be adjusted intra period without any frictions.

Since this paper considers firms in a partial equilibrium framework, output demand and input supplies are not modeled here. Consequently, I assume constant exogenous prices of output p_y and wage w, and set them to 1. Similar for intermediate goods price p_m in the later subsection.

Pre-Order for Intermediate Goods Firms order intermediate goods m_{t+1} one period in advance (pre-order), when choosing next period capital k_{t+1}. For m_{t+1} intermediate goods, firms pay ω fraction, ωm_{t+1} at time t, and the remaining $(1-\omega)m_{t+1}$ at time $t+1$.

In this environment, firms need working capital to pay for intermediate goods before sales revenue is collected. If the realization of next period productivity z_{t+1} is relatively low, the pre-ordered level m_{t+1} could be too high. In this case, firms can choose $\bar{m}_{t+1} < m_{t+1}$ to maximize profit at time $t+1$ by selling off the extra intermediate good $m_{t+1} - \bar{m}_{t+1}$. However, if the pre-ordered intermediate goods level m_{t+1} is too low to be optimal in a high productivity realization z_{t+1}, firms cannot adjust the intermediates goods beyond m_{t+1}. In other words, firms choose $\bar{m}_{t+1} \leq m_{t+1}$ to maximize the profit Π_{t+1} after payment for intermediate goods:

$$\Pi_{t+1} = \max_{\bar{m}_{t+1} \leq m_{t+1}} \pi_{t+1}(z_{t+1}, k_{t+1}, \bar{m}_{t+1}) - (1-\omega)m_{t+1} + (m_{t+1} - \bar{m}_{t+1}) \tag{3.4}$$

Capital Adjustment Costs Firms can adjust capital stock using two technologies that involve different adjustment costs: a maintenance type mt that can be used for small changes in capital stock, and a construction type ct that allows large investment/divestment.

In particular, while next period capital k_{t+1} can be any value in the construction type ct, in the maintenance type mt, next period capital k_{t+1} is restricted to a small range around the depreciated current capital stock $(1-\delta)k_t$, i.e. $k_{t+1} \in [(1-\delta-\zeta)k_t, (1-\delta+\zeta)k_t]$. There is no fixed cost ξk_t with the maintenance type

mt, while a convex cost with parameter θ exists in both two types. Thus, the cost structure $C(k_t, k_{t+1}|x)$ under type $x \in \{ct, mt\}$ is

$$C(k_t, k_{t+1}|x) = \begin{cases} \xi k_t + \frac{\theta((1-\delta)k_t - k_{t+1})^2}{2(1-\delta)k_t} & \text{if } k_{t+1} \notin [(1-\delta-\zeta)k_t, (1-\delta+\zeta)k_t], x = ct \\ \frac{\theta((1-\delta)k_t - k_{t+1})^2}{2(1-\delta)k_t} & \text{if } k_{t+1} \in [(1-\delta-\zeta)k_t, (1-\delta+\zeta)k_t], x = mt \end{cases} \tag{3.5}$$

Note that there is no cost when firms choose to let capital depreciates and do nothing, i.e. $k_{t+1} = (1-\delta)k_t$.

This capital adjustment cost structure in (3.5) is similar to that in Cooper and Haltiwanger (2006) with an addition of the fixed cost free investment range $[(1-\delta-\zeta)k_t, (1-\delta+\zeta)k_t]$.

Borrowing Constraints Firms can save or borrow at financial intermediaries. To save, they purchase bonds issued by financial intermediaries with a competitive price q^1_{t+1} at time t and get paid back \$1 at time $t+1$. The saving interest rate is thus $r_1 = \frac{1}{q^1_{t+1}} - 1$.

When firms borrow, they issue one-period corporate bonds. As detailed later in the section on financial intermediaries, the price of corporate bonds, $q^2_{t+1}(z_t, b_{t+1}, k_{t+1}, m_{t+1})$, depends on firms' fundamentals: current productivity z_t, future debt b_{t+1}, future capital stock k_{t+1} and future intermediate goods m_{t+1}. The price of bonds q^2_{t+1} decreases with the expected default probability, implying a higher interest rate for borrowing. In the extreme case with no default probability, debt price $q^2_{t+1} = \frac{1}{1+r_2}$ where r_2 is called the prime borrowing interest rate.

The prime borrowing interest rate r_2 is greater than the saving interest rate r_1 by assuming intermediation costs for financial intermediaries. With the spread $r_2 - r_1$, firms never find it optimal to have savings and borrowings at the same time. Therefore, I collapse borrowings and savings into one variable b_{t+1}. When $b_{t+1} > 0$, firms borrow at the price $q_{t+1} = q^2_{t+1}$. When $b_{t+1} < 0$, firms save at the price $q_{t+1} = q^1_{t+1}$.

Since firms cannot issue new equity, dividend by the end of each period d_t shall be nonnegative after payments for next period intermediate goods ωm_{t+1},capital investment $k_{t+1} - (1-\delta)k_t + C(k_t, k_{t+1}|x)$, repayment of debt or saving b_t, new borrowings or savings $q_{t+1}(z_t, b_{t+1}, k_{t+1}, m_{t+1})b_{t+1}$ and operation cost c_o:

$$d_t = \Pi_t(z_t, k_t, m_t) + (1-\delta)k_t - \omega m_{t+1} - k_{t+1} - C(k_t, k_{t+1}|x) - b_t + q_{t+1}(z_t, b_{t+1}, k_{t+1}, m_{t+1})b_{t+1} - c_o \geq 0 \tag{3.6}$$

The above constraint is endogenous, since the price of corporate bonds depends on how much firms borrow.

Value of Continuation At the beginning of each period, firms choose to continue operation or exit. If a firm continues, given state variables (z, b, k, m) and the bond price schedules $q'(z, b', k', m')$, firms' problem is to maximize its value of continuation

$$V^c(z,b,k,m) = \max_{b',k',m',x\in\{ct,mt\}} \Pi(z,k,m) + (1-\delta)k - \omega m' - k' - C(k,k'|x) - b + q'(z,b',k',m')b'$$
$$- c_o + \beta E_{z'|z} V(z',b',k',m') \tag{3.7}$$
$$s.t. \quad \Pi(z,k,m) + (1-\delta)k - \omega m' - k' - C(k,k'|x) - b + q'(z,b',k',m')b' - c_o \geq 0$$
$$b' \leq \bar{b} \qquad \text{(No-Ponzi Game)}$$

where β is the discounting factor, and $\bar{b}$ is a debt limit to prevent firms from playing Ponzi Games. Note that β cannot be greater than $\frac{1}{1+r_2}$, because otherwise firms borrow indefinitely and store the cash. If $\beta < \frac{1}{1+r_2}$, firms only borrow when investment on capital and intermediate goods has a return greater than $\frac{1}{\beta} - 1$. To make sure that firms borrow whenever the return is greater than the prime borrowing rate, I set $\beta = \frac{1}{1+r_2}$.

Solutions to problem (3.7) are policy functions

$$\text{Demand of debt/saving: } b'^d = b'^d(z,b,k,m;q') \tag{3.8}$$
$$\text{Next period capital: } k' = k'(z,b,k,m;q') \tag{3.9}$$
$$\text{Next intermediate goods: } m' = m'(z,b,k,m;q') \tag{3.10}$$

The indicator function $\mathbb{T}(z,b,k,m;b',k',m')$ describes the transition of current state (z,b,k,m) to the choice partition of future state (b',k',m') and is defined as:

$$\mathbb{T}(z,b,k,m;b',k',m') = \begin{cases} 1 & \text{if (3.8), (3.9), and (3.10) hold for firm } (z,b,k,m) \\ 0 & \text{otherwise} \end{cases} \tag{3.11}$$

Exit and Default The value of exit is straightforward. At the end of production in the current period, net worth of a firm equals cash $\Pi(z,k,m) - b\mathbb{1}(b \leq 0)$ plus depreciated capital $(1-\delta)k$ minus debt $b\mathbb{1}(b > 0)$. Once a firm decides to exit, $(1-\gamma_2)$ fraction of cash $\Pi(z,k,m) - b\mathbb{1}(b < 0)$, and $(1-\gamma_1)$ fraction of capital $(1-\delta)k$ evaporates, $\gamma_2 < \gamma_1$. In other words, exit is costly and consumes some resources. Under limited liability, value of exiting $V^x(z,b,k,m)$ for the firm

$$V^x(z,b,k,m) = \max\{\gamma_2\Pi(z,k,m) - b[\mathbb{1}(b \leq 0)\gamma_2 + \mathbb{1}(b > 0)] + \gamma_1(1-\delta)k, 0\} \tag{3.12}$$

An endogenous exit decision $\chi(z,b,k,m)$ is made by comparing the continuation value V^c and the exiting value V^x:

$$\chi(z,b,k,m) = \mathbb{1}\{V^x(z,b,k,m) > V^c(z,b,k,m)\} \tag{3.13}$$

Therefore, the value function $V(z,b,k,m)$ before the exit decision is made is:

$$V(z,b,k,m) = \max\{V^x(z,b,k,m), V^c(z,b,k,m)\} \tag{3.14}$$

Default on debt repayment only happens when firms exit, while debts are rolled over when firms choose to continue.[7] However, firms may exit without default. First, if firms save $b \leq 0$, there is no meaningful default discussion. Second, if the liquidation value of capital and cash $\gamma_2\Pi(z,k,m) + \gamma_1(1-\delta)k$ is greater than debt repayment b, firms repay all.

Thus, the only case when an exiting firm defaults is that it borrows and the liquidation value of asset is smaller than the debt. Loss for lenders in this case is:

$$b - \gamma_2\Pi(z,k,m) - \gamma_1(1-\delta)k \tag{3.15}$$

3.2.2 Entrants and Firm Size Distribution

In each time period t, there are a mass of $\mu_{ent}M_t$ entrants. Each entrant draws an initial productivity z_0 and an initial wealth $b_0 < 0$ independently. Entrants do not differ from incumbents in the unconditional productivity distribution, i.e. $z_0 \sim N(\mu_z, \sigma_z^2)$. The initial wealth $b_0 < 0$ is from a Pareto distribution with a density function $g(-b_0)$:

$$g(-b_0) = \begin{cases} \frac{\alpha a_{\min}^{\alpha}}{(-b_0)^{\alpha+1}} & if - b_0 \geq a_{\min}, \\ 0 & if - b_0 < a_{\min}. \end{cases} \tag{3.16}$$

where a_{min} is the minimum wealth. Note that firms have zero initial capital stock and intermediate goods, i.e. $k_0 = 0, m_0 = 0$.

By the above assumptions, firms do not enter and produce right away. There

[7] See, for instance, Bai, Lu, and Tian (2016) for default when firms continue.

exists a preparation period for entrants to build up capital stock and intermediate goods out of scratch, according to their initial productivity z_0 and wealth b_0. Since $k_0 = 0$, entrants do not pay capital adjustment costs. Their choices of debt/saving $b'_{ent}(z_0, -b_0, 0, 0)$, capital $k'_{ent}(z_0, -b_0, 0, 0)$, and intermediate goods $m'_{ent}(z_0, -b_0, 0, 0)$ are given by maximizing the value function of entrants $V_{ent}(z_0, b_0, 0, 0)$: [8]

$$V_{ent}(z_0, b_0, 0, 0) = \max_{b',k',m'} -\omega m' - k' - b_0 + q'(z, b', k', m')b'$$

$$- c_o + \beta E_{z'|z_0} V(z', b', k', m') \tag{3.17}$$

$$s.t. - \omega m' - k' - b_0 + q'(z, b', k', m')b' - c_o \geq 0 \tag{3.18}$$

$$b' \leq \bar{b} \qquad \text{(No-Ponzi Game)}$$

where the next period z' evolves in the same AR(1) process as for incumbents in Equation (3.2), and V is the value function of incumbents in Equation (3.14). This implies that from $t + 1$ on, the mass of $\mu_{ent} M_t$ entrants start production and behave the same as incumbents.

As before, the indicator function $\mathbb{T}_{ent}(z, b, 0, 0; b', k', m') = 1$ if policy functions with state $(z, b, 0, 0)$ give next period choices b', k', m', and 0 otherwise. Note that an initial misallocation on entrants arises due to the constraint of (3.18). A high productivity entrant with a low draw of wealth finds quite hard to finance the first period capital stock and intermediate goods without retained earnings from the past.

3.2.3 Financial Intermediaries

There exists a continuum of risk-neutral competitive intermediaries that take deposits and lend.[9] For every dollar of intermediation, the cost includes a deposit interest rate r_1 and an intermediate cost c_I.

Given a debt price function $q'(z, b', k', m')$, the problem for a competitive lender

[8]Unlike most literature in entry, exit and industrial dynamics, e.g. Hopenhayn (1992), Cooley and Quadrini (2001), Bento and Restuccia (2015), this paper does not model endogenous entry. In general equilibrium, an endogenous entry means that there is an equilibrium mass of entrants, such that the value of entry equates entry costs for the marginal entrants. If there are more entrants than the equilibrium, value of entry is too low because output price is lower, and input prices are higher than those in equilibrium, and vice versa. Since this paper adopts a partial equilibrium framework, there are no price channels to shape the endogenous entry. The approach here is rather to take some equilibrium mass of entrants, as well as their distribution of productivity and wealth as given.

[9] In this paper, intermediaries are not restricted to financial institutions. One can view the competitive lender as a representation that includes intermediate goods suppliers and other lenders as well. What is implicitly assumed is that the competitive lender can see all borrowing, including trade credit in the form of account payables for example.

is to choose a supply function $b'^s = b'^s(z, k', m'; q')$ to maximize its expected profit:

$$\max_{b'}(1 - E_{z'|z}\chi'(z', b', k', m'))b' + E_{z'|z}\{\chi'(z', b', k', m')(b' - \gamma_2\Pi(z', k', m') \\ - \gamma_1(1-\delta)k')\} - (1 + r_1 + c_I)q'b' \tag{3.19}$$

The first term here presents debt repayment b'^s with the probability that firms continue and debt is rolled over $1 - E_{z'|z}\chi'(z', b'^s, k', m')$. The second term gives an expected loss when the firm defaults when productivity is below some threshold.

3.2.4 Equilibrium

A recursive equilibrium is a debt price function $q'(z, b', k', m')$, policy functions of incumbent firms $b'^d(z, b, k, m; q')$, $k'(z, b, k, m; q')$ and $m'(z, b, k, m; q')$, a transition indicator function for incumbents $\mathbb{T}(z, b, k, m; b', k', m')$, policy functions of entrants $b'_{ent}(z_0, -b_0, 0, 0; q')$, $k'_{ent}(z_0, -b_0, 0, 0; q')$ and $m'_{ent}(z_0, -b_0, 0, 0; q')$, an exit rule $\chi(z, b, k, m)$, a transition indicator function for entrants $\mathbb{T}_{ent}(z, b, 0, 0; b', k', m')$, a supply function of funds $b^s(z, k', m'; q')$, a debt price function $q(z, b', k', m')$, an endogenous mass of firms M' and a distribution of firms $f'(z', b', k', m')$ such that

1. given the debt price function $q'(z, b', k', m')$, policy functions of $b'^d(z, b, k, m; q')$, $k'(z, b, k, m; q')$ and $m'(z, b, k, m; q')$ solve the problem of firms in (3.7), and the exit rule $\chi(z, b, k, m)$ solves the exiting problem (3.13).

2. given the debt price function $q'(z, b', k', m')$, the supply function of funds $b^s(z, k', m'; q')$ solves lenders' problem (3.19).

3. the debt price function $q'(z, b', k', m')$ clears supply and demand of funds at the firm-level, if $b' > 0$:

$$\mathbb{T}(z, b, k, m; b', k', m')b'^d(z, b, k, m; q') = b'^s(z, k', m'; q') \text{ for incumbents} \tag{3.20}$$

and

$$\mathbb{T}_{ent}(z, b_0, 0, 0; b', k', m')b'^d_{ent}(z, b, 0, 0; q') = b'^s(z, k', m'; q') \text{ for entrants} \tag{3.21}$$

with a special case $q'(z, b', k', m') = \frac{1}{1+r_1+c_I}$ when there is zero expected default probability. The consequent interest rate $r_2 = r_1 + c_I$ is the prime borrowing interest rate previously named.

4. the distribution and mass of firms f' and M' evolve recursively as in (3.22) and (3.23), respectively, given an initial mass M_0, an initial firm distribution f_0, mass of entrants μ_{ent}, an exit rule $\chi(z, b, k, m)$ and policy functions of incumbents and entrants:

$$f'(z', b', k', m') = \chi'(z', b', k', m')\{\int_z\int_b\int_k\int_m f(z, b, k, m)\mathbb{T}(z, b, k, m; b', k', m')\pi(z'|z)dzdbdkdm$$
$$+ \mu_{ent}\int_z\int_b \phi(z)g(-b)\mathbb{T}_{ent}(z, b, 0, 0; b', k', m')dzdb\} \quad (3.22)$$

$$M' = M(1 - \int_{z'}\int_{b'}\int_{k'}\int_{m'}\chi(z', b', k', m')f(z', b', k', m')dz'db'dk'dm' + \mu_{ent}) \quad (3.23)$$

with a growth rate $\mu_{ent} - \int_{z'}\int_{b'}\int_{k'}\int_{m'}\chi(z', b', k', m')f(z', b', k', m')dz'db'dk'dm'$. A stationary distribution is defined as $f'(z, b, k, m) = f(z, b, k, m)$ for any state (z, b, k, m).

3.3 Quantitative Analysis

This section discusses how I map the model in Section 3.2 into the CIES data in order to quantify the role of intermediate goods frictions in shaping misallocation. I calibrate the model to the CIES data and compare the measured misallocation generated in the model to that in the CIES data. Several counterfactual experiments are then implemented to quantitatively assess misallocation generated by each friction.

To implement the calibration and counterfactual experiments, simulated data are sampled from the model implied stationary distribution. Despite not capturing the non-stationary part of China's growth,[10] the goal is to understand how intermediate goods frictions affect firm-level investment and production decisions in a stationary distribution approach.

My calibration strategy takes into account the threshold sales of 5 million yuan in the CIES data. Section 3.3.1 elaborates entry and exit patterns in the CIES data,

[10]Unlike developed countries, China's market based economy starts since 1980s and is hard to be described as an economy in its steady state or on some balanced growth path. For countries that experience reforms, Jeong and Townsend (2007) and Buera and Shin (2013) provide a framework of transitional dynamics analysis to understand how reforms gradually change resource misallocation over time.

and the necessity of modeling the threshold sales. The model is then parametrized and calibrated to replicate key moments in firm-level debt, productivity, as well as the market share distribution in the CIES data. Out of model fit suggests similar firm dynamics in the model and in the CIES data for a given cohort.

Section 3.3.2 compares the measured misallocation in the model and in the data, using Hsieh and Klenow (2009)'s definition. If marginal products of intermediate goods, capital and labor were equalized, gross output could be increased by 96% in the model and 140% on average in the CIES data over 1998-2007. In other words, the model accounts for 69% of measured misallocation in the CIES data. Section 3.3.3 decomposes the misallocation generated by the model into contributions by borrowing constraints and real frictions on intermediate goods and capital. I find that intermediate goods frictions account for about a half and 34% of the misallocation in the model and in the CIES data, respectively. Borrowing constraint on intermediate goods is quantitatively more important than pre-order, and accounts for a third and 23% of the misallocation in the model and in the CIES data. While borrowing constraints on capital generates small misallocation, capital adjustment costs account for the other half of misallocation in the model, and 35% in the CIES data. With intermediate goods frictions, the capital accumulation is much slower than without. This leads to a larger amount of misallocation that is hard to get with only frictions on capital.

3.3.1 Parametrization

The CIES data includes only the top 20% firms in the manufacturing sector because of the minimum sales of 5 million yuan threshold (see Table 2.3 in Chapter 2). This impacts the mapping between model and data in firm dynamics and measured misallocation. A significant fraction of entrants and exiters in the CIES are incumbent firms crossing the 5 million threshold sales. If the CIES was taken as the manufacturing sector, several model parameters, e.g. standard deviation of firm-level productivities, would be misspecified. Therefore, I first outline how I take into account the threshold sales before discussion on model parametrization. For more data details, see Section 2.2 in Chapter 2.

Entry and Exit in the CIES Because of the 5 million threshold sales, I name entrants in the CIES data as *data entrants* to distinguish from new firms in the manufacturing sector. Similarly, *data exiters* refer to firms that disappear from the CIES data. To evaluate the potential impact of the threshold sales on the model calibration, I compare the fraction and market share of these firms in the CIES to that in the U.S. manufacturing sector (Dunne, Roberts, and Samuelson, 1988).

Since the U.S. literature uses census data that are 5-year apart, I study entry and exit in the CIES over a 5-year horizon. With the birth year information, I classify data entrants into two groups depending on whether these firms are younger or older than 5 years old. I name firms who are more than 5 years old as *old data entrants*, and those less as *young data entrants*.

Table 3.1 presents fractions, market shares and relative sizes of data entrants over 1998-2003 and over 2002-2007.

Table 3.1: CIES Data Entrants over a 5-Year Horizon

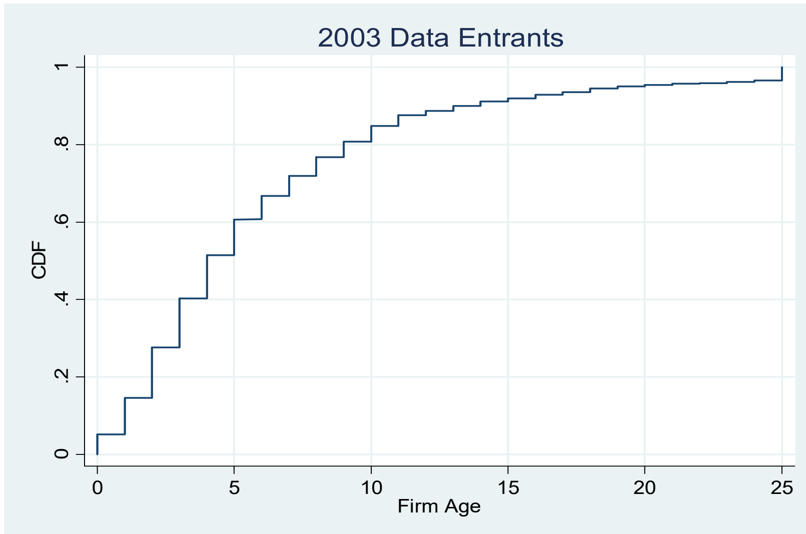
	1998-2003	2002-2007
Number of Firms		
Incumbents	32.11%	24.37%
Data Entrants (Age > 5)	29.06%	30.47%
Data Entrants (Age ≤5)	38.83%	45.16%
Total Market Share		
Incumbents	50.86%	44.42%
Data Entrants (Age > 5)	22.45%	24.85%
Data Entrants (Age ≤5)	26.69%	30.73%
Relative Size of Output		
Data Entrants (Age > 5)	0.49	0.45
Data Entrants (Age ≤ 5)	0.43	0.37

Incumbents are defined as firms who are in CIES data for both year t and $t+5$. Data entrants are defined as firms that are not in the CIES data at t, but appear in the data at $t+5$. Age is computed as observation year $t+5$ minus the birth year.

Data entrants over a 5-year horizon are the majority of firms in the CIES data. Between 1998 to 2003, 66% of firms in 2003 enter into the dataset after 1998, among which 43% are old data entrants more than 5 years old. The fraction of data entrants increases during 2002-2007, with a similar fraction of old data entrants. In terms of market share, data entrants in 2003 and 2007 produce 49% and 56% of gross output in the economy. This suggest that a CIES data entrant produces, on average, 37% to 49% of the average gross output level of CIES incumbents in 2003 and 2007.

Compared to their U.S. counterparts, the fraction and the average size of data entrants are much larger in the CIES. According to Dunne, Roberts, and Samuelson (1988), entrants over a 5-year horizon are 52% of U.S. census firms, producing

Figure 3.1: Age Distribution of Data Entrants in 2003



Data entrants are defined as firms are not in the CIES data in 1998, but appear in the data in 2003, same to the definition in Table 3.1.

a market share of 17%. This suggests that an average data entrant is at least two times large as its counterpart in the U.S. census data.[11]

The threshold sales mean that a large fraction of new firms in the manufacturing sector are unobserved. Over time, some of these firms exit, while others accumulate capital, net worth and grow in sales. The age distribution of data entrants over 1998-2003 in Figure 3.1 suggests that about 30% of these firms take 5 to 15 years to grow above 5 million yuan. This implies that either the threshold sales is too high in the sales distribution, or the growth of firms is too slow. The steepness of this CDF provides an identification to gauge the role of frictions on intermediate goods and capital in slowing down firms growth. Logically, if frictions are small, productive firms grow rapidly and quickly surpass the threshold sales. The resulting age distribution among data entrants in Figure 3.1 should have a large density on ages smaller than 5, and the CDF should be steep. The degree of age dependence in entering the CIES in Figure 3.1 reveals information on frictions of intermediate goods and capital. [12]

The interpretation of exit from the CIES data is complicated by the fact that

[11]According to Hsieh and Klenow (2014), average employment among 5- to 9- year-old firms is twice as that of firms who age from 1 to 5.

[12]The idea of age and size dependence has been studied in Davis, Haltiwanger, and Schuh (1996) and Cooley and Quadrini (2001) with the latter using financial frictions as an explanation.

one cannot distinguish between firms whose sales fall below 5 million yuan and firms who truly exit. Table 3.2 shows that over a 5-year horizon, 59% firms in the 1998 CIES data are no longer there in 2003, with a slightly lower number 56% over 2002-2007. These 5-year rates imply an annualized exit rate around 15% to 16%, and is much higher than the annual 8% rate from a survival analysis report by State Administration for Industry and Commerce in the manufacturing sector.[13] Similar to entrants, the relative size of exiters are about 3 times large as their counterparts in U.S. census data (Dunne, Roberts, and Samuelson, 1988).

Table 3.2: CIES Data Exiters over a 5-year Horizon

	1998-2003	2002-2007
Number of Firms		
Stayers	40.67%	44.34%
Data Exiters	59.33%	55.66%
Total Market Share		
Stayers	59.91%	61.23%
Data Exiters	40.09%	38.77%
Relative Size of Output		
Data Exiters	0.46	0.50

Stayers are defined as firms that are in CIES data for both year t and $t+5$. Data exiters are defined as firms that are in CIES data for year t and absent for year $t+5$.

The above analysis suggests that accounting for the threshold sales is important to quantify how frictions shape misallocation through entry, exit and firm dynamics. Several problems arise if one takes the CIES data as the entire manufacturing sector in China. First, since firms that stay in the CIES data are relatively large, the volatility of firm-level productivities is underestimated by using the standard deviation of productivity in the cross-sectional data. This could lead to a smaller misallocation in the simulated model (Asker, Collard-Wexler, and De Loecker, 2014) that can hardly match that in the CIES data. Second, if old data entrants are treated as new firms, misallocation on these firms is misleadingly attributed to an initial misallocation on new firms. This bias is likely to be large, since old data entrants are about a quarter in numbers and market shares.

Assigned Parameters The parameterization procedure takes two steps: assigned parameters from literature or other direct sources, and calibrated ones to

[13]See http://www.saic.gov.cn/zwgk/tjzl/zxtjzl/xxzx/201307/P020130731318661073618.pdf

match key moments in firm-level debt and productivity, as well as age and market share distributions in China Industrial Enterprise Survey Data (CIES).

Capital adjustment costs are parameterized from Cooper and Haltiwanger (2006) with a fixed cost parameter $\xi = 0.039$ and a convex adjustment cost parameter $\theta = 0.049$. The fixed cost free range for investment parameter $\zeta = 0.09$, which is equal to the capital deprecation rate $\delta = 0.09$.

Firms' discount factor β is set to 0.94, which implies an average prime borrowing interest rate $r_2 = \frac{1}{\beta} - 1 = 0.06$ according to People's Bank of China (PBOC) annual reports from 1998 to 2007. Similarly, saving interest rate is set $r_1 = 0.03$ to match the average deposit rate in PBOC reports.

Capital recovery rate γ_2 upon default is set to 30%, which is the average asset recovery rate of non-performing loans from 2001 to 2006 in China (Fan and Morck, 2013, p. 85). The cash recovery rate γ_1 is lower than that of capital and equals 10%, closer to the lower bound of cash recovery rate 6.90% reported in Fan and Morck (2013).

Calibrated Parameters Given assigned parameters in the first step, the remaining parameters are calibrated to match key moments in firm-level debt, productivity, as well as age and market share distributions in the CIES.

In the gross output production function, the labor share α_l is set 0.05, which is the wage bill fraction of gross output revenue (see Table B.1 in appendix). The intermediate goods share α_m is set to 0.7, between the number in Table 2.4 and 68% reported in Jones (2011). Since the capital share is unobserved, I calibrate the return to scale parameter η to match the fact that 84.5% of total gross output is produced by the top 10% firms in the manufacturing sector, which are equivalently the top 50% firms in the CIES. The idea is that as η increases, gross output is more concentrated on top producers in the sales distribution. This gives $\eta = 0.85$ and consequently $\alpha_k = 0.10$.[14]

The threshold sale $y_c = 436.30$ is chosen to have 20% of firms above this level in the simulated gross output distribution. In other words, the subgroup of firms with sales above y_c in the simulated data is the model equivalent of the CIES data.

The productivity process parameters are calibrated to match the productivity moments in the simulated top 20% firms to those in the CIES data.[15] In particular,

[14]This is higher than an average capital share 0.06 when using an imputed rental rate 13% in appendix A.2.

[15]In computation, the support of productivity is $[\mu_z - 4.5\sigma_z, \mu_z + 4.5z]$ and discretized into 15 grids. The AR(1) process is then discretized into a 15×15 transition matrix by Tauchen (1986)'s method.

its persistence $\rho_z = 0.7$ is chosen to match the fact that 40.67% of firms in the 1998 CIES data remain by 2003. Since more than 50% of data exiters are estimated as continuing firms, it is mainly the persistence of productivity ρ, not the operating cost c_o, that shoves firms out of the CIES data. The mean $\mu_z = 0.9$ and volatility $\sigma_z = 0.7$ are calibrated so that the average and standard deviation for firm-level productivity are 1.82 and 0.45 in the simulated top 20% firms as in the CIES data.

The debt limit $log(\bar{b})$ determines access to credit. With a higher limit in the future, the chance that firms are unable to roll over debt after a low productivity shock is lower. Consequently, the incidence of exit and default is less likely, which induces more borrowing by firms with q low net worth. Therefore, the level of $log(\bar{b})$ is chosen to match the fraction of firms access to credit in the simulated data to that in the World Bank Data Survey 2012.[16] For the empirical underpinnings, I interpret firms' debt in a more general form that includes borrowings from bank and non-bank financial institutions, trade credit from suppliers and other borrowings from friends, relatives and moneylenders in the World Bank data. This gives a 34.29% of firms with debt and a calibrated debt limit $log(\bar{b}) = 6.096$.

The population exit rate differs from the exit rate in the CIES, and is largely determined by the operating cost c_o. The level is set to match the population exit rate 8% during 2008-2012, according to the survival analysis report of firms by State Administration for Industry and Commerce. The annual growth rate in the manufacturing population during this period is approximately 9%, according to censuses 2004 and 2008. To match this growth rate, the relative mass of entrants μ_{ent} is chosen to be 17%.

Entrants draw productivity from the unconditional distribution $\phi(z)$, and wealth from the Pareto wealth distribution with a shape parameter α and a minimum wealth a_{min}. The productivity distribution of entrants is the same as that of incumbents. The shape parameter α and minimum wealth a_{min} determines the distribution of first-period output for entrants post entry. The fraction of intermediate goods paid a period ahead ω impacts how fast a firm grows after birth, and therefore the relative market share over different ages and different percentiles of sales. Thus, the three parameters α, a_{min} and ω are jointly pinned down to match a 67.89% market share of entrants in 2003, a 59% market share of data exiters in 1998, and

[16] The liability on accounting tables of the CIES is too noisy to be a good measure of debt, since it includes several accounting items such as wages and pension payable, interest payable, customer deposits etc. Instead, the World Bank Enterprise Survey 2012 asks firms specifically whether they have borrowed for financing working capital and fixed capital investment, from banks, non-bank financial institutions, suppliers and other sources.

Table 3.3: Parametrization

Parametrized			**Calibrated**		
Parameter		Value	Parameter		Value
Discounting factor	β	0.94	Return to Scale	η	0.85
Depreciation rate	δ	0.09	Labor share	α_l	0.05
Capital Adjustment Cost			Intermediate goods share	α_m	0.70
Fixed cost	ξ	0.039	fraction of intermediate goods in advance	ω	60%
Fixed cost free band	ζ	0.09	Debt limit	$log(\bar{b})$	6.10
Convex cost	θ	0.049	Threshold sales	y_c	436.30
Interest Rates			Operating cost	c_o	5.00
Saving rate	r_1	0.03	*Productivity Process*		
Prime borrowing rate	r_2	0.06	Population persistence of productivity	ρ_z	0.70
Recovery Rates			Population S. D. of productivity	σ_z	0.70
Cash	γ_2	0.10	Unconditional mean	μ_z	0.90
Capital	γ_1	0.30	*Initial Wealth Distribution of Entrants*		
			Mass of entrants	μ_{ent}	0.17
			Pareto Shape	α	0.60
			Min. Wealth	a_{min}	20.00

a market share of new firms as 1.3 times as that of old data entrants in 2003.

Table 3.4: Targeted Moments

Moments	Data	Model
Market share by firms of top 10% sales	84.5%	87.5%
Population exit rate*	8%	5.8%
Frac. of firms above threshold	20%	20%
Frac. of firms with debt*	34.29%	34.60%
Mean productivity (Top 20% firms)	1.82	1.80
SD of productivity (Top 20% firms)	0.45	0.43
5-year Horizon		
Frac. of CIES data exiters in t	59%	57.88%
Frac. of CIES data entrants in $t+5$	67.89%	72.20%
Market share of CIES data exiters in t	40.09%	41.29%
Market share of CIES data entrants in $t+5$	49.14%	56.38%
$\frac{\text{Market Share of Young Data Eantrants}}{\text{Market Share of Old Data Entrants}}$	1.3	1.3

Moments except for * are from the CIES data 1998-2003. The population exit rate is from a survival analysis of firms in China by State Administration for Industry and Commerce. The fraction of firms with debt is from World Bank Enterprise Survey 2012.

Table 3.3 lists all calibrated parameters and their levels, and Table 3.4 shows the differences of targeted moments in the model and in the data. The model overall replicates these key moments in the data with some room to improve on the exit rate and the market share of CIES data entrants in $t+5$.

Out of Model Fit To check whether the calibrated model reasonably captures the frictions that firms in the CIES data face, I compare the dynamics of a given cohort in the simulation to that in the CIES data for 6 years after their entry.

The specific dynamics presented here is how productivity, capital and sales of a given cohort converge to those of all firms, both in the CIES data and in the simulated top 20% subgroup. Since the CIES data only includes the top 20% of manufacturing firms, I assume that only the top 20% of the simulated data are observable every period. This suggests that firms of a given cohort that are observable may not be the same set of firms every year. For example, in the 1998 CIES data, 5024 firms report that their birth years are 1998. In 1999, 801 of the 5024 firms disappear from the CIES data, while another 3038 firms of the 1998 cohort who are below the threshold sales in 1998 enter into the 1999 CIES data.[17]

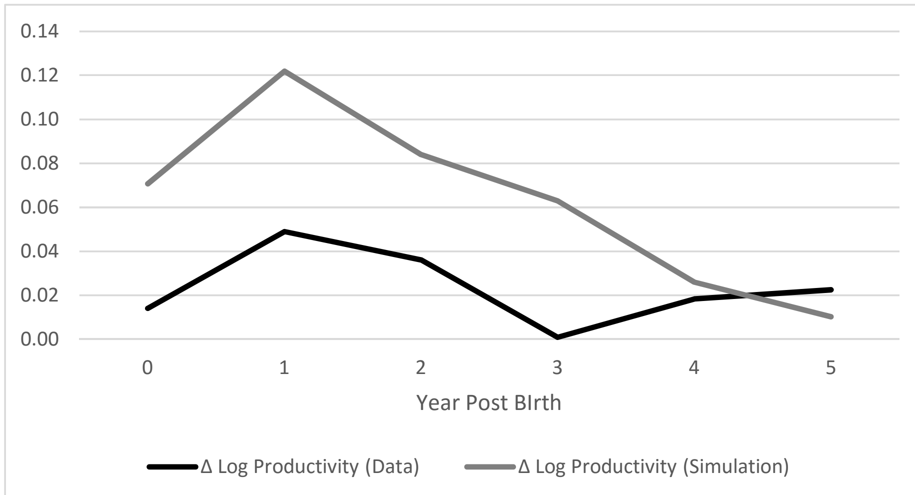
I focus on the 1998 cohort of the CIES since it provides the longest observation window. Figure 3.2, 3.3 and 3.4 plot the differences in average log productivity, log capital and log sale between a birth cohort and all firms, both in the CIES data and in the simulated top 20% firms over 6 years post birth. In Figure 3.2, the model replicates two features of productivity for the 1998 cohort in the CIES data. First, firms of the birth cohort that produce above threshold sales outperform other firms in productivity. Since most entrants start from a low capital stock (see Figure 3.3), those producing more than the threshold sales must have a higher productivity. Second, such a productivity advantage decays over time. The decaying effect is due to the Markovian process of productivity, and a lower level productivity needed to be included in the CIES data or top 20% firms by a growing average capital for the birth cohort, as evident (see Figure 3.3). Combining productivity and capital, log sales of birth firms first grow then get stable in Figure 3.4.[18] Overall, the model delivers a similar pattern of firm dynamics to that observed in CIES data.[19]

[17] See Table B.4 in appendix for more details. Table B.4 also compares entry and exit of the simulated top 20% subgroup for a given cohort over time to that in the CIES data.

[18] I also look at 1999-2006 cohorts. The general pattern of decaying productivity, growing capital stock and sales hold in other cohorts. However, from 1998 to 2006, the relative log productivity of a birth cohort compared to all firms in the CIES data is declining over time, while the average firm-level productivity in the CIES grows. One explanation is that firm-level productivity growth, e.g. through R&D, takes place on incumbent firms rather than on new firms.

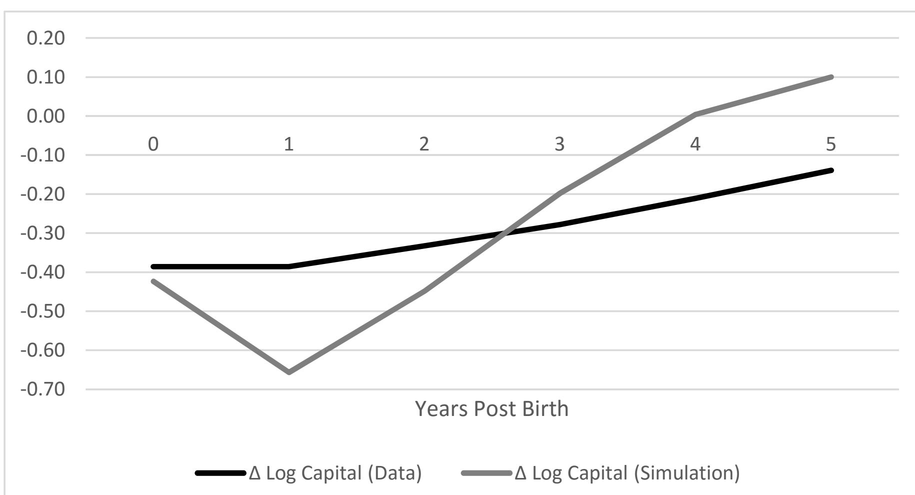
[19] The dip of log capital in Figure 3.3 is due to intermediate goods frictions. Note that these firms enter at time $t=-1$, and prepare in $t=-1$ before productions in $t=0$. At time $t=-1$, the birth cohort have no retained earnings from the past and are constrained in intermediate goods m_0. At time $t=0$, firms have retained earnings from the production and are therefore less constrained in intermediate goods choice m_1 than in $t=-1$. As a result, firms enter into $t=1$ with more pre-

Figure 3.2: Diff. in Log Productivity of a Given Cohort Compared to All Firms above the Threshold Sales, Data vs Model



The cohort for the CIES data is 1998 cohort. ∆ log productivity is defined as the difference of log productivity between 1998 cohort and all firms in the CIES data. A similar definition applies for the simulated top 20% firms in the model.

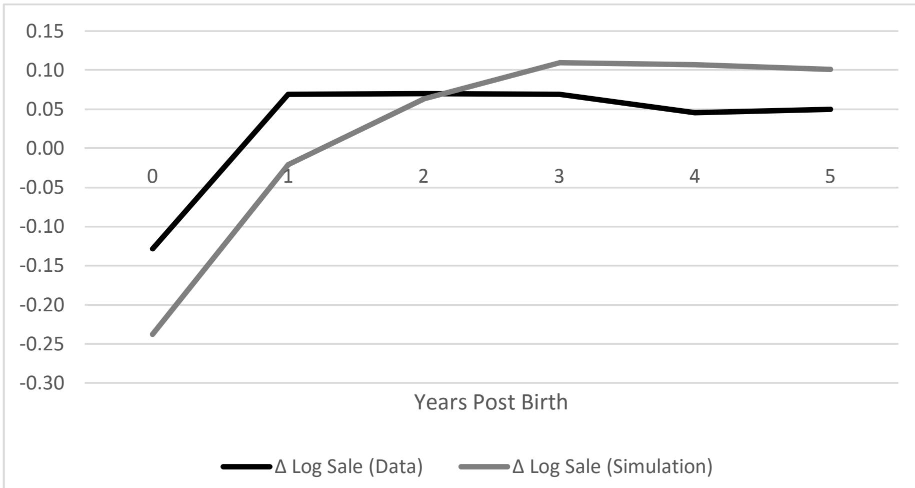
Figure 3.3: Diff. in Log Capital of a Given Cohort Compared to All Firms above the Threshold Sales, Data vs Model



The cohort for the CIES data is 1998 cohort. ∆ log capital is defined as the difference of log capital between 1998 cohort and all firms in the CIES data. A similar definition applies for the simulated top 20% firms in the model.

ordered intermediate goods, and hence have lower capital stocks on average to produce more than the threshold sales. A similar argument applies to the spike in productivity (see Figure 3.2).

Figure 3.4: Diff. in Log Sale of a Given Cohort Compared to All Firms above Threshold Sales, Data vs Model



The cohort for the CIES data is 1998 cohort. Δ log sale is defined as the difference of log sale between 1998 cohort and all firms in the CIES data. A similar definition applies for the simulated top 20% firms in the model.

3.3.2 Misallocation: Model vs Data

This section compares measured misallocation in the calibrated model with that in the CIES data. I compute measured misallocation as the percentage of gross output gain if marginal products of intermediate goods, capital and labor were equalized across firms. I find that the gross output gain averages 140% of actual gross output in the CIES data over 1998-2007, and 96% in the simulated top 20% firms. Overall, the model accounts for 69% of misallocation that is measured in the CIES data.

Measured Misallocation Following Hsieh and Klenow (2009) and Midrigan and Xu (2014), measured misallocation quantifies the potential output gain by reallocating inputs to equalize marginal products across firms, given a fixed distribution distribution of firm-level productivity. My measure of misallocation differs, however, in that the output refers to gross output rather than value-added.

Given n firms with capital stock k_i, labor employment l_i, intermediate goods m_i and actual gross output y_i for each firm i, a hypothetically efficient aggregate Y_{eff} is calculated by reallocating k_i, l_i and m_i to equate marginal products of capital MP_k, labor MP_l and intermediate goods MP_m, holding aggregate capital $\sum_i k_i$, intermediate goods $\sum_i m_i$ and labor $\sum_i l_i$ constant. This potential output Y_{eff} is achievable when all inputs are adjustable intra period without frictions, assuming that firms purchase inputs in a competitive markets without any distortions. In other words, Y_{eff} is the first-best gross output if there is a social planner that maximizes the sum of firm-level gross output, facing no real and financial frictions. Measured misallocation is calculated:

$$\text{Measured Misallocation} = \frac{Y_{eff} - Y}{Y} \tag{3.24}$$

while $Y = \sum_i y_i$ is the actual aggregate gross output in the data.

Misallocation in the CIES Data Using Equation (3.24), I compute the measured misallocation in the CIES data each year over 1998-2007. Since there are a range of industries in the CIES data, the efficient output Y_{eff} is computed to equalize marginal products within each 2-digit CIC industry.

Several pre-treatments on firm-level productivities are required for two reasons. First, the distribution of firm-level productivities in the CIES data exhibits thicker left and right tails than the log normal distribution in the model. Since the hypothetical efficient output Y_{eff} is sensitive to tail values in the productivity distribution,[20] I trim the productivity distribution in the CIES to make the data and the model share the same support of firm-level productivities. Second, there is a growth on average firm-level productivities in the CIES data, from 1.82 in 1998 to 2.12 in 2007, which is absent in the model. Therefore, the trimming scheme is adapted to take this growth into account. Specifically, in 1998, the range of productivity in CIES data is trimmed to be $[\mu_z - 4.5\sigma_z, \mu_z + 4.5\sigma_z]$ to match that in the model computation. In later years, the ranges are adjusted to be $[\mu_z + \Delta z - 4.5\sigma_z, \mu_z + \Delta z + 4.5\sigma_z]$, while Δz is the difference in productivity of a given year compared to 1998 (see Table B.2 in Appendix).

Table 3.5 presents the measured misallocation in the CIES data over 1998-2007. The average 1.4 suggests that if marginal products of inputs were hypothetically equalized, gross output in the CIES data could be 2.4 times that of actual gross output. A general declining trend in output gain exists from 1998 to 2003, and from 2004 to 2007. There is an increase of 25 percentage points in 2004, which is

[20] In the most extreme case, if there is a small fraction of firms with extremely high productivities yet not much inputs, the gross output gain would be tremendous. Therefore, one must take a stand whether these firms are outliers compared to firms in the model. This is more a problem when the return to scale gets closer to 1. My view here is that the productivity range in the CIES must be consistent with the calibrated model.

Table 3.5: Gross Output Gain by Equalizing Marginal Products within 2-digit CIC Industry, CIES

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	**Average**
Gross Output Gain	1.68	1.43	1.54	1.29	1.33	1.15	1.40	1.35	1.26	1.24	**1.40**

potentially caused by the fact that the data entry rate in this year is more than 50%, much higher than 20% in other years during 1998-2007.[21]

Misallocation in the Model How does the model perform in accounting for the measured misallocation in the CIES data in Table 3.5? I compute the gross output gain in the simulated top 20% firms, which is the model equivalent of the CIES data.[22] Since the model does not distinguish industries, marginal products are equalized across all top 20% firms to calculate the efficient gross output. Table 3.6 presents that the gross output gain is 96% in the model, suggesting that the gross output would almost double if marginal products were equalized. Compared to the CIES data, the model accounts for close to 70% of misallocation in the data.

Table 3.6: Gross Output Gain, CIES Data vs Top 20% Firms in Model Simulation

CIES Data 1998-2007	Model	Model % of Data
1.40	0.96	69.34%

3.3.3 Decomposing Misallocation

The benchmark model has four frictions: borrowing constraints and real frictions on both intermediate goods and capital. How much does each friction account for the measured misallocation in the calibrated model? More importantly, do frictions on intermediate goods help to account for more misallocation, on top of a standard investment model with capital frictions? To answer these questions, I carry out

[21] Such a high entry rate is possibly due to the fact that the First Economic Census takes place in 2004, which includes a set of firms that have sales barely at 5 million yuan, and avoid reporting to National Bureau of Statistics prior 2004 (Holz, 2013)

[22] The measured misallocation in the simulated top 20% firms does not differ much from that in the simulated manufacturing population (see Table B.3 in Appendix). This is sensible since the CIES data produces 90% of manufacturing gross output in 2004.

Table 3.7: Model Specifications of Counterfactual Experiments

Friction	Benchmark	Exp. 1	Exp. 2	Exp. 3
Real Frictions on Capital	✓	✓	✓	✓
Borrowing Constraints on Capital	✓	✓	✓	
Real Frictions on Intermediate Goods	✓	✓		
Borrowing Constraints on Intermediate Goods	✓			

counterfactual experiments that remove the frictions one by one from the benchmark model. The counterfactual experiments proceed in the following order:

- Experiment 1 (Exp. 1) removes the borrowing constraint on intermediate goods. Specifically, payment for next period intermediate goods, $\omega m'$, enters into firms' objective function, but not into the endogenous borrowing constraint (see Equation 3.7). One interpretation of this counterfactual is that firms cannot default on any borrowings for intermediate goods purchases. Or, whenever firms default on these borrowings, intermediate goods can be seized and sold without any costs by the lenders, who are most likely intermediate goods suppliers in practice.

- Experiment 2 (Exp. 2) further removes pre-order on intermediate goods by allowing firms to choose the static optimal amount of intermediate goods after productivity shocks. In other words, firms solve optimal labor l and intermediate goods m intra period:

$$\max_{l,m} y(z, k, l, m) - m - l \tag{3.25}$$

  The resulting intermediate goods m_{static} is static optimal. This specification is close to the modern sector part of Midrigan and Xu (2014) but with capital adjustment costs.

- Experiment 3 (Exp. 3) further removes borrowing constraints on capital by allowing negative dividends or new equity issuance. This is essentially Asker, Collard-Wexler, and De Loecker (2014) or a gross output version of Cooper and Haltiwanger (2006) but with entry and exit.

Table 3.7 lists which frictions are included under benchmark and counterfactual specifications. In each experiment, the counterfactual model is simulated to obtain a new stationary distribution given calibrated parameters from the benchmark model. The threshold sales y_cs are recalculated so that there are always 20% firms above the threshold. Given these top 20% firms under each counterfactual experiment with their actual output y_i, I calculate measured misallocation as in Equation (3.24). This exercise answers the question how much gross output gain would be if firms in the CIES behaved as in Exp. 1, 2 and 3, if marginal products of all inputs were equalized. Results are presented in Table 3.8.

Borrowing Constraints on Intermediate Goods In Exp. 1, gross output gain would be 0.64 times the actual output Y, if marginal products of capital, labor and intermediate goods were equalized among the top 20% simulated firms. Compared to the gain of 0.96 in the benchmark model, the gross output gain in Exp. 1 is 0.32 lower. Therefore, measured misallocation in Exp. 1 accounts for 46.69% of that in the CIES data, which is 22.65 percentage points lower than that in the benchmark model. This implies that borrowing constraints on intermediate goods account for 22.65% of measured misallocation in the CIES data.

Borrowing constraints on intermediate goods generate measured misallocation through three channels. First, since intermediate goods is 70% of revenue, a positive down-payment for intermediate goods increases the borrowing need and tightens the borrowing constraint. Roughly, for each \$1 of expected sales in the next period, \$0.42 of intermediate goods have to be financed on top of capital investment. This increases the amount of borrowing that are subject to the default risk from the lenders' view. Further, this increase of borrowing need is recurrent since intermediate goods depreciate in one-period.

Second, at time t, capital investment $k_{t+1} - (1-\delta)k_t$ could be crowded out because of the borrowing need for intermediate goods m_{t+1}. Since capital serves as better collateral than intermediate goods, the level of collateral decreases, causing a more tightened constraint from time t and on.

Third, while pre-order on intermediate goods reduces profits and lowers capital investment, borrowing constraints on intermediate goods prolong this negative effect. At time t, firms cannot achieve the static optimal intermediate goods m_{static} if the pre-ordered level m_t is too low. With a lower current profit, next period capital k_{t+1} is lower than without pre-order. If there is no borrowing constraints on intermediate goods, firms could react to order more intermediate goods m_{t+1} for $t+1$ by the persistence of productivities. This increases profit in $t+1$ and consequently alleviates the negative effect on capital from time $t+2$ and on. With borrowing

Table 3.8: Simulated Output Gain by Equalizing Marginal Products, Benchmark vs Counterfactuals

	Model Specifications							
	Benchmark		Exp. 1		Exp. 2		Exp. 3	
	Potential Gain	% of Data	Potential Gain	% of Data	Potential Gain	% of Data	Potential Gain	% of Data
	0.96	69.34%	0.64	46.69%	0.49	35.62%	0.48	34.75%
Each Friction:								
B.C. on Intm. Goods	0.32	22.65%						
R.F. on Intm. Goods			0.15	11.07%				
B.C. on Capital					0.01	0.87 %		
R.F. on Capital							0.48	34.75%

B.C. on Intm. Goods: Borrowing constraints on intermediate goods; R.F. on Intm. Goods: Real frictions on intermediate goods; B.C. on Capital: Borrowing constraints on capital; R.F. on Capital: Real Frictions on Capital.
Benchmark includes all four frictions. Exp. 1 removes borrowing constraints on intermediate Goods. Exp 2 further removes real frictions on intermediate goods. Exp 3 lastly removes borrowing constraints on capital.

Table 3.9: Effect of Inadequate Intermediate Goods β_3 on Future Capital, Benchmark vs Exp. 1

	Benchmark	Exp. 1
$t+1$	-0.5722	-0.3196
$t+2$	-0.5389	-0.0012
$t+3$	-0.4331	0.0222
$t+4$	-0.3270	0.0241
$t+5$	-0.2727	-0.0073

Benchmark includes all four frictions. Exp. 1 removes borrowing constraints on intermediate Goods.

constraints on intermediate goods, however, firms are constrained in the amount of intermediate goods m_{t+1} they can order. Therefore, the status of inadequate intermediate goods persists, keeps the constraint tight in $t+1$, and so forth. In other words, the negative effect on capital persists over time.

To illustrate the second and the third channels at time t, I regress capital in the following periods $logk_{i,t+\Delta}$, $\Delta \geq 1$, on a dummy of inadequate intermediate goods, Dummy($m_t < m_{static}$) in t, controlling for starting state variables of productivity z_{it}, capital $logk_{i,t}$, and debt/saving level $b_{i,t}$:

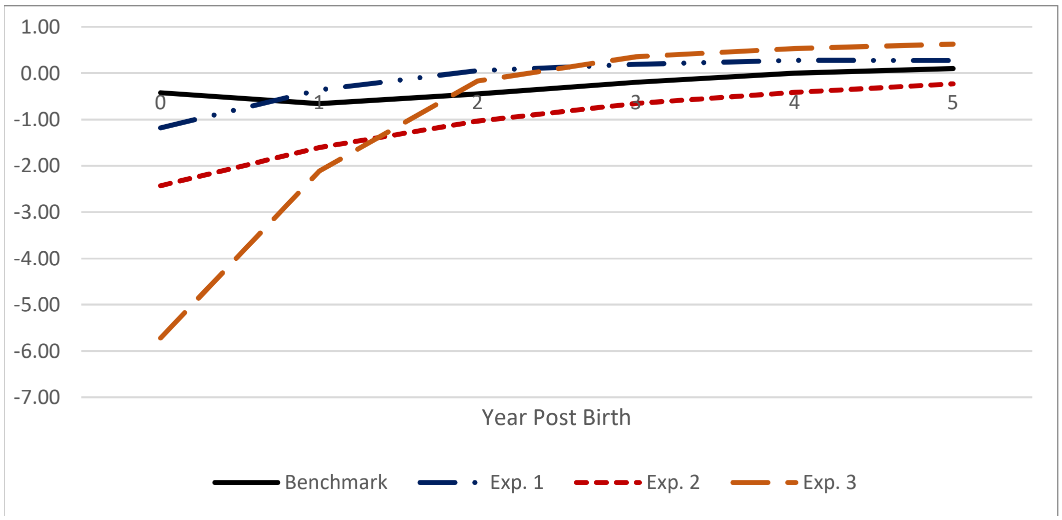
$$\text{log}k_{i,t+\Delta} = \beta_0 + \beta_1 z_{i,t} + \beta_2 logk_{i,t} + \beta_2 b_{i,t} + \beta_3 \text{Dummy}(m_t < m_{static})_{i,t} + residual \tag{3.26}$$

among the top 20% simulated firms in the benchmark model and in Exp. 1. The idea is that without intermediate goods frictions, firms with the same state variables (z_t, b_t, k_t) choose the same capital k_{t+1} in $t+1$. With intermediate goods frictions, a firm that has a lower productivity z_{t-1} and capital k_{t-1} in $t-1$, and chooses a lower intermediate goods m_t is in a disadvantage situation in choosing capital k_{t+1}. Equation (3.26) quantifies the impact of such a disadvantage on future capital.

Table 3.9 gives estimates of β_3 in both Benchmark and Exp. 1. In the benchmark model, the impact of inadequate intermediate goods lowers next period capital by 57%, 25 percentage points higher than in Exp.1 due to the second channel. The negative effect in Benchmark prolongs to time $t+5$ with a negative impact of 27%, due to the third channel. In contrast, next period capital only decreases by 32% in Exp. 1 and such a negative effect almost vanishes in one period.

Because of the three channels, the capital accumulation process is consequently slower when borrowing constraints on intermediate goods are on. In Figure 5,

Figure 3.5: Diff. in Log Capital of A Birth Cohort Compared to All Firms above Threshold Sales, Benchmark vs Counterfactuals



Benchmark includes all four frictions. Exp. 1 removes borrowing constraints on intermediate Goods. Exp. 2 further removes pre-order on intermediate goods. Exp. 3 lastly removes borrowing constraints on capital.

Table 3.10: Log Gross Output and Log Capital Distributions, Benchmark vs Counterfactuals

Percentile	Model Specifications							
	Benchmark		Exp. 1		Exp. 2		Exp. 3	
	log y	log k	log y	log k	log y	log k	log y	log k
10%	6.27	4.06	6.92	6.02	7.36	5.04	7.64	6.50
25%	6.40	5.53	7.51	6.34	7.70	6.34	8.20	7.80
50%	7.29	6.34	8.20	7.15	8.70	8.45	9.18	9.27
75%	8.14	6.50	9.74	8.94	9.74	9.27	10.00	9.92
90%	9.24	7.48	10.85	9.75	11.24	9.92	11.37	9.92
Mean	7.46	6.04	7.98	7.49	9.04	7.84	9.33	8.83

Distribution for each specification is within the subgroup of top 20% firms. Benchmark includes all four frictions. Exp. 1 removes borrowing constraints on intermediate Goods. Exp. 2 further removes pre-order on intermediate goods. Exp. 3 lastly removes borrowing constraints on capital.

average log capital for a birth cohort takes 6 years to converge to the average of all top 20% firms in the benchmark model, and only 3 years in the Exp. 1. The resulting firm size distribution also differs. Table 3.10 illustrates that compared to Exp. 1, the benchmark has uniformly lower log capital on all percentiles with 145% lower average capital. Consequently, average sales is 52% lower in the benchmark model than in Exp. 1.

Pre-Order on Intermediate Goods Exp. 2 further removes the real friction of pre-order on intermediate goods. Table 3.8 shows that if firms in the CIES data behaved as in the model of Exp. 2, the potential output gain would be 0.49 times the actual output Y, which accounts for 35.62% of measured misallocation in the CIES. The difference of 11.07% between Exp. 1 and Exp. 2 is therefore attributed to pre-order on intermediate goods.

In Exp. 1, since firms cannot purchase intermediate goods above the pre-ordered level, there exists a pre-cautionary motive for investment in intermediate goods. Consequently, the pre-ordered level of intermediate goods m' is on average 16 times of that used in the production.

Despite the pre-cautionary motive, the choice of next period intermediate goods m' is influenced by an expected return that is determined by the distribution of future productivity, and capital stock in the next period. Therefore, firms may have a pre-ordered level of intermediate goods lower than the static optimal level after high productivity shocks are realized. Table 3.11 shows that there are 31.12% of these firms among the top 20% group, and 16.57% among all simulated firms. These firms have a unsurprisingly higher productivity, 42.7% higher among the top 20% firms and 10.42% higher among all simulated firms. They also tend to have a lower capital stock, 86.86% lower among top 20% firms and 42.19% lower among all simulated firms, which rationalizes a lower marginal return and a lower option value of ordering intermediate goods a period ahead.

Real frictions also lowers the capital investment through a lower current profit. However, as shown in Table 3.9, this effect is short-lived. This is because firms react to the high productivity by investing a high level of next period intermediate goods, which soon increases the next period profit and capital investment. As a result, the average capital in Exp. 1 is 35% lower than in Exp. 2, smaller than the difference of 145% between Benchmark and Exp. 1.

Borrowing Constraints on Capital Exp. 3 further removes borrowing con-

Table 3.11: Intermediate Goods Usage Compared to the Static Optimal, Simulated Exp. 2

	Frac. of	Average Log Productivity		Average Log Capital	
	Smaller	Equal	Smaller	Equal	Smaller
Top 20% Firms	31.12%	1.7474	2.1744	7.7653	6.8967
All Firms	16.57%	0.8838	0.9880	6.1316	5.7097

Subgroup Smaller (Equal): firms with intermediate goods usage in Exp. 2 smaller than (equal) the static optimal level. Exp 2 removes borrowing constraints and real frictions on intermediate goods.

straints on capital. Table 3.8 presents that in a model with only one-period time-to-build and capital adjustment costs, the gross output gain is 0.48 of actual gross output Y, if marginal products of capital, labor and intermediate goods were equalized across firms. This is close to the 0.49 in Exp. 2, and suggests that borrowing constraints on capital only account for 0.87% of measured misallocation in the CIES data.

One may argue that the small output loss is driven by restricting attention to firms in the top 20% of sales distribution, as these firms are relatively unconstrained. There may still exist many firms that are constrained and below the threshold sales such that removing borrowing constraints on capital increases output significantly among all firms. Yet if I calculate measured misallocation among all simulated firms in Exp. 2 and Exp. 3, the results are similar to that in Table 3.8. The potential output gain is 0.52 in Exp. 2, and 0.51 in Exp. 3.

The small output loss from borrowing constraints on capital, however, does not indicate firms are unconstrained. In Figure 3.5, capital accumulation with the constraint takes more than 5 years to converge to the average log capital among top 20% firms, slower than the case without the constraint. As a result, percentiles of log capital distribution in Table 3.10 with the constraint are lower than those without the constraint, except for the 90 percentile. The seeming contradiction of a small output loss and a lower average capital stock is reconciled by the fact that the most productive firms have high capital stocks with and without the constraint. Since most output is produced by these top productive firms,[23] reallocating inputs does not increase more output gain despite an average lower capital with the constraint. Table 3.12 illustrates this point. If I calculate the fraction of capital stock owned by firms above log productivity 2.25 (75 percentile of the productivity distri-

[23]This property is called granularity, see Xavier (2011).

bution), the number is 53% in Exp. 2 with borrowing constraints on capital, and a similar fraction 57% in Exp. 3 without. In terms of output, they produce a very similar fraction of aggregate output at 97% in Exp. 2 and Exp. 3. A qualitatively similar result holds for the 99 percentile of productivity 3.15.

Table 3.12: Fraction of Gross Output Produced and Owned by Top Productive Firms, Top 20% Simulated Firms

		Exp. 2		Exp. 3	
Log Productivity Above	Productivity Percentile	Frac. of Agg. Output	Frac. of Agg. Capital	Frac. Of Agg. Output	Frac. Of Agg. Capital
3.15	99%	71.3%	14.5%	71.3%	11.5%
2.25	75%	97.1%	53.4%	97.4%	56.7%

Exp 2 removes borrowing constraints and real frictions on intermediate goods. Exp 3 further removes borrowing constraints on capital.

Compared to the literature, this finding of a small output loss from borrowing constraints on capital is unsurprising. Midrigan and Xu (2014) and Moll (2014) address that self-financing can undo misallocation caused by borrowing constraints on capital, which is also true in this paper as firms can save and productivities are persistent. As Hopenhayn (2014) points out, the impact of financial constraints on misallocation may be larger, given a high entry with young and constrained firms. However, this paper shows that with capital adjustment costs slowing down the growth of firms, a high entry rate of 17% in China, about two times in Midrigan and Xu (2014), does not magnify the role of borrowing constraints on capital in accounting for misallocation. Unless there are other frictions that further slow down the growth of firms, it could be hard to get a sizable misallocation by simply feeding in more entrants.

Real Frictions on Capital The remained frictions in Exp. 3 are one-period time-to-build and adjustment costs for capital. The channels of these two frictions are well addressed in the investment literature, e.g. Cooper and Haltiwanger (2006) and Khan and Thomas (2008), and in the misallocation literature, e.g. Asker, Collard-Wexler, and De Loecker (2014). First, because of one-period time-to-build, capital stock determined a period ahead is not static optimal after stochastic productivity shocks. Second, due to the fixed and convex adjustment costs, firms that would like to adjust capital to a certain level may find it too costly to do.

Table 3.8 presents that in Exp. 3, gross output gain is 0.48 if marginal products of capital, labor and intermediate goods were equalized across firms, which accounts for about 35% of measured misallocation in the CIES data. Without frictions on intermediate goods and borrowing constraints on capital, the capital threshold for top 20% firms in sales are effectively lower, and capital accumulation is pretty fast. In Figure 3.5, a birth cohort starts from an average capital much lower than the mean of all firms in the top 20% group, and converges to the mean in 3 years. This accumulation speed is much faster than that in the CIES data in Figure 3.3.

3.3.4 Bias in Value Added Misallocation

In parallel to my discussion of biased value added misallocation in Section 2.4 in Chapter 2, this section explores whether value added approach in literature also underestimates value added misallocation in the model.

Table 3.13 presents value added gains under the value added approach and gross output approach, for both simulated data and China's data. The calculation is the same as in Section 2.4 in Chapter 2. In the model, when I reallocate capital and labor based on firm-level value added productivity, value added gain is 67%. However, if intermediate goods is reallocated along with capital and labor, value added gain is 203%. In short, value added approach in the literature also underestimates the magnitude of value added misallocation in the model.

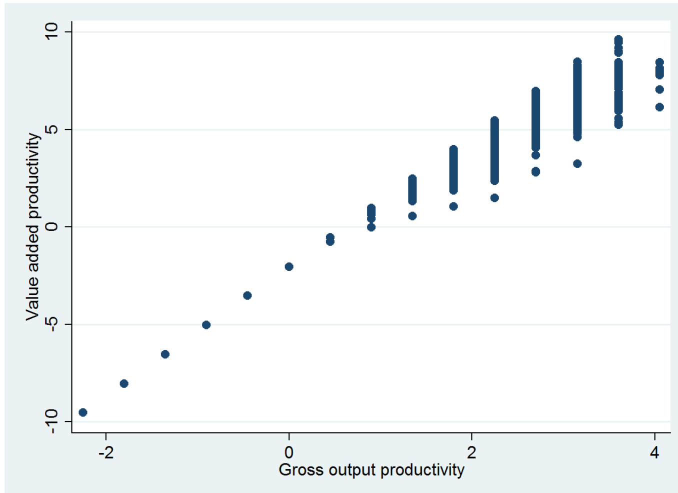
Table 3.13: Value Added Gains under Value Added Approach and Gross Output Approach, Model vs Data

	Value added approach in literature		Gross output approach	
	Data	Model	Data	Model
Value added gain	98.12%	66.76%	550.74%	203.48%

Source of the bias in value added approach partly comes from biased value added productivity. As discussed in Section 2.4, the correlation between intermediate goods distortions and gross output productivity matters for the direction of either an underestimated or an overestimated value added productivity for firms with high gross output productivities. While there may exist intermediate goods distortions that work in both directions in the data,[24] mechanisms in this model

[24]For example, financial frictions on intermediate goods lead to a wedge on intermediate goods that is positively correlated with gross output productivity. In other scenarios, productive firms in gross output could have an advantage in access to more intermediate goods suppliers and thus have a lower wedge on intermediate goods.

Figure 3.6: Non-linear Mapping Between Log Value Added and Log Gross Output Productivities, Model



lead to a positively correlation between intermediate goods distortion and gross output productivity. Specifically, because of financial constraints on intermediate goods, firms with a high gross output productivity desire to invest a large volume of intermediate goods but are constrained to do so. For this reason, value added productivity is lower for constrained firms than their unconstrained counterparts with the same level of gross output productivity. This can be seen evidently in Figure 3.6, where the mass of constrained firms is nonzero when log gross output productivity is greater than 0.9.

Summary This section first compares gross output misallocation in model and data. Overall, the model accounts for 69% of gross output misallocation in CIES data.

This section further decomposes the model simulated misallocation in Section 3.3.2 into contributions by borrowing constraints and real frictions on intermediate goods and capital. Out of 69% CIES measured misallocation by the model, I find that real frictions on capital and borrowing constraints on intermediate goods account for the most measured misallocation in the CIES data by 34.75% and 22.65%, respectively. Real frictions on intermediate goods account for another 11.07% while borrowing constraints on capital account for a negligible 0.87%. In other words, new frictions proposed in this chapter on intermediate goods can account for an extra 33.72% of misallocation in China, on top of the misallocation caused by capital frictions that has been well studied in the literature.

Lastly, this section compares value added misallocation measures in value added approach and gross output approach in the model. I find that consistent with the findings in data, value added misallocation under the value added approach is 67%, and underestimates the true misallocation 203% under the gross output approach.

3.4 Conclusion

This chapter introduces borrowing constraints and pre-order on intermediate goods and quantifies their role in accounting for measured misallocation in the CIES data. I incorporate these intermediate goods frictions, as well as borrowing constraints on capital, into a standard firm investment model Cooper and Haltiwanger (2006) with entry and exit. When calibrated to key moments in China, I find that the model generates substantial misallocation, and accounts for 69% of that in the CIES data. A further decomposition shows that frictions on intermediate goods are quantitatively important. They account for about a half of the model generated misallocation, and about 34% of misallocation in the CIES data. In particular, borrowing constraints account for a third and around a quarter of misallocation in the model and in the CIES data, respectively. Real frictions on capital are also important and account for a similar magnitude of misallocation as intermediate goods frictions. Consistent with the literature, I find that borrowing constraints on capital generate little misallocation.

There are several future extensions for this project. First, in the current version, costs in intermediate goods and capital are in one borrowing constraint. It can be, however, separated into two with one working capital constraint on current period intermediate goods, and one borrowing constraints on capital. This alternative setting will keep the negative effect of borrowings for intermediate goods on capital, as long as competitive lenders observe both types of borrowings. An upper limit for the current period intermediate goods would be the result of the working capital constraint. This benefits in relaxing the assumption that firms cannot buy intermediate goods above the pre-ordered level in the current version. Second, the partial equilibrium stand on intermediate goods, labor and output markets can be extended into a general equilibrium analysis. Output of each firm can be used as final goods consumed by households, and also intermediate goods for other firms, as in Basu (1995). If inputs are misallocated at the first place by mechanisms studied in this paper, most productive firms cannot produce and supply a large amount of output in the intermediate goods market. The extra effect, studied by Jones

(2011) and Bartelme and Gorodnichenko (2015), could increase prices of intermediate goods, which could further create misallocation by worsening the financing problem for productive constrained firms.

Bibliography

Amaral, Pedro S. and Erwan Quintin. 2010. "Limited Enforcement, Financial Intermediation and Economic Development: A Quantitative Assessment." *International Economic Review* 51(3):785–811.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic Inputs and Resource (Mis)Allocation." *Journal of Political Economy* 122(5):1013–1063.

Bai, Yan, Dan Lu, and Xu Tian. 2016. "Do Financial Frictions Explain Chinese Firms' Saving and Misallocation?" *2016 Meeting Papers. No. 998. Society for Economic Dynamics, 2016* .

Bartelme, Dominick and Yuriy Gorodnichenko. 2015. "Linkages and Economic Development." Working Paper 21251, National Bureau of Economic Research.

Bartlesman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review* 103(1):305–334.

Basu, Susanto. 1995. "Intermediate Goods and Business Cycles: Implications for Productivity and Welfare." *The American Economic Review* 85(3):512–531.

Bento, Pedro and Diego Restuccia. 2015. "Misallocation, Establishment Size, and Productivity." *2015 Meeting Papers. No. 1391, Society for Economic Dynamics* .

Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten. 2016. "Firm Entry and Regional Growth Disparities: the Effect of SOEs in China." *Mimeo, University of Toronto* .

Brandt, Loren, Trevor Tombe, and Xiaodong Zhu. 2013. "Factor Market Distortions across Time, Space and Sectors in China." *Review of Economic Dynamics* 16 (1):39 – 58.

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

Buera, Francisco and Yongseko Shin. 2013. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy* 121(2):221–272.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin. 2015. "Entrepreneurship and Financial Frictions: A Macrodevelopment Perspective." *Annual Review of Economics* 7 (1):409–436.

Caselli, Francesco and Nicola Gennaioli. 2013. "Dynastic Management." *Economic Inquiry* 51 (1):971–996.

Castro, Rui and Pavel Sevcik. 2016. "Occupational Choice, Human Capital, and Financial Constraints." Tech. Rep. 20162, University of Western Ontario, Centre for Human Capital and Productivity (CHCP).

Cooley, Thomas F. and Vincenzo Quadrini. 2001. "Financial Markets and Firm Dynamics." *American Economic Review* 91(5):1286 – 1310.

Cooper, Russell W. and John C. Haltiwanger. 2006. "On the Nature of Captial Adjustment Costs." *Review of Economic Studies* 73:611–633.

Da-Rocha, Jose-Maria, Marina Mendes Tavares, and Diego Restuccia. 2016. "Firing Costs, Misallocation, and Aggregate Productivity." *No. w23008. National Bureau of Economic Research, 2016* .

Da-Rocha, Jos´e-Mar´ıa, Marina Mendes Tavares, and Diego Restuccia. 2015. "Firing Costs, Misallocation, and Aggregate Productivity." *Working Paper, University of Toronto* .

Davis, Steven J., John C. Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction*. MIT Press.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. 1988. "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries." *The RAND Journal of Economics* 19(4):495–515.

Fan, Joseph P.H. and Randall Morck. 2013. *Capitalising China*. The University of Chicago Press.

Fazzari, Steven M. and Bruce C. Petersen. 1993. "Working Capital and Fixed Investment: New Evidence on Financing Constraints." *The RAND Journal of Economics* 24 (3):328–342.

Gabler, Alain and Markus Poschke. 2013. "Experimentation by Firms, Distortions, and Aggregate Productivity." *Review of Economic Dynamics* 16 (1):26 – 38.

Ghate, Chetan and Kenneth Kletzer. 2012. *Financial Frictions and Monetary Policy Transmission in India*. Oxford University Press.

Holz, Carsten A. 2013. "Chinese Statistics: Classification Systems and Data Source." *Eurasian Geography and Economics* 54 (5-6):532–571.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and firm Dynamics in Long Run Equilibrium." *Econometrica* 60(5):1127–1150.

———. 2014. “Firms, Misallocation, and Aggregate Productivity: A Review.” *Annual Review of Economics* 6:735–770.

Hopenhayn, Hugo A. and Richard Rogerson. 1993. “Job Turnover and Policy Evaluations: A General Equilibrium Analysis.” *Journal of Political Economy* 101:915–938.

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

———. 2014. “The Life Cycle of Plants in India and Mexico.” *The Quarterly Journal of Economics* 129(3):1035–1084.

Jeong, Hyeok and Robert M. Townsend. 2007. “Sources of TFP Growth: Occupational Choice and Financial Deepening.” *Economic Theory Special Edition Honoring Edward Prescott 32* 1:189–221.

Jones, Charles I. 2011. “Intermediate Goods and Weak Links in the Theory of Economic Development.” *American Economic Journal: Macroeconomics* 3(2):1–28.

Jose, Manuel L., Carol Lancaster, and Jerry L. Stevens. 1996. “Corporate Returns and Cash Conversion Cycles.” *Journal of Economics and Finance* 20(1):33–46.

Khan, Aubhik and Julia K. Thomas. 2008. “Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics.” *Econometrica* 76(2):395–436.

Mendoza, Enrique G. and Vivian Z. Yue. 2012. “A General Equilibrium Model of Sovereign Default and Business Cycles.” *The Quarterly Journal of Economics* 127 (2):889–946.

Midrigan, Virgiliu and Daniel Yi Xu. 2014. “Finance and Misallocation: Evidence from Plant-Level Data.” *American Economic Review* 104(2):422–458.

Moll, Benjamin. 2014. “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation.” *American Economic Review* 104(10):3186–3221.

Mukoyama, Toshihiko and Sophie Osotimehin. 2016. “Barriers to Reallocation and Economic Growth: the Effects of Firing Costs.” *Mimeo, University of Virginia* .

Pratap, Sangeeta and Carlos Urrutia. 2012. “Financial Frictions and Total Factor Productivity: Accounting for the Real Effects of Financial Crises .” *Review of Economic Dynamics* 15 (3):336 – 358.

Restuccia, Diego and Richard Rogerson. 2013. “Misallocation and Productivity.” *Review of Economic Dynamics* 16:1–10.

Tauchen, George. 1986. "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions." *Economics Letters* 20 (2):177 – 181.

Tombe, Trevor and Xiaodong Zhu. 2015. "Trade, Migration and Productivity: A Quantitative Analysis of China." *Manuscript, University of Toronto* .

Xavier, Gabaix. 2011. "The Granular Origins of Aggregate Fluctuations." *Econometrica* 79 (3):733–772.

Ziebarth, Nicolas L. 2013. "Are China and India Backward? Evidence from the 19th Century U.S. Census of Manufactures." *Review of Economic Dynamics* 16 (1):86 – 99.

Chapter 4

Left-Censored CIES Data and Source of China's TFP Growth

4.1 Introduction

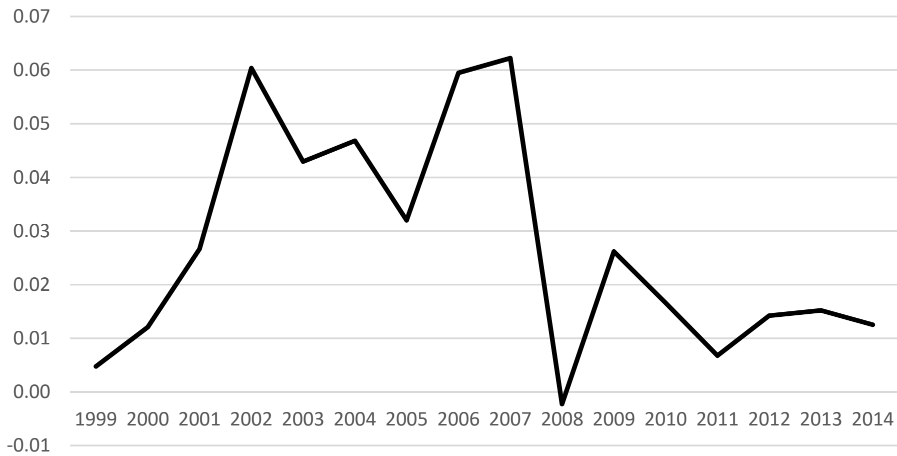
China's total factor productivity (TFP) grew at a rate of 3.9% during 1999-2007 before the 2008 financial crisis (see Figure 4.1).[1] Most of the TFP growth comes from the industrial sector (Brandt and Zhu, 2010; Brandt, Van Biesebroeck, and Zhang, 2012), which has undergone a massive entry of new firms and policy reforms.[2] One natural question is about the source of this high TFP growth. Is it the case that Chinese firms are becoming more productive over time? Or is the growth from resource reallocations among existing firms? Or is it driven by the entry of new and more productive firms?

This paper builds on the aggregate productivity decomposition literature which has been used to study the U.S. manufacturing sector (Baily, Hulten, and Campbell, 1992; Foster, Haltiwanger, and Krizan, 2001). Using the China Industrial Enterprise Survey (CIES) data, I define aggregate productivity as the gross output weighted sum of firm-level productivity. The aggregate productivity is then decomposed into three components: technological growth, intensive reallocation and extensive reallocation. The first component, technological growth, measures the average firm-level productivity growth, while the second and the third measure

[1] An alternative estimate of China's TFP growth during 1998-2007 is 4.68% in Zhu (2012), where the agricultural sector grows at 4.13%, the non state non-agricultural sector grows at 3.67%, and the state non-agricultural sector grows at 5.5 %.

[2] Number of firms in China Industrial Enterprise Survey (CIES) data increased from 147,690 to 304,599 during 1998-2007. Major reforms include privatization of loss-making state owned firms, drops in inter-provincial migration and trade costs, bankruptcy law reforms, interest rate liberalization and accession to WTO (Xu, 2011; Zhu, 2012).

Figure 4.1: Aggregate TFP Growth in China, 1999-2014



Source: Penn World Table 9.0

changes in firm-level weights through resource reallocations. The intensive reallocation evaluates reallocations among continuing firms, and the extensive one is about whether more productive entrants replace less productive exiters over time.

The decomposition exercise, however, faces a potential bias when applied directly in the CIES data. Unlike the U.S. census data used in the literature, the CIES includes all state owned firms, and non state owned firms with sales more than 5 million yuan. According to The First Economic Census 2004 (see Table 2.3 in Chapter 2), there are about 80% manufacturing firms, all non state owned, that are not included in the dataset. Therefore, the entry and exit of firms in the CIES are not defined in the same way as in the U.S. census data.

Does this left-censoring data limitation cause a quantitatively significant bias on decomposing manufacturing productivity growth in the CIES? To answer these questions, this paper develops a method to recover the true technological growth, intensive and extensive reallocations in China's manufacturing sector. I first impose assumptions that the firm-level productivity process is AR(1), that firms close exogenously, that the correlation between market share and productivity is constant over the sales distribution, and that log sales in the manufacturing sector follows a Pareto distribution. Under these assumptions, I show that the three components can be recovered using CIES and several aggregate statistics on firms whose sales are less than 5 million yuan.

To investigate whether the productivity decomposition is biased in the CIES, I first take CIES as the manufacturing universe, and carry out the Baily, Hulten, and Campbell (1992)'s version of decomposition exercise (henceforth BHC decomposition). Similar exercises have bee done in Van Biesebroeck (2008), Brandt, Van Biesebrock, and Zhang (2014) and Ding, Guariglia, and Harris (2016) using other methodologies. Then I use my method to correct for the left-censoring problem and re-do the decomposition. In preparing for the correction exercise, the AR(1) productivity process is consistently estimated, taking account of the selection problem in the left-censored CIES data. I also calibrate the log sales distributions for all firms and firms younger than 5 years old.

The direct productivity decomposition in the CIES suggests that manufacturing productivity grows 18% over 1998-2003 and 16% over 2002-2007. Extensive reallocation through net entry contributes 93% and 144% of these growths for the two periods, respectively. Intensive reallocation, i.e. reallocation of inputs across existing firms, contributes -10% and -93% to the aggregate productivity growth over 1998-2003 and 2002-2007. The rest 17% and 49% are contributed by technological growth at the firm-level. This contrasts to the estimates based on U.S. data that intensive reallocation contributes at least 20% of 5-year aggregate productivity growth across business cycles, while extensive reallocation barely contributes to the growth (Baily, Hulten, and Campbell, 1992). This China-U.S. difference cannot be accounted by more entrants in the CIES, neither can the existence of state owned firms.

The left-censoring problem partially explains the difference in productivity decomposition between China and U.S.. After corrected for the left-censoring, I find that the fraction of aggregate productivity growth accounted by extensive reallocation drops to 74% and 85% in 1998-2003 and 2002-2007, respectively. Most of the drops are picked up by higher intensive reallocations. In 1998-2003, intensive reallocation increases the aggregate productivity by 4.53%, and contributes to 24% of its growth. In 2002-2007, while the intensive reallocation is still negative, it lowers aggregate productivity by -7%, which is half of -14% in the direct CIES decomposition, and contributes -40% of its growth. Corrected results for technological growth are more mixed. In 1998-2003, because firms that decline in sales and exit from CIES experience more negative productivity changes, the contribution from technological growth drops to 2%. The opposite happens in 2002-2007 and technological growth contributes to 54% of aggregate productivity growth.

The corrected decomposition differs from the naive estimates in the CIES for two reasons. First, manufacturing firms with sales less than 5 million are neglected in the direct CIES decomposition. These firms actually increase aggregate productivity by 1% through intensive reallocation and at least another 1% through technological growth in 1998-2003 and 2002-2007. Second, incumbent firms that grow and decline in sales, and cross the 5 million cutoff are mis-classified as entrants and exiters in the direct CIES decomposition. Therefore, the extensive reallocation is biased upward and the intensive reallocation is biased downward if one takes CIES as the manufacturing universe. The movements of these firms across the 5 million minimum sales increase aggregate productivity by about 4% in both 1998-2003 and 2002-2007.

To check the robustness of my procedure, I create a hypothetical dataset with 10 million as the minimum sales. In particular, these measures of technological growth, intensive and extensive reallocations in the manufacturing sector are close to that in the CIES when I redo the correction exercise in the hypothetical data. Furthermore, the qualitative results of upward biased extensive reallocation and downward biased intensive reallocation also hold for the non state owned manufacturing sector.

This paper is related to the literature about micro-level channels of aggregate productivity growth in China. Brandt, Van Biesebroeck, and Zhang (2012) estimated firm-level productivity in the CIES, and finds that net entry drives most output-weighted aggregate productivity growth during the same period 1998-2007. Similar results can be found in Van Biesebroeck (2008) and Ding, Guariglia, and Harris (2016). Built on earlier studies, this paper contributes to the discussion by finding that the left-censoring problem in China's data biases the importance of net entry in growth upward.

This paper also relates to the aggregate productivity decomposition literature that mainly focuses on the U.S. manufacturing sector. Baily, Hulten, and Campbell (1992) developed the methodology of output-weighted aggregate productivity growth used in this paper. Similar decomposition methods were followed by Foster, Haltiwanger, and Krizan (2001) and Bartelsman, Haltiwanger, and Scarpetta (2004), of which the latter studies cross-country differences in intensive reallocation.[3] Compared to these studies, this paper develops a new method that intends to recover the decomposition in the manufacturing sector when only left-censored data are available.

[3]Other popular decomposition methods include Olley and Pakes (1996) and Petrin and Levinsohn (2012). Each decomposition method has its own advantage. For example, Petrin and Levinsohn (2012) has a better macroeconomic consistency, since their aggregated firm-level productivity equals to the one used in the representative firm framework. This paper uses Baily, Hulten, and Campbell (1992)'s approach because it is more intuitive and provides detailed stylized facts in the U.S. manufacturing sector. I leave the differences of aggregate productivity decomposition by different methods for future research.

This paper is lastly related to the misallocation literature of China. Hsieh and Klenow (2009) documented that there is substantial static misallocation in China's firm-level data. Several papers, Brandt, Tombe, and Zhu (2013), Tombe and Zhu (2015), Brandt, Kambourov, and Storesletten (2016) and Bai, Lu, and Tian (2016) among many others, proposed explanations of financial frictions, entry costs and trade and migration costs to account for the misallocation. With no intent to provide any explanation of misallocation, this paper provides a dynamic lens to look at China's misallocation and how it is affected by the left-censoring problem in China's data.

Although this paper works on China's data, there are many other countries whose firm-level data are left-censored. For example, the commonly used Chilean and Colombian datasets have a cutoff of 5 employees. France, Italy and Argentina also have a sales cutoff as in China (Bartelsman, Haltiwanger, and Scarpetta, 2004). The results of this paper suggest that the left-censoring problem shall be taken into account when carrying out cross-country analysis in the source of aggregate productivity growth.

The rest of this paper is organized as follows. Section 2 introduces measurement of productivity and the methodology of decomposing output-weighted aggregate productivity. Section 3 adds more relevant data details on top of Section 2.2 in Chapter 2. Section 4 presents the direct aggregate productivity decomposition, taking CIES as the manufacturing universe. Section 5 proposes the method to correct for the left-censoring problem and presents the corrected aggregate productivity decomposition. Section 6 concludes.

4.2 Measurement and Methodology

Consider an economy with heterogeneous firms that has firm-level productivity z_{it} and market share θ_{it} in year t. Aggregate productivity is defined as market share weighted average of firm-level productivity $\sum_i \theta_{it} z_{it}$. According to Baily, Hulten, and Campbell (1992) and Foster, Haltiwanger, and Krizan (2001), the growth of aggregate TFP growth from year $t-\tau$ to year t can be decomposed as:

$$\begin{aligned}
&\sum_{it} \theta_{it} z_{it} - \sum_{it-\tau} \theta_{it-\tau} z_{it-\tau} \\
&= \sum_{i \in \text{stayer}} \theta_{it-\tau}(z_{it} - z_{it-\tau}) + \sum_{i \in \text{stayer}} (\theta_{it} - \theta_{it-\tau}) z_{it} + \sum_{i \in \text{entrant}} \theta_{it} z_{it} - \sum_{i \in \text{exiter}} \theta_{it-\tau} z_{it-\tau} \qquad (4.1)
\end{aligned}$$

The first term, *Technological Growth*, in Equation (4.1) captures market share weighted growth of productivity in stayers, which can be a result of exogenous technological growth or endogenous productivity improvement from firms' learning (Jovanovic, 1982).[4]

The second term, *Intensive Reallocation*, evaluates whether relatively more productive firms at time t gain market shares from $t-\tau$ to t.[5] In an economy with distortions, this reallocation of inputs could be slow and there are little changes of market shares towards more productive firms. Therefore, this second term is arguably smaller in an economy with higher distortions.

The third term, *Extensive Reallocation*, evaluates the aggregate productivity gain from replacing less productive exiters by more productive entrants. This term would be positive when entrants are more productive or are larger in market share, or both, than exiters.

Because of the co-existence of state owned firms and non state owned firms in Chinese data, I further decompose Equation (4.1) by ownerships:

$$
\begin{aligned}
&\sum_{it} \theta_{it} z_{it} - \sum_{it-\tau} \theta_{it-\tau} z_{it-\tau} \\
&= \sum_{i \in \text{SOE stayer}} \theta_{it-\tau}(z_{it} - z_{it-\tau}) + \sum_{i \in \text{NonSOE stayer}} \theta_{it-\tau}(z_{it} - z_{it-\tau}) \\
&+ \sum_{i \in \text{SOE stayer}} (\theta_{it} - \theta_{it-\tau}) z_{it} + \sum_{i \in \text{NonSOE stayer}} (\theta_{it} - \theta_{it-\tau}) z_{it} \\
&+ \Big(\sum_{i \in \text{SOE entrant}} \theta_{it} z_{it} - \sum_{i \in \text{SOE exit}} \theta_{it-\tau} z_{it-\tau} \Big) \\
&+ \Big(\sum_{i \in \text{NonSOE entrant}} \theta_{it} z_{it} - \sum_{i \in \text{NonSOE exit}} \theta_{it-\tau} z_{it-\tau} \Big) \\
&+ \Big(\sum_{i \in \text{NonSOE to SOE}} \theta_{it} z_{it} - \sum_{i \in \text{NonSOE to SOE}} \theta_{it-\tau} z_{it-\tau} \Big) \\
&+ \Big(\sum_{i \in \text{SOE to NonSOE}} \theta_{it} z_{it} - \sum_{i \in \text{SOE to NonSOE}} \theta_{it-\tau} z_{it-\tau} \Big)
\end{aligned} \tag{4.2}
$$

Within each ownership type, Equation (4.2) is the same as in Equation (4.1) in decomposing sector level productivity growth into technological growth, intensive and extensive reallocation. Extra terms come from the fact that some firms switch

[4] Since I use revenue productivity as the productivity measure, productivity growth among stayers can also comes from shifts of firm level demand curves. While interesting, this paper keeps an agnostic stand about the source of productivity growth.

[5] In Foster, Haltiwanger, and Krizan (2001), it is further decomposed into a covariance term, and a product of market share changes times productivity in $t-\tau$: $\sum_{i \in \text{stayer}} (\theta_{it} - \theta_{it-\tau}) z_{it} = \sum_{i \in \text{stayer}} (\theta_{it} - \theta_{it-\tau})(z_{it} - z_{it-\tau}) + \sum_{i \in \text{stayer}} (\theta_{it} - \theta_{it-\tau}) z_{it-\tau}$

ownership status.

4.3 Data

This section introduces the productivity measure used in this chapter, and more comparisons between entry and exit firms in the CIES and in the U.S. census data. Most of the data analysis here builds on earlier data Section 2.2 in Chapter 2 and Section 3.3.1 in Chapter 3.

Productivity Measure Firm-level productivity measure in this chapter z_{it} is computed as:

$$\begin{aligned} z_{it} &= log(\text{Gross Output}_{it}) - \alpha_l log(\text{Total Wage}_{it}) - \alpha_m log(\text{Intermediates}_{it}) \\ &\quad - (1 - \alpha_l - \alpha_m) log(\text{Capital}_{it}) \end{aligned} \tag{4.3}$$

for firm i at time t, with a constant return to scale Cobb-Douglas production function in mind. Shares of labor and intermediate goods, α_l and α_m, are the median of firm-level labor and intermediates shares within 2-digit CIC industries.

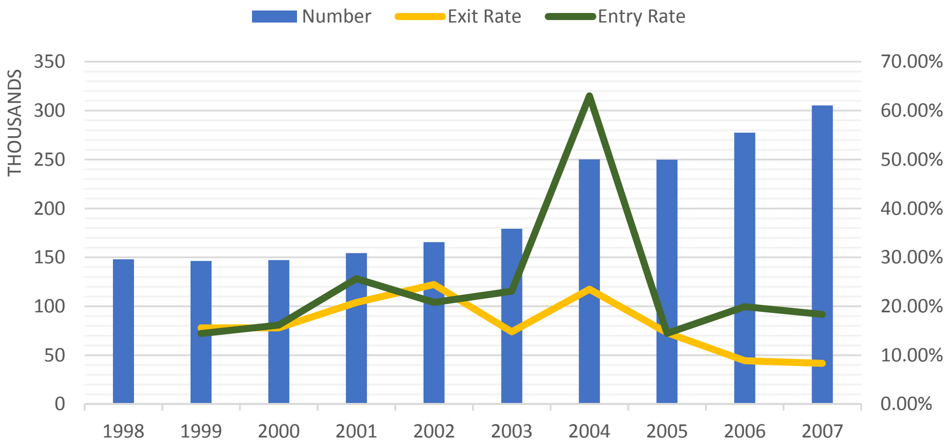
Entry and Exit in CIES With the unbalanced panel constructed in Chapter 2, I can identify firms that enter, exit and stay in the CIES data. Figure 4.2 plots numbers of firms from 1998 to 2007, as well as the entry (green) and exit (yellow) rates. Entry rate averages about 24%, with an exceptionally high rate 60% in 2004.[6] Exit rate is also high and averages 15% over 1998-2007.

These values are much higher than the entry and exit rates in the entire manufacturing sector. According to a survival analysis report by the State Administration for Industry and Commerce (henceforth, the Report), exit rate in China's manufacturing sector is about 8%.[7] While the total number of manufacturing firms

[6] According to Holz (2013), such a high rate could be a result of two facts. First, there is an economic census in 2004 that has better coverage than normal years in the sense that firms who used to under-report their sales are caught in 2004. Second, subsidiary firms who used to report financial statements under one firm are separated in 2004.

[7] See `http://www.saic.gov.cn/zwgk/tjzl/zxtjzl/xxzx/201307/P020130731318661073618.pdf`. I choose to rely on the exit rate in the administrative report, since estimates of the manufacturing exit rate in China are rare and most studies are based on the CIES data. To the best of my knowledge, there is one paper, Yang and Zhang (2009), that provides an alternative estimate of 9%. This exit rate is defined as the fraction of firms who report a closing or a bankrupt status in the Economic Census 2004.

Figure 4.2: Entry and Exit Rate in CIES Data



Numbers are on the left axis. Entry and exit rates are on the right axis.

grows at 8% annually over 2004-2008, the entry rate in manufacturing sector is about 16%. This highlights the importance of distinguishing the notions of birth and closure of firms, from entry and exit in the CIES data. This can also be seen from the age distribution for firms that are not in the CIES in t,[8] and enter in $t+1$. More than 75% of firms are older than 1 year among CIES entrants, and 90% of them are older than 0 year.

Table 4.1: Percentiles of Age Distribution among CIES Entrants

Percentile	10%	25%	50%	75%	90%
Age	0	1	4	9	20

Aggregate productivity decomposition is usually implemented over 5 years (Baily, Hulten, and Campbell, 1992; Foster, Haltiwanger, and Krizan, 2001). Table 4.2 illustrates the entry pattern of CIES data in sub-periods of 1998-2003 and 2002-2007. Over 60% of CIES firms in 2003 and 2007 do not exist in the data 5 years ago, and they produce a market share over 50% (see Table 4.2). Among them, more than a half are older than 5, meaning that these firms already exist and have a sales below 5 million in 1998.

These facts indicate that entry into the CIES data provides a biased estimate of the entry of new firms in the manufacturing sector. As a result, the role of entry on aggregate productivity growth could be biased. On one hand, incumbents who grow

[8] Age_{it} = Year t – Opening year$_{it}$

from sales smaller than 5 million to sales greater than 5 million are misclassified as entrants, and hence bias the entry channel upwards. On the other hand, there are missing new firms in 2003 and 2007 with sales below 5 million and not captured in the CIES data. This biases the entry channel downwards. The former is likely to dominate because CIES entrants with age greater than 5 have a much larger total market share than their counterparts with sales less than 5 million yuan.

Table 4.2: CIES Data Entrants over a 5-Year Horizon

	China		U.S. Average
	1998-2003	2002-2007	1967-1972, 1972-1977, 1977-1982
Number of Firms, End Year			
Data Incumbents	32.11%	24.37%	48.30%
Data Entrants (Age > 5)	29.06%	30.47%	-
Data Entrants (Age ≤5)	38.83%	45.16%	51.7%
Market Share, End Year			
Data Incumbents	50.86%	44.42%	83.10%
Data Entrants (Age > 5)	22.45%	24.85%	-
Data Entrants (Age ≤5)	26.69%	30.73%	16.90%

U.S. numbers are averages of statistics over three periods: 1967-1972, 1972-1977 and 1977-1982 from Dunne, Roberts, and Samuelson (1988).

Table 4.3: CIES Data Exiters over a 5-Year Horizon

	China		U.S. Average
	1998-2003	2002-2007	1967-1972, 1972-1977, 1977-1982
Number of Firms, Beginning Year			
Data Stayers	40.67%	44.34%	52.00%
Data Exiters (Age>5)	36.66%	27.76%	48.00%
Data Exiters (Age≤5)	22.67%	27.90%	
Total Market Share, Beginning Year			
Data Stayers	59.91%	61.23%	82.57%
Data Exiters (Age>5)	24.60%	27.01%	17.43%
Data Exiters (Age≤5)	15.49%	11.76%	

U.S. numbers are averages of statistics over three periods: 1967-1972, 1972-1977 and 1977-1982 from Dunne, Roberts, and Samuelson (1988).

While one can distinguish between new firms and growing incumbents in the CIES, it is impossible to distinguish whether an exit from the CIES is due to closing, merger or a decline in sales. Over 5-year horizons, about 60% of firms exit from CIES data, which produce about 40% of market share at the beginning of 5-year periods (see Table 4.3). Compared to U.S., the number of entrants and exiters in the CIES over 5 years are 1.5 times that in U.S. census data (Dunne, Roberts, and Samuelson, 1988). However, because of the minimum sales, both entrants and exiters in the CIES are more than twice as large as their counterparts in the U.S. census data (Dunne, Roberts, and Samuelson, 1988). Therefore, if there is any mis-classification that treats declining firms as closures, the productivity growth from closing less productive firms is also biased.

Summary This section discusses several features of the CIES data. First, number of manufacturing firms increases at a rate of 8% over 1998-2007, with a declining share of state-owned firms. Second, CIES data only covers top 20% manufacturing firms in the sales distribution. Third, there may exist a bias in extensive reallocation, since growing and declining incumbents that cross the 5 million sales are misclassified as entrants and exiters in the CIES.

4.4 Direct Decomposition in CIES

In this section, I take the CIES data as the manufacturing universe, and decompose the aggregate output-weighted productivity growth into technological growth, intensive reallocation and extensive reallocation, using the approach of Baily, Hulten, and Campbell (1992).

To compare those results in Baily, Hulten, and Campbell (1992) and Foster, Haltiwanger, and Syverson (2008), I divide the 1998-2007 window into two overlapping sub periods: 1998-2003 and 2002-2007. Table 4.4 shows that during these two periods, output-weighted gross output productivity in the CIES increases by 17.61% and 15.57%, respectively. This translates into an annualized rate about 3.54 to 3.20%, and is close to the estimated gross output productivity growth 2.85% for CIES stayers over two consecutive years in Brandt, Van Biesebroeck, and Zhang (2012).

In the decomposition exercise, extensive reallocation through replacing less productive exiters by more productive entrants is the main driving force of aggregate productivity growth in the CIES. During 1998-2003, 93% of aggregate productivity growth is accounted by the extensive reallocation. This number is about 144% in 2002-2007. The second engine of growth is from productivity growth for firms who stay in data over 5-year periods.

Surprisingly, the reallocation of inputs among existing firms contributes negatively to aggregate productivity growth. For both two periods, if I shut down extensive reallocation and technological growth, China's aggregate productivity would decline by 1.77 to 14.53% because of negative intensive reallocations. In other words, among firms that stay in the CIES over 1998-2003 and 2002-2007, inputs are reallocated to less productive firms rather than more productive ones.

Table 4.4: Direct BHC Decomposition in CIES

	Tech. Growth	Inten. Reallocation	Exten. Reallocation	Aggregate Growth
China CIES				
1998-2003	0.0298	-0.0177	0.1641	0.1761
	(16.90%)	(-10.05%)	(93.15%)	(100.00%)
2002-2007	0.0768	-0.1453	0.2242	0.1557
	(49.53%)	(-93.32%)	(143.97%)	(100.00%)
U.S. Census				
1972-1977	0.0504	0.0212	0.0001	0.0717
	(70.29%)	(29.56%)	(0.14%)	(100%)
1977-1982	-0.0109	0.0253	0.0095	0.0239
	(-45.61%)	(106.56%)	(39.75%)	(100%)
1982-1987	0.1352	0.0315	-0.0105	0.1553
	(86.50%)	(20.15%)	(-6.72%)	(100%)

Percentages of contribution in aggregate productivity growth are in parenthesis. U.S. numbers are from Baily, Hulten, and Campbell (1992).

The above results contrast to those in the U.S. manufacturing sector. According to Baily, Hulten, and Campbell (1992), intensive reallocation increases 5-year aggregate productivity growth by 2 to 3% across business cycles (see the U.S. panel in Table 4.4). Extensive reallocation has a negligible role on aggregate productivity growth in the U.S., compared to its dominant role in China. The magnitude of technological growth for the two countries are more comparable than the intensive and extensive reallocations.[9]

The low intensive and high extensive reallocations are not accounted by the existence of state owned firms. Table 4.5 decomposes aggregate productivity growth into technological growth, intensive reallocation, and extensive reallocation both in the state owned sector and the non state owned sector, as well as from transitions between ownerships. Although there are differences in the decomposition of growth between state owned and non state owned sectors, the contribution of intensive reallocation remains small in the non state owned sector. Extensive reallocation and technological growth in non state owned sector are the two major driving forces in aggregate productivity growth.

One may argue that the fast growing number of firms in the CIES contributes to

[9]One exception is 1977-1982 that includes the 1981-1982 recession.

Table 4.5: BHC Decomposition in CIES Data, By Ownership

	1998-2003	2002-2007
Agg. Growth	0.1770	0.1581
SOE Sector		
Tech. Growth	0.0071	0.0106
Inten. Reallocation	-0.0183	-0.0238
Exten. Reallocation	-0.0871	-0.0605
Switchers		
NonSOE to SOE	-0.0006	-0.0006
SOE to NonSOE	0.0027	-0.0015
Non SOE sector		
Tech. Growth	0.0150	0.0598
Inten. Reallocation	0.0067	-0.1124
Exten. Reallocation	0.2626	0.2209

the low intensive and high extensive reallocations in Table 4.4. Indeed, annualized net growth rate of firms in the CIES averages 8%, much higher than the approximate rate of 1% during 1977-1987 in the U.S. manufacturing sector. This may lower intensive reallocation because more entrants take up more market share, and thus dilute the market share of stayers at the end of 5-year horizons. As a result, the intensive reallocation is low and negative because $\sum_i \theta_{it}$ is smaller than $\sum_i \theta_{it-5}$ for CIES stayers, but not because inputs are reallocated to less productive firms. To exclude this possibility, I redo the decomposition if (i) the net entry in CIES is at the U.S. rate of 1%; (ii) the size of CIES entrants and exiters shrink to that of their U.S. counterparts.

In the first experiment, I rank CIES entrants from the highest to lowest sales during 1998-2003 and 2002-2007. I hypothetically treat CIES entrants below a certain threshold sales to be nonexistent, which guarantees a 1% net entry in the CIES.[10] For example, during 1998-2003, there are 105,973 CIES entrants in 2003 and 79,440 CIES exiters. If the net entry rate in CIES is 1% as in the U.S., the net growth in number during 1998-2003 would be the total number of firms in 1998, 143,595, times $(1.01^5 - 1)$, i.e. 22,871 firms. This means that in a scenario of 1% net entry rate, only top 86,765 firms above the sales 7.11 million in the sales distribution are treated as hypothetical entrants (see Table 4.6 Experiment 1 panel).

[10]The rationale is driven by models of endogenous entries, e.g. Hopenhayn (1992) and Melitz (2003), where larger firms enter into production.

Table 4.6: Hypothetical BHC Decomposition in CIES, Controlled for Net Entry Rates and Sizes of Entrants and Exiters

	Experiment 1: 1% Net Entry Rate as in U.S.				
	No. of CIES Entrants	No. of CIES Exiters	Hypothetical No. of CIES Entrants	Threshold Sales (thousand yuan)	Output of Hyp. CIES Entrants / Output of CIES Entrants
1998-2003	105,973	79,440	86,765	7,111	98.48%
2002-2007	216,435	72,625	80,559	33,069	87.44%
	Tech. Growth	Inten. Reallocation	Exten. Reallocation	Aggregate Growth	
1998-2003	0.0266	-0.0104	0.1612	0.1774	
	(15.02%)	(-5.87%)	(90.85%)	(100.00%)	
2002-2007	0.0726	-0.1122	0.1956	0.1561	
	(40.92%)	(-63.20%)	(110.27%)	(100.00%)	

	Experiment 2: Entrants and Exiters Same Relative Size as in U.S.			
	Tech. Growth	Inten. Reallocation	Exten. Reallocation	Aggregate Growth
1998-2003	0.0309	0.0334	0.0999	0.1642
	(21.32%)	(23.05%)	(68.91%)	(100.00%)
2002-2007	0.0827	-0.0720	0.1342	0.1449
	(57.03%)	(-49.64%)	(92.62%)	(100.00%)

Percentages of contribution to aggregate productivity growth are in parenthesis.

These firms produce 98.48% of output among CIES entrants. The top panel in Table 4.6 suggests that the low intensive and high extensive reallocations are robust to different net entry rates between China and U.S.. In both periods, the intensive reallocation increases slightly but is still negative. In other words, there is not much diluting effect on market share at least induced by more entries in the CIES.

The second is to control for the market share differences between CIES entrants and those in U.S. census data. I hypothetically shrink the size of CIES entrants and exiters to that of U.S. entrants and exiters. In particular, the relative size of U.S. entrants and exiters is 0.21 and 0.23 (Dunne, Roberts, and Samuelson, 1988), compared to existing firms. If CIES entrants and exiters had the same relative sizes as in U.S. entrants, total market share of CIES entrants would be 25.96% in 2003 and 26.37% in 2007. According to Table 4.2 and Table 4.3, each CIES entrant shall shrink its output to 53% and 47% of the observed level in 2003 and 2007, respectively. Similarly, each exiter shall shrink its output to 63% and 58% of its observed level in 1998 and 2002, respectively.

The lower panel of Table 4.6 lists the result of this experiment. It suggests that there is a larger market share diluting effect because average entrant and exiter are large. Intensive reallocation would increase to 0.0334 in 1998-2003, and -0.07 in 2002-2007, if CIES entrants and exiters have the same relative size as their U.S. counterparts. Notice that in this exercise, all CIES entrants and exiters, regardless of their ages, shrink in output. Thus, this puts an upper bound on the impact of size differences between CIES and U.S. entries and exits, since the correct thought experiment shall only shrink entrants younger than 5 years old and exiters that close if they could be identified from declining firms. Given this limiting case, the low intensive and high extensive reallocations are not likely caused by the market share dilution effect on stayers induced by larger entrants and exiters in the CIES.

Summary Taking the CIES data as the manufacturing universe, this section finds a low intensive reallocation among existing firms and a high extensive reallocation through net entry in China. This contrasts to the stylized facts of U.S., where intensive reallocation is the main contributor of aggregate productivity growth.

This section also illustrates that the China-U.S. difference is not induced by the diluting effect from more and larger CIES entrants that crowd out the market share of stayers at the end of 5-year horizons. The canonical explanation by the existence of state owned firms also cannot account for the similar low intensive and high intensive reallocations within non state owned sector.

4.5 BHC Decomposition Corrected for the Cutoff Sales

Because of the minimum sales in the CIES data, 80% of manufacturing firms in China are not included in the aggregate productivity decomposition in Section 4.4. This could mistakenly attribute some of the reallocation among existing firms to net entry in the manufacturing sector, and hence bias the source of aggregate productivity growth.

To quantify how large this bias is, this section first discusses the method to correct the left-censoring problem and then repeats the aggregate productivity decomposition. Second, I present the corrected decomposition results using the CIES data and aggregate information in China Economic Censuses 2004 and 2008, and compare to results in Section 4.4.

4.5.1 Methodology to Correct the Left-censoring

To correct the left-censoring problem, one needs to impute how productivity co-moves with market share, as well as how productivity and market share change over time for firms that are unobserved in the CIES. A dynamic firm-level model (e.g. one based on Cooper and Haltiwanger, 2006) could serve this purpose. This subsection outlines an alternative approach that builds on the fact that the BHC decomposition relies on a few key moments in market share and productivity.

To begin with, I rewrite technological growth, intensive reallocations and extensive reallocations into parts that are observed and unobserved in the CIES data separately. Following the same methodology as in Equation (4.1) for a 5-year horizon, technological growth in the manufacturing sector can be decomposed as:

$$\begin{aligned}
&\sum_{y_{it}\neq .,y_{it-5}\neq .} \theta_{it-5}(z_{it} - z_{it-5}) \\
&= \sum_{y_{it}\geq y_c, y_{it-5}\geq y_c} \theta_{it-5}(z_{it} - z_{it-5}) - \sum_{y_{it}<y_c, y_{it-5}\geq y_c} \theta_{it-5} z_{it-5} + \sum_{y_{it}<y_c, y_{it-5}\geq y_c} \theta_{it-5} z_{it} \\
&+ \sum_{y_{it}\neq ., y_{it-5}<y_c} \theta_{it-5} z_{it} - \sum_{y_{it}\neq ., y_{it-5}<y_c} \theta_{it-5} z_{it-5}
\end{aligned} \tag{4.4}$$

where $y_{it} \neq .$ denotes that firm i operates in time t, while $y_{it} < y_c$ ($y_{it} > y_c$) denotes that firm i operates under (above) minimum sales y_c. Note Equation (4.4) decomposes three sources of technological growth: firms observed in the CIES at $t-5$ and t (first term), and firms who decline in sales and fall out of the CIES over $t-5$ to t (second and third terms), and firms that are below minimum sales in $t-5$ (fourth and fifth terms). Only the first and second terms are observable in the CIES.

Similar to technological growth, intensive and extensive reallocations in the manufacturing sector are re-written in the following Equation (4.5) and Equation (4.6)

$$\begin{aligned}
&\sum_{y_{it}\neq .,y_{it-5}\neq .} (\theta_{it} - \theta_{it-5}) z_{it} \\
&= \sum_{y_{it}\geq y_c, y_{it-5}\geq y_c} (\theta_{it} - \theta_{it-5}) z_{it} + \sum_{y_{it}\geq y_c, y_{it-5}<y_c} \theta_{it} z_{it} \\
&- \sum_{y_{it}\neq ., y_{it-5}<y_c} \theta_{it-5} z_{it} + \sum_{y_{it}<y_c, y_{it-5}\neq .} \theta_{it} z_{it} - \sum_{y_{it}<y_c, y_{it-5}\geq y_c} \theta_{it-5} z_{it}
\end{aligned} \tag{4.5}$$

$$\begin{aligned}
&\sum_{y_{it}\neq .,y_{it-5}= .} \theta_{it} z_{it} - \sum_{y_{it}= .,y_{it}\neq .} \theta_{it-5} z_{it-5} \\
&= \sum_{y_{it}> y_c, y_{it-5}= .} \theta_{it} z_{it} + \sum_{y_{it}< y_c, y_{it-5}= .} \theta_{it} z_{it} - \sum_{y_{it}= ., y_{it-5}\geq y_c} \theta_{it-5} z_{it-5} \\
&- \sum_{y_{it}= ., y_{it-5}<y_c} \theta_{it-5} z_{it-5}
\end{aligned} \tag{4.6}$$

To compute equations (4.4), (4.5) and (4.6), one needs to impute moments of $\sum_{y_{it}<y_c, y_{it-5}\geq y_c} \theta_{it-5} z_{it}$, $\sum_{y_{it}<y_c} \theta_{it} z_{it}$, $\sum_{y_{it}<y_c, y_{it-5}=.} \theta_{it} z_{it}$, and $\sum_{y_{it}\neq ., y_{it-5}<y_c} \theta_{it-5} z_{it}$. I argue that these moments and consequently the BHC decomposition in manufacturing sector can be imputed under several assumptions about productivity, exiting behavior and firm size distribution.

There are four assumptions needed for the imputation process I use.

Assumption 4.5.1. *The deviation of firm-level productivity z_{it} from economy-wide average $\{\mu_t\}_{t=1998}^{2007}$, $z_{it} - \mu_t$, follows an AR(1) productivity process:*

$$z_{it+1} - \mu_{t+1} = \rho(z_{it} - \mu_t) + \epsilon_{it+1} \tag{4.7}$$

Assumption 4.5.2. *Firms close exogenously with an annual rate χ. The rate is χ_{cies} for firms above the minimum sales y_c, and χ_{below} for firms below the minimum sales*

Under Assumption 4.5.1 and Assumption 4.5.2, firms whose sales decline and exit from the CIES from $t-5$ to t have

$$\begin{aligned}
\sum_{y_{it}<y_c, y_{it-5}\geq y_c} \theta_{it-5} z_{it} &= \rho^5 \sum_{y_{it}<y_c, y_{it-5}\geq y_c} \theta_{it-5} z_{it-5} + \sum_{y_{it}<y_c, y_{it-5}\geq y_c} \theta_{it}\{(\mu_t - \rho^5 \mu_{t-5}) + \epsilon_t + \ldots + \rho^4 \epsilon_{it-4}\} \\
&= \rho^5 (1-\chi_{cies})^5 \sum_{y_{it-5}\geq y_c} \theta_{it-5} z_{it-5} + \rho^5 (1-\chi_{cies})^5 \sum_{y_{it-5}\geq y_c} \theta_{it}(\mu_t - \rho^5 \mu_{t-5}) \\
&+ \sum_{y_{it}<y_c, y_{it-5}\geq y_c} \theta_{it}\{\epsilon_t + \ldots + \rho^4 \epsilon_{it-4}\}
\end{aligned} \tag{4.8}$$

, while firms that have sales less than 5 million at $t-5$ and keep operating at t have

$$\begin{aligned}
\sum_{y_{it}\neq ., y_{it-5}<y_c} \theta_{it-5} z_{it} &= \sum_{y_{it}\neq ., y_{it-5}<y_c} \theta_{it-5}(\rho^5 z_{it-5} + (\mu_t - \rho^5 \mu_{t-5}) + \epsilon_t + \ldots + \rho^4 \epsilon_{it-4}) \\
&= \rho^5 (1-\chi_{below})^5 \sum_{y_{it-5}<y_c} \theta_{it-5} z_{it-5} + \rho^5 (1-\chi_{below})^5 \sum_{y_{it-5}<y_c} \theta_{t-5}(\mu_t - \rho^5 \mu_{t-5})
\end{aligned} \tag{4.9}$$

Equations (4.8) and (4.9) show that it is important to consistently estimate the productivity persistence ρ, average firm-level productivity $\{\mu_t\}_{t=1998}^{2007}$, and shocks $\{\epsilon_{is}\}_{s=t-4}^{t}$, for computing technological growth, intensive and extensive reallocations. In particular, for firms that are observed in $t-5$ and exit the CIES in t, it is crucial to quantify the unobserved shocks $\{\epsilon_{is}\}_{s=t-4}^{t}$ when z_{it} are unobserved. If these firms exit the data because of low productivities, i.e. negative ϵs, this suggests more intensive reallocation in the manufacturing sector compared to that in the CIES. Vice versa.

Computing equations (4.8) and (4.9) further needs the sum of contemporaneous product between market share and productivity. This sum of contemporaneous product $\sum_{y_{it}<y_c} \theta_{it} z_{it}$ can be rewritten as:

$$
\begin{aligned}
&= \sum_{y_{it}<y_c} (\theta_{it} - \bar{\theta}_t^s)(z_{it} - \bar{z}_t^s) + N^s \bar{\theta}_t^s \bar{z}_t^s \\
&= \frac{N_t^s \rho_{\theta z,t}^s \sigma(y_t^s)\sigma(z_t^s)}{Y_t} + N_t^s \bar{\theta}_t^s \bar{z}_t^s
\end{aligned} \tag{4.10}
$$

where $\bar{\theta}_t^s$ ($\bar{z}_t^s$) is the average market share (productivity) for N_t^s below cutoff firms, $\sigma(y_t^s)(\sigma(z_t^s))$ is the standard deviation of sales (productivity) for below cutoff firms, Y_t is total sales in manufacturing sector, and $\rho_{\theta z,t}^s$ is the correlation between market share and productivity for firms below the minimum sales. Since $\rho_{\theta z,t}^s$ is unobservable, I approximate it by the correlation between market share and productivity for firms in the CIES, i.e.

Assumption 4.5.3. *Correlation between productivity z and market share θ, $\rho_{\theta z}$, is constant over the distribution of sales y*

Under Assumption 4.5.3, there are $\sigma(y_t^s)$, $\bar{z}_t^s$ and $\sigma(z_t^s)$ in Equation (4.10) that remains to be estimated, while $\bar{\theta}_t^s$, N_t^s, and Y_t could be obtained from Census 2004 data. To compute $\sigma(y_t^s)$, I impose the fourth assumption

Assumption 4.5.4. *Log sales $log(y)$ follows a Pareto distribution with density function f*

$$
f(log(y)) = \begin{cases} \frac{\alpha(log(y)_{\min})^\alpha}{(log(y))^{\alpha+1}} & \textit{if } log(y) \geq log(y)_{\min}, \\ 0 & \textit{if } log(y) < log(y)_{\min}. \end{cases}
$$

where $log(y)_{min}$ is the minimum log sales, and α is the Pareto shape.

The mean productivity for firms below the minimum sales, $\bar{z}_t^s$, is

$$
\bar{z}_t^s = \mu_t - (1 - \frac{N_t^s}{N_t})\bar{z}_t^{cies} \tag{4.11}
$$

and the standard deviation $\sigma(z_t^s)$ is approximated by that in the CIES data. [11]

There are three steps to further compute the BHC decomposition. First, I estimate the AR(1) productivity process to back out consistent estimators of ρ, $\{\mu_t\}_{t=1998}^{2007}$

[11] This approximation is quantitatively unimportant, since the first term in Equation (4.10) is in a much smaller magnitude than the second. This would be illustrated latter.

and consequently $\{\bar{z}_t^s\}_{t=1998}^{2007}$. Second, I calibrate the parameters, α and $log(y)_{min}$, for Pareto distribution of log sales to match the mean and the variance of log sales in the CIES so that $\{\sigma(y_t^s)\}_{t=1998}^{2007}$ can be quantified. Lastly, I compute the correlation between market share and productivity in the CIES, $\{\rho_{\theta z,t}\}_{t=1998}^{2007}$.

Estimating Productivity Process Due to the minimum sales, there is selection of firms that remain in the CIES over two consecutive years. In other words, the expected value of ϵ_{t+1} for firms that stay in data from t to $t+1$ could be non-zero. This is important to quantify, since whether high ϵ_{t+1} firms stay and low ϵ_{t+1} ones leave the CIES represents some efficiency in intensive reallocation. Therefore, OLS estimation that assumes the zero expected shocks for staying firms is not appropriate here.

To overcome this selection problem, I employ the following 2-step estimation. Suppose firm i is observed in the CIES at t and $t+1$. Formally, selection shows up as

$$\begin{aligned} y_{it+1} &= F(exp(z_{it+1}), k_{it+1}, m_{it+1}, l_{it+1}) > y_c \\ &= F(exp(\rho z_{it} + \mu_{t+1} - \rho\mu_t + \epsilon_{it+1}), k_{it} + \Delta k_{it+1}, m_{it} + \Delta m_{it+1}, l_{it} + \Delta l_{it+1}) > y_c \end{aligned}$$

Changes of inputs $\Delta k_{it+1}, \Delta m_{it+1}$, and Δl_{it+1}, could be a function of current level of inputs k_{it}, m_{it}, l_{it}, and current productivity z_{it}, future innovations in productivity ϵ_{it+1}, as well as other unknown shocks.

Whether $E(\epsilon_{it+1}|y_{it+1} > y_c)$ is greater or smaller than 0 is not obvious. On one hand, a higher shock ϵ_{it+1} increases productivity and sales at $t+1$. This induces a positive $E(\epsilon_{it+1}|y_{it+1} > y_c)$. On the other hand, firms with high levels of capital stock k_{it}, intermediate goods m_{it} and labor l_{it}, are more likely to maintain sales above the cutoff regardless of a low innovation ϵ_{it+1}.[12] Thus, larger firms are more likely to stay than small firms with the same shock, implying a negative $E(\epsilon_{it+1}|y_{it+1} > y_c)$.

This leads to an agnostic specification

$$log(y)_{it+1} = \phi(log(z)_{it}, log(k)_{it}, log(m)_{it}, log(l)_{it}) + \nu_{it+1}(\epsilon_{it+1}) \tag{4.12}$$

where $\phi(\cdot)$ is a polynomial of log productivity $log(z)_{it}$, log capital $log(k)_{it}$, log intermediate goods $log(m)_{it}$ and log labor $log(l)_{it}$, and ν_{it+1} is a (nonlinear) function of ϵ_{it+1}.

Let $\tilde{\nu}_{it+1}$ be the demeaned ν_{t+1}. Since $\tilde{\nu}_{it+1}$ contains information of ϵ_{it+1}, I proxy

[12] This argument hinges on some quasi-fixed property of inputs. If inputs are fully flexible, sales y_{it+1} would be an increasing function of ϵ_{it+1} only.

$E(\epsilon_{it+1}|y_{it+1} > y_c)$ by polynomials of $\tilde{\nu}_{it+1}$, i.e.

$$E(\epsilon_{it+1}|y_{it+1} > y_c) = E(\epsilon_{it+1}|\tilde{\nu}_{it+1} > \bar{\tilde{\nu}}(k_{it}, m_{it}, l_{it}, z_{it})) = \alpha_1\tilde{\nu}_{it+1} + \alpha_2(\tilde{\nu}^2_{it+1} - E(\tilde{\nu}^2_{it+1})) \\ + ... + \alpha^n(\tilde{\nu}^n_{it+1} - E(\tilde{\nu}^n_{it+1}))$$

where $\bar{\tilde{\nu}}(k_{it}, m_{it}, l_{it}, z_{it})$ is the threshold residuals for staying in the CIES at $t+1$. Note the above polynomial specification guarantees that $E(\epsilon_{it+1}|\tilde{\nu}_{it+1} > \bar{\tilde{\nu}}(k_{it}, m_{it}, l_{it}, z_{it})) = 0$ when $\bar{\tilde{\nu}}(k_{it}, m_{it}, l_{it}, z_{it}) = -\infty$, which is the no selection case.

Following the above rationale, I first run a Probit model: regressing an indicator of firms observed at t and $t+1$, $\mathbb{1}(y_{it+1} \geq y_c, y_{it} \geq y_c)$, on log sales $log(y)_{it}$, log capital $log(k)_{it}$, log intermediate goods $log(m)_{it}$, and log wage bill $log(l)_{it}$, as well as an indicator for state owned firms SOE_{it}.[13] Residuals from the Probit model $\hat{\nu}_{it}$ and its polynomial are then used as the proxy for $E(\epsilon_{t+1}|y_{it+1} > y_c)$ to augment the AR(1) estimation. Specifically, I run the following regression

$$z_{t+1} = \rho z_t + (\mu_{t+1} - \rho\mu_t) + E(\epsilon_{t+1}|\tilde{\nu}_{it+1} > \bar{\tilde{\nu}}_{it+1}) + \\ \{\epsilon_{it+1} - E(\epsilon_{t+1}|\tilde{\nu}_{it+1} > \bar{\tilde{\nu}}_{it+1})\} \tag{4.13}$$

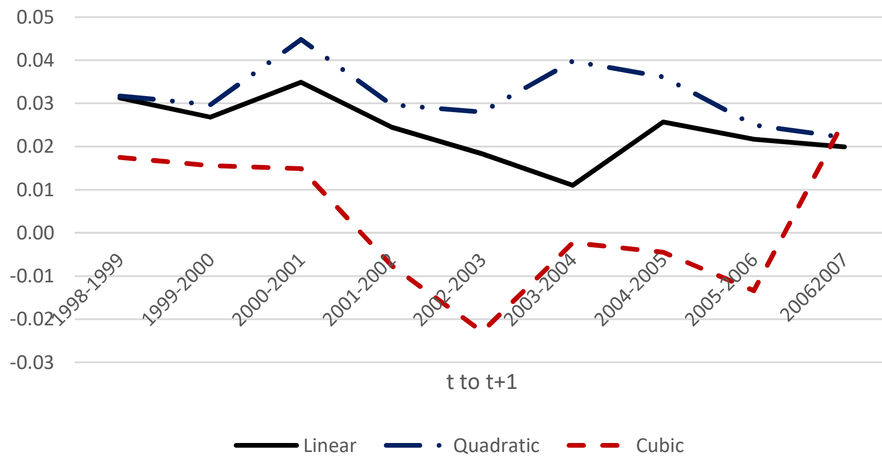
where $E(\epsilon_{t+1}|y_{it+1} > y_c)$ is set to (i) $\alpha_1\tilde{\nu}_{it}$ for a linear approximation; (ii) $\alpha_1\tilde{\nu}_{it} + \alpha_2(\tilde{\nu}^2_{it} - E(\tilde{\nu}^2_{it}))$ for a quadratic approximation; (iii) $\alpha_1\tilde{\nu}_{it} + \alpha_2(\tilde{\nu}^2_{it} - E(\tilde{\nu}^2_{it})) + \alpha_3(\tilde{\nu}^3_{it} - E(\tilde{\nu}^3_{it}))$ for a cubic approximation.

Table 4.7 reports the result for the productivity process estimation. Under all specifications of $E(\epsilon_{t+1}|y_{it+1} > y_c)$, the persistence of productivity ρ is about 0.664 and slightly higher than that obtained from OLS estimation, 0.652. The average shock ϵ_{it+1} for CIES stayers is positive for all years under linear and quadratic specifications of $E(\epsilon_{t+1}|y_{it+1} > y_c)$, and for some years under cubic specification (See Figure 4.3). This indicates that firms that stay in the CIES data on average have a higher productivity level z_{it} than those who exit. For the rest of the paper, I use quadratic specification of $E(\epsilon_{t+1}|y_{it+1} > y_c)$.

The estimated shocks $\hat{\epsilon}_{it+1}$ for those observable CIES stayers help quantify the shocks of unobserved CIES exiters who remain production. Specifically, by Assumption 4.5.2, firms who survive to operate in t and $t+1$, below or above cutoff

[13]See Table C.1 in appendix for estimation results, as well as the Probit results when the dependent variables also include quadratic terms in $logy_{it}$, $logk_{it}$, $logl_{it}$ and $logm_{it}$ and their cross terms.

Figure 4.3: Average Productivity Shock $\hat{\epsilon}_{it+1}$ Conditional on Staying, t to $t+1$



Linear/ quadratic/ cubic means that $E(\epsilon_{it+1}|y_{it+1} > y_c)$ is approximated by a linear/ quadratic/ cubic function of Probit residual v_{it+1}.

sales, have mean zero shocks ϵ_{it+1} (unweighted and market share weighted), i.e.

$$\sum_{y_{it+1}\neq .,y_{it}>y_c} \epsilon_{it+1} = 0 \quad \text{and} \quad \sum_{y_{it+1}\neq .,y_{it}>y_c} \theta_{it}\epsilon_{it+1} = 0$$

Therefore, for firms who exit from the CIES data and decline in sales, their average productivity shocks, for both unweighted and market share weighted:

$$\sum_{y_{it+1}<y_c,y_{it}>y_c} \epsilon_{it+1} = -\sum_{y_{it+1}\geq y_c,y_{it}>y_c} \epsilon_{it+1} \quad \text{and} \quad \sum_{y_{it+1}<y_c,y_{it}>y_c} \theta_{it}\epsilon_{it+1} = -\sum_{y_{it+1}\geq y_c,y_{it}>y_c} \theta_{it}\epsilon_{it+1}$$

By this property, the second and third parts of Equation (4.8) equals to

$$\rho^5(1-\chi_{cies})^5 \sum_{y_{it-5}\geq y_c} \theta_{it-5}\{(\mu_t - \rho^5\mu_{t-5})\} - \sum_{y_{it}>y_c,y_{it-5}\geq y_c} \theta_{it-5}\{\epsilon_t + \ldots + \rho^4\epsilon_{it-4}\}$$

Since $\mu_t - \rho^5\mu_{t-5} = \mu_t - \rho\mu_{t-1} - \rho(\mu_{t-1} - \rho\mu_{t-2}) - \ldots - \rho^4(\mu_{t-4} - \rho\mu_{t-5})$ and $\mu_t - \rho\mu_{t-1}$ can be estimated from the productivity process estimation, $\sum_{y_{it}<y_c,y_{it-5}\geq y_c} \theta_{it-5}z_{it}$ can be computed as well as its impact on technological growth and intensive reallocation.

Table 4.7: Estimating AR(1) Productivity Process

	Linear	Quadratic	Cubic
Lag Productivity	**0.665*****	**0.664*****	**0.663*****
	(992.41)	(991.44)	(986.96)
Probit Residual			
1999	0.210***	0.204***	0.292***
	(32.84)	(22.08)	(12.35)
2000	0.187***	0.142***	0.175***
	(28.10)	(14.46)	(7.05)
2001	0.183***	0.0877***	0.106***
	(29.37)	(10.54)	(4.16)
2002	0.194***	0.0627***	0.175***
	(32.31)	(6.20)	(7.44)
2003	0.146***	-0.0358***	0.134***
	(25.19)	(-4.14)	(6.83)
2004	0.0560***	0.0704***	0.266***
	(7.09)	(8.88)	(13.00)
2005	0.178***	0.0943***	0.226***
	(32.40)	(13.75)	(14.67)
2006	0.253***	0.150***	0.307***
	(38.29)	(13.37)	(14.84)
2007	0.261***	0.209***	0.140***
	(38.72)	(21.01)	(7.91)
Probit Residual Square			
1999		-0.0153	-0.139***
		(-1.02)	(-4.09)
2000		-0.0980***	-0.145***
		(-6.32)	(-4.12)
2001		-0.254***	-0.270***
		(-17.23)	(-10.53)
2002		-0.245***	-0.413***
		(-16.14)	(-11.74)
2003		-0.371***	-0.579***
		(-28.45)	(-23.02)
2004		-0.242***	-0.367***
		(-16.24)	(-19.11)
2005		-0.235***	-0.410***
		(-20.18)	(-18.86)
2006		-0.179***	-0.550***
		(-11.48)	(-12.52)
2007		-0.101***	0.0638
		(-7.12)	(1.69)

	Linear(cont'd)	Quadratic(cont'd)	Cubic(cont'd)
Probit Residual Cubic			
1999			-0.304***
			(-4.06)
2000			-0.113
			(-1.49)
2001			-0.0550
			(-0.75)
2002			-0.379***
			(-5.30)
2003			-0.518***
			(-9.64)
2004			-0.582***
			(-10.35)
2005			-0.452***
			(-9.54)
2006			-0.646***
			(-9.03)
2007			0.284***
			(4.71)
Constant	0.148***	0.148***	0.149***
	(110.99)	(104.88)	(105.16)
N	1399406	1399406	1399406

t statistics in parentheses

="* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$"

	Linear	Quadratic	Cubic
$corr(\epsilon_{it+1}, z_{it} \mid y_{it} > y_c, y_{it+1>y_c})$	**-0.1158***	**-0.0974***	**-0.0109***

	OLS
Lag Productivity	**0.652*****
	(976.53)
Constant	0.183***
	(170.29)
Year FE	Y

Another result of the estimation of productivity process is to back out average productivity in the manufacturing sector, $\{\mu_t\}_{t=1998}^{2007}$, and consequently average productivity for firms with sales less than 5 million yuan, $\{\bar{z}_t^s\}_{t=1998}^{2007}$ according to Equation (4.11). Unfortunately, only consistent estimators of $\{\mu_{t+1} - \rho\mu_t\}_{t=1998}^{2006}$ are attainable in Equation (4.13). To estimate $\{\mu_t\}_{t=1998}^{2007}$, I assume that average productivity $\{\mu_t\}_{t=1998}^{2007}$ grows at the same rate as that in the CIES $\{\bar{z}_t^{cies}\}_{t=1998}^{2007}$, which equals to 3.49% on average during 1998-2007. Results of estimated $\{\mu_t\}_{t=1998}^{2007}$ and $\{\bar{z}_t^s\}_{t=1998}^{2007}$ are reported in Table 4.8. It turns out that during 1998-2007, average productivity in the CIES is 15% higher than that in the manufacturing sector, and 33% than that among firms with sales less than 5 million yuan.

Table 4.8: Average Productivity for Firms whose Sales Smaller than 5 Million Cutoff

Year	$\mu_{t+1} - \rho\mu_t$	Average Productivity in CIES $\bar{z}_t^{cies}$	Average Manu. Productivity μ_t	Average Productivity for Firms below 5 Million $\bar{z}_t^s$
1998	0.1710	0.5384	0.4050	0.3716
1999	0.1820	0.5586	0.4395	0.4097
2000	0.1489	0.6177	0.4734	0.4373
2001	0.1649	0.6488	0.4627	0.4162
2002	0.2304	0.6685	0.4717	0.4225
2003	0.2844	0.7335	0.5432	0.4956
2004	0.3022	0.7825	0.6446	0.6101
2005	0.2114	0.8174	0.7296	0.7076
2006	0.2931	0.8593	0.6950	0.6540
2007		0.8877	0.7539	0.7204
Average Growth Rate		0.0349	0.0349	0.0349

Pareto Distribution of Log Sales and Size-Productivity Correlation In order to impute the co-movement between market share and productivity for firms below the cutoff sales, I impose a Pareto distribution structure on log sales for all manufacturing firms. I then calculate the correlation coefficient between market share (size) and productivity for CIES firms, which equals to its counterpart for firms with sales less than 5 million yuan by Assumption 4.5.3.

There are two distributions to be calibrated. One for all manufacturing firms, and the other for firms with age smaller than 5. While the first is used for intensive reallocation, the latter is for extensive reallocation.

For each distribution, there are two parameters, the Pareto shape α and minimum log sales $log(y)_{min}$, which are calibrated to match the average log sales and its fraction of firms with sales above 5 million yuan. According to the CIES data and 2004 Census information, average log sales is 9.8792 in the CIES and the fraction

Table 4.9: Calibrating *Logy* Distribution, All Firms and Firms with Age< 5

	Moments		Parameters		Implied S.D. of Sales $< y_c$
	Average Log Sales	Frac. Above y_c	α	lny_{min}	$\sigma(y^s)$
All Firms	9.8792	79.76%	7.2533	6.8223	964.60
Age < 5 Firms	9.7747	45.63%	10.1781	6.7802	872.21

of all firms above the cutoffs sales is 80%. For firms younger than 5 years old in the CIES, average log sales is 9.7747. Since the 2004 census does not give total number of manufacturing firms with age smaller than 5, I impute the number through multiplying the total number of firms in 1998 and 2002 by 0.08^5, with a 16% birth rate and an 8% closure rate annually. The fraction of firms in the CIES is therefore the number of firms with age smaller than 5 in the CIES divided by the imputed total number.

Table 4.9 shows the calibration results. For young firms, the distribution shows a fatter left tail, indicating that there are far more small firms for younger cohorts. Although there is a small difference between the two minimum sales (4.21%), the two minimum log sales are roughly the same. Given these parameters, the estimate for the sales dispersion among unobserved firms with sales less than 5 million is 964.60 thousand yuan for all firms, and 872.21 thousand yuan for firms younger than 5 years old. These two standard deviations are used in calculating Equation (4.10).

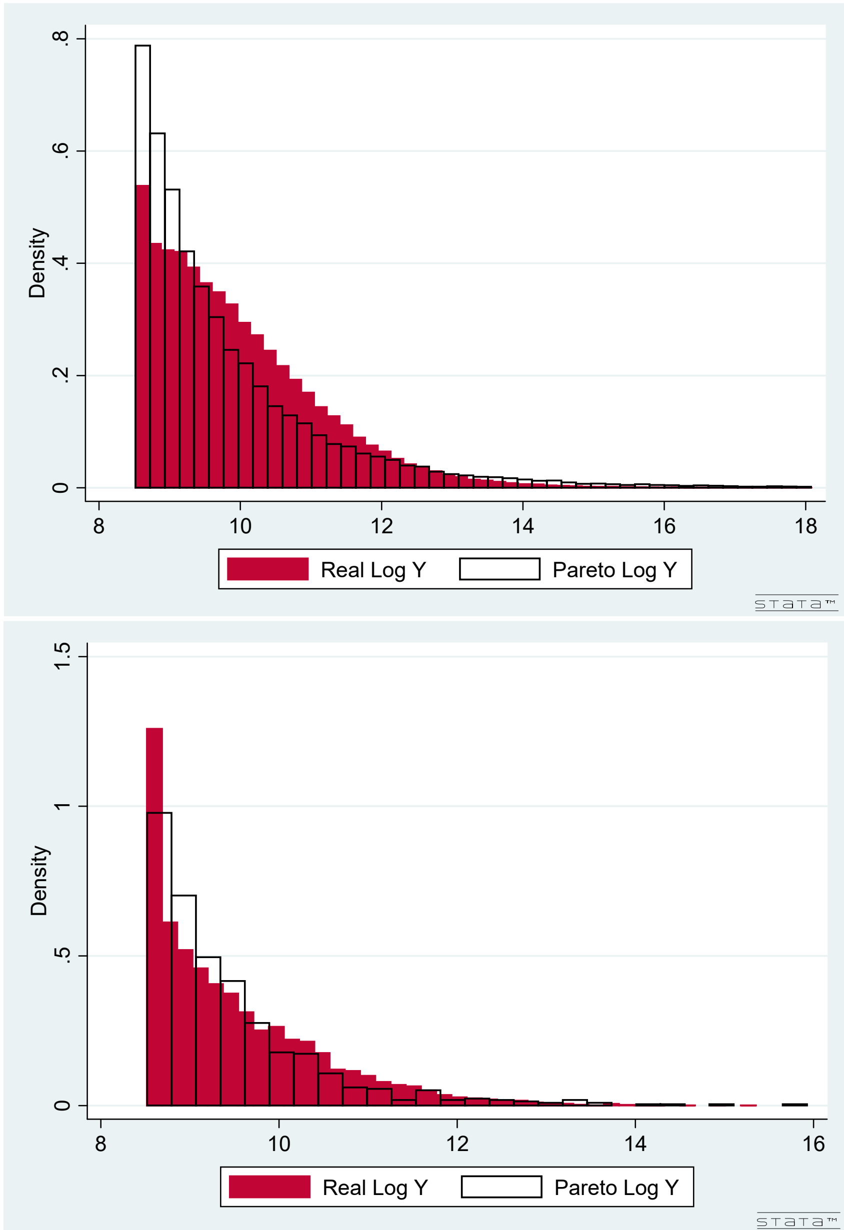
Out of model fit suggests that the Pareto distributions match the average sales of below cutoff firms and the truncated CIES log sales distributions well. In particular, the fitted distribution for all firms generate an average sales of 1.87 million for below cutoff firms, which is close to 1.86 million yuan according to the 2004 census.

Further, I compare the histogram of log sales from the truncated Pareto distribution to that in the CIES. Specifically, the Pareto is truncated at Log (5000) with the same scale as in the CIES. Figure 4.4 suggests that the Pareto distribution (transparent bars) fits the CIES data (pink bars) well, especially for firms with age smaller than 5. For all firms, there is a larger mass near Log (5000) in the Pareto distribution than that in the CIES.[14]

The next parameter for computing $\sum_{y_{it-5}<y_c} \theta_{it-5} z_{it-5}$ (recall Equation 4.10) is the

[14]One explanation could be that firms tend to under-report their sales as suggested in Holz (2013). I do find evidence along this line. In particular, if I plot histograms year by year, the closest match between the real log y and the Pareto log y is in year 2004 when there is a better coverage because of the census.

Figure 4.4: Histogram Fit for Truncated Pareto Distribution vs CIES Data, All (Top), Age< 5 (Bottom), All Years



correlation between market share (size) and productivity, $\rho_{\theta z,t}$. Once again, there are two correlation coefficients needed, one for all firms in computing technological growth and intensive reallocation, and the other for firms younger than 5 years old in computing extensive reallocation.

Table 4.10 suggests that there is a very small correlation between market share and productivity, consistent with the misallocation story told in Hsieh and Klenow (2009) in China's data. Overall, the contemporaneous sum of productivity and market share are about 0.352 to 0.535 for all firms, and 0.0086 to 0.0133 for firms younger than 5 years old.In particular, the first covariance component for firms outside of CIES $\sum_{y_{it-5}<y_c} \theta_{it-5} z_{it-5}$, $\sum_{y_{it}<y_c} (\theta_{it} - \bar{\theta}_t^s)(z_{it} - \bar{z}_t^s)$, turns out to be fairly small and averages 8.87*10^{-6}, if I assume that standard deviation of productivity of these firms are the same as in the CIES. This magnitude is much smaller than the overall sum of products between market share and productivity for firms with sales less than 5 million. Thus, by looking at Equation (4.10), even if the correlation $\rho_{\theta z,t}$ or the standard deviation of productivity σ_t^s increases, say, by 10 folds, the fourth and the fifth columns are unlikely to change.

Table 4.10: Correlation between Market Share and Productivity in CIES, and Sum of Products of Market Share and Productivity for Firms Smaller than 5 Million Sales

	Correlation		Sum of Products	
	All Firms	Age<5	All Firms	Age<5
1998	0.0185	0.0284	0.0352	0.0086
1999	0.0203	0.0270	0.0325	0.0072
2000	0.0166	0.0082	0.0338	0.0072
2001	0.0115	0.0012	0.0317	0.0078
2002	0.0110	0.0100	0.0337	0.0081
2003	0.0088	0.0145	0.0383	0.0095
2004	0.0069	0.0123	0.0394	0.0125
2005	0.0052	0.0244	0.0443	0.0123
2006	0.0007	0.0181	0.0497	0.0133
2007	-0.0052	0.0077	0.0535	0.0133

Closure Rate for CIES Firms According to Assumption 4.5.2, firms above the minimum sales close at a rate χ_{cies}. Unfortunately, the closure rate χ_{cies} is not observable, since exiters in the CIES include both firms who close and those whose sales go below 5 million yuan.

To get around this problem, I try several different levels for χ_{cies}, 0%, 2%, 4%, 6% and 8%. The upper bound 8% is the closure rate in the manufacturing sector. According to the aforementioned survival report, the closure rate equals to 2% for firms above 10 million sales, and increases to 4% for firms between 1 to 10 million sales. Since half of the CIES firms have sales above 10 million, I choose 2% as the benchmark case for the corrected BHC decomposition.

Note that the closure rate among firms whose sales are smaller than 5 million

χ_{below} is

$$\chi_{below} = \frac{8\% - \text{Share of CIES Firms} \times \chi_{cies}}{1 - \text{Share of CIES Firms}}$$

4.5.2 Corrected BHC Decomposition Results

Table 4.11 presents the benchmark results. For both periods of 1998-2003 and 2002-2007, the corrected extensive reallocation is 2 to 6 percentage points lower than that directly computed in the CIES, while the corrected intensive reallocation is 6 to 7 percentage points higher than that in the CIES. This implies overstated extensive reallocation when the CIES is treated as the manufacturing universe. The opposite happens to intensive reallocation. In particular, the contribution of extensive reallocation drops from 93% to 74% in 1998-2003, and from 144% to 86% in 2002-2007. Most of the drops are picking up by intensive reallocation in both periods.

For technological growth, the change of direction depends on the period examined. In 1998-2003, corrected technological growth drops to almost 0, and contributes only 2% to aggregate productivity growth. The opposite happens in 2002-2007. Technological growth increases to 0.10. The difference between the two periods comes from the difference in productivity changes of declining firms who exit from the CIES. It can be clearly seen in Figure 4.5 that declining incumbents experience more declines in productivity and thus contribute negatively to technological growth in 1998-2003. This also suggests that there is a larger mis-attributed intensive reallocation from the exiting channel in the CIES during 1998-2003.

Table 4.11: Direct and Corrected BHC Decomposition with χ_{cies} = 2%

	Tech. Growth	Inten. Reallocation	Exten. Reallocation	Aggregate Growth
		1998-2003		
CIES	0.0298	-0.0177	0.1641	0.1761
	(16.90%)	(-10.05%)	(93.15%)	(100.00%)
χ_{cies} = 2%	0.0031	0.0453	0.1373	0.1857
	(1.65%)	(24.38%)	(73.97%)	(100.00%)
		2002-2007		
CIES	0.0768	-0.1453	0.2242	0.1557
	(49.53%)	(-93.32%)	(143.97%)	(100.00%)
χ_{cies} = 2%	0.1021	-0.0754	0.1599	0.1865
	(54.72%)	(-40.43%)	(85.72%)	(100.00%)

Percentages of contributions in parenthesis.

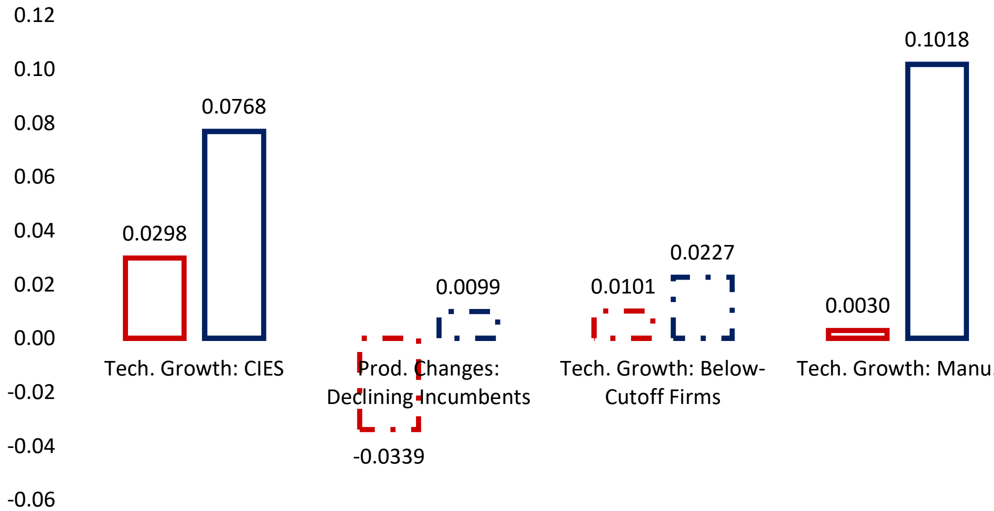
Compared to numbers in the U.S. manufacturing sector, the dominant role of extensive reallocation in growth remains. This points to the key role of new firms in China. From Table 4.6 and Figure 4.7, most of this China-U.S. difference comes from the fact that top productive new firms are larger, rather than there being more new firms in China. There also exists a large heterogeneity in sizes of new firms. Although new firms in the CIES are large, an average new firm in the manufacturing sector is imputed to be a fifth of its counterpart in the U.S.. Thus, outside of the CIES, extensive reallocation barely contributes to the aggregate productivity growth (see Figure 4.7).

The level of corrected intensive reallocation in 1998-2003 is higher than that in U.S., and contributes a similar fraction to aggregate productivity growth. However, this observation is not robust over time, unlike the U.S. case in which intensive reallocation contributes more than 20% of growth across business cycles (Baily, Hulten, and Campbell, 1992). In 2002-2007, the level of intensive reallocation is -0.07, and still contributes negatively to the aggregate productivity growth as in the direct CIES decomposition. Therefore, I conclude that the corrected intensive reallocation is larger than that implied in the CIES, but still smaller than that in U.S..

To further understand where the changes come from, Figure 4.5, 4.6 and 4.7 illustrate how the corrected decomposition in the manufacturing sector differs from that in the CIES, by looking at different groups of firms. To read these figures, using the period of 1998-2003 in Figure 4.5 for illustration, one can first multiply the first red bar, 0.0298, by the total gross output share of CIES firms, 0.9. This is because the total gross output denominator of market share now includes sales from smaller firms outside of the CIES. Then I add technological growth from declining firms, -0.0339, and that from firms who stay below minimum sales, 0.0101. This process finally gives the corrected technological growth for the manufacturing sector, 0.0030. The blue bar is for 2002-2007 and reads similarly. So do Figure 4.6 and Figure 4.7.

The mis-classification problem in the CIES and the absence of firm with sales smaller than 5 million cause quantitatively important biases on the strength of technological growth, intensive reallocation and extensive reallocation. First, firms that stay below minimum sales generate non-negligible technological growth and intensive reallocation. For example, in 1998-2003, technological growth for this group is 0.01, greater than that in the manufacturing sector (0.003). During the same time period, the intensive reallocation of this group of firms is about 0.01, which is about a quarter of the corrected intensive reallocation in the manufactur-

Figure 4.5: Sources of Differences in Technological Growth, Direct CIES and Corrected Manu.



$\chi_{cies} = 2\%$. Red 1998-2003; Blue 2002-2007

ing sector.

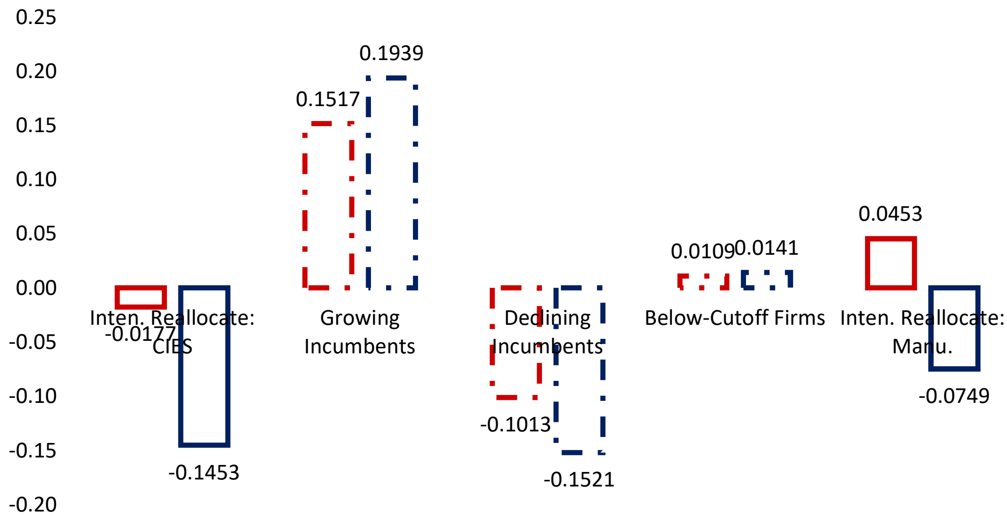
Second, a significant channel of aggregate productivity growth comes from intensive reallocation between firms that grow and enter into CIES, and firms that shrink and exit from CIES over 5 years. The magnitude of this channel is about 0.04 in both periods, which is about a quarter of the 5-year aggregate productivity growth. However, this channel is misclassified as extensive reallocation in the direct CIES decomposition.

The above results are based on a hypothetical $\chi_{cies} = 2\%$. Table 4.12 illustrates how each component of BHC decomposition changes when I vary χ_{cies}. When χ_{cies} increases, the role of intensive reallocation in aggregate productivity growth increases, while extensive reallocation decreases, as well as the technological growth.[15]

Since the true closure rate of CIES firms is in the range of 2% to 4%, the true intensive reallocation in the manufacturing sector is within 0.05 to 0.08 in 1998-2003 and -0.08 to -0.03 in 2002-2007, while the true extensive reallocation is within 0.11 to 0.14 in 1998-2003 and 0.12 to 0.16 in 2002-2007. Similarly, the true techno-

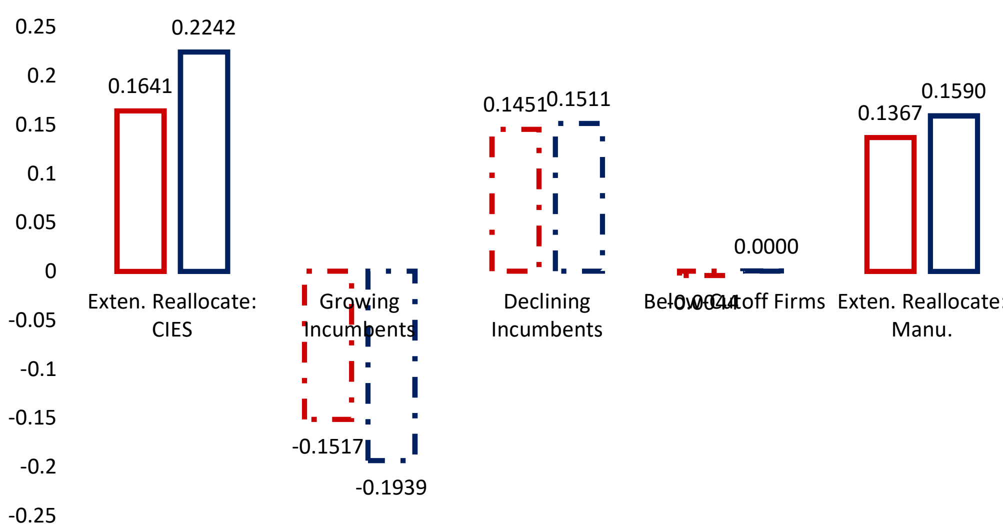
[15] Since I hold the exit rate in manufacturing sector constant at 8%, more CIES firms close when χ_{cies} increases, replacing those closing firms below the 5 million sales. This induces less efficiency in closure and thus lowers extensive reallocation. The reason why intensive reallocation increases is more mechanical. Recall that the declining incumbents enter into intensive reallocation in a negative term $\sum_{y_{it}<y_c, y_{it-5}>y_c}(\theta_{it} - \theta_{it-5})z_{it}$. When χ_{cies} increases, less firms drop into the bin of sales smaller than 5 million, and proportionally lowers the absolute value of this negative term. Therefore, intensive reallocation increases.

Figure 4.6: Sources of Differences in Intensive Reallocation, Direct CIES and Corrected Manu.



$\chi_{cies} = 2\%$. Red 1998-2003; Blue 2002-2007

Figure 4.7: Sources of Differences in Extensive Reallocation, Direct CIES and Corrected Manu.



$\chi_{cies} = 2\%$. Red 1998-2003; Blue 2002-2007

Table 4.12: Corrected BHC Decomposition in Manufacturing Sector, Varying χ_{cies}

	Tech. Growth	Inten. Reallocation	Exten. Reallocation	Aggregate Growth
		1998-2003		
CIES	0.0306	-0.0177	0.1641	0.1770
	(17.30%)	(-10.01%)	(92.71%)	(100.00%)
$\chi_{cies} = 0\%$	0.0046	0.0114	0.1672	0.1832
	(2.49%)	(6.23%)	(91.28%)	(100.00%)
$\chi_{cies} = 2\%$	0.0031	0.0453	0.1373	0.1857
	(1.65%)	(24.38%)	(73.97%)	(100.00%)
$\chi_{cies} = 4\%$	0.0017	0.0764	0.1098	0.1879
	(0.91%)	(40.66%)	(58.43%)	(100.00%)
$\chi_{cies} = 6\%$	0.0005	0.1050	0.0846	0.1901
	(0.26%)	(55.25%)	(44.49%)	(100.00%)
$\chi_{cies} = 8\%$	-0.0006	0.1312	0.0614	0.1743
	(15.24%)	(49.79%)	(33.45%)	(100.00)
		2002-2007		
CIES	0.0770	-0.1453	0.2242	0.1559
	(43.52%)	(-82.11%)	(126.68%)	(100.00%)
$\chi_{cies} = 0\%$	0.1078	-0.1252	0.1985	0.1811
	(59.53%)	(-69.12%)	(109.59%)	(100.00%)
$\chi_{cies} = 2\%$	0.1021	-0.0754	0.1599	0.1865
	(54.72%)	(-40.43%)	(85.72%)	(100.00%)
$\chi_{cies} = 4\%$	0.0968	-0.0296	0.1244	0.1915
	(50.54%)	(-15.47%)	(64.93%)	(100.00%)
$\chi_{cies} = 6\%$	0.0921	0.0124	0.0918	0.1962
	(46.91%)	(6.33%)	(46.76%)	(100.00%)
$\chi_{cies} = 8\%$	0.0878	0.0510	0.0619	0.1791
	(53.37%)	(13.23%)	(34.62%)	(100%)

Percentages of contribution are in parenthesis.

logical growth is less than 0.01 in 1998-2003, and around 0.1 in 2002-2007.

Sensitivity Despite its seemingly restrictive assumptions, does the current methodology successfully recover the BHC decomposition in China's manufacturing sector? Although this question could not be answered directly without the census data, an indirect way is to see if similar BHC decomposition results remain after correction when the CIES data is left-censored at a higher sales cutoff. To do that, I choose a new cutoff of 10 million sales and create a hypothetical CIES data that are truncated at 10 million. In this hypothetical CIES data, I redo the direct decomposition as if the data covers all manufacturing firms, and the corrected decomposition under the same set of assumptions. Figure 4.8 shows that starting with the hypothetical CIES data, the method recovers very close numbers in each component of BHC decomposition. In particular, corrected extensive reallocation for the 10 million data is 0.1304, close to 0.1373 in the CIES. For intensive reallocation, the direct number is about 1.4 to 1.8 percentage points lower, reflecting a more severe misclassification problem. Yet, the corrected measure in the 10 million data is very close to that in the CIES. Lastly for technological growth, the corrected number is around 0 in 1998-2003 and about 0.10 in 2002-2007, regardless of the minimum sales. Therefore, I argue that the method can reasonably correct the left-censored data limitation and recover the true BHC decomposition in China's manufacturing sector.

Another robustness check is to see if similar changes of BHC decomposition apply to non state owned firms only. To do that, I drop all state owned firms every year. Figure 4.9 suggests that a more severe bias exists if I take non state owned firms in the CIES as the universe of all non state owned firms. In particular, extensive reallocation is more biased upward (13 percentage points) in both 1998-2003 and 2002-2007, due to a relatively higher entrant-to-incumbent ratio in the non state owned version of CIES.[16] Similarly, intensive reallocation is more biased downward (15 percentage points) in the non state owned sector.

The findings above are likely not restricted to China. According to Bartelsman, Haltiwanger, and Scarpetta (2004), firm-level data in many countries is left-censored by minimum sales or minimum employment. For example, two popular datasets in the trade literature, Chilean and Colombian datasets, only include firms with employment more than 10. Countries like Korea and Venezuela also

[16]Entrant-to-incumbent ratio equals to number of CIES entrants divided by number of CIES incumbents.

Figure 4.8: Direct and Corrected BHC Decomposition in 10 Million Hypothetical CIES , Compared to CIES

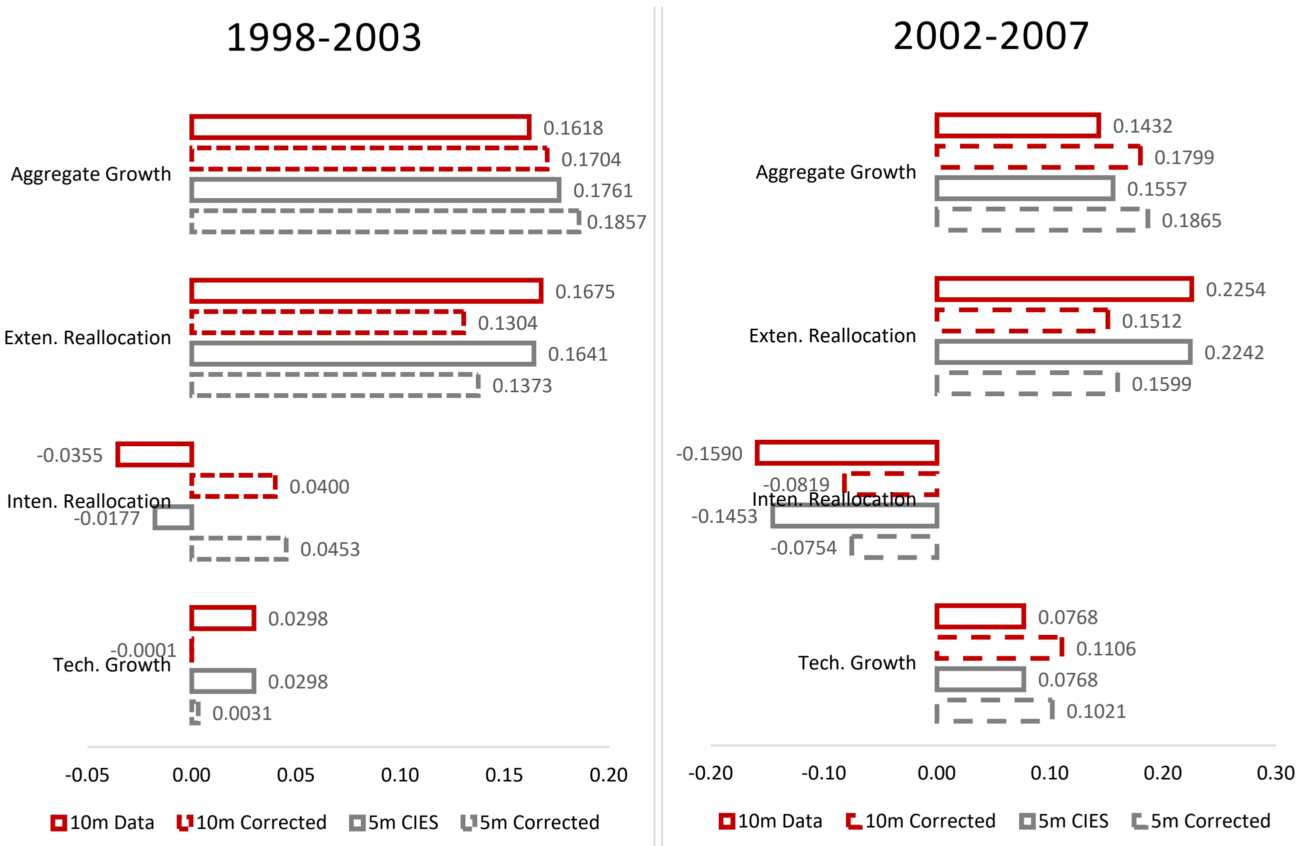
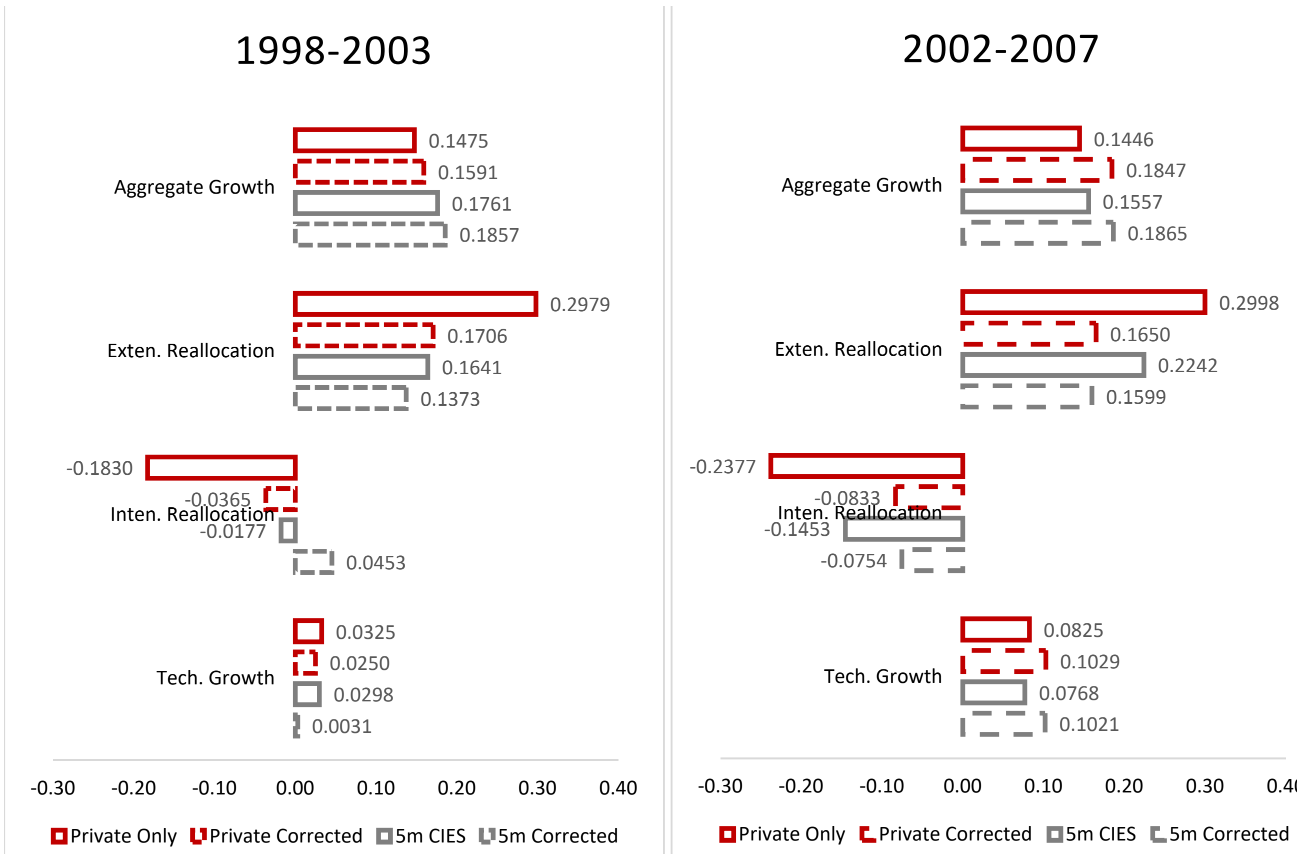


Figure 4.9: Direct and Corrected BHC Decomposition within Non State Owned Sector, Compared to CIES



have firm-level data left-censored at employment of 5. For left-censoring in sales, French data only includes firms with sales more than 2 million Euro. This suggests that when comparing channels of aggregate productivity growth across countries, one needs to take the left-censoring problem in mind.

Summary This section finds that the direct BHC decomposition in left-censored CIES data tends to overstate the role of extensive reallocation in China's aggregate manufacturing productivity growth, and understate that of intensive reallocation. After correcting for the left-censoring problem, 5-year manufacturing aggregate productivity grows at 18% for both periods of 1998-2003 and 2002-2007. Extensive reallocation contributes to 74% of this growth in 1998-2003, and 86% in 2002-2007, which are smaller than 93% and 144% in the direct CIES decomposition, respectively. The decreased extensive reallocation is mainly picked up by the increased intensive reallocation. During 1998-2003 and 2002-2007, intensive reallocation contributes 24% and -40% to aggregate productivity growth, much higher than -10% and -93% in the direct CIES decomposition.

Corrected results for technological growth are mixed, and depends on whether firms exit from the CIES data experience large productivity decline. In 1998-2003, corrected technological growth contributes almost zero to growth, while in 2002-2007, this number is 55%. These contrast to 16% and 50% in the direct CIES decomposition.

I also find that the method proposed here can recover similar magnitudes of technological growth, intensive reallocation and extensive reallocation, regardless of the minimum sales being 5 or 10 million yuan. Lastly, there is a more severe bias in the direct BHC decomposition if one treats non state owned firms in the CIES as the universe of non state owned manufacturing firms in China.

4.6 Conclusion

China's TFP growth reaches 3.9% annually during 1999-2007 before the 2008 financial crisis. This paper answers the quantitative roles of technological growth at the firm-level, intensive reallocation of inputs across existing firms, and extensive reallocation through net entry in accounting for this growth. Since the firm-level data in China is left-censored at 5 million sales, this paper develops a new method to correct the data limitation. The purpose is to understand whether the source of aggregate manufacturing productivity growth is biased in the left-censored data.

I first carry out the BHC decomposition in China Industrial Enterprise Survey Data (CIES) 1998-2007, assuming that the data covers all manufacturing firms in China. During 1998-2003 and 2002-2007, aggregate productivity growths are 17.61% and 15.57% in the CIES. 93% and 144% of these growths are from extensive reallocation in the two periods, while -10% and -93% are from intensive reallocations. The rest are accounted by technological growth at the firm level. This result contrasts greatly to that in U.S. manufacturing sector, where intensive reallocation contributes robustly at least 20% of aggregate productivity growth across business cycles. Meanwhile, the role of extensive reallocation in growth in U.S. is almost 0. I find that a higher net entry rate and the existence of state owned firms in China cannot account for such a difference between China and U.S..

The left-censoring problem in China's data can partially account for the China-U.S. difference. With several assumptions about firm productivity process, closure behavior and log sales distribution, I propose a method to correct for the minimum sales, and recover technological growth, intensive and extensive reallocations in China's manufacturing sector. I find that extensive reallocation is overstated in the direct CIES decomposition, and the opposite happens to intensive reallocation. There are two reasons for the result. First, firms that are below the 5 million cutoff are neglected in the direct CIES decomposition, yet they contribute a quarter of intensive reallocation in the corrected case. Second, firms that cross the 5 million sales are mis-classified as entrants and exiters in the direct decomposition, which overstates the extensive reallocation.

I further find the method can reasonably recover the true aggregate productivity decomposition in China's manufacturing sector. Given a hypothetical minimum sales of 10 million, the method recovers similar magnitudes of true technological growth, intensive and extensive reallocations. The qualitative results of overstated extensive reallocation and understated intensive reallocation in the direct CIES decomposition also hold within the non state owned sector.

The method developed here is applicable to other countries of which micro-level datasets are also censored, such as Chile, Colombia and France etc. Results of this paper indicate that the censoring problem should be taken into account when comparing the aggregate decomposition across countries.

Bibliography

Bai, Yan, Dan Lu, and Xu Tian. 2016. "Do Financial Frictions Explain Chinese Firms' Saving and Misallocation?" *2016 Meeting Papers. No. 998. Society for Economic Dynamics, 2016* .

Baily, Martin Neil, Charles Hulten, and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity. Microeconomics* 1992:187–267.

Bartelsman, Eric J., John Haltiwanger, and Stefano Scarpetta. 2004. "Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries." *IZA Discussion Paper No. 1374* .

Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten. 2016. "Firm Entry and Regional Growth Disparities: the Effect of SOEs in China." *Mimeo, University of Toronto* .

Brandt, Loren, Trevor Tombe, and Xiaodong Zhu. 2013. "Factor Market Distortions across Time, Space and Sectors in China." *Review of Economic Dynamics* 16 (1):39 – 58.

Brandt, Loren, Johannes Van Biesebrock, and Zhang. 2014. "Challenges of Working with the Chinese NBS Firm-level data." *China Economic Review* 30:339–352.

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

Brandt, Loren and Xiaodong Zhu. 2010. "Accounting for China's growth." *IZA Discussion Paper No. 4764* .

Cooper, Russell W. and John C. Haltiwanger. 2006. "On the Nature of Captial Adjustment Costs." *Review of Economic Studies* 73:611–633.

Ding, Sai, Alessandra Guariglia, and Richard Harris. 2016. "The Determinants of Productivity in Chinese Large and Medium-Sized Industrial firms, 1998-2007." *Journal of Productivity Analysis* 45 (2):131–155.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. 1988. “Patterns of Firm Entry and Exit in U.S. Manufacturing Industries.” *The RAND Journal of Economics* 19(4):495–515.

Foster, Lucia, John C. Haltiwanger, and Cornell John Krizan. 2001. “Aggregate Productivity Growth: Lessons from Microeconomic Evidence.” *New Developments in Productivity Analysis* :203–372.

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

Holz, Carsten A. 2013. “Chinese Statistics: Classification Systems and Data Source.” *Eurasian Geography and Economics* 54 (5-6):532–571.

Hopenhayn, Hugo A. 1992. “Entry, Exit, and firm Dynamics in Long Run Equilibrium.” *Econometrica* 60(5):1127–1150.

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

Jovanovic, Boyan. 1982. “Selection and the Evolution of Industry.” *Econometrica* :1008–1038.

Melitz, Marc. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica* 71 (6):1695–1725.

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

Petrin, Amil and James Levinsohn. 2012. “Measuring Aggregate Productivity Growth Using Plant-level Data.” *The RAND Journal of Economics* 43 (4):705–725.

Tombe, Trevor and Xiaodong Zhu. 2015. “Trade, Migration and Productivity: A Quantitative Analysis of China.” *Manuscript, University of Toronto* .

Van Biesebroeck, Jo. 2008. “Aggregating and Decomposing Productivity.” *Review of Business and Economics* 53 (2):122–146.

Xu, Chenggang. 2011. “The Fundamental Institutions of China’s Reforms and Development.” *Journal of Economic Literature* 49(4):1076–1151.

Yang, Tianyu and Lei Zhang. 2009. “ An Analysis of Influential Factors of the Entry into and Withdrawal from China’ s Manufacturing Industries.” *Management World (in Chinese)* 6:82–90.

Zhu, Xiaodong. 2012. “Understanding China’s Growth: Past, Present, and Future.” *Journal of Economic Perspectives* 26 (4):103–24.

Chapter 5

Conclusion

During 1998-2007, China experienced its historically fastest growth in the industrial sector when significant policy reforms including privatization and trade liberalization take place. This thesis seeks to understand input misallocation and aggregate productivity growth from reallocation during this time period, and provide several policy implications.

Chapter 2 finds that input misallocation are not restricted to the conventional capital and labor misallocation in the literature. Intermediate goods, with its 74% of gross output revenue, are also misallocated across firms within industries in China's firm-level data. In fact, eliminating intermediate goods misallocation in Hsieh and Klenow (2009)'s approach gives the highest gross output and value added gains, compared to that of capital and labor. This novel finding invalidates the value added approach in quantifying misallocation in the literature. Since there exists input complementaries between capital-labor and intermediate goods, value added approach ignores the extra gain by reallocating intermediate goods directly and indirectly through the complementarity. Potential explanations of intermediate goods misallocation include pre-order friction and borrowing constraints in intermediate goods.

Chapter 3 quantifies the role of pre-order and borrowing constraints on intermediate goods in accounting for the measured gross output misallocation in China. Quantitative results suggest that a model with both intermediate goods and capital frictions, i.e. adjustment costs and borrowing constraints, performs better in accounting for misallocation than a model with only capital frictions. This chapter indicates that intermediate goods, not only capital investment, could also be distorted, causing output loss. This novel finding partly resolves the puzzle how misallocation persists in a quantitative model with financial frictions in the literature (see Moll, 2014, Buera and Shin, 2013 for example).

Chapter 4 decomposes aggregate productivity growth into sources of technological growth, intensive and extensive reallocations. The conventional wisdom in the literature is that extensive reallocation drives the most of TFP growth in the left-censored China Industrial Enterprise Survey (CIES) data. Yet, when one takes the left-censoring problem into account, the role of extensive reallocation in China's manufacturing sector is overstated. In contrast, the intensive reallocation of resources among manufacturing firms are stronger than what is implied in the CIES data, although it is still worse than the U.S. case.

Results of this thesis point to several policy implications for China. First, financial market reforms could have a more profound impact on firm-level growth when its impact on intermediates reallocation is also considered. This suggests that welfare analysis of these reforms shall take the extra intermediate goods channel into account. Second, a more sustained aggregate productivity growth needs a better intensive reallocation of inputs across firms, when China transits into a lower growth regime. Reforms along this line include financial market reforms as well as the integration of domestic markets.

Bibliography

Buera, Francisco and Yongseko Shin. 2013. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy* 121(2):221–272.

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

Moll, Benjamin. 2014. "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation." *American Economic Review* 104(10):3186–3221.

Bibliography

Amaral, Pedro S. and Erwan Quintin. 2010. "Limited Enforcement, Financial Intermediation and Economic Development: A Quantitative Assessment." *International Economic Review* 51(3):785–811.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic Inputs and Resource (Mis)Allocation." *Journal of Political Economy* 122(5):1013–1063.

Asturias, Jose, Sewon Hur, Timothy J. Kehoe, and Kim J. Ruhl. 2017. "Firm Entry and Exit and Aggregate Growth." *Research Department Staff Report, Federal Reserve Bank of Minneapolis* .

Bai, Yan, Dan Lu, and Xu Tian. 2016. "Do Financial Frictions Explain Chinese Firms' Saving and Misallocation?" *2016 Meeting Papers. No. 998. Society for Economic Dynamics, 2016* .

Baily, Martin Neil, Charles Hulten, and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity. Microeconomics* 1992:187–267.

Bartelme, Dominick and Yuriy Gorodnichenko. 2015. "Linkages and Economic Development." Working Paper 21251, National Bureau of Economic Research.

Bartelsman, Eric J., John Haltiwanger, and Stefano Scarpetta. 2004. "Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries." *IZA Discussion Paper No. 1374* .

Bartlesman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review* 103(1):305–334.

Basu, Susanto. 1995. "Intermediate Goods and Business Cycles: Implications for Productivity and Welfare." *The American Economic Review* 85(3):512–531.

Bento, Pedro and Diego Restuccia. 2015. "Misallocation, Establishment Size, and Productivity." *2015 Meeting Papers. No. 1391, Society for Economic Dynamics* .

Bils, Mark, Peter J. Klenow, and Cian Ruane. 2017. "Misallocation or Mismeasurement?" *Working Paper* .

Boedo, Hernan J Moscoso and Toshihiko Mukoyama. 2012. "Evaluating the Effects of Entry Regulations and Firing Costs on International Income Differences." *Journal of Economic Growth* 17 (2):143–170.

Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten. 2016. "Firm Entry and Regional Growth Disparities: the Effect of SOEs in China." *Mimeo, University of Toronto* .

Brandt, Loren, Trevor Tombe, and Xiaodong Zhu. 2013. "Factor Market Distortions across Time, Space and Sectors in China." *Review of Economic Dynamics* 16 (1):39 – 58.

Brandt, Loren, Johannes Van Biesebrock, and Zhang. 2014. "Challenges of Working with the Chinese NBS Firm-level data." *China Economic Review* 30:339–352.

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

Brandt, Loren and Xiaodong Zhu. 2010. "Accounting for China's growth." *IZA Discussion Paper No. 4764* .

Buera, Francisco and Yongseko Shin. 2013. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy* 121(2):221–272.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin. 2015. "Entrepreneurship and Financial Frictions: A Macrodevelopment Perspective." *Annual Review of Economics* 7 (1):409–436.

Caselli, Francesco and Nicola Gennaioli. 2013. "Dynastic Management." *Economic Inquiry* 51 (1):971–996.

Castro, Rui and Pavel Sevcik. 2016. "Occupational Choice, Human Capital, and Financial Constraints." Tech. Rep. 20162, University of Western Ontario, Centre for Human Capital and Productivity (CHCP).

Cooley, Thomas F. and Vincenzo Quadrini. 2001. "Financial Markets and Firm Dynamics." *American Economic Review* 91(5):1286 – 1310.

Cooper, Russell W. and John C. Haltiwanger. 2006. "On the Nature of Captial Adjustment Costs." *Review of Economic Studies* 73:611–633.

Da-Rocha, Jose-Maria, Marina Mendes Tavares, and Diego Restuccia. 2016. "Firing Costs, Misallocation, and Aggregate Productivity." *No. w23008. National Bureau of Economic Research, 2016* .

Da-Rocha, Jos´e-Mar´ıa, Marina Mendes Tavares, and Diego Restuccia. 2015. "Firing Costs, Misallocation, and Aggregate Productivity." *Working Paper, University of Toronto* .

Davis, Steven J., John C. Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction*. MIT Press.

Di Giovanni, Julian, Andrei A Levchenko, and Jing Zhang. 2014. "The Global Welfare Impact of China: Trade Integration and Technological Change." *American Economic Journal: Macroeconomics* 6 (3):153–183.

Dias, Daniel A., Carlos Robalo Marques, and Christine Richmond. 2016. "Misallocation and Productivity in the Lead Up to the Eurozone Crisis." *Journal of Macroeconomics* 49:46–70.

Ding, Sai, Alessandra Guariglia, and Richard Harris. 2016. "The Determinants of Productivity in Chinese Large and Medium-Sized Industrial firms, 1998-2007." *Journal of Productivity Analysis* 45 (2):131–155.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. 1988. "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries." *The RAND Journal of Economics* 19(4):495–515.

Fan, Joseph P.H. and Randall Morck. 2013. *Capitalising China*. The University of Chicago Press.

Fazzari, Steven M. and Bruce C. Petersen. 1993. "Working Capital and Fixed Investment: New Evidence on Financing Constraints." *The RAND Journal of Economics* 24 (3):328–342.

Foster, Lucia, John C. Haltiwanger, and Cornell John Krizan. 2001. "Aggregate Productivity Growth: Lessons from Microeconomic Evidence." *New Developments in Productivity Analysis* :203–372.

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

Gabler, Alain and Markus Poschke. 2013. "Experimentation by Firms, Distortions, and Aggregate Productivity." *Review of Economic Dynamics* 16 (1):26 – 38.

Ghate, Chetan and Kenneth Kletzer. 2012. *Financial Frictions and Monetary Policy Transmission in India*. Oxford University Press.

Holz, Carsten A. 2013. "Chinese Statistics: Classification Systems and Data Source." *Eurasian Geography and Economics* 54 (5-6):532–571.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and firm Dynamics in Long Run Equilibrium." *Econometrica* 60(5):1127–1150.

———. 2014. "Firms, Misallocation, and Aggregate Productivity: A Review." *Annual Review of Economics* 6:735–770.

Hopenhayn, Hugo A. and Richard Rogerson. 1993. "Job Turnover and Policy Evaluations: A General Equilibrium Analysis." *Journal of Political Economy* 101:915–938.

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

———. 2014. "The Life Cycle of Plants in India and Mexico." *The Quarterly Journal of Economics* 129(3):1035–1084.

Hsieh, Chang-Tai and Ralph Ossa. 2016. "A Global View of Productivity Growth in China." *Journal of International Economics* 102:209–224.

Jeong, Hyeok and Robert M. Townsend. 2007. "Sources of TFP Growth: Occupational Choice and Financial Deepening." *Economic Theory Special Edition Honoring Edward Prescott 32* 1:189–221.

Jones, Charles I. 2011. "Intermediate Goods and Weak Links in the Theory of Economic Development." *American Economic Journal: Macroeconomics* 3(2):1–28.

Jose, Manuel L., Carol Lancaster, and Jerry L. Stevens. 1996. "Corporate Returns and Cash Conversion Cycles." *Journal of Economics and Finance* 20(1):33–46.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica* :1008–1038.

Khan, Aubhik and Julia K. Thomas. 2008. "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics." *Econometrica* 76(2):395–436.

Melitz, Marc. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6):1695–1725.

Mendoza, Enrique G. and Vivian Z. Yue. 2012. "A General Equilibrium Model of Sovereign Default and Business Cycles." *The Quarterly Journal of Economics* 127 (2):889–946.

Midrigan, Virgiliu and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review* 104(2):422–458.

Moll, Benjamin. 2014. "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation." *American Economic Review* 104(10):3186–3221.

Mukoyama, Toshihiko and Sophie Osotimehin. 2016. "Barriers to Reallocation and Economic Growth: the Effects of Firing Costs." *Mimeo, University of Virginia* .

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

Petersen, Mitchell A. and Raghuram G. Rajan. 1997. "Trade Credit: Theories and Evidence." *Review of Financial Studies* 10 (3):661–691.

Petrin, Amil and James Levinsohn. 2012. "Measuring Aggregate Productivity Growth Using Plant-level Data." *The RAND Journal of Economics* 43 (4):705–725.

Pratap, Sangeeta and Carlos Urrutia. 2012. "Financial Frictions and Total Factor Productivity: Accounting for the Real Effects of Financial Crises ." *Review of Economic Dynamics* 15 (3):336 – 358.

Restuccia, Diego and Richard Rogerson. 2013. "Misallocation and Productivity." *Review of Economic Dynamics* 16:1–10.

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

Tauchen, George. 1986. "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions." *Economics Letters* 20 (2):177 – 181.

Tombe, Trevor and Xiaodong Zhu. 2015. "Trade, Migration and Productivity: A Quantitative Analysis of China." *Manuscript, University of Toronto* .

Van Biesebroeck, Jo. 2008. "Aggregating and Decomposing Productivity." *Review of Business and Economics* 53 (2):122–146.

Xavier, Gabaix. 2011. "The Granular Origins of Aggregate Fluctuations." *Econometrica* 79 (3):733–772.

Xu, Chenggang. 2011. "The Fundamental Institutions of China's Reforms and Development." *Journal of Economic Literature* 49(4):1076–1151.

Yang, Tianyu and Lei Zhang. 2009. " An Analysis of Influential Factors of the Entry into and Withdrawal from China' s Manufacturing Industries." *Management World (in Chinese)* 6:82–90.

Yang, Yang. 2016. "Transport Infrastructure, City Productivity Growth and Sectoral Reallocation: Evidence from China." *Job Market Paper,University of California Los Angeles* .

Young, Alwyn. 2000. "The Razor's edge: Distortions and Incremental Reform in the People's Republic of China." *The Quarterly Journal of Economics* 115 (4):1091–1135.

Yu, Miaojie. 2010. "Processing Trade, Firm Productivity, and Tariff Reductions: Evidence from Chinese Products." *Economic Journal* 125(585):943–988.

Zhu, Xiaodong. 2012. “Understanding China’s Growth: Past, Present, and Future.” *Journal of Economic Perspectives* 26 (4):103–24.

Ziebarth, Nicolas L. 2013. “Are China and India Backward? Evidence from the 19th Century U.S. Census of Manufactures.” *Review of Economic Dynamics* 16 (1):86 – 99.

Appendix A

Chapter 1 Appendices

A.1 Supplementary Details for China Industrial Enterprise Survey (CIES) 1998-2007

Table A.1: 2-Digit China Industrial Classification Code (CIC), Manufacturing

2-Digit	GB/T 4754-1994	Changes in GB/T 4754-2002
13	Food processing	N/A
14	Manufacture of foods	N/A
15	Manufacture of beverages	N/A
16	Manufacture of tobacco	N/A
17	Manufacture of textiles	N/A
18	Garments and other fiber products	N/A
19	Leather, furs, down and related products	N/A
20	Timber processing, bamboo, cane, palm fiber and straw products	N/A
21	Manufacture of furniture	N/A
22	Papermaking and paper products	N/A
23	Printing and recorded media	N/A
24	Cultural, educational and sports goods	N/A
25	Petroleum processing and coking	N/A
26	Raw chemical materials and chemical products	N/A
27	Medical and pharmaceutical products	N/A
28	Chemical fiber	N/A
29	Rubber products	N/A
30	Plastic products	N/A
31	Nonmetal mineral products	N/A
32	Smelting and pressing of ferrous metals	N/A
33	Smelting and pressing of nonferrous metals	N/A
34	Metal products	N/A

35	Ordinary machinery	N/A
36	Special purpose equipment	N/A
37	Transport equipment	N/A
39	Weapons and ammunition manufacturing	Merged into 36, Relabeled: Electric equipment and machinery
40	Electric equipment and machinery	Relabeled: Electronic and telecommunications equipment
41	Electronic and telecommunications equipment	Relabeled: Instruments, meters, cultural and office equipment
42	Instruments, meters, cultural and office equipment	Relabeled: Other manufacturing
43	Other manufacturing	New: Industrial Waste Recycling

Source: National Bureau of Statistics. There are four versions of Industrial Classification Codes. GB/T 4754-1984 was used during 1985 to 1994. GB/T 4754-1994 was used during 1995 to 2002. GB/T 4754-2002 was used during 2003 to 2011. The newest code system is GB/T 4754-2011 and has been used since 2012.

A.2 Alternative Capital Share

This appendix introduces an alternative way of computing 2-digit industry capital share α_k^s and corresponding measures of misallocation.

Imputed Capital Share α_k^s Within CIES data, rental rate of capital is unobservable and needs imputation from elsewhere. I utilize rental cost information from World Bank Enterprise Survey Data (WBES, 2011).

In WBES(2011), firms are asked

For fiscal year 2011, please provide the following information about this establishment by refering directly to your income statement:
(1) Total annual rental costs of vehicles, machinery and equipment
(2) Total annual rental costs of land, buildings

Rental rate could be computed as a ratio of rental costs to total value of capital stock rented. Unfortunately, only value of land and buildings rented is readily computable from WBES(2011) from the following two questions:

1. Of the land occupied by this establishment, what percent is:
(1) owned by this establishment
(2) rented or leased by this establishment
(3) other

2. Referring directly to your balance sheet, at the end of fiscal year 2011, what was the net book value, that is the value of assets after depreciation, of the following:
(1) machinery, vehicles, and equipment
(2) land and buildings

The median rental rate for firms that rent land and buildings is 8%. Assuming 9% depreciation rate for general capital and 4% on structures, rental rate for capital is estimated to be 13%.

Given 13% rental rate, capital share α_k^s is computed as the median of firm-level capital share $13\% \times \frac{k_{is}}{y_{is}}$ within 2-digit CIC industries. On average, capital share is 4%.

Magnitude of Misallocation With this alternative capital share, I recompute magnitude of gross output and value added misallocation by reallocating one input in Table 2.5. Results in A.2 suggest that analysis in Section 2.2 is robust under alternative capital share.

Table A.2: Dispersion in Marginal Products and Output Gains by Reallocating One Input within CIC 2-digit Industries, Output Weighted, 1998-2007 Average, with Imputed Capital Share

	Intermediate Goods	Capital	Labor
All Firms			
CV	0.63	12.84	3.74
Gross output gain	41.28%	2.37%	2.36%
Value added gain	175.15%	10.04%	9.94%
Private Owned Firms			
CV	0.52	11.53	3.19
Gross output gain	36.75%	2.27%	2.29%
Value added gain	154.34%	9.57%	9.60%

A.3 Biased Value Added Productivity

Suppose firm i face distortion τ_i on intermediate goods. This distortion can come from taxes or subsidies, or could be a reduced form representation of some constraints on intermediate goods purchases.[1] With observed choices of capital k_i and labor l_i, firm's choice of intermediate goods is determined by the following static problem

$$max_{m_i} exp(z_i) k_i^{\alpha_k} m_i^{\alpha_m} l_i^{\alpha_l} - (1+\tau_i) p_m m_i$$

First order condition implies optimal choice of m_i

$$m_i^* = [\frac{(1+\tau_i)p_m}{\alpha_m exp(z_i) k_i^{\alpha_k} l_i^{\alpha_l}}]^{\frac{1}{\alpha_m - 1}}$$

Given observation of m_i^*, value added is

$$Y - m^* = exp(z_i)[\frac{\alpha_m exp(z_i)}{p_m(1+\tau_i)}]^{\frac{\alpha_m}{1-\alpha_m}} k_i^{\frac{\alpha_k}{1-\alpha_m}} l_i^{\frac{\alpha_l}{1-\alpha_m}}$$

Therefore, log value-added productivity =

$$\frac{1}{1-\alpha_m} z_i \overbrace{- \frac{\alpha_m}{1-\alpha_m} log(1+\tau_i)}^{BiasTerm} + \frac{\alpha_m}{1-\alpha_m} log(\frac{\alpha_m}{p_m})$$

If the distortion τ_i is positively correlated with z_i, a high productivity z_i firm could have a low log value-added productivity. In contrast, a low productivity z_i firm is labeled with a high value-added productivity. Vice versa for a negative correlation between τ_i and z_i.

[1]For example, financial constrained firms have a higher shadow price of intermediate goods than p_m.

Appendix B

Chapter 2 Appendices

Table B.1: Constant Shares of Intermediate Goods and Labor

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Labor Share (All)	6%	6%	6%	5%	5%	4%	4%	4%	4%	4%
Labor Share (State-Owned)	7%	7%	7%	7%	6%	6%	6%	5%	5%	5%
Labor Share (Private-Owned)	5%	5%	5%	5%	5%	4%	4%	4%	4%	4%
Intermediate Goods Share	78%	77%	75 %	76%	76%	75%	73%	71%	70%	70%
Intermediate Goods Share (State-Owned)	77%	76%	74%	75%	75%	74%	74%	73%	72%	73%
Intermediate Goods Share (Private-Owned)	79%	78%	76%	77%	77%	75%	73%	71%	69%	69%

Table B.2: Growth in Mean Productivities in the CIES, 1998-2007

Year	Mean Productivity $\bar{z}$	Mean Productivity Relative to 1998 Δz
1998	1.82	0
1999	1.84	0.02
2000	1.89	0.07
2001	1.90	0.08
2002	1.92	0.10
2003	1.97	0.15
2004	1.98	0.16
2005	2.04	0.22
2006	2.09	0.27
2007	2.12	0.30

Source: China Industrial Survey Data 1998-2007

Table B.3: Gross Output Gain in Top 20% Firms vs Simulated All Firms, Benchmark

Simulated Period	Top 20%	All	Extra TFP Gain from All Firms
0	1.01	1.09	0.08
1	0.95	1.03	0.08
2	0.94	1.02	0.08
3	0.95	1.03	0.08
4	0.95	1.02	0.08
5	0.96	1.04	0.08

Table B.4: Data Entry and Exit in Above Threshold Sample: Simulated Data vs CIES Data

Model: Given One Birth Cohort							
Years Post Birth		No. of Firms in Data	Data Exit	Exit %	Data Enter	Enter %	Enter-Exit
0		2246					
1		3075	656	29.21%	1485	66.12%	36.91%
2		3862	860	27.97%	1647	53.56%	25.59%
3		4177	1232	31.90%	1547	40.06%	8.16%
4		4374	1264	30.26%	1461	34.98%	4.72%
5			1298	29.68%	1411	32.26%	2.58%
Data: 1998 Cohort							
Years Post Birth	Year	No. of Firms in Data	Data Exit	Exit %	Data Enter	Enter %	Enter-Exit
0	1998	5024					
1	1999	7430	801	15.94%	3038	60.47%	44.53%
2	2000	9622	1031	13.88%	2888	38.87%	24.99%
3	2001	11886	1871	19.45%	3706	38.52%	19.07%
4	2002	12870	1425	11.99%	2395	20.15%	8.16%
5	2003	13308	1736	13.49%	2079	16.15%	2.67%
6	2004	15892	3017	22.67%	5567	41.83%	19.16%
7	2005	14749	2188	13.77%	1089	6.85%	-6.92%
8	2006	14712	1238	8.39%	1183	8.02%	-0.37%
9	2007		1155	7.85%	1083	7.36%	-0.49%

Source: China Industrial Survey Data 1998-2007

Appendix C

Chapter 3 Appendices

Table C.1: Probit Estimates of Staying in CIES Data in t and $t+1$, All Years

	Benchmark	Quadratic
Log Capital$_t$	0.0380***	0.237***
	(35.89)	(34.13)
Log Intermediates$_t$	0.0174***	0.0958***
	(5.34)	(5.29)
Log Labor$_t$	0.0820***	0.0600***
	(52.66)	(5.82)
Log Sales$_t$	0.130***	-0.0753***
	(37.13)	(-3.90)
State Owned Dummy$_t$	-0.128***	-0.116***
	(-39.71)	(-35.25)
Log Capital Square$_t$		-0.0166***
		(-38.78)
Log Intermediates Square$_t$		-0.00463***
		(-4.63)
Log Labor Square$_t$		-0.0162***
		(-16.91)
Log Sales Square$_t$		-0.00878***
		(-5.29)
Log Capital$_t$ $\times$ Log Labor$_t$		0.00210*
		(2.04)
Log Capital$_t$ $\times$ Log Intermediates$_t$		-0.0122***
		(-6.23)
Log Capital$_t$ $\times$ Log Sales$_t$		0.0160***
		(7.55)
Log Intermediates$_t$ $\times$ Log Labor$_t$		-0.00176
		(-0.67)
Log Intermediates$_t$ $\times$ Log Sales$_t$		0.0110***
		(5.40)
Log Labor$_t$ $\times$ Log Sales$_t$		0.0255***
		(8.88)
Industry FE	Y	Y
Year FE	Y	Y
No. of Observation	1614141	1614141
Correctly Classified	86.19%	86.19%

Curriculum Vitae

Name:	Wenya Wang
Post-Secondary Education and Degrees:	University of Western Ontario Ph.D. in Economics 2011 - 2017 (Expected)
	University of Waterloo Waterloo, Canada 2010 - 2011 M.A. in Economics
	Shanghai University of Finance and Economics Shanghai, China 2008 - 2010 M.A. in Finance
	Zhejiang University Hangzhou, China 2004-2008 B.A. in Economics
Honours and Awards:	Sir Arthur Currie Memorial Scholarship, 2015 Graduate Teaching Assistant of the Year, 2013 Western Graduate Student Scholarship, 2011-2015 Graduate Student Scholarship, 2010 Third Class of People's Scholarship, 2009 National Scholarship, 2007 First Class of Academic Scholarship, 2007
Related Work Experience:	Research Assistant University of Western Ontario 2012 - 2017
	Course Instructor University of Western Ontario 2013,2015,2017
	Teaching Assistant University of Western Ontario 2011 - 2015 University of Waterloo 2010-2011