Essays on Firm Heterogeneity and Heterogeneous Effects of Economic Policies

Graduate School of Economics, Keio University

Kenta Yamanouchi

2019

A dissertation submitted to the Graduate School of Economics of Keio University in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Ph.D.) in Economics

Acknowledgements

I would like to express my profound gratitude to everybody that supported and encouraged me. Foremost, I would like to express my deepest thankfulness to my academic supervisor, Professor Fukunari Kimura, for his continuous support and encouragement. This dissertation would not be possible without his constructive comments.

I would like to express my sincere gratitude to Professors Kozo Kiyota, Toshihiro Okubo, Toshiyuki Matsuura, Yoshimasa Shirai, Mitsuyo Ando, Masahiro Endoh, and Nobuaki Yamashita for their insightful comments and warm encouragement in the hardest times. I am also thankful to Kazunobu Hayakawa and Professor Taiyo Yoshimi for their invaluable advice on the submission of the thesis. I am deeply grateful to Professors Kaoru Hosono, Miho Takizawa, and Hiroko Okudaira for a lot of helpful comments to improve the quality of the thesis. They are also co-authors of the revised versions of Chapters 3 and 4 in the thesis.

Research Institute of Economy, Trade and Industry made the thesis possible by providing me with data avalability. I am also grateful to Mitsubishi Economic Research Institute. The working experience at the institute was really helpful.

I could not forget to mention the support from all colleagues with special mention to Mateus Silva Chang, Doan Thi Thanh Ha, Hayato Kato, Chin-Ho Lin, and Mina Taniguchi. Last but not least, I thank my family for providing me with the constant support and patience.

Abstract

This Ph.D. thesis consists of an introduction chapter and four essays on firm heterogeneity and heterogeneous effects of economic policies. In the conventional view of economics, firms, plants, or other kinds of production units are considered to be homogeneous. They are assumed to produce a homogeneous product using the same constant-returns-to-scale technology in a perfectly competitive market. In this setting, firms face the same input prices and impose the same price on the product. This is quite different from the real world. The firms are producing diffentiated products in different locations, using different technologies even within narrowly defined industries. The extremely simplified setting is used when the studies focus on the issues not closely related to firms. Many studies related to firm issues assume that the firms produce differentiated goods and operate in the monopolistically competitive market. While the setting is more realistic, the underlying firm-level parameters like productivity are assumed to be constant across firms. As a result, their activities are similar to each other and they uniformly change their activities in response to the changes in economic environment. While the simple setting is valid if the use of more complicated but realistic settings does not change the major results quantitatively and qualitatively, we should investigate to what extent the results are affected and what kind of changes would be observed when we explicitly consider firm heterogeneity in terms of their productivity and other parameters.

The objective of the thesis is twofold. First, I explore heterogeneity across firms and plants and estimate the effects of heterogeneity on the aggregate economy. In particular, I focus on the effects of heterogeneous productivity, markup, and factor price distortions on the resource misallocation across producers. I address this issue in Chapters 2 and 3. In these chapters, I mainly quantify the effects of resource misallocation as the change in aggregate productivity when resource is appropriately reallocated. I also explore the roles of uncertainty as a cause of misallocation across producers in Chapter 3. Second, I investigate the heterogeneous effects of economic policies and institutions, related to firm heterogeneity. This issue is addressed in Chapters 4 and 5. More specifically, I focus on the effects of minimum wage in Chapter 4 and Free Trade Agreements (FTAs) in Chapter 5 as examples of the economic policies. The main finding of the thesis is that heterogeneity across firms and plants is large and it largely affects the performance of the aggregate economy. It is also found

that the variations in the effects of the economic policies are large and those heterogeneous effects are attributed to the heterogeneous producers to some extent.

Chapter 1 is the introduction of the thesis. In the chapter, I first explain heterogeneity of producers as the main topic connecting the chapters of the thesis. Then I summarize the literatures of methodological issues and implications of large firm heterogeneity. The estimation of firm-level parameters is required to consider the effects of firm heterogeneity. The progress of the method to estimate the production function at firm level is, therefore, closely related to the thesis. I also take some examples for the heterogeneous effects of the economic policies. The heterogeneous effects can be attributed to the heterogeneity of producers at least to some extent. Finally, I summarize the methods and the main findings of each chapter.

In Chapter 2, I explore the resource misallocation across Vietnamese manufacturing firms. The framework of Hsieh and Klenow (2009) is applied to measure the degree of resource misallocation to compare it with other Asian countries. Hsieh and Klenow (2009) construct a tractable model to estimate the distortion of output price and factor prices. In their model, Cobb-Douglas production function and demand function with constant elasticity of substitution are imposed. Under these assumptions, appropriate reallocation of inputs across firms would increase aggregate productivity if Revenue-based Total Factor Productivity (TFPR) is different across firms. This result implies that the dispersion of TFPR can be interpreted as a measure of resource misallocation across firms. Hsieh and Klenow (2009) compare the resource misallocation in the U.S., China, and India. Using their results and other studies using the same framework, I compare the degree of resource misallocation in Vietnam with other Asian countries. As a result, the resource misallocation in Vietnam is comparable to China, India, and Thailand, and larger than the U.S. and Japan. In addition, the simulation results show that the aggregate productivity would rise by 30% if the allocative efficiency of Vietnam is improved to the level of the U.S. Finally, I explore the change in firm size distribution when market distortion is totally removed. The simulation result shows that large firms are facing to disadvantageous distortion and those firms would be larger if they were equally treated in the market. On the other hand, small firms face advantageous distortion and they would be smaller if the resources are appropriately allocated.

In Chapter 3, I focus on the effects of uncertainty and competition on the allocative efficiency, using a large panel dataset of manufacturing plants in Japan. In general, uncertainty reduces irreversible investment. This effect is related to capital misallocation across plants because the plants facing large uncertainty do not undertake enough investment even when their productivity levels are largely deviated from the optimal levels in terms of static view. This mechanism is suggested by Asker, Collard-Wexler, and De Loecker (2014). I first measure the degree of capital misallocation and volatility at industry-year level. The degree of capital misallocation is defined as the dispersion of marginal revenue productivity of capital. Volatility is defined as the dispersion of the change rates in plant-level productivity within industry. I found a positive relationship between volatility and capital misallocation. This result implies that the plants do not adjust their capital inputs to the optimal levels immediately. I also investigate the role of competition in the relationship between uncertainty and misallocation. Theoretically, the effect of uncertainty on misallocation depends on the degree of market competition. I explore the roles of competition in the simulation analysis and estimate the effects of competition on the relationship between uncertainty and misallocation. It is found that the effect of uncertainty is larger if the output market is more competitive. The results of the structural estimation and simulation analysis show that the aggregate productivity increases by 0.7% on average for all industries if the volatility of TFPR shocks reduces by half. For more competitive industries, the effect on the aggregate productivity is as high as 2.1%.

In Chapter 4, I focus on the labor market and explore the heterogeneous effects of minimum wage, using Japan's plant data. While previous studies have reached little consensus on the employment effect of the minimum wage, the effect theoretically depends on the monopsonistic power of the producers in the local labor market. In the chapter, therefore, I first estimate the monopsonistic power of plants by applying the methods of estimating the production function and markups. Monopsonistic power or surplus in the local labor market is defined as a function of output elasticities with respect to intermediates and labor and the shares of intermediates and labor to sales. Then I estimate the employment effects of minimum wage by monopsonistic power in the local labor market. I found the negative effects of minimum wage when all plants are included in the sample. But the impact of the minimum wage is concentrated in specific markets. The employment effect of an increase in the minimum wage is significantly negative for the plants with little surplus in the local labor market. The minimum wage increases have little employment effect on plants with a relatively high surplus, even when they have a significant number of minimum wage employees.

In Chapter 5, I investigate the heterogeneous effects of FTAs. While FTAs have positive effects on the bilateral trade values by reducing trade costs, Baier, Bergstrand, and Clance (2018) and Baier, Yotov, and Zylkin (2019) found that the trade creation effects of FTAs have a large variation across agreements or country pairs. I focus on the trade creation effects of Japan's FTAs and its variation across partner countries using bilateral trade data. I also estimate the gravity model specified in many ways and compare the results with the most reliable specification. The main finding is that the effects of Japan's FTAs are not clearly observed when the gravity model is specified with three types of fixed effects, i.e. exporter-year fixed effects, importer-year fixed effects, and country-pair fixed effects. In fact, the effects of FTAs vary substantially among trade partners and around half of the FTAs increase Japan's trade values. The estimation results also suggest that FTAs with small trade partners tend to have large effects on Japan as well as other countries. Recently enforced FTAs, however, increase Japan's import values more rapidly. In this chapter, I focus on the heterogeneous effects across country pairs instead of producers due to the data restriction. But some of the differences in the effects of FTAs must be attributed to firm heterogeneity because the changes in the numbers of exporting firms play important roles to explain the changes in bilateral trade values.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Main topic of the thesis and related literature

In the conventional view of economics, firms, plants, or other kinds of production units are considered to be homogeneous. They are assumed to produce a homogeneous product using the same constant-returns-to-scale technology in a perfectly competitive market. In this setting, firms face the same input prices and impose the same price on the product. This is quite different from the real world. The firms are producing differentiated products in different locations, using different technologies even within narrowly defined industries. The extremely simplified setting is used when the studies focus on the issues not closely related to firms. Many studies related to firm issues assume that the firms produce differentiated goods and operate in the monopolistically competitive market. While the setting is more realistic, the underlying firm-level parameters like productivity are assumed to be constant across firms. As a result, their activities are similar to each other and they uniformly change their activities in response to the changes in economic environment. While the simple setting is valid if the use of more complicated but realistic settings does not change the major results quantitatively and qualitatively, we should investigate to what extent the results are affected and what kind of changes would be observed when we explicitly consider firm heterogeneity in terms of their productivity and other firm-level parameters.

To tackle the problem, I specify two issues to be explored as the objectives of the thesis. First, I explore heterogeneity across firms and plants and estimate the effects of heterogeneity on the aggregate economy. In particular, I focus on the effects of heterogeneous productivity, markup, and factor price distortions on the resource misallocation across producers. I address this issue in Chapters 2 and 3. In these chapters, I mainly quantify the effects of resource misallocation as the change in aggregate productivity when resource is appropriately reallocated. I also explore the roles of uncertainty as a cause of misallocation across producers in Chapter 3. Second, I investigate the heterogeneous effects of economic policies and institutions related to firm heterogeneity. This issue is addressed in Chapters 4 and 5. More specifically, I focus on the effects of minimum wage in Chapter 4 and Free Trade Agreements (FTAs) in Chapter 5 as examples of the economic policies. The main finding of the thesis is that heterogeneity across firms and plants is large and it largely affects the performance of the aggregate economy. It is also found that the variations in the effects of the economic policies are large and those heterogeneous effects are attributed to the heterogeneous producers to some extent.

Since 1990s, the availability of microdata, particularly firm-level and plant-level data, has been drastically improved and it opened the possibility of many kinds of research topics in macroeconomics, industrial organization, labor economics, and international economics. As an early study using microdata, for example, Bernard and Jensen (1995) show that the number of exporting plants is small in the U.S. Their paper and subsequent studies mainly attributed the difference in export status to the difference in productivity. The relationship between the export status and firm productivity is theoretically established in the seminal model of Melitz (2003). The model incorporates firm heterogeneity into Krugman (1980) model and is a natural extension of Hopenhayn (1992). In the model, export status of the firm is not uniform and only high productive firms export their products. Now, productivity is one of the central concepts in the issues of heterogeneity across producers. Syverson (2011) is the survey of the studies on productivity and describes large productivity heterogeneity even within narrowly defined industries as one of the most important findings of the studies on productivity.

The estimation methods of production function at firm level were progressed in parallel with the improvement in the data availability. In the estimation of the production function, unobservable productivity works as a source of endogeneity because managers hire more inputs if they observe favorable shocks in productivity. Olley and Pakes (1996) suggest that the method to deal with the endogeneity by controlling for the productivity shocks using past investment. Levinsohn and Petrin (2003) suggest the use of intermediate goods instead of past investment to deal with the case of no investment. Wooldridge (2009) suggests the use of Generalized Method of Moments instead of the sequential estimation. They are subsequent works to improve the method of Olley and Pakes (1996). Recently, Ackerberg, Caves, and Frazer (2015) point out that the functional dependence problems for the method of Olley and Pakes (1996) and Levinsohn and Petrin (2003). Gandhi, Navarro, and Rivers (2013) also point out the identification issue of intermediate goods and suggest the use of first order conditions for the flexible inputs to estimate the production function consistently. The methodological development of estimating the production function provided the necessary background for the productivity analysis.

Recently, other dimensions than productivity are also explored in the literature. Hsieh and Klenow (2009) construct a tractable model to estimate the distortion of output price and factor prices. In the model, Cobb-Douglas production function and demand function with constant elasticity of substitution are imposed. Under these assumptions, they show factor intensity must be constant across firms within industry if the firms face the same input prices. In addition, appropriate reallocation of inputs across firms would increase aggregate productivity if Revenue-based Total Factor Productivity (TFPR) is different across firms. It implies that the dispersion of TFPR can be interpreted as a measure of resource misallocation across firms. Using the microdata of three countries, the U.S., China, and India, they show that the effects of resource misallocation across firms on the aggregate productivity are huge in China and India and it can explain large differences in TFP at the aggregate economy across countries. A number of studies are subsequent to Hsieh and Klenow (2009), and most of those papers explore the causes for large misallocation. Hopenhayn (2014) is a survey of those misallocation studies. Hopenhayn (2014) points out that most of the policies and institutions have limited impacts on the resource misallocation. Hopenhayn (2014) also describes the roles of the adjustment costs suggested by Asker, Collard-Wexler, and De Loecker (2014) can be large among the causes of resouce misallocation across firms, although Hopenhayn (2014) avoids the final verdict.

Another important dimension of heterogeneity is markup imposed by firms and plants. In the theoretical model, markup heterogeneity can be generated in several ways. One of the tractable modelling techniques is the use of quadratic utility function as in Melitz and Ottaviano (2008). On the other hand, De Loecker and Warzynski (2012) suggest an empirical method to estimate markup at firm level. They apply the method of Hall (1988) to firm data and find that the exporters have higher markups than domestic producers. If firms minimize their costs for intermediate inputs, the share of intermediate expenditure to sales must be equal to the output elasticity with respect to intermediates divided by markup. In other words, markup can be calculated as the ratio of output elasticity with respect to intermediates to the share of intermediate expenditure. Many subsequent studies make much effort to measure and analyze the differences of markup empirically, although De Loecker and Warzynski (2012) is severely criticized due to the use of sales data to estimate the output elasticity with respect to intermediates. Recently the method is extended to estimate the market power of large producers in the local labor market instead of the traditional output market. Lu, Sugita, and Zhu (2017) call the market power “wage markdown” and explore the effects of foreign direct investment liberalization on the labor market in China. Markup and similar kinds of market power are major candidates to apply the methods of estimating the production function.

One of the implications of the large heterogeneity across firms and plants is that economic growth and aggregate productivity growth can be generated by reallocation of inputs across firms and plants. As explained above, Hsieh and Klenow (2009) mainly explore the relationship between the levels of aggregate productivity and allocative efficiency. This idea can be applied to the relationship between changes in aggregate productivity and allocative efficiency. In the literature of productivity analysis, researchers tried to decompose the aggregate productivity growth into average productivity growth of producers and the rise in allocative efficiency. Baily, Hulten, and Campbell (1992) and Foster, Haltiwanger, and Krizan (2001) define the aggregate productivity as the weighted average of firm-level productivity. They found the large role of allocation improvement across producers. Similarly, Olley and Pakes (1996) suggest another decomposition method of the aggregate productivity growth. While their definition of the aggregate productivity, or more precisely industry productivity, is the same as Baily, Hulten, and Campbell (1992), they decomposed it into mean of productivity at plant level and covariance between productivity and market share. Melitz and Polanec (2012) is a survey of these decomposition methods. While both of the decomposition methods are clear and intuitive, Petrin and Levinsohn (2012) point out drawbacks of these methods. One of the key points is that if the aggregate productivity is defined as the average productivity of individual producers, it has no links with the aggregate productivity defined by macroeconomic variables.

Another interesting implication of large heterogeneity across firms and plants is the heterogeneous effects of economic policies. Many kinds of economic policies and institutions like regulations of output, capital, and labor markets determine the economic environment where the firms operate. Therefore, the effects of economic policies and their changes depend on the parameters of producers. If firms have different parameters, the economic policies affect them in different ways. In the Melitz model, for example, high productivity firms are exporting, and they can expand the size when the market is more integrated with other economies. On the other hand, low productivity firms are forced to be smaller due to tougher competition and some firms must exit from the market.

Firm heterogeneity also has some impacts on the interpretation for the coefficients in the gravity equation. In the Krugman model, firms are homogeneous and have the same export status. In this model, therefore, all of changes in bilateral trade values can be explained by the changes in the trade values of exporters and importers. The coefficient for bilateral trade costs with respect to the bilateral trade value can be interpreted as the elasticity of substitution. On the other hand, Chaney (2008) derives the gravity equation based on the Melitz model. In this model, some of the changes in bilateral trade values are attributed to the changes in the numbers of exporters and importers. The interpretation of the coefficient changes to a parameter of Pareto distribution of firm productivity. The parameter of Pareto distribution is closely related to the dispersion of the realized values. The important role of the number of firms in the bilateral trade is empirically studied in Bernard, Jensen, Redding, and Schott (2007).

The increase in minimum wage, one of the most important regulations for the labor market, also has heterogeneous effects on the employment levels of producers. As shown in Manning (2003), the effects of the increase in minimum wage depend on the market power of local firms if the labor market is monopsonistic. While Card and Krueger (1994) concluded the rise in minimum wage has no effects on employment by clear identification strategy, the evidences thereafter are mixed and there is no concensus on the employment effects of minimum wage. The absence of the concensus itself implies the heterogeneous effects of the policy.

In the thesis, I explore the effects of firm heterogeneity and the heterogeneous effects of economic policies. To this end, I first explore the heterogeneity of firms and plants. Using the microdata of producers, I estimate the effects of resource misallocation in various ways. I also explore the underlying factors for resource misallocation across producers. Then I investigate the heterogeneous effects of economic policies. In particular, I focus on the effects of minimum wage and FTAs as examples of the economic policies. The roles of policy uncertainty in the allocative efficiency across producers can also be considered as the heterogenesous effects of the economic policies. As one of the main results, I found large variations in the effects of the economic policies. The heterogeneous effects are attributed to the heterogeneity of firms and plants to some extent. I have to admit that I focus on the heterogeneous effects across country pairs instead of producers in the case of FTAs due to the data restriction. But I consider that some of the differences in the effects of FTAs are attributed to firm heterogeneity because the changes in the numbers of exporting firms play important roles to explain the changes of bilateral trade values. The results obtained in the thesis imply that we need to consider the distributions of the firm-level parameters more seriously. The large heterogeneity of the effects of economic policies are important for both of policymakers and academic researchers.

1.2 Outline of the thesis

In Chapter 2, I explore the resource misallocation across Vietnamese manufacturing firms. The framework of Hsieh and Klenow (2009) is applied to measure the degree of resource misallocation in Vietnam to compare it with other Asian countries. As described above, Hsieh and Klenow (2009) compare the resource misallocation in the U.S., China, and India. Their framework is employed for many subsequent works to analyze resource misallocation. For Asian countries, Hosono and Takizawa (2013) apply the framework to Japan and Dheera-Aumpon (2014) to Thailand. Using their results, I compare the degree of resource misallocation in Vietnam with other Asian countries. The main findings are as follows. First, the resource misallocation in Vietnam is comparable to China, India, and Thailand, and larger than the U.S. and Japan. Second, the aggregate productivity would rise by 30% if the allocative efficiency of Vietnam is improved to the level of the U.S. Finally, I simulate the model and consider the change of firm size when market distortion is totally removed. The result shows that large firms are facing to disadvantageous distortion and those firms would be larger if they were equally treated in the market. On the other hand, small firms face advantageous distortion and they would be smaller if the resources are appropriately allocated.

In Chapter 3, I focus on the effects of uncertainty and competition on the allocative efficiency, using a large panel dataset of manufacturing plants in Japan. In general, uncertainty reduces irreversible investment. This effect is related to capital misallocation across plants because the plants facing large uncertainty do not undertake enough investment even when their productivity levels are largely deviated from the optimal levels in terms of static view. This mechanism is suggested by Asker, Collard-Wexler, and De Loecker (2014). I first measure the degree of capital misallocation and volatility at industry-year level, and found a positive relationship between the measures of misallocation and volatility. I also consider the role of competition in the relationship between uncertainty and misallocation. Theoretically, the effect of uncertainty on misallocation depends on the degree of market competition. I explore the roles of competition in the simulation analysis and estimate the effects of competition on the relationship between uncertainty and misallocation. It is found that the effect of uncertainty is larger if the output market is more competitive. The results of the structural estimation and simulation analysis show that the aggregate productivity increases by 0.7% on average for all industries if the volatility of TFPR shocks reduces by half. For more competitive industries, the effect on the aggregate productivity is as high as 2.1%.

In Chapter 4, I focus on the labor market and explore the heterogeneous effects of minimum wage, using Japan's plant data. While previous studies have reached little consensus on the employment effect of the minimum wage, the effect theoretically depends on the monopsonistic power of the producers in the local labor market. In the chapter, therefore, I first estimate the monopsonistic power of plants by appling the methods of estimating the production function and markups. The method is developed in Dobbelaere and Mairesse (2012) and Petrin and Sivadasan (2013). Then I estimate the employment effects of minimum wage by monopsonistic power in the local labor market. I found that that local labor markets are heterogeneous and the impact of the minimum wage is concentrated in specific markets. The employment effect of an increase in the minimum wage is significantly negative for the plants with little surplus in the local labor market. The minimum wage increases have little employment effect on plants with a relatively high surplus, even when they have a significant number of minimum wage employees.

In Chapter 5, I investigate the heterogeneous effects of FTAs. While FTAs have positive effects on the bilateral trade values by reducing trade costs, Baier, Bergstrand, and Clance (2018) and Baier, Yotov, and Zylkin (2019) found that the trade creation effects of FTAs have a large variation across agreements or country pairs. I focus on the trade creation effects of Japan's FTAs and its variation across partner countries using bilateral trade data. I also estimate the gravity model specified in many ways and compare the results with the most reliable specification. The main finding is that the effects of Japan's FTAs are not clearly observed when the gravity model is specified with three types of fixed effects, i.e. exporter-year fixed effects, importer-year fixed effects, and country-pair fixed effects. In fact, the effects of FTAs vary substantially among trade partners and around half of the FTAs increase Japan's trade values. The estimation results also suggest that FTAs with small trade partners tend to have large effects on Japan as well as other countries. Recently enforced FTAs, however, increase Japan's import values more rapidly. In this chapter, I focus on the heterogeneous effects across country pairs instead of producers due to the data restriction. But some of the differences in the effects of FTAs must be attributed to firm heterogeneity because the changes in the numbers of exporting firms play important roles to explain the changes in bilateral trade values.

Chapter 2 Misallocation and Productivity: The Case of Vietnamese Manufacturing[1]

This chapter attempts to measure the effect of resource misallocation on aggregate manufacturing total factor productivity (TFP), focusing on Vietnamese manufacturing firms for the period 2000–2009. One of the major findings of this chapter is that there would have been substantial improvement in aggregate TFP in Viet Nam in the absence of distortions. The results imply that potential productivity gains from removing distortions are large in Vietnamese manufacturing. I also find that smaller firms tend to face advantageous distortions, while larger firms tend to face disadvantageous ones. Moreover, the efficient size distribution is more dispersed than the actual size distribution. These results together suggest that Vietnamese policies may constrain the largest and most efficient producers and coddle its small and least efficient ones.

2.1 Introduction

Differences in per capita income across countries result mainly from differences in total factor productivity (TFP).[2] Therefore, clarifying the underlying causes of low productivity in developing countries is one of the central concerns in various fields of economics such as development economics, international economics, and macroeconomics. Given the fact that production efficiency is heterogeneous across firms, some recent studies on this issue argue that aggregate TFP depends not only on the TFP of individual firms but also on the allocation of resources across firms.[3] In other words, low productivity in developing countries can be attributable to the misallocation of resources across heterogeneous firms.

---

[1] The revised version of this chapter is in *Asian Development Review* 33 (2), 94-118, co-authored with Doan Thi Thanh Ha and Kozo Kiyota.

[2] "Large differences in output per worker between rich and poor countries have been attributed, in no small part, to differences in total factor productivity" (Hsieh and Klenow, 2009, p.1403); "cross-country income differences mostly result from differences in total factor productivity" (Waugh, 2010, p. 2095). McMillan and Rodrik (2011) also argued for the importance of resource reallocation in enhancing productivity growth in developing countries.

[3] See Restuccia and Rogerson (2013) and Hopenhayn (2014) for a survey.

How do I measure the misallocation of resources? One way to answer this question is to focus on distortions that reflect the difference between the actual and efficient outcomes. Such distortions are called "wedges" in the literature. A seminal paper is Hsieh and Klenow (2009), which estimates wedges from data on value added and factor inputs for manufacturing establishments in People's Republic of China, India, and the United States. They found that the distortions were much larger in People's Republic of China and India than in the United States. Moreover, as mentioned above, Hsieh and Klenow (2009) found that the removal of distortions has a significant effect on aggregate TFP in People's Republic of China and India. Following Hsieh and Klenow (2009), several studies have provided a similar picture: large TFP gains could be expected from the removal of distortions.[4]

Along this line of literature, this chapter extends the analysis of Hsieh and Klenow (2009) to Vietnamese manufacturing between 2000 and 2009 and asks the following four questions:

i. To what extent are resources misallocated in Viet Nam?

ii. How large would the productivity gains have been in the absence of distortions?

iii. Are the distortions related to firm size?

iv. What would the distribution of firm size have been in the absence of distortions?

Answering these questions have important implications for the potential growth, because reallocation would lead to productivity gains that would accelerate potential growth in transition towards the improved inter-firm resource allocation.

The study is closely related to Bach (2014), which also examined resource misallocation in Viet Nam using firm-level data. His study addressed the first two questions but did not compare resource misallocation in Viet Nam with misallocation in other Asian countries. Moreover, his study did not address the last two questions. From a policy perspective, the last two questions are important because many countries give preferential treatment to small and medium-sized enterprises (SMEs). Indeed, size-dependent policies, which limit the size of firms, could be an important source of misallocation (Restuccia and Rogerson, 2013). In answering the four questions above, this chapter goes one step

[4] See, for example, Camacho and Conover (2010) for the case of Colombia; Busso et al. (2012) for Latin America; Bellone and Mallen-Pisano (2013) for France; Hosono and Takizawa (2013) for Japan; de Vries (2014) for Brazil; Dheera-Aumpon (2014) for Thailand; Bach (2014) for Viet Nam; and Calligaris (2015) for Italy.

further by providing a deeper understanding of the potential productivity gains from removing distortions in Viet Nam.[5]

The rest of this chapter is organized as follows. In Section 2.2, I describe the methodology of Hsieh and Klenow (2009). Section 2.3 describes the Vietnamese firm-level data used in the study. Section 2.4 presents the results. Concluding remarks and policy implications are presented in Section 2.5.

2.2 Measurement of misallocation

Hsieh and Klenow (2009) formulated an analytical framework to estimate misallocation. Although some studies such as Bartelsman et al. (2013) developed an alternative framework, this chapter employs Hsieh and Klenow's framework for the following two reasons. First, their framework is tractable in the sense that it is simple and its data requirements are minimal. This is a significant advantage in estimating misallocation in Viet Nam because of the limited data availability, as I will discuss in the next section. Second, the framework allows us to decompose the source of misallocation into distortions in output markets and those in capital markets. Such decompositions are useful if the distortions come from different sources. The Hsieh and Klenow (2009) methodology is summarized below.

Assume that a representative firm produces a single final good Y in a perfectly competitive final goods market. The firm produces Y, using the output Y_s of S manufacturing industries, with the following Cobb–Douglas production technology:

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

and θ_s is the output share of each industry s.

[5] Another important difference between his study and my study is that his study did not control for the skill differences of workers across firms in measuring quantity-based TFP (hereafter, TFPQ) and revenue based TFP (hereafter, TFPR).

Each industry produces output, Y_s, using M_s differentiated goods produced by individual firm i with a constant elasticity of substitution technology ($s = 1, \ldots, S$). Output in industry s is then given by:[6]

$$Y_s = \left(\sum_{i=1}^{M_s} Y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad \sigma > 1, \tag{2}$$

where σ is the elasticity of substitution between varieties and Y_{si} is the output of the differentiated good produced by firm i in industry s, using capital and labor, based on the following Cobb–Douglas technology:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}, \tag{3}$$

where A_{si}, K_{si}, and L_{si} denote productivity, capital, and labor of firm i in industry s, respectively; α_s represents the capital share, which is different across industries but the same across firms within an industry.

To assess the extent of misallocation, Hsieh and Klenow (2009) followed Foster et al. (2008) in making a distinction between physical productivity, denoted by TFPQ, and revenue productivity, denoted by TFPR:

$$TFPQ_{si} \triangleq A_{si} = \frac{Y_{si}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}} \tag{4}$$

and

$$TFPR_{si} \triangleq P_{si} A_{si} = \frac{P_{si} Y_{si}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}}, \tag{5}$$

respectively, where P_{si} represents the firm-specific output price.

In addition to firm heterogeneity in terms of productivity (as in Melitz, 2003), firms potentially face different output and capital distortions. More specifically, Hsieh and Klenow (2009) incorporated two types of firm-level wedges into this framework. One raises the marginal product of capital and labor by the same proportion, which is denoted by τ_{Ysi}. The other increases the marginal product

[6] I suppress the time subscript to avoid heavy notation, although I utilize firm-level panel data in the empirical analysis.

of capital relative to labor, which is denoted by τ_{Ksi}. These wedges are given from the firm's viewpoint, and I do not make any assumptions about what generates them.[7]

An example of such distortions is subsidized credit. If two firms have identical technologies but one of the firms can borrow at a lower interest rate (and the other firm can borrow at a higher interest rate from the financial market), the marginal product of capital of the firm that can access the subsidized credit will be lower than that of the other firm. This results in the misallocation of capital because one firm enjoys a lower interest rate even though the two firms have the same technologies. In other words, in Hsieh and Klenow (2009)'s framework, the differences in factor prices mean the existence of distortions.

With these wedges, the expected profits of the firm are written as:[8]

$$\pi_{si} = (1 - \tau_{Ysi})P_{si}Y_{si} - wL_{si} - (1 + \tau_{Ksi})RK_{si}, \tag{6}$$

where w and R denote the common wages and rental costs facing all firms, respectively. Firms maximize their profits under the following constraint:

$$Y_{si} = Y_s \left(\frac{P_s}{P_{si}}\right)^{\sigma}, \tag{7}$$

where

$$P_s \equiv \left(\sum_{i=1}^{M_s} P_{si}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}. \tag{8}$$

In the presence of distortions, firms will produce a different quantity compared with what they would produce without these wedges (i.e., the efficient case).

[7] Distortions can be generated by various factors such as trade policies and credit market imperfections. In Ha and Kiyota, 2015, I examined the determinants of distortions in Vietnamese manufacturing. Leon-Ledesma and Christopoulos (2016) examined the effects of access to finance obstacles on misallocation. Using the firm-level data covering about 45 countries, they found that access to finance obstacles and private credit increase the dispersion of distortions. However, they also found that the financial variables explain a small part of the dispersion of factor market and size distributions.

[8] Distortions to output and to capital relative to labor are an observationally equivalent characterization of those to the absolute levels of capital and labor. For more details, see Hsieh and Klenow (2009, Appendix III)

Solving the profit maximization problem under a monopolistic competition framework and the equilibrium allocation of resources across industries, I have:

$$P_{si} = \frac{\sigma}{\sigma - 1}\left(\frac{R}{\alpha_s}\right)^{\alpha_s} \left(\frac{w}{1-\alpha_s}\right)^{1-\alpha_s} A_{si}^{-1} \frac{(1+\tau_{Ksi})^{\alpha_s}}{1-\tau_{Ysi}}, \tag{9}$$

$$1 - \tau_{Ysi} = \frac{\sigma}{\sigma - 1} \frac{wL_{si}}{(1-\alpha_s)P_{si}Y_{si}}, \tag{10}$$

$$1 + \tau_{Ksi} = \frac{\alpha_s}{1-\alpha_s} \frac{wL_{si}}{RK_{si}}. \tag{11}$$

From equation (9), I have:

$$TFPR_{si} = \xi_s \frac{(1+\tau_{Ksi})^{\alpha_s}}{1-\tau_{Ysi}}, \tag{12}$$

where

$$\xi_s = \frac{\sigma}{\sigma - 1}\left(\frac{R}{\alpha_s}\right)^{\alpha_s} \left(\frac{w}{1-\alpha_s}\right)^{1-\alpha_s}. \tag{13}$$

Noting that ξ_s is different across industries but constant within an industry, equation (12) implies:

$$TFPR_{si} \propto \frac{(1+\tau_{Ksi})^{\alpha_s}}{1-\tau_{Ysi}}. \tag{14}$$

This equation means that the large deviation of firm TFPR from ξ_s is a sign that the firm faces large distortions.

Denote industry TFP as TFP_s. Define industry TFP as a weighted geometric average of firm i's $TFPQ_{si}$:

$$TFP_s \triangleq \left[\sum_{i=1}^{M_s} \left(TFPQ_{si} \frac{\overline{TFPR}_s}{TFPR_{si}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}}, \tag{15}$$

where $\overline{TFPR}_s$ is the geometric average of the average marginal revenue product of labor and capital in industry s:

$$\overline{TFPR}_s \triangleq \frac{\sigma}{\sigma-1}\left[\frac{R}{\alpha_s \sum_{i=1}^{M_s} \frac{1-\tau_{Ysi}}{1+\tau_{Ksi}} \frac{P_{si}Y_{si}}{P_sY_s}}\right]^{\alpha_s}\left[\frac{w}{(1-\alpha_s)\sum_{i=1}^{M_s}(1-\tau_{Ysi})\frac{P_{si}Y_{si}}{P_sY_s}}\right]^{1-\alpha_s}. \tag{16}$$

There are two remarks regarding equation (15). First, the higher the dispersion in TFPR, the lower the industry TFP will be. Hsieh and Klenow (2013) showed that when TFPQ and TFPR are jointly log-normally distributed and when there is only variation in $\log(1-\tau_{Ysi})$, aggregate TFP can be expressed as follows:[9]

$$\mathrm{logTFP_s} = \frac{1}{\sigma-1}[\log M_s + \log E(TFPQ_{si}^{\sigma-1})] - \frac{\sigma}{2} var(\log TFPR_{si}). \tag{17}$$

This equation suggests that industry TFP will decline if the elasticity of substitution σ and/or TFPR dispersion increase.

Second, TFPR will be equalized across firms within industry s if τ_{Ksi} and τ_{Ysi} are equalized. For example, from equation (12), $TFPR_{si} = \xi_s \ \forall i$ if $\tau_{Ksi} = \tau_{Ysi} = 0$. This in turn implies that $TFPR_{si} = \xi_s = \overline{TFPR}_s \ \forall i$.[10] Denote industry TFP without any distortions as $\overline{TFPQ}_s$. From equation (15), I can obtain:

$$\overline{TFPQ}_s \triangleq \bar{A}_s = \left(\sum_{i=1}^{M_s} A_{si}^{\sigma-1}\right)^{\frac{1}{\sigma-1}}, \tag{18}$$

which is called "efficient" industry TFP.

Note that in order to obtain "efficient" TFP, one needs information on firm-level TFPQ (i.e., A_{si}). One problem is the limited availability of firm-level price data, P_{si}, which are not available in many countries including Viet Nam.[11] Hsieh and Klenow (2009) rewrote equation (4) as:

[9] A similar property is obtained even when there is variation in $\log(1+\tau_{Ksi})$, although the equation becomes more complicated. For more details, see Hsieh and Klenow (2013).

[10] Note that even when TFPR is equalized across firms, TFPQ can be different across firms because more productive firms charge lower prices (see equation (9)). In other words, if $A_{si} > A_{sj}$ and $P_{si} < P_{sj}$, $P_{si}A_{si}$ could be equal to $P_{sj}A_{sj}$ for $i \neq j$.

[11] There are a few countries in which firm-level (or plant-level) price data are available. For example, Eslava et al. (2004) utilized plant-level price data for Colombia to estimate plant-level TFPQ.

$$TFPQ_{si} = A_{si} = \kappa_s \frac{(P_{si}Y_{si})^{\frac{\sigma}{\sigma-1}}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}}, \text{where } \kappa_s = w^{1-\alpha_s} \frac{(P_s Y_s)^{-\frac{1}{\sigma-1}}}{P_s}. \quad (19)$$

Noting that κ_s is a scaling constant by industry and does not affect the relative differences between firms within industry s, it can be normalized to unity (i.e., $\kappa_s = 1$). This manipulation enables us to estimate TFPQ without firm-level price data. Note that from equations (5) and (19), $TFPQ_{si} > TFPR_{si}$ if $\kappa_s = 1$ and $P_{si}Y_{si} \geq 1$. Therefore, in the Hsieh and Klenow (2009) framework the dispersion of TFPQ tends to be larger than that of TFPR.

2.3 Data

2.3.1 Source

This chapter utilizes firm-level data from the *Annual Survey on Enterprises* collected by the General Statistics Office (GSO) of Viet Nam.[12] The survey was conducted for the first time in 2000 and then annually thereafter to provide researchers and policy-makers with comprehensive information on Vietnamese firms. These data cover registered firms operating in all industries, including agriculture, industry and construction, and services.

The survey covers all state-owned enterprises and foreign-owned firms without any firm size threshold. However, as for domestic private firms, those with fewer than 10 workers are chosen by random sampling.[13] Household business activities are also not covered in this survey.[14] The survey information includes the type of ownership, assets and liabilities, number of employees, sales, capital stock, the industry that the firm belongs to, and obligations to the government (for example, taxes) from January to December of that year.

The data have some disadvantages. Some of the input data, such as materials, are only available for some years. Information on working hours and capital utilization rates is also unavailable. Firms'

[12] I use the same data used in Ha and Kiyota (2014). This section is based on Section 3 of Ha and Kiyota (2014). Note also that the use of firm-level data is more consistent with the theory than the use of plant-level data. This is because, as Nishimura et al. (2005) pointed out, resource allocation within a firm is determined by managerial decisions. Moreover, research and development or headquarter activities are typically classified as service activities, which are not covered in the manufacturing survey.

[13] This threshold was used in surveys before 2010. From 2010, different regions set different firm size thresholds.

[14] The survey covered 62.2 percent of total employment in manufacturing in 2009. The data on total employment in manufacturing are obtained from the GSO online database on population and employment at http://www.gso.gov.vn

year of establishment and export status are not available every year. This chapter uses firms with information on inputs, outputs, and cost shares. There are some reentry firms that disappeared and then reappeared later, which are omitted from the analysis. Some firms changed industry and/or ownership during the sample period.[15] I drop firms with fewer than 10 employees, regardless of their ownership, to avoid the effects of the random sampling.

2.3.2 Variables and parameters

The main variables that I use are the two-digit Viet Nam Standard Industry Classification (VSIC) industry code, ownership type, value added, employment, total labor costs, and capital stock. Following Hsieh and Klenow (2009), I use wage bills instead of the number of workers to capture the potential differences in employee quality.[16] Capital stock is measured as total fixed assets recorded at the end of each year. Both wage bills and capital stock are deflated by the manufacturing GDP deflator.[17]

To compute dispersion, I follow other research in setting the key parameters σ and R as follows. I assume that the elasticity of substitution σ equals 3 and R is 10 percent, comprising a 5 percent depreciation rate and a 5 percent interest rate. I also follow Hsieh and Klenow (2009) to set α_s equal to one minus the labor share in the corresponding industry in the United States. Under Hsieh and Klenow's framework, the output elasticities of capital and labor (i.e., α_s and $1 - \alpha_s$) do not embed distortions. Given the assumption that the United States economy is less distorted than the Vietnamese economy, the use of the United States shares can be justified.

The United States labor share is obtained from the NBER-CES Manufacturing Industry Database, which is a joint product of the National Bureau of Economic Research and the United States Census Bureau's Center for Economic Studies.[18] Industry classifications are based on the North American Industry Classification System (NAICS) version 1997. Based on the data, I first match the NAICS code with the four-digit VSIC code using concordance tables between NAICS, International Standard

[15] If a firm has switched industry, the industry to which the firm belonged for the majority of the surveyed years is regarded as that firm's industry. If a firm belonged to more than one industry for equal amounts of time, I assign the industry code of the industry that the firm belonged to most recently.
[16] The use of wage bills as a measure of labor input implies that $w = 1$. See Camacho and Conover (2010, p. 10).
[17] As Aw et al. (2001) pointed out, it is preferable to utilize the investment goods price deflator rather than the manufacturing GDP deflator to obtain the real capital stock. However, as Ha and Kiyota (2014) discussed, the investment goods price deflator is not available for my data set.
[18] Data can be downloaded from the NBER's website at http://www.nber.org/nberces/

Industry Classification revision 3, and VSIC. I then aggregate total payroll and total value added by two-digit VSIC sectors. To compute the labor share, I take the ratio of total payroll over total value added by sector. Because total payroll in the database does not include fringe benefits and employer's contribution to social security, this labor share only reflects two-thirds of the aggregate labor share in the whole manufacturing sector. Therefore, I follow Hsieh and Klenow (2009) to inflate the labor shares by 3/2 to obtain United States labor elasticities.

As firms' output prices are not available, I have obtained TFPQ by raising nominal output to the power of $\sigma/(\sigma-1)$, assuming that normal demand relationships hold. If a firm's real output is high, one would expect its price to be low so that consumers demand more output. Following Ziebarth (2013), the dispersion of TFP is defined as the deviation of the log of TFP from its industry mean: $\log(TFPR_{si}/\overline{TFPR}_s)$ and $\log\left(TFPQ_{si}\cdot M_s^{\frac{1}{\sigma-1}}/\overline{TFPQ}_s\right)$, where $\overline{TFPR}_s$ and $\overline{TFPQ}_s$ are from equations (16) and (18), respectively.[19] I trim 2 percent of firm productivity and distortions by removing values below the 1st percentile and above the 99th percentile from the distribution of $\log(TFPR_{si}/\overline{TFPR}_s)$ and $\log\left(TFPQ_{si}\cdot M_s^{\frac{1}{\sigma-1}}/\overline{TFPQ}_s\right)$. Then, I recalculate $\overline{TFPR}_s$, $\overline{TFPQ}_s$ and $\overline{TFP}_s$. As robustness checks, Section 2.5 examines whether the results are sensitive to the values of σ, α_s, and the threshold level of trimming.

2.4 Results

2.4.1 To what extent are resources misallocated in Viet Nam?

This section addresses the first question: To what extent are resources misallocated in Viet Nam? To answer this question, I compare the dispersions of TFP in Viet Nam with those in People's Republic of China, India, Japan, Thailand, and the United States. The dispersions of TFPR are reported in Table 2:1, while those of TFPQ are reported in Table 2:2. Both tables present standard deviations, differences between the 90th and 10th percentiles, differences between the 75th and 25th percentiles, and average per capita GDP during the sample period.[20] Figures for People's Republic of China, India,

[19] Note that some of the effects of the changes in prices are controlled for by taking the ratio.

[20] Noting that both TFPR and TFPQ are divided by their industry means, these statistics can be interpreted as the coefficients of variation.

and the United States are directly retrieved from Hsieh and Klenow (2009); for Japan from Hosono and Takizawa (2013); and for Thailand from Dheera-Aumpon (2014).

Table 2:1 Dispersion of TFPR in People's Republic of China, India, Japan, Thailand, the United States, and Viet Nam

	Viet Nam 2000–2009	Thailand 2006	People's Republic of China 1998–2005	India 1987–1994	Japan 1981–2008	United States 1977–1997
SD	0.79	0.85	0.68	0.68	0.55	0.45
75–25	0.97	1.04	0.89	0.80	0.70	0.47
90–10	2.00	2.09	1.72	1.66	1.40	1.08
GDP per capita	685	2,813	1,304	400	31,101	30,533

Note: Figures for Thailand are directly retrieved from Dheera-Aumpon (2014, Table 3); for People's Republic of China from Hsieh and Klenow (2009, Table 2, arithmetic averages); for Japan from Hosono and Takizawa (2013). TFPR is calculated from equation (5) and then scaled by the geometric mean of $TFPR_{si}$ across all firms in an industry s. Industries are weighted by value added shares. For more details, see the main text. GDP per capita is the annual average over each sample period (constant 2005 United States dollars).

Source: Hsieh and Klenow (2009), Hosono and Takizawa (2013), Dheera-Aumpon (2014), and authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam. Per capita GDP is obtained from World Bank (2014).

Table 2:2 Dispersion of TFPQ in People's Republic of China, India, Japan, Thailand, the United States, and Viet Nam

	Viet Nam 2000–2009	Thailand 2006	People's Republic of China 1998–2005	India 1987–1994	Japan 1981–2008	United States 1977–1997
SD	1.42	1.59	1.00	1.19	0.98	0.83
75–25	2.01	2.18	1.34	1.56	1.27	1.16
90–10	3.70	4.12	2.57	3.03	2.48	2.15

Note: Figures for Thailand are directly retrieved from Dheera-Aumpon (2014, Table 2); for People's Republic of China, India, and the United States from Hsieh and Klenow (2009, Table 1, arithmetic averages); for Japan from Hosono and Takizawa (2013, Table 1). TFPQ is calculated from equation (19) and then scaled by the geometric mean of $TFPQ_{si}$ across all firms in an industry s. Industries are weighted by value added shares. For more details, see the main text.

Source: Hsieh and Klenow (2009), Hosono and Takizawa (2013), Dheera-Aumpon (2014), and authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam.

These tables indicate that the standard deviation of TFPR for Viet Nam is 0.79, which is comparable to those for People's Republic of China (0.68), India (0.68), and Thailand (0.85), and is larger than those for Japan (0.55) and the United States (0.45). Similar patterns are also confirmed for the differences between the 75th and 25th percentiles and those between the 90th and 10th percentiles.[21] Although more careful examination is needed in the form of a direct comparison, the results suggest that distortions in developing countries, including Viet Nam, tend to be large relative to those in developed countries.

2.4.2 How large would the productivity gains be without distortions?

This section addresses the second question of this chapter: How large would the productivity gains have been in the absence of distortions? To answer this question, I estimate TFP gains when the marginal products of labor and capital are equalized across firms within each industry. For each industry, the gains are computed as the ratio of actual TFP obtained from equation (15) to the "efficient" TFP obtained from equation (18). I then aggregate the gains across industries using industry value added shares as the weights. In particular, I compute:

$$\begin{aligned} \frac{Y}{Y^*} &\triangleq \prod_{s=1}^{S}\left(\frac{Y_s}{Y_s^*}\right)^{\theta_s} = \prod_{s=1}^{S}\left(\frac{TFP_s}{\overline{TFPQ_s}}\right)^{\theta_s} \\ &= \prod_{s=1}^{S}\left\{\frac{1}{\overline{TFPQ_s}}\left[\sum_{i=1}^{M_s}\left(TFPQ_{si}\frac{\overline{TFPR_s}}{TFPR_{si}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}}\right\}^{\theta_s} \\ &= \prod_{s=1}^{S}\left[\sum_{i=1}^{M_s}\left(\frac{A_{si}}{\bar{A}_s}\frac{\overline{TFPR_s}}{TFPR_{si}}\right)^{\sigma-1}\right]^{\frac{\theta_s}{\sigma-1}}, \end{aligned} \tag{20}$$

where Y^* is the "efficient" output that corresponds to the "efficient" TFP and θ_s is the value added share of industry s ($\sum_s \theta_s = 1$). The first equality (i.e., $Y_s/Y_s^* = TFP_s/\overline{TFPQ_s}$) is obtained when K_s

[21] The difference between the 75th and 25th percentile firms is 0.97, which corresponds to a TFP ratio of $e^{0.97} = 2.63$. Similarly, the difference between the 90th and 10th percentile firms is 2.00, which corresponds to a TFP ratio of $e^{2.00} = 7.39$. These figures are much larger than those of the United States. For more details, see Syverson (2011).

and L_s are given. As the total amount of inputs is fixed, the output gains come solely from the reallocation of resources in the absence of distortions.

Table 2:3 presents the TFP gains from equalizing TFPR across firms within each industry. The gains are measured relative to the TFP gains in the United States in 1997.[22] To report the percentage TFP gains in Viet Nam relative to those in the United States, I take the ratio of Y^*/Y to the United States equivalent in 1997, subtract 1, and multiply by 100. If Viet Nam hypothetically moves to "United States efficiency," substantial gains are expected: 30.7 percent. The gains are smaller than those for People's Republic of China (39.2 percent), India (46.9 percent), and Thailand (73.4 percent), but larger than those for Japan (3.0 percent). One may be concerned that the dispersion of TFPR is larger (Table 2:1), whereas the gains are smaller (Table 2:3) in Viet Nam than in People's Republic of China and India. Noting that the gains are computed from the inverse of equation (20) (i.e., $(Y^*/Y - 1) \times 100$), Y^*/Y will be small if $A_{si}/\bar{A}_s$ and/or $\overline{TFPR}_s/TFPR_{si}$ become large. The results suggest that, on average, $A_{si}/\bar{A}_s$ is larger in Viet Nam than in People's Republic of China and India. Similarly, I find large TFP gains for Thailand, which is possibly attributed to a small $A_{si}/\bar{A}_s$ for Thailand.[23] Although these are hypothetical exercises and thus should not be taken literally, the results suggest that substantial productivity gains are expected in Viet Nam by the kind of reallocation considered here.

Table 2:3 TFP gains from equalizing TFPR relative to 1997 United States gains

	Viet Nam 2000–2009	Thailand 2006	People's Republic of China 1998–2005	India 1987–1994	Japan 1981–2008
%	30.7	73.4	39.3	46.9	3.0

Note: The data for Thailand are calculated from Dheera-Aumpon (2014, Table 4). The data for People's Republic of China, India, and the United States are arithmetic averages of Hsieh and Klenow (2009, Table 6). The data for Japan are calculated from Hosono and Takizawa (2013, Table 2).

[22] Hsieh and Klenow (2009) called this comparison a conservative analysis because the United States gains are largest in 1997.

[23] Indeed, Figure 1 in Dheera-Aumpon (2014) suggests that the distribution of TFPQ in Thailand moves to the left and its mean takes a negative value. Although it is not clear why the distribution moves to the left, this may be a reason why the large TFP gains are expected in Thailand.

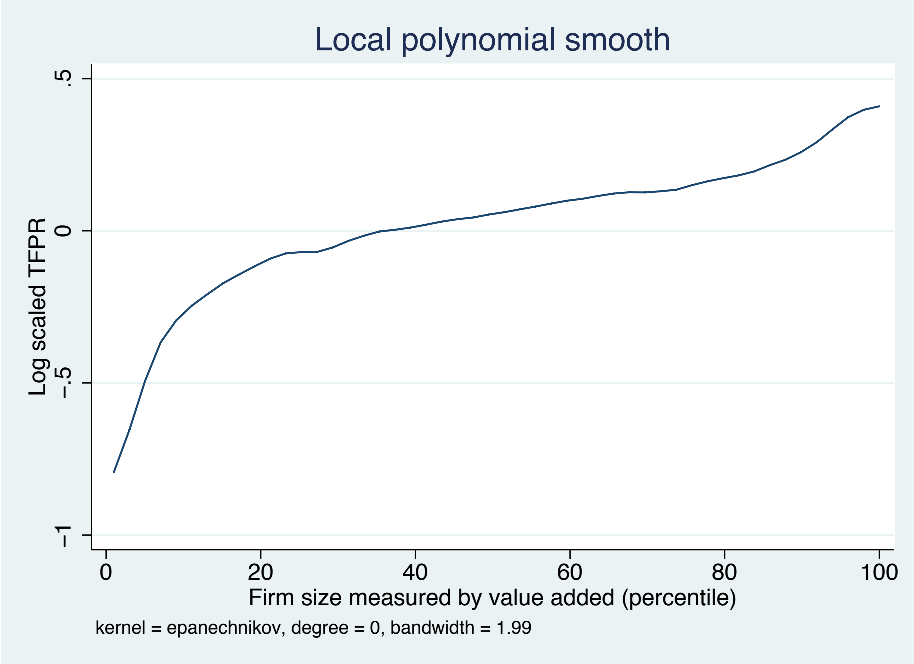
Source: Hsieh and Klenow (2009), Hosono and Takizawa (2013), Dheera-Aumpon (2014), and authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam.

2.4.3 Are the distortions related to firm size?

This section examines whether the distortions are related to firm size. This question has important policy implications because, for example, many countries give preferential treatment to SMEs. If SMEs tend to face larger disadvantageous distortions, preferential treatment to SMEs can be justified. Following Hsieh and Klenow (2009) and Ziebarth (2013), I examine the relationship between firm size and TFPR.

Figure 2:1 presents the relationship between firm size percentile as measured by value added and scaled TFPR relative to a given industry. Figure 2:1 indicates that TFPR is strongly increasing in percentiles of firm size. Noting that TFPR is proportional to the distortions (equation 14), this result implies that smaller firms tend to face advantageous distortions, whereas larger firms tend to face disadvantageous ones. This result is similar to that in India (Hsieh and Klenow, 2009, Figure 6) and the United States in the 19th century (Ziebarth, 2013, Figure 3).

Figure 2:1 TFPR and size

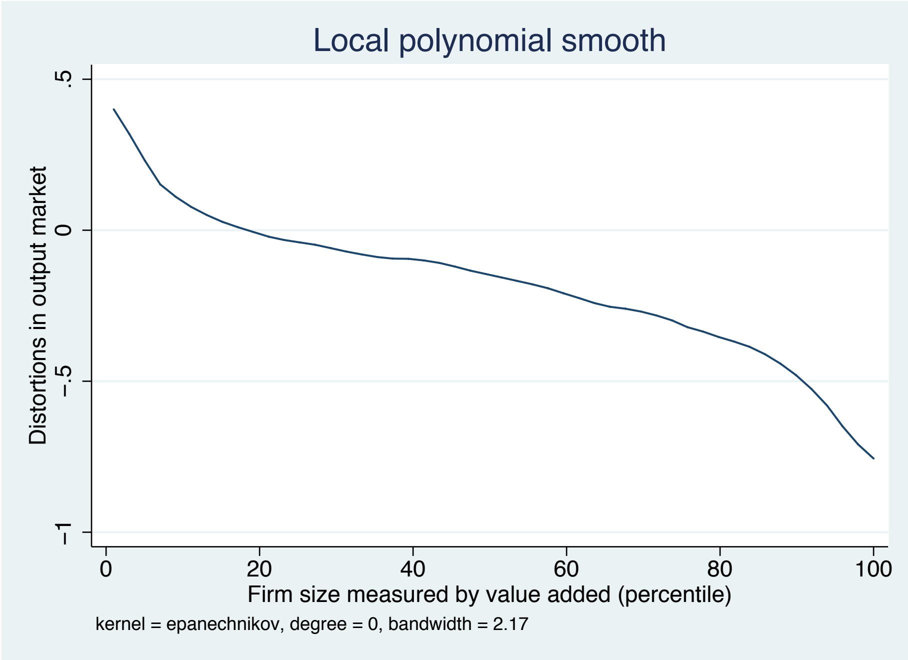


Note: This figure presents the relationship between scaled TFPR relative to a given industry and size percentile as measured by value added.

Source: Authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam.

Interestingly, this correlation with firm size is different for the distortions in output and the distortions in capital markets. Figure 2:2 presents the relationship between the distortions in output markets and firm size (in terms of value added). Figure 2:2 indicates that the distortions in output markets are strongly decreasing in percentiles of firm size. Noting that the distortions in output markets are measured by $(1 - \tau_Y)$, this result is similar to that in TFPR: smaller firms tend to face advantageous distortions, whereas larger firms tend to face disadvantageous ones.

Figure 2:2 Distortions in output markets and size

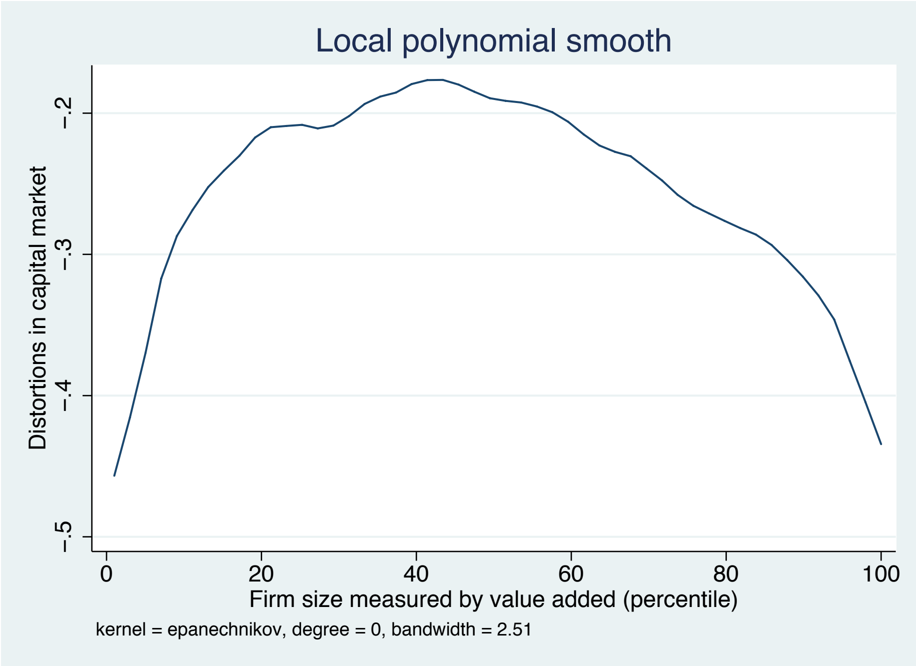


Note: This figure presents the relationship between scaled $1 - \tau_Y$ relative to a given industry and size percentile as measured by value added.

Source: Authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam.

Figure 2:3 presents the relationship between the distortions in capital markets and firm size. In contrast to the distortions in output markets, Figure 2:3 presents an inverse U-shaped relationship. Noting that the distortions in capital markets are measured by $(1 + \tau_K)$, this result suggests that both small and large firms tend to face advantageous distortions. In contrast, middle-sized firms tend to face disadvantageous distortions. This pattern is different from those of TFPR and distortions in output markets. This may be because small firms are treated preferentially, whereas large firms can diversify their capital procurement.

Figure 2:3 Distortions in capital markets and size



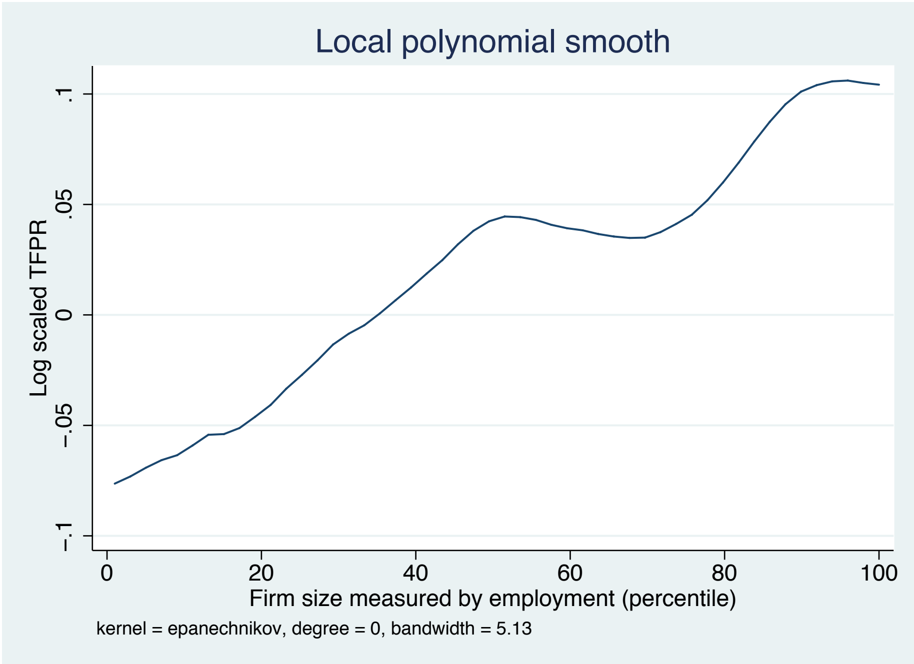
Note: This figure presents the relationship between scaled $1 + \tau_K$ relative to a given industry and size percentile as measured by value added.

Source: Authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam.

It is also interesting to note that the result for TFPR mainly reflects that of distortions in output markets. This result implies that the distortions in output markets have stronger effects on TFPR than those in capital markets. This result is consistent with the result of Midrigan and Xu (2014) which showed that financial frictions, measured by borrowing constraints, had relatively small impacts on productivity.

One may be concerned that the measurement of firm size, following Hsieh and Klenow (2009), is based on value added rather than employment. However, in reality, SMEs are defined by the number of employees, not by the size of their value added, in many countries. To address this concern, I examine the relationship between distortions and firm size measured by employment. The results are presented in Figures 2:4, 2:5, and 2:6. The results are different from, but qualitatively similar to, those measured by value added: TFPR is increasing in percentiles of firm employment size, the distortions in output markets are decreasing, and the distortions in capital markets are inverse U-shaped except for the top 20 percentiles. Noting that the results for TFPR mainly reflect the distortions in output markets, I can conclude that the main messages remain unchanged even when firm size is measured by employment.

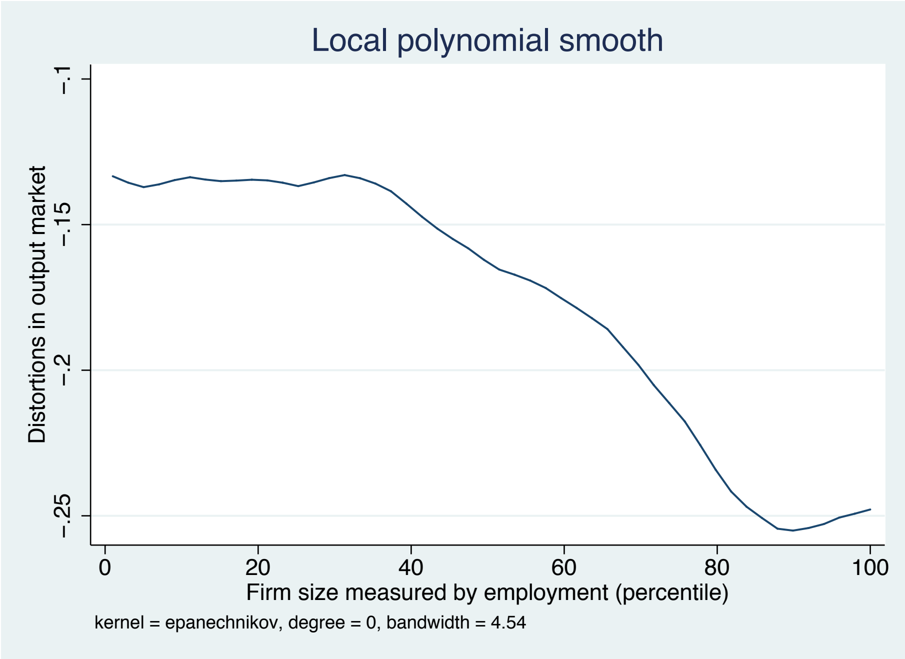
Figure 2:4 TFPR and employment size



Note: This figure presents the relationship between scaled TFPR relative to a given industry and size percentile as measured by employment.

Source: Authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam.

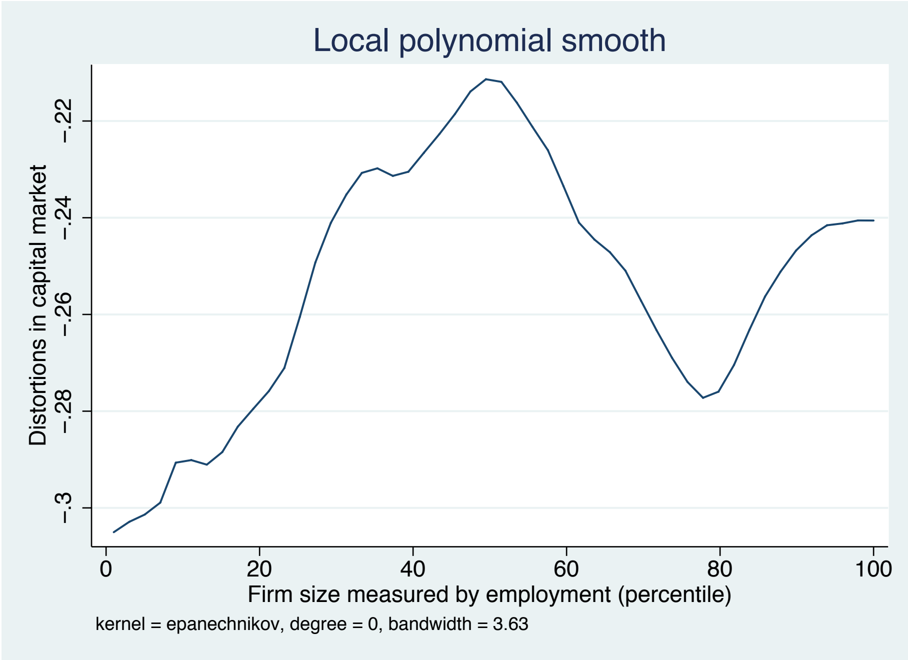
Figure 2:5 Distortions in output markets and employment size



Note: This figure presents the relationship between scaled $1 + \tau_Y$ relative to a given industry and size percentile as measured by employment.

Source: Authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam.

Figure 2:6 Distortions in capital markets and employment size



Note: This figure presents the relationship between scaled $1 + \tau_K$ relative to a given industry and size percentile as measured by employment.

Source: Authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam.

2.4.4 What would the distribution of firm size have been in the absence of distortions?

The model also has an implication for the distribution of firm size. Equation (7) is rewritten as:

$$P_{si}Y_{si} = Y_{si}^{\frac{\sigma-1}{\sigma}} P_s Y_s^{\frac{1}{\sigma}}. \tag{21}$$

From equations (7) and (9), I have:

$$Y_{si} = \left[\frac{\sigma-1}{\sigma}\left(\frac{\alpha_s}{R}\right)^{\alpha_s}\left(\frac{1-\alpha_s}{w}\right)^{1-\alpha_s}\right]^{\sigma} P_s^{\sigma} Y_s \left[\frac{A_{si}(1-\tau_{Ysi})}{(1+\tau_{Ksi})^{\alpha_s}}\right]^{\sigma}. \tag{22}$$

Similar to equation (14), from equations (21) and (22), I have:

$$P_{si}Y_{si} \propto \left[\frac{A_{si}(1-\tau_{Ysi})}{(1+\tau_{Ksi})^{\alpha_s}}\right]^{\sigma-1}. \tag{23}$$

Equation (23) suggests that without distortions, more (less) productive firms tend to be larger (smaller). When A_{si} and $1-\tau_{Ysi}$ are correlated negatively, more productive firms tend to be smaller than the efficient size. Similarly, if A_{si} and $1+\tau_{Ksi}$ are correlated positively, less productive firms tend to be larger than the efficient size. Both cases result in smaller size dispersion. This in turn implies that when distortions are large, the efficient size distribution is more dispersed than the actual size distribution.

To examine this implication, I compare the actual firm size distribution with the efficient firm size distribution. The size is measured as the value added of the firms, following Hsieh and Klenow (2009). Let $P^*_{si}Y^*_{si}$ be the efficient firm size. The efficient sizes relative to actual sizes are:

$$\frac{P^*_{si}Y^*_{si}}{P_{si}Y_{si}} = \frac{Y^*}{Y}\left(\frac{Y_s}{Y^*_s}\right)^{\sigma-1}\left[\frac{(1+\tau_{Ksi})^{\alpha_s}}{1-\tau_{Ysi}}\right]^{\sigma-1}, \tag{24}$$

where the efficient firm size is obtained when τ_{Ksi} and τ_{Ysi} are equalized within industry s; Y^*/Y and Y_s/Y^*_s are obtained from equation (20), respectively.[24] I compute the actual and efficient sizes from this equation, by year, and then take averages over the period.

Table 2:4 and Figure 2:7 present the results. In Table 2:4, the rows are the actual firm size quartiles with equal numbers of firms. The columns are the bins of efficient firm size relative to actual firm size. I classify firms into four bins. For example, 0%–50% means that the firm size would be less than half of the actual firm size if all distortions are removed. Similarly, 200+% means that the firm size would be more than double without distortions. The entries are the shares of firms (averaged over the period). The rows sum to 25 percent, and the rows and columns together to 100 percent.

Examining Table 2:4, I highlight two results. First, although average output rises substantially (as I confirmed in Section 2.4.2), many firms of all sizes would shrink. Second, the largest quartile indicates the largest expansion among the firm sizes (8.7 percent). This result means that initially large firms are less likely to shrink and more likely to expand. This finding is also confirmed from Figure 2:7.

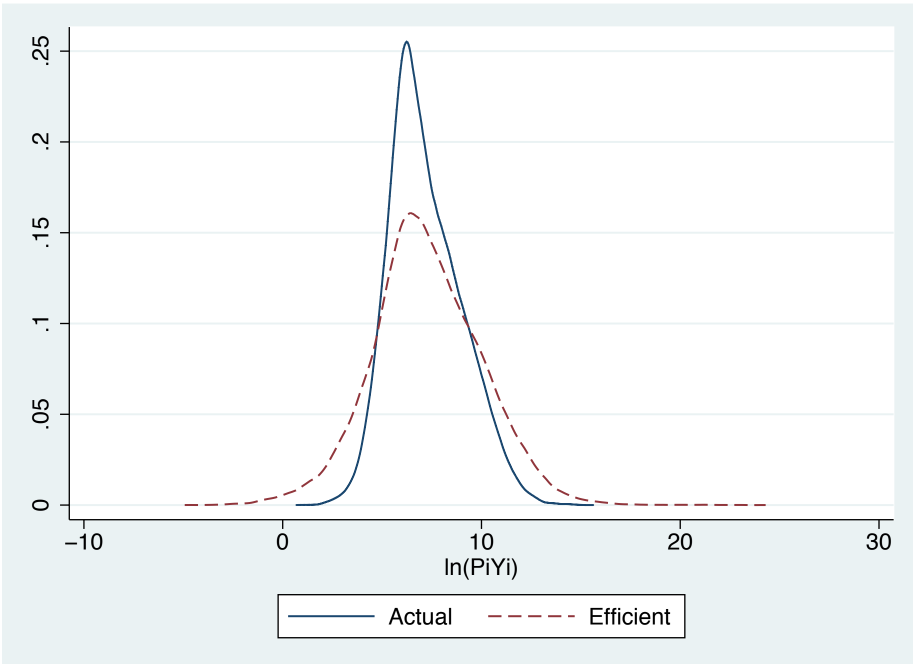
[24] For the derivation of equation (24), see the Appendix.

Table 2:4 Actual size vs efficient size

	Efficient firm size relative to actual firm size				
2000–2009 (average)	**0%–50%**	**50%–100%**	**100%–200%**	**200%+**	**Total**
Actual firm size					
Top quartile	5.1	5.5	5.7	8.7	25.0
Second quartile	8.0	5.6	4.6	6.8	25.0
Third quartile	9.1	6.3	4.4	5.2	25.0
Bottom quartile	13.7	5.1	3.0	3.1	25.0
Total	**36.0**	**22.4**	**17.6**	**23.9**	**100.0**

Notes: The rows are the actual firm size quartiles with equal numbers of firms. The columns are the bins of efficient firm size relative to actual firm size. I classify firms into four bins, by the value added of firms. For example, 0%–50% means that the firm size would be less than half of the actual firm size if all distortions were removed. Similarly, 200+% means that the firm size would be more than double without distortions. The entries are the shares of firms (averaged over the period).

Figure 2:7 Distribution of firm size



Note: The solid line indicates the actual size distribution, whereas the dashed line indicates the efficient size distribution.

Source: Authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam.

As the model suggests, the efficient size distribution is more dispersed than the actual size distribution. This result is consistent with the finding of the previous section. Like the case of India (Banerjee and Duflo, 2005, p.507), Viet Nam's policies may constrain its large and most efficient producers and coddle its small and least efficient ones. Indeed, Vietnamese SMEs are supported by various policies such as government supporting funds (Tran et al., 2008, pp. 347–359). These results for Viet Nam are similar to those for People's Republic of China, India, and the United States in Hsieh and Klenow (2009).[25]

2.4.5 Robustness check: different parameter values

One may be concerned that the analysis is sensitive to the choice of parameter values and sample selection because the results are based on specific parameter values such as $\sigma = 3$. To address this concern, I reconduct all the analyses using different parameter values. Because it is tedious to examine all the results, this section examines i) how sensitive the estimated TFPR and TFP gains (reported in Section 2.4.2 and in Table 2:3) are to the choice of parameter values and sample selection, and ii) the correlation between alternative and baseline TFPR. In this robustness check, I report absolute TFP gains rather than relative TFP gains (to the United States) because I only change the parameter values in Viet Nam (not in the United States).

We first examine whether the results are sensitive to the value of the elasticity of substitution: σ. In the baseline analysis, following Hsieh and Klenow (2009), I set $\sigma = 3$. This implies that the markup is 1.5 ($= 3/(3-1)$). As a robustness check, I set $\sigma = 2$ and $\sigma = 6$, and the corresponding markups are 2 ($= 2/(2-1)$) and 1.2 ($= 6/(6-1)$), respectively. The second and third columns in Table 2:5 present the results. The TFP gains are somewhat sensitive to the value of the elasticity of substitution. The TFP gains are 65.3 percent when $\sigma = 2$ and 161.9 percent when $\sigma = 6$, while the baseline TFP gains are 86.8 percent.[26]

[25] Indeed, the Vietnamese government had launched various schemes to improve the performance of SMEs, such as establishing credit funds and providing worker training (Tran et al., 2008, pp. 347–359). However, unlike India, where size-related policies are explicitly imposed by law, such policies in Viet Nam are only guidelines. I cannot identify from the data which individual firm is eligible for support or has received any form of support. It is thus difficult for us to conduct an analysis similar to Hsieh and Klenow (2009, Part VI).

[26] This result is consistent with equation (17), which implies that the TFP gains will be large if the elasticity of substitution is large.

Nevertheless, the estimated TFPR is qualitatively similar to the baseline results. Table 2:5 also reports the correlation with baseline TFPR, which is 0.997 when $\sigma = 2$ and 0.994 when $\sigma = 6$. These high correlations suggest that the results are quantitatively different from, but qualitatively similar to, the baseline results.[27] The standard deviation of lnTFPR is 0.78 when $\sigma = 2$ and 0.79 when $\sigma = 6$, both of which are similar to that of the baseline model (0.79).

Table 2:5 Robustness check: TFP gains from equalizing TFPR

	(1) Baseline	**(2) Robustness Check 1**	**(3) Robustness Check 2**	**(4) Robustness Check 3**	**(5) Robustness Check 4**	**(6) Robustness Check 5**	**(7) Robustness Check 6**	**(8) Robustness Check 7**
Elasticity: σ	$\sigma = 3$	$\sigma = 2$	$\sigma = 6$	$\sigma = 3$	$\sigma = 3$	$\sigma = 3$	$\sigma = 3$	$\sigma = 3$
Technology: α	United States	United States	United States	1/3	Viet Nam	Firm specific	United States	United States
Trim	1%	1%	1%	1%	1%	1%	2%	2%
N	100,601	100,601	100,612	100,848	100,832	100,947	97,263	10,186
SD(TFPR)	0.79	0.78	0.79	0.64	0.64	0.61	0.71	0.68
TFP gains (%)	86.8	65.3	161.9	70.1	68.0	40.0	75.7	64.5
Correlation with baseline TFPR	1.000	0.997	0.994	0.927	0.889	0.794	0.995	0.948
Panel structure	Unbalanced Panel	Unbalanced Panel	Unbalanced Panel	Unbalanced Panel	Unbalanced Panel	Unbalanced Panel	Unbalanced Panel	Balanced Panel

Note: The baseline is obtained from Table 2:3.

Source: Authors' calculations, based on the *Annual Survey of Enterprises* by the GSO of Viet Nam.

We also examine the sensitivity of the results to the value of the technology parameter (i.e., capital share α_s). I examine two different technologies. One is $\alpha_s = 1/3$ as in Ziebarth (2013) and the other is the Vietnamese cost share, which is defined as the industry- year average capital share of the sample firms. The results are presented in the fourth and fifth columns in Table 2:5. The TFP gains are 70.1 percent when $\alpha_s = 1/3$ and 68.0 percent when I assume Vietnamese technology. The correlation with the baseline TFPR is 0.927 when $\alpha_s = 1/3$ and 0.889 when I assume Vietnamese technology. The standard deviation of lnTFPR is 0.64 for both cases. Similar to the value of the elasticity of substitution, the results are quantitatively different from, but qualitatively similar to, the baseline results.

One may also be concerned that the technology parameter α_s is heterogeneous across firms even within industries. To address this concern, I use the firm-level capital share so that the capital share

[27] It may also be important to allow the elasticities to vary across industries. Although Broda et al. (2006) estimated the elasticity of substitution for various countries, Viet Nam is not covered in their analysis. I thus leave this exercise for future research.

can vary across firms.[28] The results are presented in the sixth column in Table 2:5, and are similar to the baseline results, although the TFP gains are somewhat sensitive to the technology parameters. The TFP gains are 40.0 percent. The correlation with the baseline TFPR is 0.794. The standard deviation of lnTFPR is 0.61. These results together suggest that the main messages remain unchanged even when I use different values for the technology parameter.

Another concern may be that the data are not precise, and thus Vietnamese firm-level data are subject to measurement error problems. Although I cannot rule out arbitrary measurement error, I can try to gauge whether the results are attributable to some specific forms of measurement error. I focus on two forms of measurement error. First, serious measurement error, possibly because of reporting error, tends to appear as outliers. I trimmed 2 percent from the tails (below the 2nd percentile and above the 98th percentile), instead of 1 percent as in the baseline analysis, and examined how sensitive the results are to the trim values. The seventh column reports the results. The TFP gains are 75.7 percent. The correlation with the baseline TFPR remains high at 0.995. The standard deviation of lnTFPR (0.71) is slightly lower than that of the baseline model (0.79).

We also estimate the TFP gains for firms that survived throughout the sample period (i.e., balanced panel). This exercise enables us to control for the effects of firm entry and exit. The eighth column presents the results. This exercise reduces the sample size substantially (N = 10,186). Nonetheless, the estimated TFP gains are large and the correlation with baseline TFP is high: 64.5 percent and 0.948, respectively. The standard deviation of lnTFPR is 0.68, which is comparable to that of the baseline model. The results suggest that about three-quarters (=64.5%/86.8%) of TFP gains come from the incumbent firms, while the rest of the gains come from entrants and exiters. I can thus conclude that the results from the balanced panel are qualitatively similar to the baseline results.

In sum, the magnitude of the TFP gains are somewhat sensitive to the choice of the values of parameters σ and α. Nonetheless, the main messages remain unchanged even if I use different parameter values or I employ different sample selection criteria: potential TFP gains from removing distortions are large in Vietnamese manufacturing.

[28] Note that ξ_s can vary across firms if the capital share is different across firms (see equation (12)). In other words, TFPR will not necessarily be proportional to the capital and output wedges. I thus present the results for reference only. Note also that, from equation (11), if the technology parameter is heterogeneous across firms (i.e., α_s $(= RK_{si}/P_{si}Y_{si})$, distortions appear only in τ_{Ysi} because τ_{Ksi} will be zero.

2.5 Conclusion

This chapter employed the Hsieh and Klenow (2009) framework to investigate misallocation and productivity linkages in Vietnamese manufacturing during the period 2000–2009 using firm-level data. The study has four major findings. First, misallocation in Viet Nam is comparable to that in People's Republic of China, India, and Thailand. This result is consistent with the common knowledge that resources in developing countries are not efficiently allocated.

Second, there would be substantial improvement in TFP if no distortions existed. If Viet Nam hypothetically moved to "United States efficiency," its TFP would be boosted by 30.7 percent. Third, smaller firms tend to face advantageous distortions, whereas larger firms tend to face disadvantageous ones. Finally, the efficient distribution of firm size is more dispersed than the actual size distribution. This result implies that Viet Nam's policies may constrain its large and most efficient producers and coddle its small and least efficient ones.

These findings have policy implications. The first finding suggests that, similar to other developing countries, resource misallocation, which is caused by the distortions, seems to be an important issue in Viet Nam. The second finding states that potential productivity gains from removing distortions are large in Vietnamese manufacturing. The result implies that reallocation would lead to productivity gains that would accelerate potential growth in transition towards the improved inter-firm resource allocation. The last two findings together imply that Viet Nam's policies may constrain its large and most efficient producers and coddle its small and least efficient ones. These results together suggest that policy-makers need to focus more on the allocation of resources. An important question, therefore, is whether or not the resources are allocated to productive firms.

2.6 Appendix

In this Appendix, I provide the derivation of equation (24). From equations (7), (8), and (9), actual firm size is written as:

$$\begin{aligned} P_{si}Y_{si} &= P_s^{\sigma} Y_s P_{si}^{1-\sigma} \\ &= P_s Y_s \left(\frac{P_{si}}{P_s}\right)^{1-\sigma} \\ &= \theta_s Y \left[\frac{(1+\tau_{Ksi})^{\alpha_s}}{A_{si}(1-\tau_{Ysi})}\right]^{1-\sigma} \Bigg/ \sum_j \left[\frac{(1+\tau_{Ksj})^{\alpha_s}}{A_{sj}(1-\tau_{Ysj})}\right]^{1-\sigma}. \end{aligned} \tag{A-1}$$

Efficient firm size is obtained when τ_{Ksi} and τ_{Ysi} are equalized within industry s (e.g., $\tau_{Ksi} = \tau_{Ks}$ and $\tau_{Ysi} = \tau_{Ys}$). From equation (A-1), the efficient firm size is written as:

$$P_{si}^{*}Y_{si}^{*} = \theta_s Y^{*}\frac{A_{si}^{\sigma-1}}{\sum_j A_{sj}^{\sigma-1}}. \tag{A-2}$$

From equations (A-1) and (A-2), I have:

$$\frac{P_{si}^{*}Y_{si}^{*}}{P_{si}Y_{si}} = \frac{Y^{*}}{Y}\left(\frac{Y_s}{Y_s^{*}}\right)^{\sigma-1}\left[\frac{(1+\tau_{Ksi})^{\alpha_s}}{1-\tau_{Ysi}}\right]^{\sigma-1}. \blacksquare \tag{A-3}$$

Chapter 3 Competition, Uncertainty, and Productivity Dispersion[1]

Uncertainty affects investment that involves adjustment costs or time-to-build, resulting in dispersion in marginal revenue productivity of capital (MRPK) and consequently in aggregate productivity, depending on the degree of product market competition. Using a simple dynamic model and a large panel dataset of manufacturing plants in Japan, I find that the effect of uncertainty on the dispersion in MRPK is stronger for industries with severer product market competition. The counterfactual experiment shows that if the volatility of revenue-based productivity (TFPR) shocks decreases by half, the dispersion in MRPK decreases by 2.4% and the aggregate productivity increases by 0.7% on average for all industries. For the more competitive industries, the effect on the aggregate productivity is as high as 2.1%.

3.1 Introduction

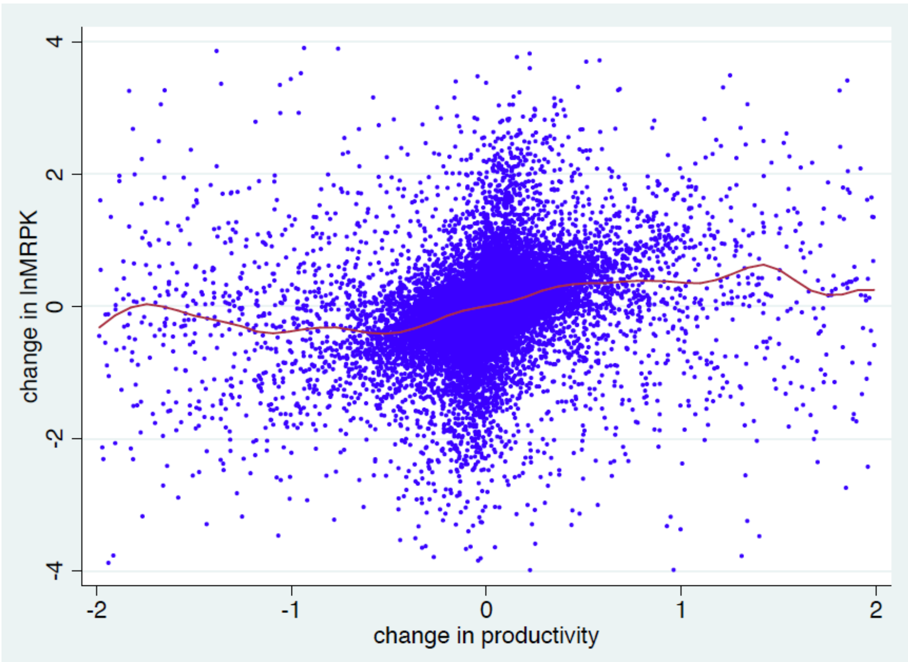
It is now well known that the dispersion in (revenue-based) productivity across firms or plants is quite large even within narrowly-defined industries. However, the reasons for the dispersion in productivity are still controversial. Many preceding studies posit that the distortions such as taxes, regulations, financial frictions, and markups that vary across producers result in the dispersion in productivity, which, in turn, cause misallocation of resources and lower aggregate productivity (Restuccia and Rogerson, 2008 Hsieh and Klenow, 2009). On the other hand, more recent studies show that uncertainty results in the dispersion in revenue-based productivity as it tends to affect investment that involves adjustment costs or time-to-build. Asker, Collard-Wexler, and De Loecker (2014: ACL hereafter), among others, show that higher time-series volatility in productivity shocks results in the greater dispersion in the marginal revenue of capital (MRPK) among plants within an industry, even if capital is allocated efficiently from the dynamic view when adjustment costs are

[1] The revised version of this chapter is in RIETI Discussion Paper 17-E-071, as “Competition, Uncertainty, and Misallocation”, co-authtored with Kaoru Hosono and Miho Takizawa.

considered. Hopenhayn (2014), in his survey of misallocation, however, stresses that there are no conclusive results as to the importance of adjustment costs as a source of misallocation.

Given such controversy, it is useful to provide new evidence on the role of uncertainty in the dispersion in revenue-based productivity. Figure 3:1 shows the cross-sectional relationship between the plant-level annual changes in the marginal revenue of capital (MRPK) and in the revenue-based productivity (TFPR) in Japan. The figure indicates that these two are positively correlated, suggesting that plants do not adjust capital in response to the TFPR shock within a year.[2]

Figure 3:1 Establishment-level changes in TFPR and MRPK



Note: This figure depicts changes in TFPR and MRPK over the period from 2012 to 2013.

Source: Authors' calculations, based on the Census of Manufacture by METI.

While uncertainty is likely to affect the dispersion in MRPK, the impact of uncertainty on investment, and hence on the dispersion in MRPK and aggregate productivity, may depend on the

[2] I describe the data in Section 3.4.1 and the definition of TFPR and MRPK in Section 3.4.2 in detail. Figure 3:1 depicts the plant-level changes in TFPR and MRPK over the period from 2012 to 2013, although I observe a similar correlation between the two over the other years in the sample.

degree of competition in the product market. Extant theoretical studies show that the effect of uncertainty on investment could be positive or negative depending on the product market competition as well as on the degree of returns to scale and the adjustment-cost asymmetries (Caballero, 1991). Theoretical studies using an options-game also posit that uncertainty is less likely to delay investment in a more competitive product market. Empirical results on the relationship between uncertainty and investment are mixed as well.[3]

This study aims at providing new evidences on to what extent uncertainty affects dispersion in MRPK and aggregate productivity depending on the product market competition. The main contribution to the extant literature is to take competition into account when I estimate the effect of uncertainty on the dispersion in MRPK. In addition, unlike the preceding studies, I estimate the effects of uncertainty on aggregate TFP as well as the dispersion in MRPK.

We use a simple dynamic model and a large dataset of manufacturing plants in Japan covering 1986 to 2013. I find that while industries with greater time-series volatility in TFPR have greater cross-sectional dispersion of MRPK, which is consistent with ACL, such an impact is stronger for the industries where product market competition is severer. The counterfactual experiment shows that if the volatility of revenue-based productivity (TFPR) decreases by half, the dispersion in MRPK decreases by 2.4% and the aggregate productivity increases by 0.7% on average for all industries. For the more competitive industries, the effect on the aggregate productivity is as high as 2.1%.

The reminder of the chapter proceeds as follows. In Section 3.2, I review the relevant literature on the impact of uncertainty on investment and dispersion in MRPK. In Section 3.3, I present a simple model to show how competition affects the relationship between volatility and dispersion in MRPK. Section 3.4 describes the dataset and methodology. Section 3.5 presents the results. Finally, I conclude the chapter with discussion of the findings in Section 3.6.

3.2 Literature review

This study investigates the effects of competition on the relationship between uncertainty and dispersion in (revenue-based) productivity. I particularly focus on the mechanism through producers' investment decision. Two strands of literature are therefore related to the study.

[3] See Bloom (2014) for a survey.

The first strand discusses the effect of uncertainty on investment. Extant theoretical studies show that the effect of uncertainty on investment depends on the product market competition as well as on the degree of returns to scale and the adjustment-cost asymmetries. Real options theory predicts that greater uncertainty should lower investment if investment is irreversible, product market is competitive, and production technology exhibits constant returns to scale (McDonald and Siegel 1986; Pindyck 1988; Bertola 1988).[4] However, uncertainty could increase investment if competition is nearly perfect and technology exhibits increasing returns to scale even under the assumption that investment is irreversible (Abel 1983; Caballero, 1991; Bar-Ilan and Strange 1996).[5] Caballero (1991) introduced the effect of competition on the uncertainty-investment relationship into the real options framework by presenting a theoretical model to show that the relationship between price uncertainty and capital investment is not robust. The negative effects require both market power and asymmetric capital adjustment costs. Another strand investigates options exercise games. Williams (1993), Kulatilaka and Perotti (1998), and Grenadier (1996, 2002) find that option values erode under fierce competition because competitors may preempt the investment opportunity. The fear of preemption leads to early investment. Theoretical studies using an options-game also posit that uncertainty is less likely to delay investment in a more competitive product market. Empirical results on the relationship between uncertainty and investment are mixed as well, although most find negative effects from uncertainty on investment.[6] Bloom, Bond, and Van Reenen (2007), for example, find that sales growth has a smaller effect on investment for publicly traded UK firms when the firms face higher volatility in stock returns.[7]

Several empirical studies provide support for the theoretical prediction that competition mitigates the negative effects of uncertainty on investment, while other studies obtain the opposite results. Porter and Spence (1982) conduct an early case study of preemptive investment in the wet corn milling industry. Guiso and Parigi (1999) explore the effect of uncertainty on investment using a measure of uncertainty based on the information on the subjective probability distribution of future demand with a database of Italian manufacturing firms. The negative effect of uncertainty on investment is smaller for firms with little market power measured by the profit margin. Using a panel data set of Italian firms, Bontempi, Golinelli, and Parigi (2010) show that an increase in competition weakens

[4] Dixit and Pindyck (1994) is an excellent introduction to real options theory.
[5] Oikawa (2010) provides a model in which firm-level uncertainty raises aggregate productivity growth.
[6] See Bloom (2014) for a survey. Bloom (2014) notes that the uncertainty literature provides "suggestive but not conclusive evidence that uncertainty damages short-run (quarterly and annual) growth, by reducing output, investment, hiring, consumption, and trade" (p.168).
[7] See Ogawa and Suzuki (2000) and Mizobata (2014) for cases in Japan.

the effect of uncertainty on investment decisions. Bulan (2005) uses a panel of U.S. manufacturing firms to explore the effects of uncertainty measured as the volatility of firms' equity returns. The study splits the sample by industry concentration ratios and finds that competition reduces the negative impact of uncertainty on investment. Bulan, Mayer, and Somerville (2009) address the same issue using the data on real estate projects in Vancouver, Canada. Their main finding is that increases in volatility lead developers to delay new real estate development, but the impact is smaller when the number of potential competitors is large. Akdogu and MacKay (2008) split a sample into three groups according to the Herfindahl-Hirschman index (HHI) and provide evidence that firm investment is less sensitive to changes in Tobin's q in monopolistic industries than in competitive industries. On the other hand, Ghosal and Loungani (2000) show the opposite result, finding that the negative effect of uncertainty on investment is higher in competitive industries. In their studies, competition is measured as the four-firm seller concentration ratio. While the results are mixed, all of these studies focus on investment at the firm or project level without considering the effect on the dispersion in MRPK at the industry level.

The second related strand investigates resource misallocation across firms and plants. Restuccia and Rogerson (2008) first establish the mechanism by which factor price distortion at the firm level reduces allocative efficiency in the aggregate economy. They calibrate U.S. data to show the large effect of resource misallocation. Hsieh and Klenow (2009) incorporate monopolistic competition into Restuccia and Rogerson's (2008) model. In Hsieh and Klenow's (2009) framework, resource misallocation depends on the dispersion of marginal revenue products. They find that the degrees of resource misallocation are larger in China and India than in the U.S.

A number of studies follow Hsieh and Klenow (2009) to specify the underlying mechanisms of the dispersion of marginal revenue products.[8] ACL is one such study, which investigates the role of productivity shocks and dynamic production factors on the static variation of marginal revenue.[9] ACL use a dynamic investment model to replicate the observed patterns in the large dispersion of MRPK. In the reduced-form estimation with nine datasets spanning 40 countries, ACL show that the higher time-series volatility of productivity shocks, measured as the variance of productivity growth rates across firms, contribute to larger resource misallocation within industries measured as the cross-

[8] See Hopenhayn (2014) and Restuccia and Rogerson (2017) for a survey. Andrews and Cingano (2014) empirically study the effects of various kinds of policies on allocative efficiency.
[9] Da Rocha and Pujolas (2011) also explore the effect of productivity shocks on resource misallocation theoretically.

sectional dispersion of MRPK. Their result suggests that welfare gains from reallocating production factors are not as large as implied by static models.

Many studies investigate capital misallocation arising from capital market frictions (Banerjee and Moll 2010; Midrigan and Xu 2014; Moll 2014), but do not focus on the roles of uncertainty and competition. Other researchers study the effect of competition on resource misallocation. Edmond, Midrigan, and Xu (2015) show that trade-induced competition causes markup harmonization through reduced market power in a monopolistic competition framework with a finite number of firms. Unlike them, I focus on the dynamic aspect of competition through uncertainty.

While existing studies reveal the various factors, including uncertainty, that induce dispersion in MRPK and TFPR, to my knowledge, no studies consider the possibility that the impact of uncertainty on the dispersion in MRPK depends on the degree of competition in the product market. This study therefore shows how product market competition affects the impact of uncertainty on the dispersion in MRPK and aggregate productivity.

3.3 Theoretical framework

This section provides a simple model to investigate the effects of the degree of competition on the relationship between the volatility of revenue productivity and MRPK dispersion. I follow ACL's dynamic investment model, which incorporates time-to-build and the adjustment cost of capital and explain their model, explicitly describing some settings they implicitly imposed and introducing the asymmetric adjustment cost between a positive and negative investment considering that preceding studies point out the importance of such an asymmetric adjustment cost in generating a negative uncertainty-investment relationship (Caballero, 1991, among others). I further explicitly introduce demand shocks to make it clear that TFP shocks are proportional to the sum of technology and demand shocks. This section starts with a static model and extending it to a dynamic setting. I then provide a simplified tractable model to show the relationship between the productivity shock and aggregate TFP. Finally, I present simulation results.

3.3.1 Static model

Following ACL, I model a profit maximizing plant facing demand (or quality) and productivity shocks. This static model is the base of the dynamic model in the next subsection as well as that of the measurement in Section 3.4.

There are a unit mass of intermediate good producers i, which I call a plant hereafter and a final good producer. The final good producer combines differentiated product Q_{it} to produce output Q_t using a constant elasticity of substitution (CES) production technology at time *t*:

$$Q_t = \left(\int (B_{it} Q_{it})^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{1}$$

where B_{it} denotes a quality shock. The demand functions for the plants is derived as follows:

$$Q_{it} = B_{it}^{\varepsilon-1} P_{it}^{-\varepsilon}. \tag{2}$$

Note that the demand is subject to shock B_{it}:[10] ε denotes demand elasticity and serves as the degree of competition and the implied markup, $\frac{1}{1-\frac{1}{\epsilon}}$, as the inverse degree of competition.

The production function of the plants is assumed to be following Cobb-Douglas type with constant returns to scale:

$$Q_{it} = A_{it} K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it}^{\alpha_M}, \tag{3}$$

where A_{it} is the (physical) productivity shock, K_{it} is the capital, L_{it} is the labor input, M_{it} is materials, and $\alpha_K + \alpha_L + \alpha_M = 1$.

The sales-generating production function can be written as

$$S_{it} = \Omega_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} M_{it}^{\beta_M}, \tag{4}$$

[10] I standardize $P_t^{\varepsilon} Q_t = 1$, where $P_t = \left(\int \left(\frac{P_{it}}{B_{it}} \right)^{1-\varepsilon} di \right)^{\frac{1}{1-\varepsilon}}$.

where $\Omega_{it} = (A_{it}B_{it})^{\left(1-\frac{1}{\varepsilon}\right)}$ is revenue productivity and $\beta_X = \alpha_X\left(1-\frac{1}{\varepsilon}\right)$ for $X \in \{K, L, M\}$.[11] Constant returns technology in terms of quantity implies decreasing returns in terms of sales, $\beta_K + \beta_L + \beta_M = 1 - \frac{1}{\varepsilon}$. I call $\omega_{it} \equiv \ln(\Omega_{it})$ TFPR. That is,

$$\omega_{it} = \left(1 - \frac{1}{\varepsilon}\right)(a_{it} + b_{it}) \tag{5}$$

where lower cases denote logs. Eq. (5) shows that TFPR shocks depend on the sum of demand and technology shocks and that a larger ε magnifies these shocks. I can rewrite the TFPR shock as

$$\omega_{it} = s_{it} - \beta_K k_{it} - \beta_L l_{it} - \beta_M m_{it}. \tag{6}$$

The $MRPK_{it}$ measured in logs is

$$MRPK_{it} = \ln(\beta_K) + s_{it} - k_{it} \tag{7}$$

It is easy to show that the optimal capital is proportional to $\varepsilon\omega_{it}$ if investment involves with no time-to-build or adjustment costs. In this hypothetical setting, therefore, the amount of optimal capital strongly depends on productivity if the product market is competitive (i.e., if ε is high). In other words, given a magnitude of TFPR shock, larger amount of investment is required to achieve the optimal level when the competition is tougher.

3.3.2 Dynamic model

In this subsection, I explain the dynamic part of ACL model. The model builds on Dixit and Pindyck (1994), Caballero and Pindyck (1996), Cooper and Haltiwanger (2006), and Bloom (2009).

[11] My definition of TFPR, Ω_{it}, is the same as ACL's, but different from Hsieh and Klenow (2009)'s. Hsieh and Klenow define TFPR as $TFPR_{HKi} = P_i A_i = \frac{S_i}{K_i^{\alpha_K} L_i^{\alpha_L} M_i^{\alpha_M}}$ (if materials are included as a production factor). The two definitions are related with each other as $TFPR_{HKi} = \Omega_i \left(K_i^{\alpha_K} L_i^{\alpha_L} M_i^{\alpha_M}\right)^{-\frac{1}{\varepsilon_S}}$. I choose the definition because Ω_{it} is composed of demand and technology shocks and hence can be safely regarded as exogenous shocks. On the other hand, Hsieh and Klenow's definition is convenient when I compute aggregate TFP, as I show in Appendix 2. Using the terminology of Foster et al. (2017), Ω_{it} is the regression-residual based TFPR while $TFPR_{HKi}$ is the cost-share based TFPR.

In each period, plants can hire labor and acquire intermediates without adjustment costs and time-to-build. This leads to a following expression for the profit,

$$\pi(\Omega_{it}, K_{it}) = \lambda \Omega_{it}^{\frac{1}{\beta_K+\varepsilon^{-1}}} K_{it}^{\frac{\beta_K}{\beta_K+\varepsilon^{-1}}} \tag{8}$$

where

$$\lambda = (\beta_K + \varepsilon^{-1}) \left(\frac{\beta_K}{p_L}\right)^{\frac{\beta_L}{\beta_K+\varepsilon^{-1}}} \left(\frac{\beta_M}{p_M}\right)^{\frac{\beta_M}{\beta_K+\varepsilon^{-1}}} \tag{9}$$

and p_L and p_M denote a wage and materials price, respectively. Capital evolves as

$$K_{it+1} = \delta K_{it} + I_{it} \tag{10}$$

where δ is one minus the depreciation rate and I_{it} is investment. Eq. (10) incorporates the assumption of one-period time to build. The investment involves adjustment costs composed of the fixed disruption costs and convex costs. Unlike ACD, I consider the possibility that both adjustment cost components are asymmetric between positive and negative investment. Specifically, the adjustment cost is

$$\begin{aligned} C(I_{it}, K_{it}, \Omega_{it}) = {} & I_{it} + C_K^{F+} 1_{\{I_{it}>0\}} \pi(\Omega_{it}, K_{it}) \\ & + C_K^{F-} 1_{\{I_{it}<0\}} \pi(\Omega_{it}, K_{it}) + C_K^{Q+} K_{it} \left(\frac{I_{it}}{K_{it}}\right)^2 1_{\{I_{it}>0\}} \\ & + C_K^{Q-} K_{it} \left(\frac{I_{it}}{K_{it}}\right)^2 1_{\{I_{it}<0\}} \end{aligned} \tag{11}$$

We specify the TFPR shock process as the AR(1) process:

$$\omega_{it} = \mu + \rho\omega_{it-1} + \sigma\kappa_{it} \tag{12}$$

Note that I assume that the volatility of TFPR shock is independent of ε despite Eq. (5) in order to compare the dispersion in MRPK across different ε's controlling for the dispersion in TFPR shock.

Defining the transition of Ω_{it} as $\phi(\Omega_{it+1} \mid \Omega_{it})$, the plant's value function is defined in recursive form as

$$V(\Omega_{it}, K_{it}) = \max \pi(\Omega_{it}, K_{it}) - C(I_{it}, K_{it}, \Omega_{it}) + \beta \int V(\Omega_{it}, \delta K_{it} + I_{it})\phi(\Omega_{it+1} \mid \Omega_{it})d\Omega_{it+1} \tag{13}$$

In the ACL model, entry and exit are because any plant can operate with a positive profit due to the decreasing returns to scale in the revenue function and the absence of fixed costs.

3.3.3 Analytical solutions of a simplified model

We cannot generally obtain analytical solutions of the model without restricting some parameters. To obtain analytical solutions, I simplify the model in this subsection. In Section 3.3.4, I derive the full model by numerically solving the model. Specifically, I assume no adjustment costs ($C_K^{F+} = C_K^{Q+} = C_K^{F-} = C_K^{Q-} = 0$) and consider only time-to-build. I further assume that TFPR shock follows a random walk:

$$\omega_{it} = \omega_{it-1} + \sigma\kappa_{it} \tag{14}$$

In this simplified model, the plant's problem is reduced to

$$\max_{K_i} \pi_i^* = E\big(\pi(\Omega_i, K_i) \mid \Omega_{-1,i}\big) - P_K K_i \tag{15}$$

where $\pi(\Omega_i, K_i)$ is defined by Eq. (8) and $\Omega_{-1,i}$ is the productivity in the previous period. I omit time subscripts for brevity. Solving this optimization and aggregating the output and inputs across plants yield aggregate productivity. Appendix 1 shows that comparing the aggregate productivity with and without time-to-build yields

$$TFPratio \equiv \frac{A}{A^*} = \left[\frac{\left\{\int u_i^{\frac{\varepsilon}{\varepsilon-(\varepsilon-1)(1-\alpha_K)}} di\right\}^{\varepsilon-(\varepsilon-1)(1-\alpha_K)}}{\int u_i^{\varepsilon} di}\right]^{\frac{1}{\varepsilon-1}} \tag{16}$$

where A and A^* respectively denote the aggregate TFPs with and without time-to-build, and $u_i = e^{\sigma\kappa_i}$. Suppose further that the shock is log-normally distributed with $Var(\ln u_i) = \sigma^2$. Then,

$$\mathrm{SD}(\ln(MRPK_i)) = \left(\frac{1}{1-\left(1-\frac{1}{\varepsilon}\right)(1-\alpha_K)}\right)\sigma \tag{17}$$

$$\ln TFPratio = -\frac{\varepsilon\alpha_K}{1-\left(1-\frac{1}{\varepsilon}\right)(1-\alpha_K)}\sigma^2 \tag{18}$$

Eqs. (17) and (18) shows that the inability to adjust the amount of inputs increases the dispersion in MRPK and lowers aggregate TFP even though plants dynamically optimize capital. In addition, the dispersion in MRPK and loss of aggregate TFP (in the static sense) is larger when the volatility of TFPR shock is higher and even more so when the competition is severer (i.e., for higher ε). [12]

3.3.4 Simulation results

We numerically solve and simulate the above dynamic models and calculate the standard deviation of the log of MRPK, which I denote by $\mathrm{SD}(MRPK_{it})$ for simplicity hereafter. The simulations aim to see the effects of various values of ε on the relation between σ and $\mathrm{SD}(MRPK_{it})$. In the following simulations, I set all parameters except adjustment cost parameters, following ACL, who estimate their model using the US Census of Manufacturers. Table 3:1 summarizes the set parameters. I set p_L and p_M to make λ=1.0 when ε=4, as in ACL. I use three alternative sets of adjustment cost parameters. In the asymmetric adjustment cost specification, I set $C_K^{F+} = C_K^{Q+} = 0$ while maintaining the same C_K^{F-} and C_K^{Q-} values as those of ACL's estimates, while in the symmetric adjustment cost specification, I set all the adjustment cost parameters at the same values of ACL's. Finally, in no adjustment cost specification, I set all the adjustment cost parameters at zero.

[12] If I extend the model to a general equilibrium one, different ε values should result in different real interest rates. However, MRPK would still be equalized across firms without frictions and a larger dispersion in MRPK would result in a larger TFP loss relative to the frictionless economy (Hsieh and Klenow, 2009).

Table 3:1 Simulation parameters

All specifications	
μ_b or μ_a	0.000
α_K	0.160
α_L	0.307
α_M	0.533
δ	0.900
β	1/(1+0.065)
p_L	0.182
p_M	0.182
ρ	0.850
Symmetric adjustment costs	
$C^{F+}{}_K$	0.090
$C^{Q+}{}_K$	8.800
$C^{F-}{}_K$	0.090
$C^{Q-}{}_K$	8.800
Asymmetric adjustment costs	
$C^{F+}{}_K$	0
$C^{Q+}{}_K$	0
$C^{F-}{}_K$	0.090
$C^{Q-}{}_K$	8.800

Note: In the no adjustment cost specification, $C_K^{F+} = C_K^{Q+} = C_K^{F-} = C_K^{Q-} = 0$.

A. Asymmetric adjustment costs

We first simulate the asymmetric adjustment cost model. In Table 3:2, columns labelled "Asym AC" show $\mathrm{SD}(MRPK_{it})$ for the simulated data from this specification. The first panel of Figure 3:2A illustrates that for each ε, $\mathrm{SD}(MRPK_{it})$ tends to increase with σ, suggesting that higher TFPR shock volatility results in grater dispersion in MRPK, which is consistent with ACL. The new finding here is that the slope is steeper as ε is higher, suggesting that increasing product market competition strengthens the effect of TFPR shock volatility on the dispersion in MRPK.

To investigate the mechanism that causes such dispersion, I decompose investment into the extensive and intensive margins. Specifically, the second and third panels of Figure 3:2A show the fraction of the plants that conduct positive and negative investment, respectively, while the bottom panel of Figure 3:2A shows the average net investment ratio of plants that conduct positive investment. The second panel shows that for each ε, the fraction of the plants that conduct positive investment decreases with σ, suggesting that higher TFPR shock volatility results in a smaller fraction of expanding plants. The negative effect of the volatility on the fraction of expanding plants tends to be smaller as ε is higher, that is, as the product market is more competitive. The third panel shows that the effects of the volatility on the fraction of shrinking plants is just the opposite to that of expanding plants. The bottom panel shows that the net investment ratio of expanding plants tends to increase as the volatility increases, and this positive effect of volatility on the intensive margin of expanding plants is stronger as the product market is more competitive. Although not reported to save space, the absolute value of the net investment ratio of shrinking plants is relatively small and do not change significantly as volatility changes. In sum, both the extensive and intensive margins seem to matter both for the volatility-MRPK dispersion relationship and the role of competition on that relationship.

Table 3:2 Simulation results: $\mathrm{SD}(MRPK)$

σ	epsilon=2			epsilon=4			epsilon=6		
	No AC	Asym AC	Sym Ac	No AC	Asym AC	Sym AC	No AC	Asym AC	Sym AC
0.1	0.169	0.186	0.326	0.265	0.330	0.426	0.327	0.439	0.535
0.2	0.338	0.421	0.636	0.530	0.745	0.869	0.654	0.966	1.084
0.3	0.507	0.681	0.930	0.795	1.170	1.319	0.982	1.504	1.646
0.4	0.676	0.950	1.209	1.061	1.603	1.775	1.311	2.055	2.217
0.5	0.846	1.220	1.479	1.329	2.044	2.236	1.642	2.614	2.794
0.6	1.015	1.490	1.750	1.598	2.491	2.701	1.978	3.176	3.375
0.7	1.186	1.762	2.027	1.869	2.943	3.171	2.320	3.743	3.960
0.8	1.356	2.039	2.310	2.144	3.397	3.642	2.669	4.313	4.546
0.9	1.528	2.318	2.598	2.423	3.855	4.116	3.028	4.885	5.133
1	1.700	2.599	2.890	2.707	4.314	4.591	3.400	5.460	5.721
1.1	1.874	2.883	3.184	2.999	4.775	5.067	3.782	6.037	6.310
1.2	2.049	3.169	3.481	3.297	5.238	5.543	4.181	6.615	6.899
1.3	2.225	3.455	3.779	3.603	5.703	6.020	4.590	7.195	7.488
1.4	2.405	3.744	4.078	3.922	6.169	6.497	5.028	7.776	8.078
1.5	2.585	4.033	4.378	4.244	6.637	6.975	5.476	8.358	8.668

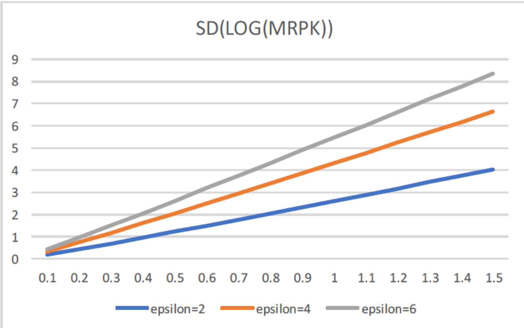
Source: Authors' calculations.

Figure 3:2 Simulation results

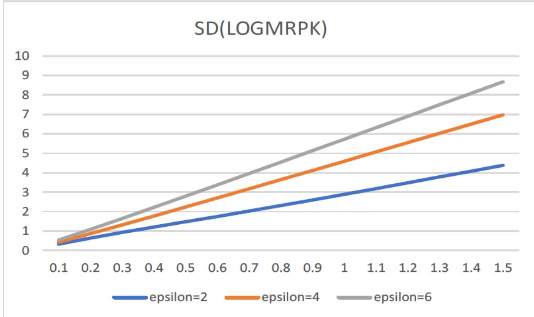
A. Asymmetric adjustment costs

B. Symmetric adjustment costs

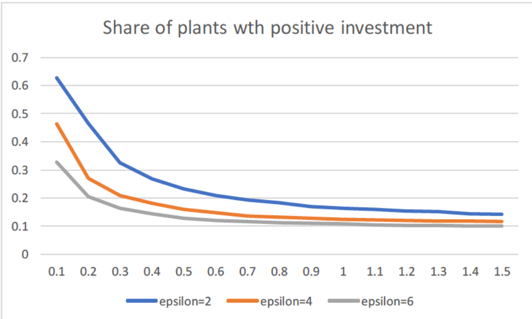
A. $\mathrm{SD}(MRPK_{it})$



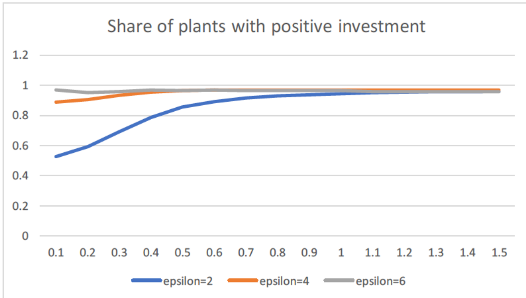
B. $\mathrm{SD}(MRPK_{it})$



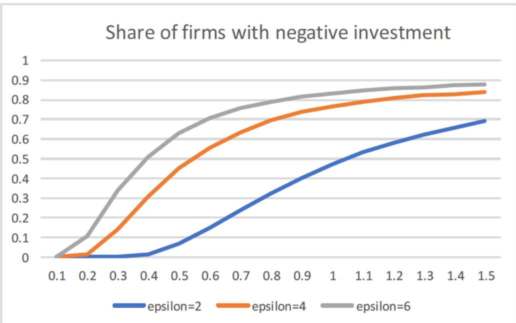
A. Share of plants with $\mathrm{I}_t > 0$



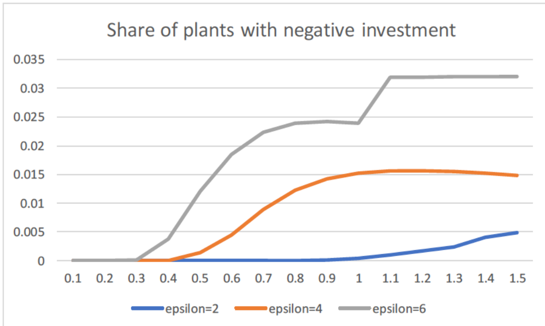
B. Share of plants with $\mathrm{I}_t > 0$



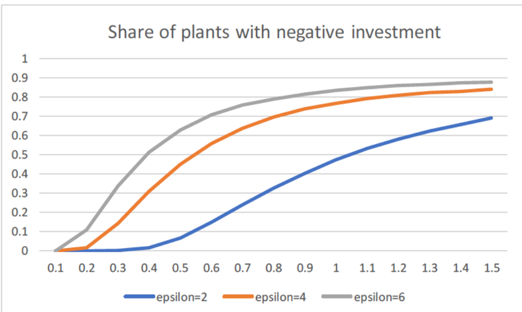
A. Share of plants with $\mathrm{I}_t < 0$



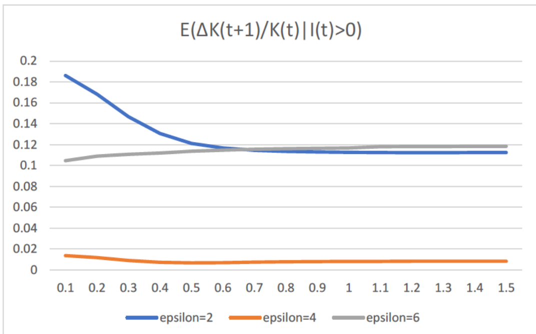
B. Share of plants with $\mathrm{I}_t < 0$



A. Average $\frac{\Delta K_{t+1}}{K_t}$ for plants with $I_t > 0$



B. Average $\frac{\Delta K_{t+1}}{K_t}$ for plants with $I_t > 0$



Source: Authors' calculations.

Finally, I investigate how the dispersion in MRPK is related to aggregate productivity. To this aim, I compute the aggregate TFP, A_t, using simulated data and compare it with the hypothetical aggregate TFP that would be realized if investment did not involve with time-to-build or adjustment costs. In this hypothetical setting, it is easy to show that aggregate TFP is

$$A_t^* = \left(\int \Omega_{it}^{\varepsilon} di \right)^{\frac{1}{\varepsilon - 1}} \tag{19}$$

Table 3:3 shows the average ratio of $\frac{A_t}{A_t^*}$. It is clear that the higher dispersion in MRPK results in lower aggregate TFP relative to the hypothetical TFP. In addition, as ε is higher, the difference in $\frac{A_t}{A_t^*}$ between low σ and high σ becomes larger.

Table 3:3 Simulation results: ratio of aggregate TFP to hypothetical TFP

σ	epsilon=2			epsilon=4			epsilon=6		
	No AC	Asym AC	Sym Ac	No AC	Asym AC	Sym AC	No AC	Asym AC	Sym AC
0.1	0.997	0.997	0.990	0.992	0.988	0.980	0.986	0.977	0.964
0.2	0.990	0.985	0.965	0.972	0.952	0.929	0.955	0.920	0.889
0.3	0.979	0.966	0.929	0.946	0.908	0.869	0.923	0.866	0.819
0.4	0.965	0.941	0.891	0.922	0.868	0.816	0.897	0.826	0.767
0.5	0.949	0.916	0.855	0.899	0.835	0.774	0.875	0.796	0.731
0.6	0.934	0.891	0.823	0.880	0.808	0.742	0.858	0.773	0.704
0.7	0.919	0.868	0.795	0.865	0.788	0.716	0.845	0.757	0.685
0.8	0.904	0.847	0.770	0.851	0.771	0.697	0.834	0.744	0.670
0.9	0.891	0.829	0.749	0.840	0.757	0.681	0.825	0.733	0.658
1.0	0.879	0.813	0.731	0.830	0.746	0.668	0.817	0.725	0.648
1.1	0.868	0.799	0.715	0.822	0.737	0.657	0.811	0.718	0.641
1.2	0.858	0.787	0.702	0.815	0.729	0.649	0.805	0.712	0.634
1.3	0.849	0.776	0.690	0.809	0.722	0.641	0.801	0.707	0.629
1.4	0.841	0.766	0.680	0.804	0.716	0.635	0.797	0.703	0.625
1.5	0.834	0.758	0.671	0.799	0.711	0.629	0.794	0.700	0.621

Source: Authors' calculations.

B. Symmetric adjustment costs

Next, I simulate the symmetric adjustment cost model. In Table 3:2, columns labelled "Sym AC" show $\mathrm{SD}(MRPK_{it})$ from the simulated data from this specification. Table 3:2 and the top panel of Figure 3:2B show that $\mathrm{SD}(MRPK_{it})$ for each ε is similar to the asymmetric adjustment cost case, although $\mathrm{SD}(MRPK_{it})$ is larger and $\frac{A_t}{A_t^*}$ is smaller for the symmetric adjustment costs than for the asymmetric adjustment costs.

The mechanism that causes such dispersion in MRPK, however, seems to be different between asymmetric and symmetric adjustment costs. The second panel of Figure 3:2B shows that the share of expanding plants is close to unity except for a low range of volatility in the case of $\varepsilon = 2$, where the share of expanding plants tends to increase as the volatility increases. On the other hand, the third panel of Figure 3:2B shows that the share of shrinking plants is very low. As for the intensive margins, the bottom panel of Figure 3:2B shows that the relationship between the net investment ratio of expanding plants and volatility is not monotonic depending on the degree of competition and the range of volatility. In addition, the level of investment ratio is smaller in the symmetric adjustment cost case than in the asymmetric adjustment cost case. Overall, when adjustment costs are symmetric, dispersion in MRPK is likely to be caused by the low investment volume by expanding plants, i.e., intensive margin.

C. No adjustment costs

Finally, I simulate the model with no adjustment costs. Note that I still assume the time-to-build: one-period lag between current-period investment and capital that serves production. I simulate this model to examine to what extent time-to-build alone accounts for the dispersion in MRPK.

Table 3:2 compares $\mathrm{SD}(MRPK_{it})$ in the case of no adjustment costs (*NoAC*) to the cases of asymmetric (*AsymAC*) and symmetric (*SymAC*) adjustment costs. It shows that $\mathrm{SD}(MRPK_{it})$ of *NoAC* accounts for 60% to 80% of those of *AsymAC* and 50-%-60% of *SymAC* for most of the parameter sets I examine, suggesting that while time-to-build accounts for a majority of the dispersion in MRPK, adjustment costs account for its significant proportion. In terms of aggregate TFP relative to the hypothetical TFP (i.e., the average ratio of $\frac{A_t}{A_t^*}$), it is lower in the case of Asym AC and Sym AC than in the case of No AC by 0-10% and 0-20%, respectively.

In addition, without adjustment costs, the share of inactive plants (i.e., those plants that do not invest or divest) is zero for most of the parameter sets I examine. These results suggest that both time-to-build and adjustment costs should play a significant role in accounting for the dispersion in MRPK.

In sum, the simulation results suggest that volatility tends to cause greater dispersion in MRPK and that competition causes larger dispersion in MRPK driven by TFPR volatility. In addition, the dispersion in MRPK results in lower aggregate TFP relative to the hypothetical TFP that would be realized without time-to-build or adjustment costs. I examine whether these simulation results are supported empirically by data from Japanese manufacturing plants below.[13]

3.4 Data and empirical methodology

3.4.1 Data

The main data source is the *Census of Manufacture* published by the Ministry of Economy, Trade, and Industry (METI) in Japan. The main purpose of this annual survey is to gauge the activities of Japanese plants in manufacturing industries quantitatively, including sales, number of employees, wages, materials and tangible fixed assets. The census covers all establishments in years ending with 0, 3, 5, and 8 of the calendar years from 1981 to 2009. For other years, the Census covers establishments with four or more employees.

The *Census of Manufacture* contains two types of surveys: one for plants with more than 30 employees (*Kou Hyou*), and the other is for plants with 29 or less employees (*Otsu Hyou*). The Otsu Hyou does not include some pieces of information, including fixed asset, especially after 2001. For this reason, I construct the panel dataset from 1986 to 2013 using the Kou Hyou.

To construct the data for output and factor inputs, first, I use each plant's shipments as the nominal gross output and then deflate the nominal gross output by the output deflator in the Japan Industrial Productivity Database (JIP) 2015 to convert it into values in constant prices (i.e., real gross output (Q_{it}) based on the year 2000. Second, I define the nominal intermediate input as the sum of raw materials, fuel, electricity, and subcontracting expenses for the plant's consigned production. Using the Bank of Japan's Corporate Good Price Index (CGPI), I convert the nominal intermediate input

[13] I use plants and establishments interchangeably throughout this chapter.

into values in constant prices (i.e., real intermediate input (M_{it})) for 2000. Third, I use each plant's total number of workers as labor input (L_{it}).

We construct the data for tangible capital stock as follows. First, I define capital input (K_{it}) as the nominal book value of tangible fixed assets from the Census multiplied by the book-to-market value ratio for each industry ($\alpha_{IND,t}$) for each data point corresponding to K_{it}. I calculate the book-to-market value ratio for each industry ($\alpha_{IND,t}$) by using the data for real capital stock ($K_{IND,t}^{JIP}$) and real value added ($Y_{IND,t}^{JIP}$) at each data point taken from the JIP database as follows:

$$\frac{Y_{IND,t}^{JIP}}{K_{IND,t}^{JIP}} = \frac{\sum_i Y_{IND,i,t}^{\text{Census}}}{\sum_i BVK_{IND,i,t}^{\text{census}} * \alpha_{IND,t}} \tag{20}$$

where $\sum_i Y_{IND,i,t}^{\text{Census}}$ is the sum of the plants' value added (i is the index of a plant), and $\sum_i BVK_{IND,i,t}^{\text{Census}}$ is the sum of the nominal book value of tangible fixed assets of industry *IND* in the Census.[14]

3.4.2 Variable measurement

Production Function

We estimate the sales-generating production function (3) for each 4-digit Japan Standard Industrial Classifications (JSIC) using the system generalized method of moments (GMM) estimator following Blundell and Bond (2000). Specifically, I estimate the following function:

$$\ln Y_{it} = \beta_K \ln K_{it} + \beta_L \ln L_{it} + \beta_M \ln M_{it} + \eta_i + year_t + \omega_{it} + \varepsilon_{it}, \tag{21}$$

where

$$\begin{gathered} \omega_{i.t} = \rho\omega_{i.t-1} + \xi_{i,t}, \qquad |\rho| < 1 \\ \varepsilon_{it}, \xi_{i,t} \sim MA(0). \end{gathered} \tag{22}$$

The left hand-side of equation (21) accounts for the natural logarithm of output produced by plant i in period t. As production inputs, $\ln K_{it}$ denotes the natural logarithm of plant i's capital input at the beginning of period t and $\ln L_{it}$ and $\ln M_{it}$ denote the natural logarithms of labor input and

[14] The real value added is negative only for the iron and steel industry in 2010. The book-to-market ratio is interpolated from the ratio as of t-1 and t+1.

intermediate goods, respectively. I measure these variables at the end of period *t*. Following the literature, I include the plant-level fixed effect η_i, year fixed effect $year_t$, and the TFPR $\omega_{i,t}$. I assume that $\omega_{i,t}$ follows the AR(1) process described by equation (11). The disturbance term, $\varepsilon_{i,t}$, represents measurement error. This model has a dynamic (common factor) presentation

$$\begin{aligned}\ln Y_{it} = {} & \beta_K \ln K_{it} - \rho\beta_K \ln K_{it-1} + \beta_L \ln L_{it} - \rho\beta_L \ln L_{it-1} \\ & + \beta_M \ln M_{it} - \rho\beta_M \ln M_{it-1} + \rho \ln Y_{it-1} + (1-\rho)\eta_i \\ & + year_t - \rho year_{t-1} + \xi_{it} + \varepsilon_{it} - \rho\varepsilon_{it-1}\end{aligned} \tag{23}$$

or

$$\begin{aligned}\ln Y_{it} = {} & \pi_1 \ln K_{it} + \pi_2 \ln K_{it-1} + \pi_3 \ln L_{it} + \pi_4 \ln L_{it-1} + \pi_5 \ln M_{it} \\ & + \pi_6 \ln M_{it-1} + \pi_7 \ln Y_{it-1} + \eta_i^* + year_t^* + \omega_{it}\end{aligned} \tag{24}$$

subject to three non-linear (common factor) restrictions: $\pi_2 = -\pi_1\pi_7$, $\pi_4 = -\pi_3\pi_7$, $\pi_6 = -\pi_5\pi_7$. I first obtain consistent estimates of the unrestricted parameter $\pi = (\pi_1, \dots, \pi_7)$ and $\mathrm{var}(\pi)$ using the system GMM (Blundell and Bond, 1998). Since $\omega_{i,t} \sim MA(1)$, I use the following moment conditions:

$$\mathrm{E}\left(x_{i,t-s}\Delta\omega_{i,t}\right) = 0 \tag{25}$$

$$\mathrm{E}\left(\Delta x_{i,t-s}(\eta_i^* + \omega_{i,t})\right) = 0 \tag{26}$$

where $x_{it} = (K_{it}, L_{it}, M_{it}, Y_{it})$ and $s \geq 3$. Next, using consistent estimates of the unrestricted parameters and their variance-covariance matrix, I impose the above restrictions by minimum distance to obtain the restricted parameter vector $(\beta_K, \beta_L, \beta_M, \rho)$.We first estimate the production function, using the data of all plants. Then I drop the 1% tails of TFPR and MRPK as outliers in each year and reestimate the production function.

Markup

From the definition of β_X and the assumption of constant returns to scale, I can derive the markup as $\frac{\varepsilon}{\varepsilon-1} = \frac{1}{\beta_K+\beta_L+\beta_M}$. Using the industry-level estimates of $(\beta_K, \beta_L, \beta_M)$, I obtain the industry-level, time-invariant markup:

$$Markup1_s = \frac{1}{\hat{\beta}_{Ks} + \hat{\beta}_{Ls} + \hat{\beta}_{Ms}} \tag{27}$$

We use this markup measure as an inverse measure of competition.[15]

Later, I use an alternative measure of markup following De Loecker and Warzynski (2012). I allow adjustment costs only for capital, suggesting that the static profit maximization condition holds for materials. Therefore, the marginal product of materials, in particular, is equal to its price, which leads to

$$\beta_{Ms} = \frac{P_{it}^M M_{it}}{S_{it}} \tag{28}$$

where P_{it}^M is the price of materials and β_{Ms} is the output elasticity of materials in industry s. Eq. (28) shows that β_{Ms} is equal to the cost share of materials in sales. Combining Eq. (28) and $\beta_{Ms} = \left(1 - \frac{1}{\varepsilon_i}\right) \alpha_{Ms}$, I obtain the markup as

$$\frac{1}{1 - \frac{1}{\varepsilon_{it}}} = \frac{\alpha_{Ms}}{\frac{P_{it}^M M_{it}}{S_{it}}} \tag{29}$$

In practice, I follow the method of replacing α_{Ms} in Eq. (29) with the estimated value of β_{Ms}, $\widehat{\beta_{Ms}}$, and take the median value of the markup among the plants within each industry:

$$Markup2_s = Median\left(\frac{\widehat{\beta_{Ms}}}{\frac{P_{it}^M M_{it}}{S_{it}}}\right) \tag{30}$$

We use this industry-level, time-invariant markup measure as a robustness test.

Volatility

[15] Many preceding studies use the markup or the Lerner index, which is a one-to-one correspondence of the markup, as a measure of competition. See, among others, Aghion et al. (2015).

To measure uncertainty, I employ two alternative measures of the volatility of productivity, ω_{it}. The first is the standard deviation of the productivity shocks across plants within an industry in a given year:

$$Volatility1_{st} = SD_{st}(\omega_{it} - \omega_{it-1}) \tag{31}$$

where s denotes the industry of plant i.

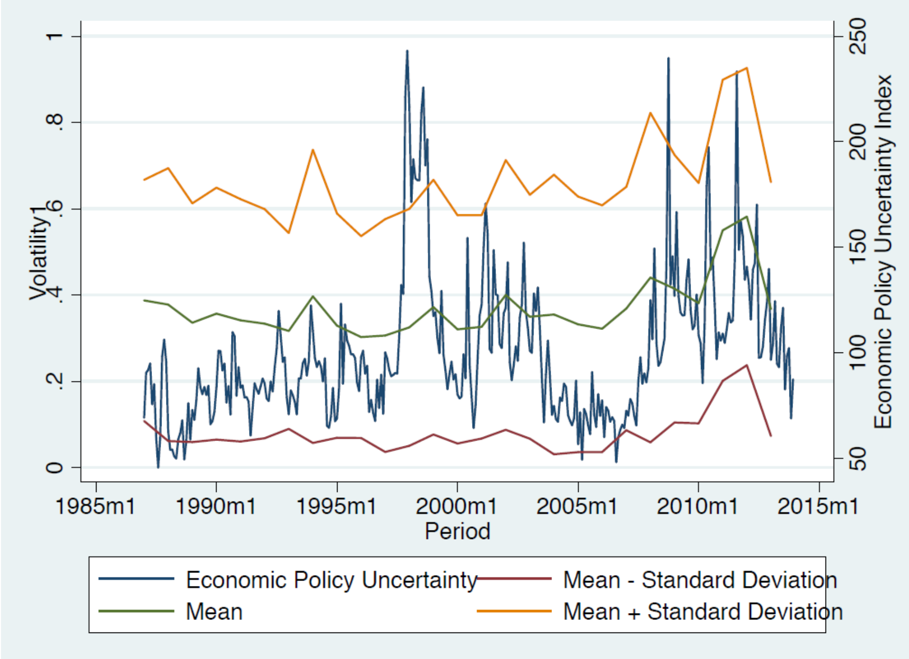
The other measure is based on the assumption that ω_{it} follows the stationary AR(1) process and is defined as

$$Volatility2_{st} = SD_{st}(\omega_{it} - \hat{\rho}\omega_{it-1}). \tag{32}$$

These volatility measures are time variant and defined at the industry level. Note also that multiplicative shocks that are common to all establishments within an industry are absorbed when I calculate the standard deviation of the log of TFPR and hence do not have effects on the volatility measures by definition. Nonetheless, it turns out that these volatility measures seem to be correlated with aggregate uncertainty shocks. Figure 3:3 depicts $Volatility1_{st}$ averaged over industries for each year and the Japan Policy Uncertainty Index.[16] Both the average $Volatility1_{st}$ and the Index spike in the late 1990s of the Japanese banking crisis, the 2008 global financial crisis, and the 2011 Tohoku Great Earthquake.

[16] This index is constructed by the Economic Policy Uncertainty Project, the Asia and Pacific Division of the International Monetary Fund (IMF), and the Research Institute of Economy, Trade, and Industry (RIETI) and available at http://www.rieti.go.jp/jp/database/policyuncertainty/. See Arbatli et al. (2017) in detail.

Figure 3:3 Average volatility and the Economic Policy Uncertainty Index



Source: Authors' calculations, based on the Census of Manufacture by METI, and Arbatli et al. (2017).

Dispersion in MRPK

We focus on the standard deviation of $MRPK_{it}$ across plants in industry s in year t: $SD_{st}(MRPK_{it})$ as a baseline measure of the dispersion in MRPK. The result below is robust to whether I use the $SD_{st}(MRPK_{it})$ or $Var_{st}(MRPK_{it})$.

Table 3:4 summarizes the descriptive sample statistics of the variables. The standard deviation of $MRPK_{it}$ across plants in all industries is 1.36, which is larger than the U.S. counterpart (0.98) but close to the French, Romanian and Mexican counterparts (1.28, 1.38, and 1.40, respectively) reported in Table 2 of ACL. I also report the sample statistics of the dispersion in the marginal revenue products of labor and materials, $SD_{st}(MRPL_{it})$ and $SD_{st}(MRPM_{it})$ to compare with $SD_{st}(MRPK_{it})$ in Table 3:4, illustrating that $SD_{st}(MRPK_{it}) > SD_{st}(MRPL_{it}) > SD_{st}(MRPM_{it})$ on average. This evidence supports the approach focusing on the adjustment cost of capital rather than that of labor or materials.[17]

[17] ACL report a similar magnitude of the standard deviation of each input for the US economy (0.81 for capital, 0.63 for labor, and 0.54 for materials) (Table 7, pp. 1036).

To see the time-series movement of the dispersion in MRPK, I depict in Figure 3:4 the standard deviation of $\mathrm{Log(MRPK}_{it})$ and $\mathrm{Log}(\mathrm{MRPK}_{it}/\overline{MRPK_{st}})$ for each year, where $\overline{MRPK_{st}}$ denotes the average MRPK in industry s, which establishment i belongs to. The former shows the overall dispersion in MRPK while the latter shows the dispersion in MRPK within the industry. Figure 3:4 shows that while the overall MRPK dispersion tends to decrease, the within-industry MRPK dispersion tends to increase over the last three decades.[18]

Table 3:4 Summary statistics

Variable name	N	mean (p1-p99)	sd (p1-p99)	min	p5	median	p95	max
Plant-level variables								
ln(Productivity)	1,391,981	5.43	1.10	-6.71	3.65	5.34	7.54	14.24
Productivity growth rate	1,243,841	0.00	0.23	-11.66	-0.35	0.00	0.36	9.11
ln(MRPK)	1,333,909	-2.66	1.36	-16.53	-5.22	-2.63	-0.26	10.22
ln(MRPK) for zero investment plants	569,506	-2.85	1.43	-16.39	-5.48	-2.84	-0.31	9.65
ln(MRPK) for positive investment plants	787,808	-2.52	1.29	-16.53	-4.99	-2.50	-0.25	10.22
ln(MRPK) for negative investment plants	34,776	-2.67	1.44	-14.39	-5.30	-2.67	-0.06	7.84
ln(MRPL)	1,387,850	6.28	0.86	-6.19	4.64	6.33	7.73	12.47
ln(MRPM)	1,389,497	-0.19	0.59	-10.42	-1.06	-0.27	1.05	11.05
Investment rate	1,392,090	0.18	0.31	-31.16	-0.01	0.07	0.85	140462.83
Industry-level variables (time-variant)								
Number of plants	13,503	88	122	1	4	41	368	2681
Volatility1	12,842	0.36	0.25	0.00	0.08	0.30	0.91	3.83
Volatility2	12,842	0.36	0.15	0.00	0.12	0.35	0.65	3.15
SD(ln(MRPK))	11,734	1.00	0.26	0.02	0.59	0.98	1.53	4.33
SD(ln(MRPL))	12,985	0.69	0.19	0.00	0.39	0.67	1.07	2.10
SD(ln(MRPM))	13,177	0.53	0.23	0.00	0.21	0.49	1.00	3.46
Fraction of zero investment plants	13,503	0.41	0.18	0.00	0.13	0.40	0.70	1.00
Fraction of positive investment plants	13,503	0.57	0.18	0.00	0.26	0.57	0.86	1.00
Fraction of negative investment plants	13,503	0.02	0.03	0.00	0.00	0.02	0.08	1.00
Industry-level variables (time-invariant)								
Markup1	491	1.42	0.51	-32.06	0.97	1.29	2.47	10.93
Markup2	491	0.80	0.29	-0.10	0.33	0.79	1.33	5.96
Output elasticity of capital	491	0.04	0.05	-0.39	-0.03	0.04	0.13	0.31
Output elasticity of labor	491	0.30	0.15	-0.85	0.05	0.31	0.55	1.41
Output elasticity of materials	491	0.40	0.14	-0.05	0.16	0.40	0.62	1.03

Source: Authors' calculations, based on the Census of Manufacture by METI.

[18] The hike in 2011-12 possibly reflect the Tohoku Earthquake on March 11, 2011.

Figure 3:4 Dispersion in MRPK

Source: Authors' calculations, based on the Census of Manufacture by METI.

3.4.3 Methodology

We examine how the time-series process of TFPR shocks affect the cross-sectional dispersion of MRPK depending on the markup levels. My working hypothesis is that while greater uncertainty reduces investment and results in the larger dispersion in MRPK, the impact of uncertainty on the dispersion in MRPK is stronger in more competitive markets. To test these hypotheses, I estimate the following baseline specifications:

$$MRPK\ Dispersion_{st} = \beta Volatility_{st} + FE_s + \varphi_{st} \tag{33}$$

$$\begin{aligned} &MRPK\ Dispersion_{st} \\ &= \beta_1 Volatility_{st} + \beta_2 Volatility_{st} * Markup_s \\ &+ FE_s + \varphi_{st} \end{aligned} \tag{34}$$

The unit of observation is industry-year. The dependent variable is the MRPK dispersion measure described above. The independent variables are one of the volatility measures or their interaction with the markup. If higher volatility results in larger dispersion in MRPK, β should be positive. On the other hand, if market competition (that is, a lower markup) increases the impact of volatility on the dispersion in MRPK, β_2 should be negative. Because I include the industry-level fixed effect, I do not include the markup measure on its own, which is time-invariant.

We further control for the previous year's dispersion in MRPK and estimate the following equation using the difference GMM in Arellano and Bond (1991):

$$\begin{aligned} &MRPK\ Dispersion_{st} \\ &= \beta_0 MRPK\ Dispersion_{st-1} + \beta_1 Volatility_{st} \\ &+ FE_s + \varphi_{st} \end{aligned} \tag{35}$$

In all estimations, I drop the industry-year observations with the volatility variable is higher than the top 1 percentile. The standard errors are clustered at the industry level.

3.5 Results

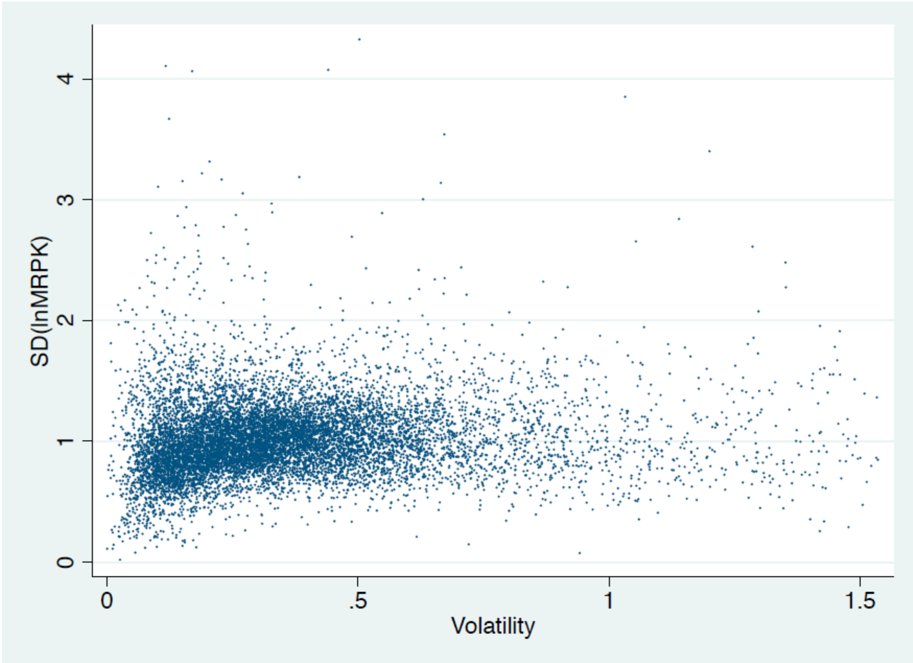
3.5.1 Dispersion in MRPK

Panel A of Figure 3:5 plots $SD_{st}(MRPK_{it})$ and $Volatility1_{st}$, indicating that there is a positive correlation between these two, which is consistent with the hypothesis that uncertainty increases MRPK dispersion.

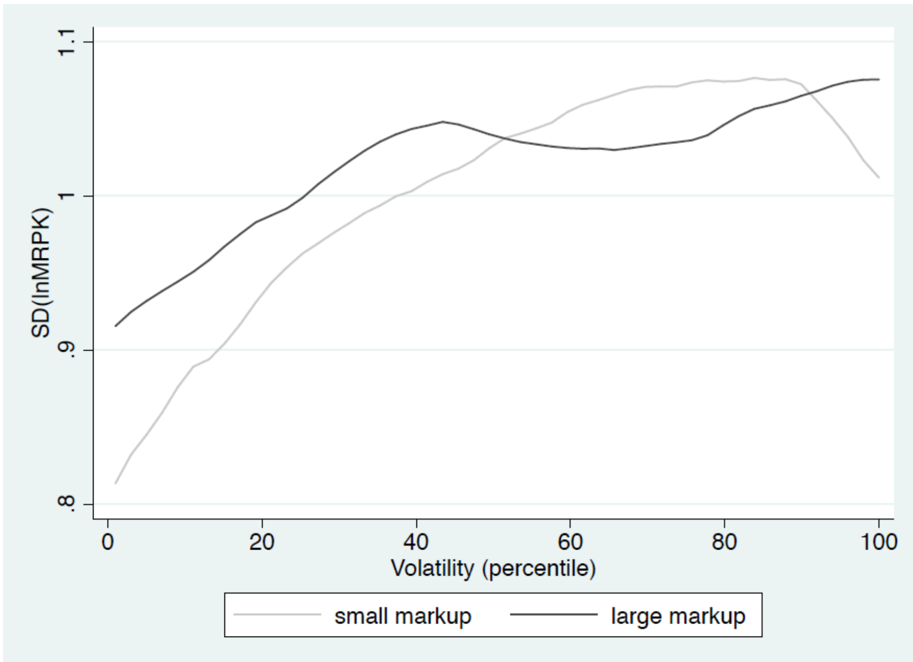
To illustrate the role of competition in the volatility-MRPK dispersion relationship, Panel B of Figure 3:5, I divide the industries into two depending on whether the markups are above or below the median, and depict the relationship between $SD_{st}(MRPK_{it})$ and the percentile of $Volatility1_{st}$. The figure shows that the slope is steeper for the lower-markup industries, suggesting that competition strengthens the volatility-MRPK dispersion relationship.

Figure 3:5 Volatility and dispersion in MRPK

A. All industries



B. High and low markup industries



Source: Authors' calculations, based on the Census of Manufacture by METI.

Table 3:5 reports the baseline estimation results when I use $SD_{st}(MRPK_{it})$ as a MRPK dispersion measure, $Volatility1_{st}$ as a volatility measure, and $Markup1_{st}$ as a markup measure. In Columns (1) and (2), I include only the current volatility measure, finding that higher TFPR volatility results in a larger MRPK dispersion regardless of whether I include industry fixed effects or not. In Columns (3) and (4), I add the lagged MRPK dispersion and estimate using GMM with industry-fixed effects. The one- to three-year lagged MRPK dispersion are all positive and significant. Importantly, even with these lagged MRPK dispersion, the current volatility still takes a positive and significant coefficient. In Column (6), I add the interaction of markup and volatility, and find that the interaction term is negative and weekly significant, suggesting that lower markup, i.e., severer competition, strengthens the adverse effect of volatility on MRPK dispersion. In Columns (7) to (12), I split the industries depending on whether the markup is higher or lower than the median. In Columns (7) and (8) I include only the current volatility measure, showing that while volatility takes positive and significant coefficients in both subsamples, the coefficient is larger for the sample with relatively lower markup. In Columns (9) and (10), I add the lagged MRPK and find that volatility takes a positive and significant coefficient only for the industries with lower markup. Finally, in Columns (11) and (12), I control for year fixed effects as well as industry fixed effects. Again I find a positive and marginally significant coefficient on volatility only for the industries with lower markup. All these results suggest that volatility increases the MRPK dispersion and that competition strengthens this volatility-MRPK dispersion relationship.

Table 3:5 Baseline estimation results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	OLS	OLS	GMM	GMM	GMM	OLS	OLS	OLS	GMM	GMM	OLS	OLS
volatility1	0.148***	0.0760***	0.0231**	0.0241**	0.0286***	0.0808***	0.0903***	0.0624***	0.0372***	0.0122	0.0462*	0.0289
	(5.176)	(4.340)	(2.488)	(2.551)	(2.925)	(4.769)	(3.482)	(2.627)	(2.706)	(0.967)	(1.903)	(1.116)
volatility_markup_crs_s						-0.00358*						
						(-1.750)						
L.sd_mrpk_all			0.430***	0.427***	0.453***				0.375***	0.550***		
			(27.07)	(26.14)	(25.65)				(18.51)	(27.81)		
L2.sd_mrpk_all				0.0585***	0.0582***							
				(5.004)	(4.774)							
L3.sd_mrpk_all					0.0292**							
					(2.491)							
Observations	11,207	11,207	10,665	10,239	9,814	11,207	5,558	5,649	5,276	5,389	5,558	5,649
Adjusted R-squared	0.013	0.441				0.441	0.367	0.512			0.398	0.532
Fixed effects	no	industry	industry	industry	industry	industry	industry	industry	industry	industry	industry+year	industry+year
Sample	all	all	all	all	all	all	markup smaller than median	markup larger than median	markup smaller than median	markup larger than median	markup smaller than median	markup larger than median

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Authors' estimations, based on the Census of Manufacture by METI.

Next, in Table 3:6, I change the volatility measure from $Volatility1_{st}$ to $Volatility2_{st}$ in Columns (1)-(3) and the markup measure from $Markup1_{s}$ to $Markup2_{s}$ in Columns (4) and (5). I report only the results for OLS estimation of Eqs. (33) and (34); the results for the GMM of Eq. (35) are virtually the same. Table 3:6 shows that the baseline results do not qualitatively change.[19]

Table 3:6 Robustness checks

	(1)	(2)	(3)	(4)	(5)
VARIABLES	OLS	OLS	OLS	OLS	OLS
volatility	volatility2	volatility2	volatility2	volatility1	volatility1
markup	markup1	markup1	markup1	markup2	markup2
volatility	0.185***	0.268***	0.111**	0.0919***	0.0621***
	(4.159)	(3.875)	(2.053)	(3.391)	(2.758)
Observations	11,317	5,626	5,691	5,549	5,658
Adjusted R-squared	0.437	0.373	0.503	0.519	0.365
Fixed effects	industry	industry	industry	industry	industry
Sample	all	markup smaller than	markup larger than median	markup smaller than	markup larger than median

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Authors' estimations, based on the Census of Manufacture by METI.

3.5.2 Plant-level evidence

To investigate the mechanism through which competition increases uncertainty-driven MRPK dispersion, I estimate the extensive and intensive margins of plant-level investment. First, to investigate the extensive margin, I run the following linear probability model of whether the plant conducts positive investment or not.

$$\begin{aligned} &1\left(\frac{I_{it}}{K_{it}} > 0.05\right) \\ &= \beta_1 \ln MRPK_{it} + \beta_2 Volatility_{st} \\ &+ \beta_3 \ln MRPK_{it} \times Volatility_{st} + FE_i + FE_t + \varphi_{it} \end{aligned} \tag{37}$$

[19] I have thus far implicitly assumed that TFPR shocks are independent across establishments. But TFPR shocks may correlate across establishments within a firm. To exclude this possibility, I restrict the sample to the firms with single establishments. Using $Volatility1_{st}$ as a volatility measure and the quartile dummies of $Markup1_{st}$ as a markup measure, I again find that the volatility is positive and significant only for the lower markup subsample.

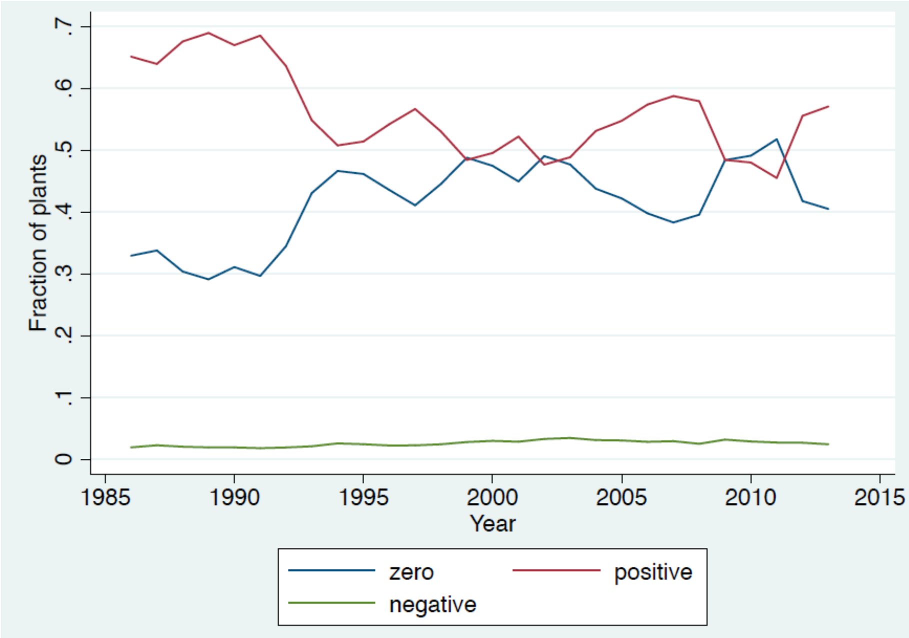
where I_{it} is gross investment measured by tangible fixed assets acquired, and K_{it} represents the tangible fixed assets at the beginning of the previous year. I use the threshold value of 0.05 rather than 0 because a very small-scaled investment is not likely to involve with time-to-build or adjustment costs. The dependent variable is a dummy for positive investment. I drop the plant-year observations with negative ($\frac{I_{it}}{K_{it}} < 0.05$) investment. I expect that β_1 takes a positive coefficient. On the other hand, I expect β_2 to take either negative or positive coefficients depending on whether the adjustment costs are asymmetric or symmetric, as the second panels of Figures 3:2A and 3:2B show. Finally, I expect β_3 to be negative if volatility weakens the plant's response to the change in (the logarithm of) MRPK. I control for fixed effects in two ways. One is to control for plant- and year-level fixed effects additively, and the other is to control for both of plant- and industry-year fixed effects. In the latter specification, I drop the single term of $Volatility_{st}$. I conduct the full sample estimation and the subsample estimation where industries are divided into the more competitive and less competitive ones depending on whether $Markup1_s$ is above or below the median.

We further estimate the linear probability model of negative investment after dropping the observations with positive ($\frac{I_{it}}{K_{it}} > 0.05$) investment as follows:

$$\begin{aligned} &1\left(\frac{I_{it}}{K_{it}} < -0.05\right) \\ &= \beta_1 \ln MRPK_{it} + \beta_2 Volatility_{st} \\ &+ \beta_3 \ln MRPK_{it} \times Volatility_{st} + FE_i + FE_t + \varphi_{it}. \end{aligned} \tag{38}$$

Figure 3:6 shows the fraction of establishments with positive, zero, and negative investment over time. While the average fractions of positive and zero investment are 0.56 and 0.41, respectively, the average fraction of negative investment is as small as 0.03.

Figure 3:6 Fraction of establishments with positive, zero, and negative investment



Note: I define zero investment as $\left|\frac{I_{t-1}}{K_{t-1}}\right| \leq 0.05$. I accordingly define positive and negative investment using the same threshold.

Source: Authors' calculations, based on the Census of Manufacture by METI.

Table 3:7 reports the results from using $Volatility1_{st}$ as a volatility measure, though using $Volatility2_{st}$ leave the results essentially unchanged. Columns (1)-(6) show the results for positive investment and Columns (7)-(12) for negative investment. In Columns (1)-(3) and (7)-(9), I control for plant- and year-fixed effects additively while in Columns (4)-(6) and (10)-(12), I control for plant- and industry-year fixed effects.

In Columns (1)-(3), β_1 is positive and significant while β_2 and β_3 are negative and significant, suggesting that while higher MRPK tends to induce positive investment, volatility reduces the likelihood of positive investment and weakens the plant's response to the change in MRPK. To compare Columns (2) and (3), I find that the absolute values of both β_2 and β_3 are larger for the industries with lower markup, suggesting that competition strengthens depressing effect of volatility on investment and on the sensitivity of investment to MRPK. In Columns (4)-(6), I control for time-varying industry fixed effects. β_1 still takes a positive and significant coefficient. β_3 takes a negative coefficient for the whole industries and the more competitive, but not for the less competitive industries. This result suggest that volatility weakens the positive response to MRPK only for relatively competitive industries.

Columns (7)-(12) show that in the case of negative investment, only β_1 is negative and significant. β_2 and β_3 are not significant, suggesting that volatility does not seem to affect the negative investment or the investment sensitivity to MRPK.

Table 3:7 Plant-level estimation for investment status

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Investment status	positive vs zero	positive vs zero	positive vs zero	positive vs zero	positive vs zero	positive vs zero	negative vs zero	negative vs zero	negative vs zero	negative vs zero	negative vs zero	negative vs zero
ln(MRPK)	0.0950***	0.109***	0.102***	0.126***	0.130***	0.127***	-0.00695***	-0.00813***	-0.00675***	-0.00814***	-0.0105***	-0.00735***
	(26.57)	(30.40)	(19.59)	(68.11)	(47.69)	(48.41)	(-5.686)	(-4.809)	(-3.514)	(-4.845)	(-4.164)	(-3.206)
volatility1	-0.0676***	-0.0839***	-0.0373**				0.00274	0.00536	-0.00435			
	(-6.253)	(-7.728)	(-2.670)				(0.366)	(0.508)	(-0.363)			
ln(MRPK)*volatility1	-0.0233***	-0.0279***	-0.0119***	-0.00824**	-0.0162***	-0.00335	0.000859	0.00350	-0.00217	-0.00350	0.000276	-0.00488
	(-6.538)	(-7.278)	(-3.039)	(-2.304)	(-2.941)	(-0.713)	(0.374)	(1.093)	(-0.601)	(-0.906)	(0.0498)	(-0.865)
Observations	1,185,147	622,315	558,119	1,185,138	622,307	558,118	254,082	126,503	123,288	252,770	125,859	122,610
Adjusted R-squared	0.274	0.280	0.279	0.289	0.292	0.292	0.262	0.269	0.258	0.268	0.273	0.267
Fixed effects	plant+year	plant+year	plant+year	plant + industry*year	plant + industry*year	plant + industry*year	plant+year	plant+year	plant+year	plant + industry*year	plant + industry*year	plant + industry*year
sample	all	markup smaller than median	markup larger than median	all	markup smaller than median	markup larger than median	all	markup smaller than median	markup larger than median	all	markup smaller than median	markup larger than median

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Authors' estimations, based on the Census of Manufacture by METI.

Next I turn to the intensive margin. I estimate the following equation for the full sample and subsamples divided by whether the plant-year conducts positive or negative investment:

$$\frac{I_{it}}{K_{it}} = \beta_1 \ln MRPK_{it} + \beta_2 Volatility_{st} + \beta_3 \ln MRPK_{it} \times Volatility_{st} + FE_i + FE_t + \varphi_{it} \tag{39}$$

We expect β_1 to be positive. As for β_2, I expect it to be positive or negative depending on the adjustment costs are asymmetric or symmetric as the bottom panel of Figures 3:2A and 3:2B show. Finally, I expect β_3 to be negative if volatility weakens the plant's response to the change in (the logarithm of) MRPK. In the estimations, I drop the observations with the investment rate is higher than the top 1 percentile.

Table 3:8 reports the estimation results for the intensive margin. Columns (1) to (6) show the results from the sample of the observations that are included regardless of whether the plant-year conducts positive, zero, or negative investment. They show that β_1 is positive and significant, while

β_2 and β_3 are negative and significant only for the smaller-markup industries, suggesting that competition strengthens the adverse effect of volatility on the intensive margin of investment sensitivity to MRPK. Columns (7) to (9) show the results from the restricted sample of plant-year observations with positive investment. The results are similar to those in Columns (4)-(6), although estimated β_1 is larger for this restricted sample. Columns (10) to (12) show the results from the restricted sample of plant-year observations with negative investment, showing that neither β_1 nor β_3 is significant.

Table 3:8 Plant-level estimation of investment ratio

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Investment Status	all	all	all	all	all	all	positive	positive	positive	negative	negative	negative
ln(MRPK)	0.0813***	0.0978***	0.0851***	0.116***	0.126***	0.113***	0.174***	0.178***	0.177***	-0.00383	-0.000326	0.0197
	(18.01)	(22.01)	(14.63)	(72.08)	(51.29)	(50.98)	(71.35)	(49.85)	(49.49)	(-0.182)	(-0.00924)	(1.332)
volatility1	-0.0535***	-0.0757***	-0.0156									
	(-5.542)	(-7.149)	(-1.420)									
ln(MRPK)*volatility1				-0.0132***	-0.0248***	-0.00258	-0.0205***	-0.0242***	-0.0110	-0.0490	-0.170	0.0561
				(-4.504)	(-5.451)	(-0.640)	(-4.419)	(-3.640)	(-1.568)	(-0.758)	(-1.165)	(1.243)
Observations	1,268,088	669,623	598,465	1,254,884	660,618	589,640	692,832	372,331	316,035	11,944	6,381	4,845
Adjusted R-squared	0.143	0.153	0.150	0.157	0.163	0.158	0.183	0.187	0.188	0.001	0.000	0.180
Fixed effects	plant+year	plant+year	plant+year	plant + industry*year	plant + industry*year	plant + industry*year	plant + industry*year	plant + industry*year	plant + industry*year	plant + industry*year	plant + industry*year	plant + industry*year
Sample	all	markup smaller than median	markup larger than median	all	markup smaller than median	markup larger than median	all	markup smaller than median	markup larger than median	all	markup smaller than median	markup larger than median

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Authors' estimations, based on the Census of Manufacture by METI.

These estimation results show that uncertainty decreases both the likelihood (i.e., extensive margin) and the extent (i.e., intensive margin) of positive investment, and product market competition strengthens these adverse effects. The negative impact of uncertainty on the likelihood and extent of positive investment, especially in severely competitive industries, seem to result in a large dispersion in MRPK.

3.5.3 Quantitative effects of uncertainty on aggregate TFP

To what extent does uncertainty affect aggregate TFP? I answer this question by conducting a counterfactual experiment to see to what extent $SD_{st}(MRPK_{it})$ and aggregate TFP would change if volatility decreases by half for each year. Specifically, based on the plant-level estimation (Column (1) in Table 3:8) of investment, I first estimate the hypothetical $\frac{I_{it}}{K_{it}}$ that would be realized if $Volatility1_{st}$ reduces by half, and construct the hypothetical K_{it} series for each plant. Using this hypothetical K_{it} and the actual TFPR (ω_{it}), I compute counterfactual aggregate TFP following Hsieh and Klenow (2009). Using the realized TFPR means that I do not consider the possibility that TFPR shocks themselves would change due to smaller volatility, which is likely to underestimate the effect of volatility. In addition, I use the actual labor and materials to compute plant-level output without considering that labor and materials would adjust to the hypothetical capital. This would further underestimate the degree to which halving volatility increases aggregate TFP. Consequently, the obtained counterfactual aggregate TFP can be regarded as the lower bound of the impacts of volatility. See Appendix 2 for the detailed procedure of this counterfactual experiment.

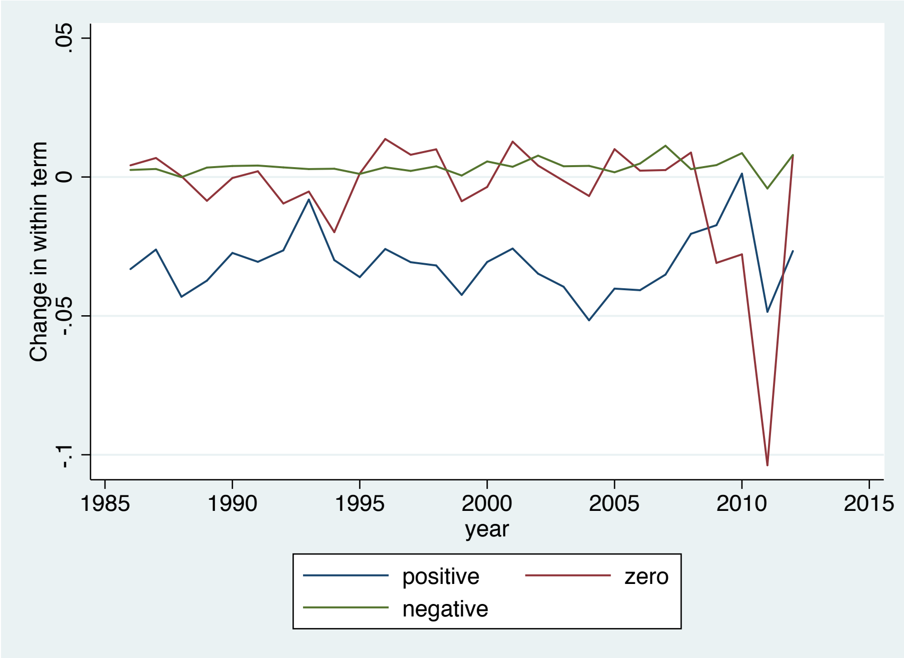
Before presenting the results from this counterfactual experiment, I see if investment, if conducted, actually reduces the dispersion in MRPK. The following decomposition is useful:

$$\begin{aligned} &\Delta Var(\ln MRPK_{it}) \\ &= \sum_{S=1}^{3} Prob(S_{it-1}) \Delta Var(\ln MRPK_{it} \,|S_{it-1}) \\ &+ \sum_{S=1}^{3} Prob(S_{it-1}) \, \Delta\big(E(\ln MRPK_{it} \,|\mathrm{S}_{it-1}) \\ &- E(\ln MRPK_{it})\big)^2 \end{aligned} \tag{40}$$

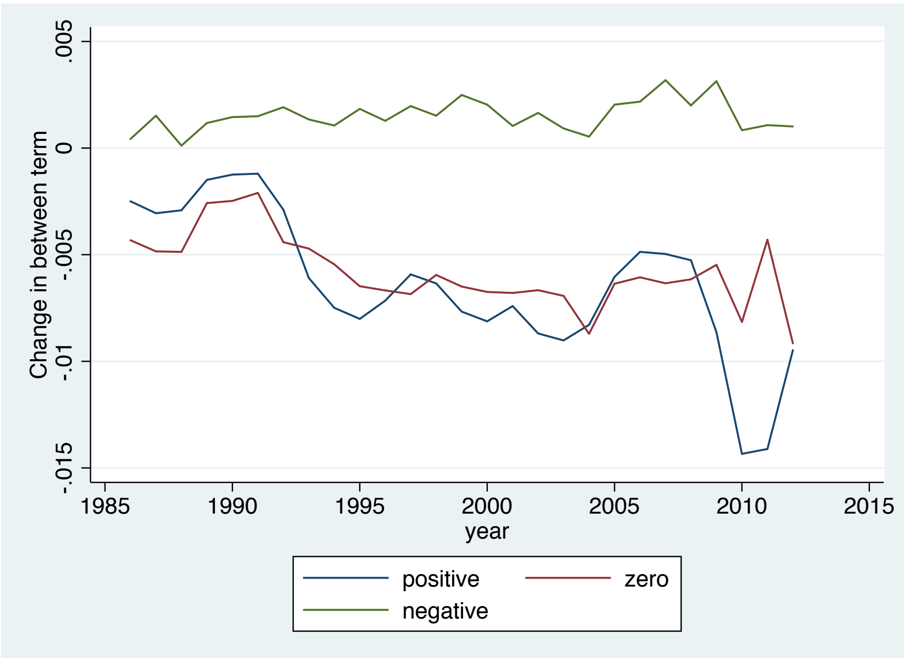
where S_{it-1} denotes the investment status of whether plant i conducted positive, zero, or negative investment in year $t-1$. Eq. (40) shows that the change in the variance of MRPK is the sum of the change in the variance of MRPK within the same status group (the first term: within effects) and that between different status groups (the second term: between effects). Figures 3:7A and 3:7B show that positive investment, if conducted, actually reduces the dispersion in MRPK both through the within and between effects.

Figure 3:7 Changes in within-industry dispersion in MRPK among establishments with positive, zero, and negative investment

A. Changes in within terms, SD $\left(\ln\left(\frac{MRPK_{it}}{\overline{MRPK_{st}}}\right) \text{conditional on investment status}\right)$



B. Changes in between terms, $\left[E\left(\ln\left(\frac{MRPK_{it}}{MRPK_{st}}\right)\text{conditional on investment status}\right) - E\left(\ln\left(\frac{MRPK_{it}}{MRPK_{st}}\right)\right)\right]^2$



Note: I define zero investment as $\left|\frac{I_{t-1}}{K_{t-1}}\right| \leq 0.05$. I accordingly define positive and negative investment using the same threshold.

Source: Authors' calculations, based on the Census of Manufacture by METI.

Now I turn to the counterfactual experiments. Table 3:9 shows the rate of change in $SD_{st}(MRPK_{it})$ and aggregate TFP in the case where $Volatility1_{st}$ reduces by half. Given that the mean of $Volatility1_{st}$ is 0.35, its half accounts for about 0.6 of its standard deviation (0.28) on average. Column 1 and 2 shows that $SD_{st}(MRPK_{it})$ decreases by 2.4% and aggregate TFP increases by 0.7% on average. Columns 3 and 4 show the results from aggregating the more competitive industries, i.e., those industries with their estimated markups are below the median. They show that in such industries, $SD_{st}(MRPK_{it})$ decreases by 2.2% and aggregate TFP increases by 2.1% on average. On the other hand, Columns 5 and 6 show the results from aggregating the less competitive industries, indicating that $SD_{st}(MRPK_{it})$ decreases by 2.7% and aggregate TFP decreases by 0.9% on average. The negative change of aggregate TFP may be a bit surprising, but it is plausible if an increase in investment due to reduced volatility results in worse allocation of capital.[20] The same reason may explain why the decrease in $SD_{st}(MRPK_{it})$ is larger for the less competitive industries than for the more competitive industries. In sum, the counterfactual experiment suggests that the effects of uncertainty on the dispersion in MRPK and the aggregate TFP are economically sizable, and that effect of uncertainty on aggregate TFP is much larger for the more competitive industries than for the less competitive industries.

Finally, I compare the above counterfactual experiment from another counterfactual experiment by which I assume no frictions or distortions exist. In this extreme experiment, the marginal revenue of inputs should be equalized at their unit input costs across plants. To conduct this counterfactual experiment, I follow Hsieh and Klenow (2009)'s methodology with the following modifications. First, I consider the intermediate inputs as one of the inputs. Second, I estimate the input elasticities of output (α_X for $X \in \{K, L, M\}$) and the price elasticity of demand (ε) for each industry as I describe in Section 3.4.2. Finally, I drop outliers in a different way from Hsieh and Klenow's method.[21] First I trim the 1% tails of $\frac{MRPK_{it}}{\overline{MRPK_{st}}}$ and $\frac{TFPR_{it}}{\overline{TFPR_{st}}}$ for each year to estimate the production function, where $\overline{MRPK_{st}}$ and $\overline{TFPR_{st}}$ are industry-level average of $MRPK_{it}$ and $TFPR_{it}$, respectively. Then I drop the plants in the industries for which the sum of the input elasticities is negative or above unity.[22]

[20] In addition, the fact that I use the realized TFPR shock to compute plant-level output might also result in negative change in aggregate TFP.

[21] Nishida et al. (2016) show that the estimated aggregate TFP losses from misallocation are substantially sensitive to the ways of dealing with outliers.

[22] I drop the industries for which the sum of the input elasticities exceeds one because the estimated markup would be negative.

Column 7 of Table 3:9 shows the estimated aggregate TFP gain that would be achieved if the marginal revenue of inputs were equalized across plants: $\frac{TFP^*}{TFP} - 1$, where TFP^* and TFP denote the counterfactual and actual aggregate TFP. Counterfactual aggregate TFP increases by 138.3% as compared to actual one, which is much larger than Hsieh and Klenow's estimates of the U.S., China, and India, but could be reasonable considering that I take into account the intermediate inputs.[23] Comparing with this counterfactual experiment with no frictions or distortions, the effects of uncertainty on aggregate TFP might seem small. However, as Hsieh and Klenow (2009) notes, counterfactual experiments with no frictions or distortions may well be overestimated due to measurement errors and the existence of outliers.[24]

[23] Nishida et al. (2017) show that ignoring intermediate inputs would lead to substantial underestimate of aggregate TFP gain.

[24] Actually, if I trim the 1% tails of $\frac{TFPR_{it}}{\overline{TFPR}_{st}}$ and $\frac{TFPQ_{it}}{\overline{TFPQ}_{st}}$ in each year, then $\frac{TFP^*}{TFP} - 1$ reduces to 47%. In fact most of huge TFP gain is explained by the hike in recent years.

Table 3:9 Rate of change in $SD(MRPK)$ and aggregate TFP when volatility of TFPR shock volatility decreases by half

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Counterfactual (half volatility) vs Actual						No frictions/ distortions vs. Actual
industries	all		more competitive		less competitive		all
	SD(LOGMRPK)	Aggregate TFP	SD(LOGMRPK)	Aggregate TFP	SD(LOGMRPK)	Aggregate TFP	Aggregate TFP
1986							0.799
1987							0.656
1988	−0.025	0.006	−0.023	0.019	−0.029	−0.011	0.563
1989	−0.021	−0.004	−0.025	0.003	−0.019	−0.015	0.585
1990	−0.028	0.023	−0.026	0.051	−0.030	−0.017	0.657
1991	−0.030	0.049	−0.028	0.043	−0.031	0.058	0.650
1992	−0.027	0.000	−0.019	0.004	−0.038	−0.005	0.808
1993	−0.019	0.010	−0.022	0.003	−0.015	0.021	0.609
1994	−0.018	0.005	−0.016	0.018	−0.021	−0.013	0.763
1995	−0.026	0.031	−0.019	0.040	−0.037	0.020	0.812
1996	−0.016	−0.003	−0.015	0.002	−0.017	−0.010	0.855
1997	−0.015	−0.012	−0.012	−0.006	−0.017	−0.020	0.505
1998	−0.030	0.021	−0.031	0.046	−0.029	−0.014	0.631
1999	−0.019	−0.002	−0.018	0.001	−0.020	−0.008	0.632
2000	−0.033	0.070	−0.041	0.126	−0.023	−0.009	0.784
2001	−0.021	0.007	−0.019	0.025	−0.024	−0.020	0.635
2002	−0.023	−0.007	−0.017	0.000	−0.032	−0.021	0.682
2003	−0.020	−0.029	−0.029	0.023	−0.013	−0.029	0.790
2004	−0.017	−0.016	−0.014	−0.009	−0.024	−0.028	0.762
2005	−0.021	0.018	−0.019	0.035	−0.021	−0.014	0.811
2006	−0.029	0.009	−0.031	0.023	−0.024	−0.017	0.709
2007	−0.027	0.023	−0.024	0.060	−0.038	−0.027	0.908
2008	−0.019	−0.004	−0.019	−0.002	−0.020	−0.007	0.886
2009	−0.025	0.021	−0.015	0.041	−0.035	−0.009	4.089
2010	−0.008	−0.019	−0.008	0.006	−0.003	−0.059	4.310
2011	−0.061	−0.006	−0.048	−0.079	−0.089	0.144	2.473
2012	−0.032	0.015	−0.033	0.078	−0.039	−0.087	4.142
2013	−0.016	−0.016	−0.010	−0.005	−0.027	−0.026	5.918
Mean (1988−2013)	−0.024	0.007	−0.022	0.021	−0.027	−0.009	1.383
Median(1988−2013)	−0.022	0.005	−0.019	0.018	−0.024	−0.014	0.773

Source: Authors' calculations, based on the Census of Manufacture by METI.

3.6 Conclusion

Uncertainty affects investment that involves adjustment costs or time-to-build, resulting in dispersion of marginal revenue productivity of capital (MRPK) and consequently in aggregate productivity, depending on the degree of product market competition. Using a simple dynamic model and a large panel dataset of manufacturing plants in Japan, I find that the effect of uncertainty on the dispersion in MRPK is stronger for industries with severer product market competition. The counterfactual experiment suggests that the effects of uncertainty on the dispersion in MRPK and the aggregate TFP are economically sizable, and that effect of uncertainty on aggregate TFP is much larger for the more competitive industries than for the less competitive industries.

The results suggest that the aggregate consequences of competition policy should not be judged based on the dispersion in MRPK or TFPR, which are often used as a measure of misallocation.

While this study sheds new lights on the role of competition in the uncertainty-productivity dispersion relationship, I have not yet explored the durability of uncertainty-driven dispersion in MRPK. If the major source of such productivity dispersion is time-to-build, then uncertainty-driven dispersion in MRPK may be short-lived. I explore this issue in future work.

3.7 Appendix

3.7.1 Analytical solution to the simplified model

In this Appendix, I first derive aggregate TFP in the case where time-to-build exists, and then compare it with the aggregate TFP in the case where capital adjusts without time lag.

In the presence of time-to-build, plant i's problem (16) leads to the following optimal inputs and output:

$$
\begin{aligned}
K_i \quad = \quad & \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_L}\left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_M}\right\}^{\epsilon} \\
& * \left\{\left(\frac{\alpha_K}{P_K}\right) E\left[u_i^{\frac{1}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}}\right]\right\}^{\epsilon\left[1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)\right]} \Omega_{-1i}^{\epsilon}
\end{aligned}
$$

$$
\begin{aligned}
L_i \quad = \quad & \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)(\alpha_L-1)+1}\left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_M}\left(\frac{\alpha_K}{P_K}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_K}\right\}^{\epsilon} \\
& * \left\{E\left[u_i^{\frac{1}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}}\right]\right\}^{\epsilon\left(1-\frac{1}{\epsilon}\right)\alpha_K} u_i^{\frac{1}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}} \Omega_{-1i}^{\epsilon},
\end{aligned}
$$

$$
\begin{aligned}
M_i \quad = \quad & \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_L}\left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)(\alpha_M-1)+1}\left(\frac{\alpha_K}{P_K}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_K}\right\}^{\epsilon} \\
& * \left\{E\left[u_i^{\frac{1}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}}\right]\right\}^{\epsilon\left(1-\frac{1}{\epsilon}\right)\alpha_K} u_i^{\frac{1}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}} \Omega_{-1i}^{\epsilon},
\end{aligned}
$$

and

$$B_i Q_i = \left(1-\frac{1}{\epsilon}\right)^{\epsilon} \left(\frac{\alpha_K}{P_K}\right)^{\epsilon\alpha_K} \left(\frac{\alpha_L}{P_L}\right)^{\epsilon\alpha_L} \left(\frac{\alpha_M}{P_M}\right)^{\epsilon\alpha_M} u_i^{\frac{\frac{\epsilon}{\epsilon-1}}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}} * \left\{E\left[u_i^{\frac{1}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}}\right]\right\}^{\epsilon\alpha_K} \Omega_{-1i}^{\frac{\epsilon^2}{\epsilon-1}} \left(PQ^{\frac{1}{\epsilon}}\right)^{-\frac{\epsilon}{\epsilon-1}},$$

where $u_i = e^{\sigma\kappa_{it}}$ and time subscript 0 denotes t-1.

Aggregating inputs and outputs across plants lead to

$$K = \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_L} \left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_M} \left(\frac{\alpha_K}{P_K}\right)^{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}\right\}^{\epsilon} * \int \Omega_{-1i}^{\epsilon} \, di \left\{E\left[u_i^{\frac{1}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}}\right]\right\}^{\epsilon-(\epsilon-1)(\alpha_L+\alpha_M)},$$

$$L = \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)(\alpha_L-1)+1} \left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_M} \left(\frac{\alpha_K}{P_K}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_K}\right\}^{\epsilon} * \int \Omega_{-1i}^{\epsilon} di \left\{E\left[u_i^{\frac{1}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}}\right]\right\}^{\epsilon-(\epsilon-1)(\alpha_L+\alpha_M)},$$

$$M = \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_L} \left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)(\alpha_M-1)+1} \left(\frac{\alpha_K}{P_K}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_K}\right\}^{\epsilon} * \int \Omega_{-1i}^{\epsilon} di \left\{E\left[u_i^{\frac{1}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}}\right]\right\}^{\epsilon-(\epsilon-1)(\alpha_L+\alpha_M)},$$

and

$$Q = \left(1-\frac{1}{\epsilon}\right)^{\epsilon} \left(\frac{\alpha_K}{P_K}\right)^{\epsilon\alpha_K} \left(\frac{\alpha_L}{P_L}\right)^{\epsilon\alpha_L} \left(\frac{\alpha_M}{P_M}\right)^{\epsilon\alpha_M} * \left\{E\left[u_i^{\frac{\epsilon}{\epsilon-(\epsilon-1)(\alpha_L+\alpha_M)}}\right]\right\}^{\frac{\epsilon}{\epsilon-1}[\epsilon-(\epsilon-1)(\alpha_L+\alpha_M)]} \left\{\int \Omega_{-1i}^{\epsilon} di\right\}^{\frac{\epsilon}{\epsilon-1}}$$

Substituting these aggregate inputs and output to the definition of aggregate TFP, $A \equiv \frac{Q}{K^{\alpha_K} L^{\alpha_L} M^{\alpha_M}}$ yields

$$A = \left\{\int \Omega_{-1i}^{\epsilon} \, di\right\}^{\frac{1}{\epsilon-1}} \left\{E\left[u_i^{\frac{1}{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}}\right]\right\}^{\frac{\epsilon}{\epsilon-1}\left[1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)\right]}.$$

Next, I turn to the case where there exists no time-to-build. In this case, the optimal inputs and output are as follows.

$$K_i^* = \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_K}{P_K}\right)^{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_L}\left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_M}\right\}^{\epsilon}\left(u_i\Omega_{-1i}\right)^{\epsilon}$$

$$L_i^* = \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)(\alpha_L-1)+1}\left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_M}\left(\frac{\alpha_K}{P_K}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_K}\right\}^{\epsilon}\left(u_i\Omega_{-1i}\right)^{\epsilon}$$

$$M_i^* = \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_L}\left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)(\alpha_M-1)+1}\left(\frac{\alpha_K}{P_K}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_K}\right\}^{\epsilon}\left(u_i\Omega_{-1i}\right)^{\epsilon}$$

Aggregating these inputs across plants yields

$$K^* = \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_K}{P_K}\right)^{1-\left(1-\frac{1}{\epsilon}\right)(\alpha_L+\alpha_M)}\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_L}\left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_M}\right\}^{\epsilon} * \int \Omega_{-1i}^{\epsilon} di \int (u_i)^{\epsilon}\, di$$

$$L^* = \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)(\alpha_L-1)+1}\left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_M}\left(\frac{\alpha_K}{P_K}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_K}\right\}^{\epsilon} * \int \Omega_{-1i}^{\epsilon} di \int (u_i)^{\epsilon}\, di$$

$$M^* = \left\{\left(1-\frac{1}{\epsilon}\right)\left(\frac{\alpha_L}{P_L}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_L}\left(\frac{\alpha_M}{P_M}\right)^{\left(1-\frac{1}{\epsilon}\right)(\alpha_M-1)+1}\left(\frac{\alpha_K}{P_K}\right)^{\left(1-\frac{1}{\epsilon}\right)\alpha_K}\right\}^{\epsilon} * \int \Omega_{-1i}^{\epsilon} di \int (u_i)^{\epsilon}\, di$$

Aggregate output must satisfy (1). Substituting plant-level optimal inputs into (1) yields

$$Q^* = \left(1-\frac{1}{\epsilon}\right)^{\epsilon}\left\{\left(\frac{\alpha_K}{P_K}\right)^{\alpha_K}\left(\frac{\alpha_L}{P_L}\right)^{\alpha_L}\left(\frac{\alpha_M}{P_M}\right)^{\alpha_M}\right\}^{\epsilon} * \left\{\int \Omega_{-1i}^{\epsilon} di\right\}^{\frac{\epsilon}{\epsilon-1}}\left\{\int (u_i)^{\epsilon}\, di\right\}^{\frac{\epsilon}{\epsilon-1}}$$

Aggregate TFP without time-to-build is

$$A^* = \left\{\int \Omega_{-1i}^{\epsilon} di\right\}^{\frac{1}{\epsilon-1}}\left\{\int (u_i)^{\epsilon}\, di\right\}^{\frac{1}{\epsilon-1}}$$

Comparing A and A^* lead to the TFP ratio in the main text.

3.7.2 Counterfactual experiment

1. I construct the counterfactual investment based on Eq. (39) as

$$I_{it}^{CF} = K_{it}\left[\beta_1 \ln MRPK_{it} + \beta_2 \frac{Volatility_{st}}{2} + \beta_3 \ln MRPK_{it} \times \frac{Volatility_{st}}{2} + FE_i + FE_t + \varphi_{it}\right],$$

or

$$I_{it}^{CF} = I_{it} - \frac{1}{2} K_{it}[\beta_2 Volatility_{st} + \beta_3 \ln MRPK_{it} \times Volatility_{st}].$$

where superscript *cf* denotes counterfactuals hereafter.

2. I construct the counterfactual capital stock as

$$K_{it}^{CF} \equiv K_{it} - I_{it-1} + I_{it-1}^{CF},$$

or

$$K_{it}^{CF} = K_{it} - \frac{1}{2} K_{it-1}[\beta_2 Volatility_{st-1} + \beta_3 \ln MRPK_{it} \times Volatility_{st-1}].$$

3. I compute the counterfactual log of sales as

$$s^{CF} = \omega_{it} + \beta_K k_{it}^{cf} + \beta_L l + \beta_M m,$$

where lower cases denote logs.

4. I compute actual and counterfactual logs of TFPR defined by Hsieh and Klenow (2009) as

$$\ln TFPR_{HK_{it}} = s - \alpha_K k - \alpha_L l - \alpha_M m,$$

and

$$\ln TFPR_{HK_{it}}^{CF} = s^{CF} - \alpha_K k^{CF} - \alpha_L l - \alpha_M m.$$

Note that

$$\ln TFPR_{HK_{it}}^{CF} = \ln TFPR_{HK_{it}} - \frac{\alpha_K}{\epsilon}(k_{it}^{CF} - k_{it}).$$

5. Using $\ln TFPR_{HK_{it}}$ and $\ln TFPR_{HK_{it}}^{CF}$, I compute industry-level TFP for each industry and year following Hsieh and Klenow (2009).

6. I aggregate industry-level counterfactual TFP using industry-level sales as a weight.

Chapter 4 Minimum Wage Effects across Heterogeneous Markets[1]

Previous studies have reached little consensus on the employment effect of the minimum wage. This chapter argues that local labor markets are heterogeneous, in that the impact of the minimum wage is concentrated in specific markets. In particular, I estimate the extent of surplus between each plant's value of marginal product of labor and wage rate and examine whether the minimum wage impact varies across markets with a differential surplus. Building on the recent policy changes in Japan, I find that the employment effect of an increase in the minimum wage is significant and negative in plants for which the value of marginal product of labor is close to the wage rate and which experienced little surplus pre-reform. The minimum wage increases have little employment effect on plants with a relatively high surplus, even when they have a significant number of minimum wage employees.

4.1 Introduction

The extent of market power in the labor market is an important topic in testing the employment effect of the minimum wage. In a simple competitive labor market model, employers have little control over wages and reduce employment levels immediately with an increase in the minimum wage. However, in a monopsonistic labor market model, employers have some power to control wages and workers are paid less than the value of the marginal product of labor. In the presence of such a surplus, the employers react to an increase in the minimum wage differently from those in a competitive labor market model; this scenario sometimes leads to a rise in the employment level (Card and Krueger, 1994; Manning, 2003). However, despite the divergent consequences that the two models predict, studies have paid little attention to identifying the extent of market power in testing the employment effect of the minimum wage. In fact, there was a lack of convincing evidence to support the existence of monopsony in the labor market (Kuhn, 2004).

---

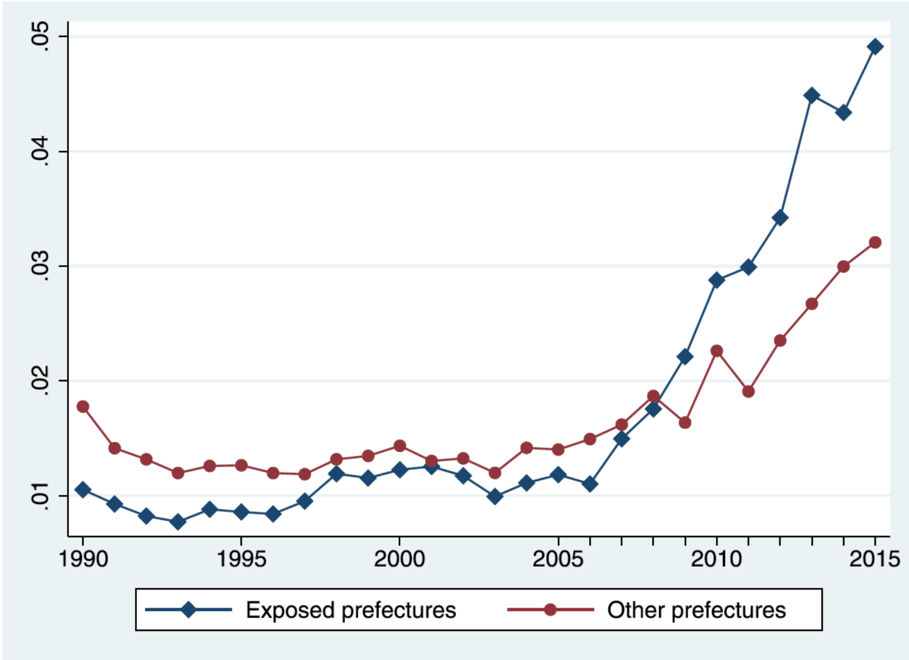
[1] The revised version of this chapter is in *Labour Economics* 59, 110-122, co-authtored with Hiroko Okudaira and Miho Takizawa.

Only recently, a growing number of studies presented the evidences of market power (Azar et al., 2017; 2018; Benmelech et al., 2018; Dube et al., 2016; Falch, 2010; Ransom and Sims, 2010; Staiger et al., 2010). For instance, Benmelech et al. (2018) constructed the Herfindahl-Hirschman Index (HHI) from the Longitudinal Business Database to measure the labor market concentration and found that the employer concentration is negatively associated with local wages. Interestingly, Azar et al. (2018) highlighted large geographical variations in the employer concentration. Analyzing a large-scale dataset from Burning Glass Technologies —which collects job vacancy information from approximately 40,000 websites —they found that the less-populated commuting zones and the zones in the Great Plains tended to have lower employer concentrations. An important takeaway on the study of minimum wage is that the local labor markets are heterogeneous in that employers can respond to an increase in minimum wage differently across local labor markets. This study directly estimates the labor market surplus by examining how far an employer is from its competitive optimal decisions and tests whether the employment effect of the minimum wage differs across regions depending on the extent of surplus. Specifically, I estimate the surplus or wage markdown that employers would face in labor markets with any frictions, which is defined as the discrepancy between the value of the marginal product of labor (VMPL) and the wage rate. The original idea comes from Petrin and Sivadasan (2013), who examined the correlation between plants surplus and the reform of firing restrictions in Chile. Taking a very similar approach, Dobbelaere and Mairesse (2013) and Dobbelaere et al. (2015) proposed a way to test market imperfections both in the labor and product markets. By applying the approach from the previous studies to Japanese manufacturing census data, I first estimate the production functions and obtain estimates for the elasticities of factor inputs to calculate the surplus in the labor market. I then use the estimated extent of the surplus to examine whether the employment effect of the minimum wage differs according to the extent of market power faced by employers in the local labor market. In the main specification, I follow the framework of Meer and West (2016) and focus on employment growth rather than on the employment level to account for the possibility that job destruction occurs gradually over time.

The identification of the minimum wage effect relies on the exogenous policy event wherein a series of Japanese government policies substantially increased regional minimum wages over a decade. The first event took place in 2007 when the Minimum Wage Act was amended to provide a legal framework to increase regional minimum wages to or above the level of welfare benefits defined for each region. The amendment disproportionately affected prefectures that initially had a high-level of welfare benefits and thus were requested to raise their regional minimum wages. Figure 4:1 presents a part of the identification variations used in the analysis. Graphs indicate the proportions of minimum-wage workers separately for prefectures that initially had high welfare benefits (Exposed prefectures) and for non-exposed prefectures (Other prefectures). The exposed prefectures experienced disproportionately higher increases in the proportion of minimum-wage workers after the amendment than the other prefectures. Importantly, regions exposed to this shock were those situated in urban areas or ones with wintry weather and not necessarily the regions that shared specific economic trends. Moreover, since such a central policy had much more influence on the determination of regional minimum wages when compared to the influence of the local authorities, it alleviates the concern that preexisting local employment trends may confound the results. The government's initiatives on minimum wage have continued in the following decade due to the wage-boosting policy of Prime Minister Shinzo Abe, which provides us with an opportunity to exploit the minimum wage variations that are less likely to reflect local economic trends over a relatively long time period. I test the exogeneity of these events by adopting the suggestion of Meer and West (2016) and find that preexisting local trends had no predictive power. I also conduct instrumental variable estimations to complement the main analysis: as instruments, I use the target amounts that the central policy sets to propose each region to determine how much minimum wage it should raise.

Consistent with a standard competitive labor market model, the main analysis revealed that plants significantly reduced employment growth in response to increases in the minimum wage. Particularly, the baseline estimates suggest that, on average, plants in exposed prefectures in Figure 4:1 experienced 3.2 to 4.2% lower employment growth from 2007 to 2014, compared to the plants in non-exposed prefectures. However, the estimated negative impact masks the heterogeneity in plants' behavioral response: an increase in minimum wage affected plants in the sample in noticeably different ways, depending on the surplus that plants face. I found that, in response to an increase in minimum wage, plants that initially experienced a large surplus did not significantly reduce their employment growth. In the preferred specifications, the estimated magnitude of the impact is reduced by at least 35% or even gets closer to zero when observations are limited to plants with the wage markdown less than 0.6. While the main data do not contain wages and hours of work for individual employees, the prediction from another source of administrative wage records confirms that the effects of minimum wage are concentrated in plants with a larger proportion of minimum-wage workers. Although the lack of individual hourly wage information prevents us from obtaining precise estimates, the results found in this study largely support the view that the local labor markets are heterogeneous and plants respond to the minimum wage shock differently, depending on the extent of surplus they can exploit in the labor market.

Figure 4:1 Proportion of minimum wage workers by extent of exposed shock



Note: The exposed prefectures are those that initially had relatively lower benefit levels than minimum wage earnings; therefore, were exposed to intense increases in minimum wage after the revision of Minimum Wage Act, which was approved in 2007.

Source: Based on authors' calculation from Basic Survey of Wage Structures (Japanese Ministry of Health, Labour and Welfare).

This study is by no means the first to directly estimate the firm's market power from production function estimations (Dobbelaere et al., 2015; Dobbelaere and Mairesse, 2013; Lu et al., 2017; Petrin and Sivadasan, 2013); however, to the best of my knowledge, this is the first study to apply this methodology to the minimum wage literature. In so doing, I take an advantage of this method. Firm's market power arises from frictions in the local labor market such as heterogeneity in worker's preferences, mobility costs due to the availability of transportation, and imperfect information or ignorance (Manning, 2003). The estimates of wage markdown have one advantage, since they are based on firms internal decisions and likely to reflect all the aforementioned sources of labor market frictions. This contrasts with concentration measurements, such as HHI adopted in recent studies (Azar et al., 2017; 2018; Benmelech et al., 2018). Importantly, the estimates also measure one aspect of labor market friction, which HHI are likely to capture. Although weakly, the estimates for the surplus are negatively associated with the number of rival plants in the same prefecture and industry.

This study also contributes to the literature by adding another explanation on the observation of the mixed employment effects of minimum wage.[2] A simple competitive labor market model predicts a negative employment effect of minimum wage. However, previous studies have already examined that firms could otherwise respond to the minimum wage shock, for instance, by reducing profits (Draca et al., 2011), increasing product prices (Aaronson and French, 2007), and substituting toward more productive labor (Horton, 2017). Recent empirical studies shed deeper light on the underlying mechanism. Harasztosi and Lindner (2019) exploited a sharp minimum wage hike in Hungary to report that firms passed most of the increased cost to consumers and also substituted labor with capital in response to the hike. Interestingly, they found a heterogeneity in firms' responses in which the disemployment effect is larger among firms in tradable sectors, wherein it is more difficult to increase product prices due to the existence of foreign competitors. In a similar vein, Cengiz et al. (2019) also found that the employment effect of minimum wage is not uniform across industries in the U.S., and that the disemployment effect is concentrated in the tradable or manufacturing sector. While these two studies look at the heterogeneous effects considering product market competitions, this study adds a new aspect to the literature by focusing on the labor market competitions. The employment effect of the minimum wage also differs depending on the extent of employer's market power in the local labor market. Thus, aggregating the employment effects across local labor markets masks heterogeneity in employers' responses, which may indicate one possible reason why previous studies have observed mixed employment effects of the minimum wage.

The remainder of the chapter is organized as follows. Section 4.2 introduces the theoretical framework used to measure the extent of labor market surplus and summarizes the institutional background. Section 4.3 describes the data and identification strategies. Section 4.4 presents the results along with some robustness tests. Section 4.5 concludes.

---

[2] Influential case studies by Card and Krueger (1994, 2000) in New Jersey and Pennsylvania found no disemployment effects, while Neumark and Wascher (1992) found significantly negative employment effects of an increase in the minimum wage among teenagers in the state-level dataset. Dube et al. (2010) constructed a dataset containing all contiguous-border-county pairs in the U.S. to generalize a case study by Card and Krueger (1994) and showed that an increase in the minimum wage has a significantly positive earnings effect but no significant employment effect. Neumark et al. (2014) criticized Dube et al. (2010)'s approach by showing that they failed to include sufficient identification variations and test the need to control for local trends. Allegretto et al. (2018) argued against Neumark et al. (2014) by adopting a synthetic control approach.

4.2 Background

4.2.1 Measuring surplus across labor markets

Although numerous studies have implied that the employment effects of minimum wage depend on a firm's ability to pass costs through to product prices Harasztosi and Lindner (2019), job-to-job turnover rates (Dube et al., 2016; Giuliano, 2013), and the degree of labor market monopsony (Card and Krueger, 1994), the majority of previous studies have treated labor markets as uniform within each nation. This chapter measures the labor market competitiveness or frictions and examines whether this presumption is plausible. To measure frictions in labor markets, I employ an approach proposed in previous studies (Dobbelaere et al., 2015; Dobbelaere and Mairesse, 2013; Lu et al., 2017; Petrin and Sivadasan, 2013), which calculates market competitiveness from production function estimates. The idea is to estimate how each plant deviates from its cost minimization behavior. In particular, a plant i at time period t has the following cost function:

$$TC(K_{it}, L_{it}, M_{it}) = C_K(K_{it}) + P_M M_{it} + W(L_{it})L_{it}, \tag{1}$$

where K_{it}, L_{it}, and M_{it} denote capital, labor, and intermediate input, respectively. I assume perfect competition for intermediate input, and so the price of intermediate input, P_M, is constant within markets. For the labor markets, I assume the employer has monopsony power and faces an upward-sloping labor supply curve. The wage rate, $W(L_{it})$, is therefore an increasing function of employment (inverse labor supply curve). $C_K(K_{it})$ is the capital cost and the functional form is not imposed. The plants choose the amounts of intermediate input and labor to minimize their production cost given a certain amount of production, $Q_{it}(K_{it}, L_{it}, M_{it}) = \bar{Q}$. The first-order condition for intermediate input is derived as follows:

$$P_M = \frac{\lambda_{it} \partial Q_{it}}{\partial M_{it}} \tag{2}$$

where λ_{it} is the Lagrange multiplier and indicates marginal cost. Transforming the above condition, output elasticity with respect to the intermediate input, $\varepsilon_{M_{it}} \equiv \frac{\partial Q_{it}/Q_{it}}{\partial M_{it}/M_{it}}$, is derived as

$$\varepsilon_{M_{it}} = \frac{P_M M_{it}}{\lambda_{it} Q_{it}} \tag{3}$$

We define markup as the ratio of the output price, P_{it}, to marginal cost. Using the above calculation, the markup can be expressed as the ratio of the output elasticity to the cost share of intermediate input, $\alpha_{M_{it}} \equiv \frac{P_M M_{it}}{P_{it} Q_{it}}$:

$$\mu_{it} \equiv \frac{P_{it}}{\lambda_{it}} = \frac{\varepsilon_{M_{it}}}{\alpha_{M_{it}}} \tag{4}$$

The condition of cost minimization for labor input is similarly derived as follows:

$$W_{it}\left(\frac{1}{1+\varepsilon_{it}^{L}}\right) = \lambda_{it}\frac{\partial Q_{it}}{\partial L_{it}} \tag{5}$$

where $\varepsilon_{it}^{L} \equiv \frac{\partial L_{it}/L_{it}}{\partial W_{it}/W_{it}}$ is the wage elasticity of the labor supply. The left hand side of Eq. (5) is the marginal cost of labor and it is equal to VMPL if I assume profit maximization instead of cost minimization for labor input. To follow Lu et al. (2017), I measure the labor market competitiveness, or surplus, by

$$\eta_{it} = \frac{W_{it}\alpha_{L_{it}}}{\lambda_{it}\dfrac{\partial Q_{it}}{\partial L_{it}}}\frac{\varepsilon_{it}^{L}}{\varepsilon_{it}^{L}+1} \tag{6}$$

Under a perfectly competitive market, $\eta_{it} = 1$. In this case, the wage rate is equalized to the marginal cost of labor. Under a monopsonistic market, on the other hand, the surplus term is strictly less than one ($\eta_{it} < 1$) and plants can lower the wage rate by reducing labor demand. Using the expression in Eq. (3), the surplus is written as follows:

$$\eta_{it} = \frac{\alpha_{L_{it}}}{\alpha_{M_{it}}}\frac{\varepsilon_{M_{it}}}{\varepsilon_{L_{it}}} \tag{7}$$

where $\alpha_{L_{it}} \equiv \frac{P_L L_{it}}{P_{it} Q_{it}}$ is the cost share of labor input and $\varepsilon_{M_{it}} \equiv \frac{\partial Q_{it}/Q_{it}}{\partial L_{it}/L_{it}} it$ is the output elasticity of labor input. I use η_{it} to measure the competitiveness of the local labor market that each plant faces. The calculation of η_{it} is straightforward, since I can directly calculate cost shares from the data and estimate output elasticities from production function estimations.

4.2.2 Minimum wage in Japan

Japanese minimum wage is determined mainly at the regional level. Japan consists of 47 prefectures, each of which sets its own regional minimum wage and considers a revision annually. It has almost complete coverage of workers. Both regular and non-regular workers are subject to the minimum wage, with only a limited number of exceptions.[3] In addition to the regional minimum wage, local labor bureaus also allow a small increment in some industries, although the number of workers covered by these industry minimums is usually small. I focus on the variations in regional minimum wages to identify the heterogeneous responses to the minimum wage increases.

The key to the analyses in this study is that regional minimum wages were raised substantially after 2007. One main reason for this rapid increase is an institutional change in wage policy. In the early 2000s, it was argued that welfare recipients in some regions receive higher benefits than workers who earn minimum wage, leading to an amendment to the Minimum Wage Act in 2007. The new Minimum Wage Act stipulates that regional minimum wages are to be consistent with the amount of welfare benefits (Art. 9, Part 3), and legally validates a further increase in minimum wage in regions with relatively high initial benefits. The Japanese government also took the initiative to continuously raise the minimum wage, attempting to boost wage standards.[4] Because prices were relatively stable over this period, these policy events increased the minimum wage not only in nominal terms but in real terms.[5]

The continuous increases in minimum wage have substantially raised the proportion of workers affected by it. As shown in Section 4.1, Figure 4:1 reveals that prefectures that initially had higher benefits (i.e., Exposed prefectures) experienced higher increases in the proportion of minimum wage workers after the amendment than the other prefectures. Similarly, Figure 4:8 in Appendix confirms similar disproportionate increases in the Kaitz index for Exposed prefectures.[6] The sharp and continuous increases in minimum wage have raised the proportion of those who work at minimum wage

[3] Exceptions are granted to workers with physical or mental disabilities and those on probation or basic training, where permitted by the local labor bureau. In Japan, workers are often distinguished as regular or non-regular, depending on the type of contract (e.g., permanent vs. fixed-term), extent of employment protection, or the number of work hours; however, regional minimum wage is applied to both groups.

[4] Examples include annual requests made by Prime Minister Shinzo Abe to the Council on Fiscal and Economic Policy.

[5] From 2008 to 2014, the GDP deflator decreased by 3.5% (Cabinet Office), while the consumer price index increased by 0.6% (Statistics Japan, Ministry of Internal Affairs and Communications).

[6] It should be noted that the Kaitz index measures the level of minimum wage in proportion to the average market wage. It does not necessarily measure the extent of minimum wage bites, or that an increasing number of plants have begun paying minimum wages, thus being affected by increases in minimum wage. It also ignores the potential spillover effect: an increase in the minimum wage itself pushes up the average market wage.

disproportionately in specific groups of workers and plants. Figures 4:9 and 4:10 in the Appendix confirm this point. To summarize the main points, the proportion of minimum wage workers has increased, especially among (1) female workers, (2) young and old workers (age < 20 or age > 60), and (3) workers in medium- or small- sized plants.[7]

In order to closely exploit the policy changes in the analysis, I use two types of target amounts as instrumental variables. These target amounts are the amounts that the Central Minimum Wage Council proposes to local authorities at each prefecture every year to determine how much it should raise the minimum wage by. The proposed target amounts are not a mandatory quota, and local authorities may still consider other factors in determining their regional minimum wage.[8] However, after the 2007 amendment, the Central Council set a much higher target level for regional minimum wages before any adjustments by local authorities, which limits the local authority's flexibility in controlling the absolute minimum wage level.

More specifically, the analysis exploits exogenous variations in regional minimum wage driven by the two types of target amounts, which I define as baseline target and benefit target. The baseline target is proposed to all prefectures but is determined by rank groups. Each rank group consists of 5 to 17 prefectures. There are four rank groups (i.e., A to D), and the same baseline target amount is proposed to prefectures in the same rank group. Because the baseline target is proposed at the rank level, not at the prefecture level, the minimum wage variations driven by the baseline target do not reflect any decisions made by each prefecture. By contrast, the benefit target is the extra target amount introduced in 2007 only to those prefectures that initially had higher benefit levels than minimum wage earnings.[9] Important to my identification strategy, the amount of benefit target in the current year does not reflect a specific economic trend in the same year. The welfare benefits were initially high not only because of high living costs, but also because of high heating costs caused by the cold

[7] The most pronounced change can be found in teenage workers; nearly 20% were working for minimum wage in 2015, which is four times the number in 2005. The proportion also more than doubled among small-sized plants. This is in sharp contrast to the modest increases observed among small-sized employers between 1982 and 2002 in an administrative household survey (Kawaguchi and Mori, 2009).

[8] More specifically, the revision process of the minimum wage involves two steps. First, the Central Minimum Wages Council proposes a target amount to which minimum wages are to be raised, after investigating overall market conditions such as prices and market wages. Second, the Regional Minimum Wages Council of each prefecture determines the extent of the minimum wage increase by taking into account the target amount proposed by the Central Council, local labor market conditions, and the standard of living in that region. The revisions take place in October or November every year. Section 2 in Kambayashi et al. (2013) provides more detail on the background.

[9] The benefit target was proposed in terms of the amount necessary to close the gap between the minimum wage and the benefit level in each prefecture, and I use this gap as one of the instruments.

weather. In fact, the exposed prefectures are located in both urban (e.g., Tokyo and Osaka) and rural areas (e.g., Hokkaido and Akita). Thus, local economic conditions in the current period are unlikely to be the primary factor that explains this differential shock. I consider that the 2007 amendment initiated sizable and exogenous increases in those affected by the minimum wage and exploit this variation to identify the employment effects of minimum wage. Section 4.6.1 in the Appendix provides further details on the target amounts.

4.3 Identification strategy

4.3.1 Main data

The main analyses draw on the plant-level administrative data set from the Census of Manufacture, which is conducted every year by the Japanese Ministry of Economy, Trade and Industry.[10] The Census of Manufacture covers nearly an entire population of plants in the manufacturing sector in Japan. It contains detailed information on factor inputs and produced outputs at each plant. I focus on annual files that cover all plants with 30 or more employees (Kou Hyou).[11] The first two panels in Table 4:1 present summary statistics for Census of Manufacture. Panel A presents summary statistics for observations used to estimate production functions, and Panel B presents those for observations used to estimate the impact of minimum wage. Although Table 4:1 presents summary statistics for the final estimation samples (i.e., observations with missing values for wage markdown are removed), the final samples are quite similar to the original samples in terms of the summary statistics of the main variables. Production functions are estimated with observations from 2001 to 2014. Section 4.3.2 explains details of production function estimations. The impacts of minimum wage are estimated with observations from 2008 to 2014, so as to avoid including endogenous variations of minimum wage prior to the 2007 amendment to the Minimum Wage Act. Section 4.3.4 discusses details of the identification of minimum wage effect.

[10] The census information is available online in English through the ministry's web page: http://www.meti.go.jp/english/statistics/tyo/kougyo/index.html.

[11] The survey also has other types of annual files that contain information on all plants with 29 or fewer employees (Otsu Hyou). Since some of these files lack information on fixed assets, which is necessary to estimate production functions, I decided not to use these files in this chapter.

Table 4:1 Summary statistics

	N	Mean	Std. Dev.	P25	P50	P75
Panel A. Census of Manufactures (2001-14)						
lnY	461,043	16.6	1.3	15.7	16.41	17.31
lnL	461,043	4.35	0.79	3.74	4.14	4.74
lnK	461,043	17.17	1.73	16.13	17.09	18.16
lnM	461,043	10.8	1.84	9.77	10.81	11.92
cost share (labor)	461,043	0.23	0.15	0.13	0.2	0.29
cost share (material)	461,043	0.39	0.22	0.21	0.37	0.55
Panel B. Census of Manufactures (2008-14)						
lnL	281,388	4.39	0.81	3.78	4.19	4.8
prefecture share of those aged 15–64	281,388	0.62	0.02	0.61	0.62	0.64
log(prefecture population)	281,388	14.96	0.79	14.41	14.86	15.79
computed proportion of MW workers ($\hat{S}_{it}$)	278,005	0.15	0.18	0.02	0.08	0.22
annual wage bill per employee ($10^4 JPY$)	278,005	378.84	169.19	267.89	361.51	465.54
plant size	278,006	132.19	308.79	43	65	120
ratio of regular workers	278,006	0.33	0.24	0.14	0.27	0.5
$base_{rt}$	281,388	7.65	5.61	4	9.5	11
$benefit_{pt}$	281,388	6.55	14.56	0	0	6
Panel C. Basic Survey on Wage Structure (2008-14, manufactures)						
proportion of MW workers (S_{it})	82,509	0.17	0.25	0	0.04	0.23
annual wage bill per employee ($10^4 JPY$)	82,509	361.18	151.92	253.65	340.98	447.41
plant size	82,509	173.69	495.13	13	38	126
ratio of regular workers	82,509	0.33	0.26	0.13	0.25	0.48

Note: Panel A presents summary statistics for observations used to estimate production functions, Panel B for observations used to estimate column 1 in Table 4:4, and Panel C for observations used to estimate models Table 4:10 in the Appendix.

Source: Authors' calculations, based on the Census of Manufacture by METI.

4.3.2 Estimating the labor market surplus

We measure the labor market surplus using the method explained in Section 4.2.1. To this end, I first estimate the production function to calculate the output elasticities of intermediate input and labor. I posit a translog production function defined as follows:[12]

$$\begin{aligned}\ln Q_{it} = {} & \beta_K \ln K_{it} + \beta_L \ln L_{it} + \beta_M \ln M_{it} + \beta_{KK}(\ln K_{it})^2 \\ & + \beta_{LL}(\ln L_{it})^2 + \beta_{MM}(\ln M_{it})^2 + \beta_{KL} \ln K_{it} \ln L_{it} \\ & + \beta_{KM} \ln K_{it} \ln M_{it} + \beta_{LM} \ln L_{it} \ln M_{it} + u_{it}.\end{aligned} \tag{8}$$

[12] I consider potential substitutions among input factors seriously and do not estimate Cobb-Douglas production functions here.

OLS estimates are not consistent because the inputs are positively correlated with unobserved productivity, $E(X_{it}u_{it}) \neq 0$, where $X_{it} \in \{K_{it}, L_{it}, M_{it}\}$. I therefore follow the method of Blundell and Bond (1998, 2000), which proposes using the system GMM to estimate the production function.[13] In this method, the unobserved productivity is decomposed into three terms: time-invariant average productivity of plant i , productivity shock, and measurement error. In addition, the productivity shock is assumed to follow AR(1) process. The production functions are estimated separately for each industry, on the assumption that plants in the same industry face the same technological parameters (β) across regions. Industry-level estimations allow us to estimate parameters efficiently. I choose industry-level estimation, not industry- prefecture-level estimation, as it allows us to avoid removing specific industries or prefectures from the sample when their sample size is too small at the industry-prefecture level. Section 4.6.2 in the Appendix describes more details about the production function estimation. Using estimated parameters, I calculate the output elasticities for each input. In the calculation, I use the median values for the inputs in each industry-prefecture group:

$$\hat{\varepsilon}_K = \hat{\beta}_K + 2\hat{\beta}_{KK} \ln \bar{K} + \hat{\beta}_{KL} \ln \bar{L} + \hat{\beta}_{KM} \ln \bar{M} \tag{9}$$

$$\hat{\varepsilon}_L = \hat{\beta}_L + 2\hat{\beta}_{LL} \ln \bar{L} + \hat{\beta}_{KL} \ln \bar{K} + \hat{\beta}_{LM} \ln \bar{M} \tag{10}$$

$$\hat{\varepsilon}_M = \hat{\beta}_M + 2\hat{\beta}_{MM} \ln \bar{M} + \hat{\beta}_{KM} \ln \bar{K} + \hat{\beta}_{LM} \ln \bar{L} \tag{11}$$

where $\bar{X}$ shows the median value for input X. The median values are taken across plant-year observations from 2001 to 2014; however, as a robustness check, I also use median values from 2000 to 2007 to exclude a potential endogeneity issue between the market scheme and the minimum wage in Section 4.6.5 in the Appendix. The median values are taken across industry-prefecture groups to account for the fact that plants in different prefectures face different production levels and therefore different output elasticities.[14] The cost shares are also aggregated into industry-prefecture groups by taking median values.[15] Finally, the markup and the labor market surplus are measured by taking the ratios of these elasticities and cost shares. I drop the observations in the markets where either labor

[13] The translog production function is estimated by the system GMM in Söderbom and Teal (2004) and Lee et al. (2013).

[14] Another reason is that it allows us to measure regional variations in labor market surpluses, which can arise from geographical proximity to rivals.

[15] The labor cost share is obtained by dividing the total wage bill by total revenue. The total wage bill includes salaries, bonuses, and severance payments. The labor cost share thus includes an important part of the adjustment cost.

or intermediate elasticity is negative. To deal with extreme values, the observations in markets with the top 5% of surplus, η, are also dropped.

Although studies have developed various ways to estimate production functions (Ackerberg et al., 2015; De Loecker and Goldberg, 2013; Gandhi et al., 2013; Levinsohn and Petrin, 2003; Olley and Pakes, 1996; Wooldridge, 2009), I adopt the system GMM approach over the other procedures for the following reasons. First, system GMM allows us to consistently estimate the parameters in the presence of plant fixed effects, which I consider a realistic specification. I aim to obtain consistent production function coefficients, rather than productivity estimates in this chapter. Second, although I also estimate the production functions with Wooldridge (2009) 's widely adopted method in Section 4.6.5 in the Appendix, it yields negative or large estimates for output elasticities in a non-trivial proportion of industry-prefecture groups. Due to these implausible values, the number of industry-prefecture groups has to be reduced from 1602 in system GMM to 799 in Wooldridge (2009) 's method. Although the results of the two methods point to similar implications, I adopt system GMM as a main framework so as not to disproportionately select specific industry-prefecture groups into the final minimum wage estimation.[16]

Table 4:2 presents summary statistics from production function estimations with this system GMM method. The estimates take plausible values. A sum of the three input elasticities ranges around unity, suggesting constant returns to scale. Summary statistics for $\hat{\eta}$ suggest that most of the industry-prefecture groups face some surplus in their local labor market.

[16] Potential biases in production function estimation includes omitted price bias. As is often the case in previous studies, I deflate the nominal revenues as well as input expenditures by the industry price index. If firms face a downwardsloping demand curve, a negative correlation might arise between firm-level price deviations and input price, thereby biasing the output elasticity estimates downward. On the other hand, other estimation issues arise if I adjust the revenues by the output price information. Estimation of a quantity-based production function without any quality adjustment again leads to downward biased parameter estimates as the product price reflects the product quality (De Loecker and Goldberg, 2013). Although this is a significant issue to be addressed in future research, I consider this to be beyond the scope of the research, and follow a standard approach to adjust the revenue with industry price information as has been done in previous studies (Petrin and Sivadasan, 2013, for example).

Table 4:2 Production function estimates

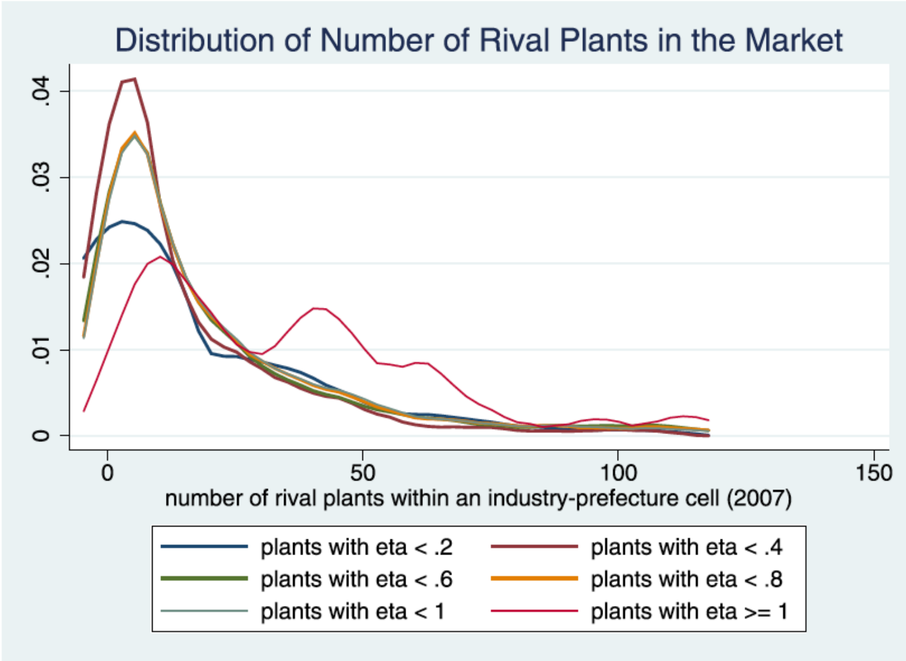
	N	Mean	Std. Dev.	P25	P50	P75
$\hat{\varepsilon}_L$	1,602	0.45	0.23	0.26	0.4	0.59
$\hat{\varepsilon}_M$	1,602	0.48	0.14	0.4	0.48	0.57
$\hat{\varepsilon}_K$	1,602	0.11	0.13	0.06	0.12	0.19
$\hat{\varepsilon}_L + \hat{\varepsilon}_M + \hat{\varepsilon}_K$	1,602	1.03	0.25	0.87	0.96	1.18
$\hat{\eta}$	1,602	0.67	0.35	0.43	0.6	0.84
$\hat{\mu}$	1,602	1.3	0.52	0.97	1.22	1.54

Note: Translog production functions are estimated separately for each industry group. The estimation procedure follows System GMM. The estimated production function estimates are then used to calculate the parameters in this table by using median values for other input variables within each prefecture-industry group.

Source: The data comes from Census of Manufactures (METI).

The estimates of plants' labor market surplus or wage markdown measure one important aspect of frictions in the labor market. To intuitively understand this point, Table 4:3 tabulates the median and average numbers of rival plants in the same prefecture-industry group by the estimated surplus. Similarly, Figure 4:2 draws kernel estimates for distributions of the number of rival plants by $\hat{\eta}$. Table 4:3 and Figure 4:2 imply that the estimated surplus $\hat{\eta}$ tends to correlate positively with the number of rival plants. In particular, plants with $\hat{\eta} < 0.4$ are likely to have smaller numbers of rivals. Tables 4:8 and 4:9 in the Appendix provide industry- or region-level summary statistics. Roughly speaking, urban regions such as Tokyo tend to have higher $\hat{\eta}$, while the same proportions are lower in rural regions such as Hokkaido. This is consistent with a monopsonistic labor market model where employers have control over wages and enjoy some surplus. Although surplus in the labor market could arise from other factors such as the heterogeneous preference of workers, adjustment costs etc. (Manning, 2003; Petrin and Sivadasan, 2013) and I by no means argue that the number of rival plants have a high predictive power, the estimates suggest that geographical proximity is one source of friction workers face in local labor markets. Table 4:14 in the Appendix, I examine the heterogeneity of minimum wage effect in terms of the number of rival plants. While the results roughly point to the same direction with the main results, the estimates are admittedly noisy. I consider that the measurement ($\hat{\eta}$) better describes the extent of plants market power than the number of rival plants does.

Figure 4:2 Number of rival plants within the same prefecture-industry group by surplus



Note: The figures show kernel estimates for the number of rival plants in the same prefecture-industry group.

Source: The data comes from Census of Manufactures (METI).

Table 4:3 Number of rival plants

$\hat{\eta_{pj}}$	< 0.1	< 0.2	< 0.3	< 0.4	< 0.5	< 0.6	< 0.7	< 0.8	< 0.9	< 1	all
median	0	7	6	7	9	10	11	11	11	11	9
mean	6.8	20.6	13.9	16.8	23.5	24.14	24.3	24.4	24.1	23.9	20.9

Note: Figures present median or mean of number of rival plants within the same prefecture-industry group by surplus or $\widehat{\eta_{pj}}$.

Source: The data comes from Census of Manufactures (METI).

4.3.3 Identifying minimum-wage plants

The Census of Manufacture provides a broad set of operational information including product prices at each plant, but unfortunately does not contain information on hours worked or wage rates for individual workers, which is necessary to measure the extent of minimum wage shock at each plant. To supplement the analysis, I use another administrative data source to compute the number of minimum wage workers each plant employs, in keeping with the spirit of Draca et al. (2011) and Aaronson et al. (2012). In particular, I draw on worker files from the Basic Survey on Wage Structure (BSWS), which contains information on individual employees' hours of work, wages, and benefits, including overtime work hours and payment at each plant.[17] The survey samples plants with 10 or more regular employees and plants with five to nine employees in private sectors only. I use pooled cross-sectional data from the BSWS to calculate the proportion of minimum wage workers. Observations are limited to those plants/workers in manufacturing sectors from 2008 to 2014.[18]

To identify the extent of the minimum wage shock to each plant in the Census of Manufacture, I first calculate the proportion of minimum wage workers at each plant i at period t from the BSWS. A worker is defined as a minimum wage worker if his or her hourly wage rate was within 120% of the minimum wage that would be effective in the following October or November.[19] Both regular and non-regular workers are subject to minimum wage in Japan, and I include all employees to calculate the proportion. I then estimate the following linear model to identify the characteristics of manufacturing plants that hire relatively large proportions of minimum wage workers:

$$S_{it} = \delta_t + z_{it}\beta^0 + \varepsilon_{it} \tag{12}$$

Covariates z_{it} it include polynomials of the plant size and annual wage bill per employee, and the ratio of regular workers. I did not include prefecture and industry fixed effects and allowed individual plant traits to predict the proportion. In so doing, I can avoid the result where the predicted proportion mostly reflects prefecture or industry variations in minimum wage, leaving sufficient minimum wage variations even after I limit the sample by the predicted initial proportions of minimum wage workers. Finally, I compute the predicted proportion of low-wage workers at each plant in the Census of Manufacture, using a common set of covariates z_{it} it and the estimated parameters. Table 4:10 in the Appendix provides the estimation results of this model. Adjusted R-squared values range above 0.6. The results there suggests that the average annual wage bill per employee and its polynomials alone

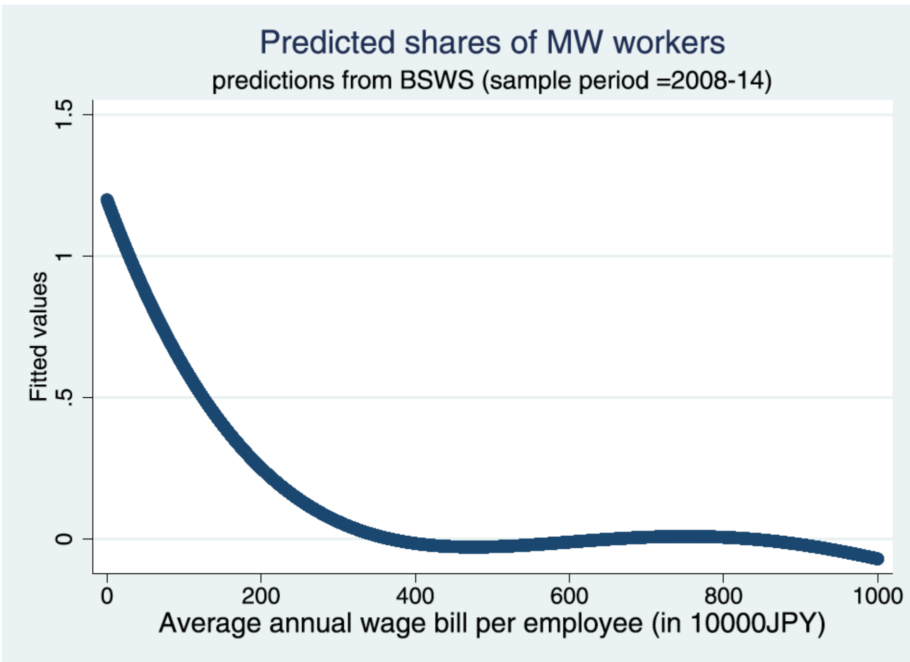
[17] Although it is possible to match worker-level information from Basic Survey on Wage Structure (BSWS) with the Manufacturing Census to construct an employer-employee data set, I chose not to do so for the following reasons. First, BSWS oversamples large plants that are less likely to be affected by an increase in minimum wage. Second, BSWS is not a population survey, and matches only 9% of the original sample in the Manufacturing Census (Kawaguchi et al., 2007).

[18] Although the main analysis analyzes plants with 30 or more employees in the Census of Manufactures, the analysis in this section also includes the smaller plants in BSWS to increase efficiency in the estimation.

[19] I chose 120% so as to accommodate the fact that the industry minimum wage could take a value more than 10% higher than the regional minimum wage.

have sufficiently high explanatory power. This is consistent with an approach in Draca et al. (2011), where they defined a treatment status by establishing the average wage to measure the impact of introducing a national minimum wage in the UK. Figure 4:3 plots fitted values from column (1) of Table 4:10 in the Appendix using the BSWS. Predicted proportions of minimum wage workers become increasingly larger when the annual wage bill per employee at the plant is less than 4 million JPY (≈36 thousand USD as of January 2018). Panel B in Table 4:1 presents the predicted proportions of minimum wage workers. The computed proportions take reasonable values, and these values are similar to those observed in the BSWS (see Panel C in Table 4:1). In the main analysis with the Census of Manufacture, I use the computed proportions of minimum wage workers in 2008 to examine whether the impact of the minimum wage is intensified in plants with initially high proportions of minimum wage workers.

Figure 4:3 Predicted proportion of minimum wage workers at each plant



Note: The figures show the fitted values from a simple regression model to predict a proportion of minimum wage workers at each plant. The covariates include; polynomials of plant size and annual wage bill per employee, and ratio of regular workers (column (1) in Table 4:10 in the Appendix).

Source: The data comes from administrative wage records, Basic Survey on Wage Structures (BSWS).

4.3.4 Testing the impact of minimum wage

Similar to some previous studies, I use differential increases in the minimum wage across regions to test its effect; however, studies have raised potential identification issues in using such regional variations over time. First, the regional minimum wage could be confounded by any pre-existing trend specific to the local area. A series of arguments clarified both the importance and difficulty of finding valid counterfactuals to control for the local pre-existing trends (Allegretto et al., 2018; Dube et al., 2010; Neumark et al., 2014). Second, despite the first point, including region-specific time trends in such a difference-in-difference (DID) type of specification can be misleading if minimum wage affects the growth, not the level, of employment (Meer and West, 2016). When the minimum wage is increased, the adjustment to this new state may take some time and may not be smooth. For instance, job destruction may occur gradually because of adjustment costs and a slow substitution for other input factors such as capital (Baker et al., 1999). Estimating the employment effect in a level specification will then be misleading, because the staggered treatment effect follows a continuous trend-shaped pattern rather than a discontinuous jump. Controlling for region-specific linear trends in such a level specification masks the true dynamic treatment effect because it cannot be identified separately from the local linear trend (Meer and West, 2016).[20]

One way to avoid such a misspecification is to check whether I ever need to control for region-specific linear trends in the first place. The local linear trends have been controlled in previous studies, since pre-existing employment trends may predict the current minimum wage (Dube et al., 2010). I follow a suggestion in Meer and West (2016) and estimate the effect of minimum wage on employment growth and examine whether there are any pre-existing local trends that could confound the changes in minimum wage. In particular, the analysis mainly focuses on the post-amendment period (e.g., 2008 to 2014) in the Census of Manufacture.[21] Then, I estimate the following model of first-differenced employment levels with leads and lags of log differences in the prefectural minimum wage:

[20] Figures 1 and 3 in Meer and West (2016) provide a graphical representation of this idea by comparing two hypothetical jurisdictions which experienced slowdowns in employment growth because of an increase in the minimum wage at different points in time. Wolfers (2006) first points out this weakness of analyzing the dynamic effect of policy shock in the level specification, for the case of
unilateral divorce laws adoption in the US. When people respond to the adoption of the new state law gradually, the separation of the dynamic treatment effect from region-specific trends is not straightforward in a standard DID framework

[21] I limit the observations to those in and after 2008, not 2007, to avoid including information from 2006, given that lagged minimum wage has high explanatory power in the preferred specification, as will be shown later.

$$\Delta \ln L_{it} = \sum_{s=-2}^{3} \gamma^s \Delta \ln mw_{p,t-s} + \delta_t I_j + f_p + \Delta x_{pt} \beta + \Delta \nu_{it} \quad (13)$$

where $\ln L_{it}$ is a logarithm of employment at plant i in year t . The employment here includes both regular and non-regular workers.[22] The model controls for industry-specific (j) year effects, prefecture-level (p) fixed effects, and some time-variant prefecture covariates. Prefecture control variables (x_{pt}) include log-population and the proportion of people aged 15 to 65. If the minimum wage change does not reflect any pre-existing trends, the estimates for the lead terms of $\Delta \ln mw_{p,t-s}$ should be insignificant and close to zero. To reflect the extent that each plant is differently bound by minimum wage, I estimate the above equation separately for plants with different exposures to the minimum wage shock ($\hat{S}_{i,t=2008} > 0.1$, for instance). Similarly, I also examine the heterogeneity of the plants' response to the shock across different market regimes. In particular, I examine whether an increase in the minimum wage brings about the same consequences on employment growth as in the competitive labor market, even when $\hat{\eta}$ is low.

While Meer and West (2016)'s method has the advantage of providing a direct and intuitive test of pre-existing trends, any other underlying factors could still confound the effect of minimum wage. Although the post-amendment variations in minimum wage were largely driven by the policy changes after 2007, the local authority could still deviate from the target amounts slightly. The test of exogeneity in Meer and West (2016) does not immediately ensure the consistency of the estimator of my interest.

For this reason, I complement the above analyses with instrumental variable estimations. In particular, I use two instrumental variables to exploit the full potential of a series of policy changes after 2007 and identify the model. These instruments are taken from the target amounts proposed by the Central Council. Specifically, the first instrument is the baseline target amount in year t for rank r ($base_{rt}$) that the Central Council proposes to all prefectures every year.[23] Because the identification variation stems from the rank-level rather than from prefecture-level information, the IV estimates are consistent, free from the endogeneity from the minimum wage revisions by local authorities. The second instrument is the benefit target amount (benefit pt), proposed only to prefectures where the

[22] I do not divide regular and non-regular employment, because I do not have separate total wage bills for each of the two groups; therefore, I cannot estimate η separately. The primary focus of this study is to examine the labor market heterogeneity in terms of overall workers.

[23] Kambayashi et al. (2013) used a similar instrument variable to identify the minimum wage effect on wage distributions, for the period 1994 to 2003.

initial benefit level was higher than the minimum wage.[24] For prefectures that did not have higher benefit levels, and thus, were not proposed with any benefit target, I set $benefit_{pt} = 0$. As Figure 4:1 indicates, the exposed prefectures with $benefit_{pt} > 0$be experienced a sharper increase in the proportion of minimum wage workers.[25] In the IV estimations, I use $base_{r,t-1}$ and $benefit_{p,t-1}$ as instruments for $\Delta \ln mw_{p,t-1}$. The first-stage F statistics are sufficiently high in all specifications.

4.4 Results

4.4.1 Test of preexisting trends

Table 4:4 shows the results of the tests of whether there are any preexisting trends in the changes in minimum wage after 2008. Specifically, Table 4:4 presents the estimation results for Eq. (13) with various combinations of leads and lags for $\Delta \ln mw_{p,t}$. These specifications are similar to the first three columns of Table 4 in Meer and West (2016). The results in Table 4:4 suggest that the elasticity of employment with respect to minimum wage is about -0.5. The impact of the first lagged change in minimum wage remains quite stable across specifications, implying that the negative impact found on the first lagged term is not driven by any preexisting employment trend. In fact, the estimates of lead terms for minimum wage in columns (3) and (4) take insignificant and small values. Thus, the current employment growth is not statistically associated with the future growth of minimum wage. The changes in minimum wage are unlikely to reflect preexisting local trends during the sample period.

[24] Although this benefit target amount was proposed as a gap between the current minimum wage and the current benefit level of the prefecture and also the Council did not request to close the gap all at once in that year, the higher gap could place more intense pressure on prefectures to increase their minimum wage. Therefore, I use the gap or the benefit target amount as it is, as the second instrument.

[25] A part of the identification is also exploited in Hara (2017), where she estimated the minimum wage impact on worker training in Japan.

Table 4:4 Test of pre-existing local trends

	(1)	(2)	(3)	(4)	(5)
$\Delta\ \ln(mw_{p,t-3})$		0.0400	0.0279	0.0487	
		(0.129)	(0.154)	(0.138)	
$\Delta\ \ln(mw_{p,t-2})$		-0.125	-0.131	-0.111	
		(0.172)	(0.183)	(0.167)	
$\Delta\ \ln(mw_{p,t-1})$	-0.538***	-0.499***	-0.500***	-0.492**	-0.518***
	(0.132)	(0.185)	(0.184)	(0.191)	(0.135)
$\Delta\ \ln(mw_{p,t})$	0.107	0.0987	0.103	0.106	
	(0.123)	(0.166)	(0.163)	(0.160)	
$\Delta\ \ln(mw_{p,t+1})$			-0.0276	-0.0284	
			(0.201)	(0.202)	
$\Delta\ \ln(mw_{p,t+2})$				0.0495	
				(0.132)	
N	281,388	281,388	281,388	281,388	281,388

Note: Robust standard errors clustered at the prefecture level in parentheses. $***$ $p < 0.01$, $**$ $p < 0.05$, $*$ $p < 0.1$. Each column controls for industry-specific linear trends and prefecture control variables. Prefecture control variables (x_{pt}) include log-population and the share of those aged 15–65.

Source: The main data comes from Census of Manufacture (METI).

Although regional minimum wages have been raised substantially in the past decades, the results in Table 4:4 blur the fact that only a subset of plants is exposed to the minimum wage shock. To confirm that increases in minimum wage are indeed concentrated in plants with a higher proportion of low-wage workers, I estimate the same models in Table 4:4, separately by the initial extent of the exposure. In particular, I divide the sample by the predicted proportions of minimum wage workers as of 2008, $\hat{S}_{i,t=2008}$, computed from administrative wage records. Table 4:5 shows the results. Columns (1) and (5) replicate the results from columns (4) and (5) in Table 4:4. Comparisons across columns confirm that increases in minimum wage indeed have stronger negative impacts on those plants with higher proportions of minimum wage workers. The employment elasticity is -0.63 in plants with more than 5% of low-wage workers, -0.64 in plants with more than 10% of low-wage workers, and -1.04 in plants with more than 20% of low-wage workers. The monotonically increasing pattern is consistent with Figure 4:3 where the predicted proportions of minimum wage workers become increasingly higher when plants average annual wage bills are less than 4 million JPY. Similar to the findings in Table 4:4, these results are robust against controlling for leads and lags of the minimum wage changes. A comparison between the first and last four columns indicates that the estimates for elasticity are mostly stable with or without the lead and lag terms for minimum wage.

Although the prediction of proportions of minimum wage workers prevents us from obtaining precise estimates for the impact of minimum wage, the results in this table suggest the plausibility of the predicted proportions in reflecting the extent of exposure to the shock. Since the first lagged term, $\Delta \ln mw_{p,t-1}$, has the largest impact in terms of magnitude, I will focus on the impact of this term in the remainder of this chapter.

Table 4:5 Minimum wage effects by predicted shares of minimum wage workers

	(1) all	(2) $\hat{S} > 0.05$	(3) $\hat{S} > 0.1$	(4) $\hat{S} > 0.2$	(5) all	(6) $\hat{S} > 0.05$	(7) $\hat{S} > 0.1$	(8) $\hat{S} > 0.2$
$\Delta\ln(mw_{p,t-3})$	0.0487	-0.143	-0.302	-0.339				
	(0.138)	(0.155)	(0.203)	(0.263)				
$\Delta\ln(mw_{p,t-2})$	-0.111	-0.0251	0.0105	0.157				
	(0.167)	(0.182)	(0.176)	(0.246)				
$\Delta\ln(mw_{p,t-1})$	-0.492**	-0.631**	-0.641**	-1.041***	-0.518***	-0.596***	-0.633***	-0.891***
	(0.191)	(0.241)	(0.272)	(0.341)	(0.135)	(0.178)	(0.225)	(0.278)
$\Delta\ln(mw_{pt})$	0.106	0.129	0.0883	0.228				
	(0.160)	(0.201)	(0.231)	(0.310)				
$\Delta\ln(mw_{p,t+1})$	-0.0284	-0.243	-0.221	-0.276				
	(0.202)	(0.257)	(0.305)	(0.328)				
$\Delta\ln(mw_{p,t+2})$	0.0495	-0.0191	0.0414	-0.188				
	(0.132)	(0.195)	(0.267)	(0.314)				
N	281,388	151,830	111,398	67,785	281,388	151,830	111,398	67,785

Note: Robust standard errors clustered at the prefecture level in parentheses. $***$ p < 0.01, $**$ p < 0.05, $*$ p < 0.1. Each column controls for industry-specific linear trends and prefecture control variables. Prefecture control variables (x_{pt}) include log-population and the share aged 15–65. $\hat{S}$ stands for $\hat{S}_{i,t=2008} > 0.1$, or a predicted share of minimum wage workers at the plant in 2008.

Source: The main data comes from Census of Manufactures (METI).

The estimated elasticities in Tables 4:4 and 4:5 are notably large, compared with those reported in previous state-level studies in the U.S. The results so far suggest that a 10% increase in minimum wage leads to about a 6% decrease in plant employment if the share of minimum wage workers, $\hat{S}_{i,t=2008}$, is higher than 10%.[26] While I admit that the estimates are imprecise in that I lack a perfect measurement of minimum wage exposure at each plant, I consider that plants in the sample

[26] Unfortunately, the predicted proportion of minimum wage workers is not a perfect measurement of the exposed shock. I also condition the estimates on the predicted share as of 2008 to allow sufficient within-plant variations in the estimations. Thus, it is possible that the plant's actual proportion of minimum wage workers is much higher than the predicted proportions indicate.

can have relatively large elasticity estimates for the following reasons. First, recessions can exacerbate the employment effect of minimum wage. The sample period includes a period of global financial crisis, followed by a period of economic recovery, although the unemployment rate was still high at the end of the sample. Negative demand shock decreases plants' value of marginal product of labor and exerts even more severe pressure to reduce employment levels in the face of increases in minimum wage.[27] Thus, plants in the sample could have experienced an intensified effect after the amendment to the Minimum Wage Act in 2007. Second, it is also important to note that, since price levels have been relatively stagnant in Japan, real minimum wage has increased, thereby restricting firms labor-demand decisions quite severely. Lastly and most importantly, the overall level of labor market concentration may be substantially different across countries. Recent empirical studies reveal that the non-trivial proportion of the labor markets in the U.S. is highly concentrated (Azar et al., 2018; Benmelech et al., 2018). As I will show in the next section, the employment effect of minimum wage vanishes when I limit the sample to those plants with relatively high market power. Thus, the relatively small or zero employment effects of minimum wage reported in the U.S. may be attributable to the fact that a significant proportion of plants are facing a labor market surplus. Unless I take into account the different extent of labor market concentrations across countries, the international comparison in the estimated elasticity may not convey a practical implication.[28]

4.4.2 Minimum wage effects across heterogeneous markets

Production function estimation in Section 4.3.2 revealed that the estimated extent of surplus or wage markdown, $\hat{\eta}$, measures important frictions in the labor market such as the one driven by geographical proximity with rival plants. This section tests a prediction of a monopsonistic labor market model, where plants do not reduce their employment level in response to increases in minimum wage.

Panel A of Table 4:6 estimates the impact of growth in minimum wage on employment growth, separately by the level of the estimated surplus. Recall from Section 4.2.1 that $\hat{\eta}$ represents the extent of wage markdown. Plants face no surplus in a competitive labor market, thus, $\hat{\eta} = 1$. Column (1)

[27] Clemens and Wither (2019) lays out a framework consistent with this idea. They found that the minimum wage increases during the financial crisis exacerbated the bite on low-skilled group's wage distributions in the U.S., implying the differential negative employment effect of minimum wage over business cycles (Clemens and Wither, 2019)

[28] Plants in manufacturing sector often face severe product market competitions with foreign manufactures, which also explain why the negative employment effect is large in manufacturing sector because plants cannot pass on the increased cost to consumers easily. Indeed, Cengiz et al. (2019) found the evidence consistent with this idea in the U.S.

replicates column (5) in Table 4:4. I do not add lead and lagged terms for changes in minimum wage, because I have obtained quite robust results in controlling for the leads and lags in previous tables, and also because I prefer to maintain the powers of the test by keeping as many observations as possible.[29]

Table 4:6 Minimum wage effects across heterogeneous labor markets (baseline estimation).

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
Panel A. all plants						
$\Delta\ln(mw_{p,t-1})$	-0.518***	0.667	-0.126	-0.256**	-0.372***	-0.414***
	(0.135)	(0.723)	(0.409)	(0.118)	(0.130)	(0.127)
N	281,388	6,966	34,491	120,173	173,345	199,693
% of MW variation	100 (base)	87.6	92.9	94.0	98.5	99.0
Panel B. plants with $\hat{S}_{i,t=2008} > 0.1$						
$\Delta\ln(mw_{p,t-1})$	-0.633***	0.0115	-0.0823	-0.497**	-0.571**	-0.556**
	(0.225)	(0.937)	(0.566)	(0.243)	(0.231)	(0.248)
N	111,398	5,504	17,423	44,936	61,272	70,272
% of MW variation	100 (base)	88.7	95.3	94.4	98.2	98.5
Panel C. Plants with $\hat{S}_{i,t=2008} \leq 0$ (placebo test)						
$\Delta\ln(mw_{p,t-1})$	-0.529	2.173	-0.0209	0.0576	-0.0731	-0.128
	(0.321)	(2.240)	(1.046)	(0.417)	(0.417)	(0.390)
N	25,526	49	2,662	12,241	17,473	19,896
% of MW variation	100 (base)	147.4	87.3	94.8	98.5	99.0

Note: Robust standard errors clustered at the prefecture level in parentheses. ∗∗∗ $p < 0.01$, ∗∗ $p < 0.05$, ∗ $p < 0.1$. Each column controls for industry-specific linear trends and prefecture control variables. Prefecture control variables (x_{pt}) include log-population and the share aged 15–65. % of MW variation is calculated by dividing standard deviation in $\Delta \ln mw_{p,t-1}$ in that sample by the same standard deviation in column (1) or all observations.

Source: The main data comes from Census of Manufactures (METI).

The results in this panel are consistent with the presence of surplus. As I limit the sample to those plants with smaller $\hat{\eta}$, the estimates become insignificant and smaller in magnitude, and even take positive values; increases in minimum wage do not significantly reduce employment growth when plants face a large extent of surplus or wage markdowns. This is in contrast to the significant and negative impact of minimum wage in the baseline case in column (1). Importantly, a plant deviates from the competitive model if it has some control over market wages due to some frictions in the

[29] Although not shown in the chapter, I also estimated the same sets of specifications with leads and lags of $\Delta \ln mw_{p,t}$. None of the lead terms were significant, and I thus did not observe any preexisting employment trend.

local labor market (Manning, 2003). The initial surplus between the value of the marginal product of labor and the wage rate imposes less pressure on the plant to reduce its employment level. A plant can even increase its profit by increasing its employment level to expand the surplus. This is consistent with the results here since I observe that the magnitude of the estimate monotonically increases as I restrict the sample to smaller $\hat{\eta}$, although they are not statistically significant.

On the other hand, plants can also deviate from the standard case if they face adjustment costs of labor. In their dynamic model, Bentolila and Bertola (1990) formalized the idea that, even in a competitive setting, the value of the marginal product of labor deviates from the equilibrium wage rate when firms face firing or hiring costs. In fact, Petrin and Sivadasan (2013) measures the extent of firing costs for manufacturers in Chile by estimating the wedge between the value of the marginal product of labor and the wage rate from production function estimations.

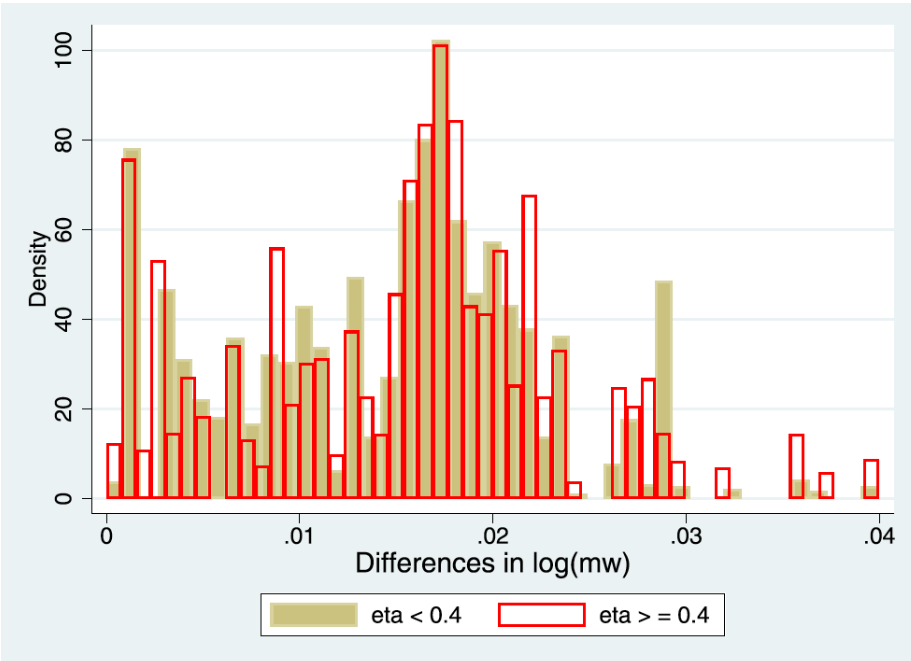
However, the results here are likely to reflect market frictions, rather than the adjustment costs of labor, for following reasons. First, when I construct $\hat{\eta}$, I use the total wage bill which includes severance payments. Thus, the surplus measured by $\hat{\eta}$ excludes an important part of the adjustment costs. Second, although the adjustment costs may still arise from expected cost of litigation and other non-pecuniary costs, the regional pattern of $\hat{\eta}$ does not match with the potential differences in such unobserved firing cost at each region.[30] Given that the extent of surplus is negatively associated with the number of rival plants in the local labor market (Figure 4:2 and Table 4:3), the results suggest that the heterogeneous estimates have mostly arisen from heterogeneity in labor market frictions.

A key to valid identification in Table 4:6 is to have sufficient identification variations in minimum wage changes by the extent of surplus. If variations in minimum wage are significantly smaller in regions or industries with smaller $\hat{\eta}$, the insignificant estimates obtained in Table 4:6 may reflect small identification variations, rather than large frictions in the labor market. I consider that this is not the case in the estimates for the following two reasons. First, despite the fact that the standard error becomes larger as I limit the sample from columns (6) to (2), the estimates also get smaller in magnitude or even positive, suggesting that the insignificant results are not driven merely by small sample size. Second, despite a relatively small sample size, I do have sufficient actual variations in

[30] Firing costs can vary across regions in Japan due to differences in local court discretion (Okudaira, 2018). However, the observed regional difference in firing costs look different from the estimated surplus by prefecture Table 4:9 in the Appendix. In particular, the Osaka District and High Courts are known to have a more stringent interpretation of the firing regulations than the courts in Tokyo (Okudaira, 2018), a pattern that does not coincide with the regional pattern of $\hat{\eta}$ Table 4:9 in the Appendix.

minimum wages even for cases with smaller $\hat{\eta}$. Figure 4:4, which shows histograms for the differences in the logarithm of regional minimum wages, confirms this point. Beige bars indicate the distribution when the estimated wage markdown is smaller: $\hat{\eta} < 0.4$. Similarly, red-lined bars indicate the same distribution for $\hat{\eta} \geq 0.4$. While plants with $\hat{\eta} \geq 0.4$ have slightly larger changes in minimum wage, the two histograms mostly overlap. At the bottom of each panel in Table 4:6, I also show the % of minimum wage variation in terms of the baseline case in column (1). The % of minimum wage variation is calculated by dividing the standard deviation in $\Delta \ln mw_{p,t-1}$ in that sample by the same standard deviation in column (1). Although the standard deviation becomes slightly smaller as $\hat{\eta}$ decreases, the minimum wage variations have not been significantly reduced. Thus, the insignificant estimates observed in Panel A of Table 4:6 are unlikely to only be driven by small identification variations. Rather, they suggest the fact that plants with small $\hat{\eta}$ did not have to immediately reduce the growth rate of employment due to the surplus they face.

Figure 4:4 Are there sufficient variations within group?



Note: The histograms represent variations in $\Delta \ln mw$ by $\hat{\eta}$.

Panel B of Table 4:6 conducts the same estimations by limiting the observations with at least 10% of minimum wage workers, $\hat{S}_{i,t=2008} > 0.1$. I observe more intensified effects when plants had an initially larger proportion of minimum wage workers. Again, the negative impact is observed only when plants have little surplus or larger $\hat{\eta}$. Plants facing some frictions in the labor market do not significantly reduce their employment growth. Similar to the results in Panel A, the estimates become substantially smaller in magnitude, suggesting that the insignificant results are not only driven by the smaller sample size. Although not shown here, similar patterns are observed when I limit observations with different values of $\hat{S}_{i,t=2008}$. The previous empirical studies have focused on the aggregate employment effect of minimum wage and ignored the potential heterogeneity in local labor markets faced by plants. The overall impact of the minimum wage often observed in the literature masks the heterogeneous response of plants operating in diverse labor markets.

Finally, Panel C of Table 4:6 conducts a placebo test to examine whether the results in Panels A and B represent the actual impact of minimum wage. In particular, I limit the observations to those plants with an initial computed proportion of minimum wage workers less than or equal to zero: $\hat{S}_{i,t=2008} \leq 0$. Since these plants did not have minimum wage workers in 2008, they were much less likely to be exposed to the minimum wage shock after 2007. Indeed, none of the estimates are significant in Panel C. Importantly, the magnitude of estimates become smaller when compared to the estimates in the same columns in Panel A and B, except in column (2). Thus, Panel C reinforces the causal interpretation of the results in Panels A and B.

4.4.3 IV estimation approach

Although I limit the observations to the post-amendment period and I also tested the existence of pre-existing trend, these approach does not immediately ensure the consistency of the estimates. To exploit the full exogenous variations from the policy changes after 2007, I complement the analyses with the instrumental variable estimations. In particular, I use $base_{r,t-1}$ and $benefit_{p,t-1}$ as instruments for $\Delta \ln mw_{p,t-1}$. Again, I do not include lead and lagged terms for changes in minimum wage in the model. Table 4:12 in the Appendix presents the first-stage results for baseline cases. According to the table, coefficients for both $base_{r,t-1}$ and $benefit_{p,t-1}$ are significantly positive, consistent with the original purpose of proposing the targets. The instruments have relatively high partial R-squared.

Table 4:7 presents results from the second stage estimations. Table 4:7 also presents some first-stage statistics, along with a standard deviation in the first-stage prediction as a proportion of the baseline standard deviation in column (1) (% of first-stage variation). P-values for over-identification test suggest that I fail to reject the null hypotheses in all cases.[31] Consistent with the previous results, I find negative elasticity estimates in cases with relatively large $\hat{\eta}$. The magnitude of the estimates in Panel A is similar but slightly larger than those reported in Panel A of Table 4:6. On the other hand, the magnitude of the estimates in Panel B is substantially larger than those reported in Panel B of Table 4:6, especially in columns (4) to (6). Thus, OLS estimates in Table 4:6 were biased upward. The IV estimates here present consistent estimates, since the instruments ensure exogenous positive shocks to the regional minimum wages in the absence of endogenous effects from local economies. Thus, the upward bias could arise if the local authority observes relatively high employment growth among plants with minimum wage workers, and also increases regional minimum wage by deviating slightly from the target amounts proposed by the Central Council. The baseline estimates in column (1) of Panels A and B suggest that, on average, plants in exposed prefectures, in Figure 4:1 , experienced 3.2 to 4.2% lower employment growth from 2007 to 2014, compared to plants in non-exposed prefectures.[32] Interestingly, the magnitude of the estimates gets smaller and insignificant as I limit the observations to those with the smaller $\hat{\eta}$ in both Panels A and B, although the estimates are admittedly noisy. For instance, in Panel A, the estimated magnitude of the impact in column (4) is reduced by 35% than the baseline magnitude in column (1). Results in Panel C confirm no significant effect in placebo plants. I consider these results still support the view that the minimum wage effect is heterogeneous across local labor markets.[33]

---

[31] In the over-identification test, I use Wooldridge's score which is robust to heteroskedasticity.

[32] On average, minimum wages growth rate over the same period in exposed prefectures are about 6.1% higher than the one in other prefectures.

[33] Table 4:13 in the Appendix shows the reduced form regression results, where I directly estimate the impact of the two instrumental variables on the employment growth. To summarize the main point, benefit target has a negative and significant impact on employment growth; however, the baseline target has no significant effect. Although both terms have significant impacts in the first- stage estimation (Table 4:12 in the Appendix), it is benefit target, rather than baseline target, which brings about the negative impact on employment growth in the second-stage estimation in Table 4:7.

Table 4:7 Minimum wage effects across heterogeneous labor markets (IV Estimation).

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
Panel A. all plants						
$\Delta\ln(mw_{p,t-1})$	-0.517***	-0.0518	-0.438	-0.334*	-0.373**	-0.478***
	(0.124)	(1.075)	(0.560)	(0.189)	(0.160)	(0.145)
N	281,388	6,966	34,491	120,173	173,345	199,693
% of first-stage variation	100 (base)	80.3	92.1	95.3	99.7	100.3
first-stage F statistic	106.2	71.3	113.7	99.2	94.5	94.3
p-value (overidentification)	0.41	0.32	0.16	0.15	0.15	0.41
Panel B. plants with $\hat{S}_{i,t=2008} > 0.1$						
$\Delta\ln(mw_{p,t-1})$	-0.695***	-0.544	-0.387	-0.786**	-0.868***	-0.854***
	(0.237)	(1.262)	(0.546)	(0.308)	(0.294)	(0.288)
N	111,398	5,504	17,423	44,936	61,272	70,272
% of first-stage variation	100 (base)	83.4	93.0	95.4	98.7	99.3
first-stage F statistics	107.8	80.1	91.2	80.3	86.8	91.7
p-value (overidentification)	0.95	0.27	0.57	0.60	0.78	0.96
Panel C. Plants with $\hat{S}_{i,t=2008} \leq 0$ (placebo test)						
$\Delta\ln(mw_{p,t-1})$	-0.684	1.289	-0.719	-0.0333	-0.0414	-0.0917
	(0.444)	(1.700)	(1.172)	(0.629)	(0.538)	(0.536)
N	25,526	49	2,662	12,241	17,473	19,896
% of first-stage variation	100 (base)	125.7	95.5	98.6	99.8	100.5
first-stage F statistics	105.1	105.1	155.8	123.1	100.9	101.4
p-value (overidentification)	0.35	0.80	0.44	0.84	0.66	0.45

Note: Table presents second-stage estimation results, where I use $base_{r,t-1}$ and $benefit_{p,t-1}$ as instruments to $\Delta \ln mw_{p,t-1}$. Robust standard errors clustered at the prefecture level in parentheses. $***$ p < 0.01, $**$ p < 0.05, $*$ p < 0.1. Each column controls for industry-specific linear trends and prefecture control variables. Prefecture control variables (x_{pt}) include log-population and the share aged 15–65. % of first-stage variation is calculated by dividing standard deviation in first-stage prediction by the same standard deviation in column (1) or baseline case.

Source: The main data comes from Census of Manufactures (METI).

One important concern on the results so far is the potential endogeneity in the labor market surplus. Because I construct the labor market parameter, $\hat{\eta}$, from production function estimates based on observations in 2001–2014 and use it in the main estimations based on observations in 2008–2014, $\hat{\eta}$ may reflect any changes in minimum wage and employment after 2008. In order to address this potential endogeneity concern, Appendix 4.6.5 estimates production functions using the information prior to the main sample period, namely, 2001–2007. I also estimate production functions by following Wooldridge (2009) to construct $\hat{\eta}$ to examine the flexibility of the framework over different estimation procedures. The results in Appendix 4.6.5 point to the similar implications to the results in Table 4:7.

4.5 Conclusion

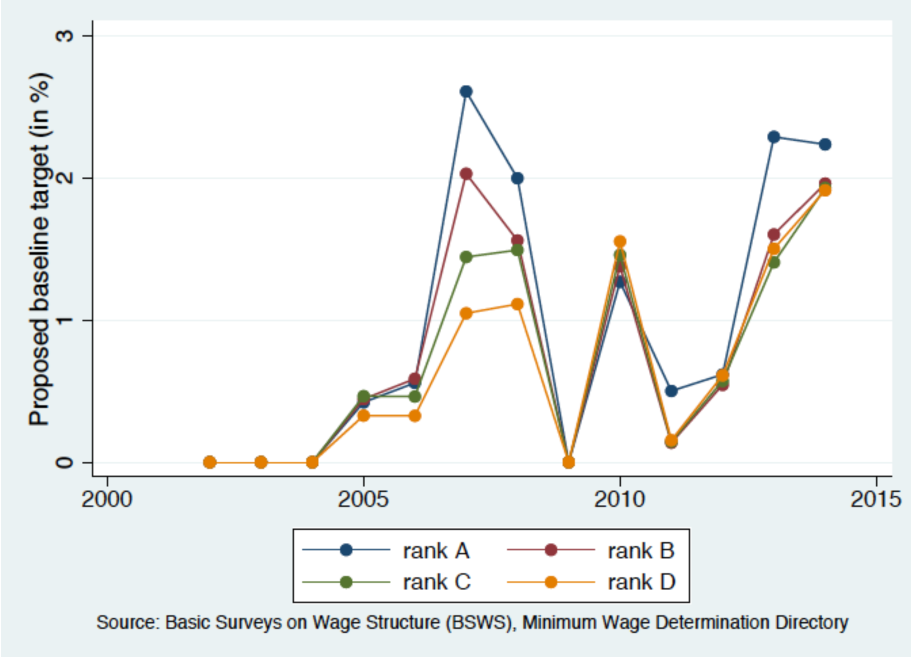
The overall evaluation of minimum wage depends on the extent to which firms bear its burden. This chapter sheds light on direct aspects of firms internal responses to increases in minimum wage, so as to examine whether the local labor markets are heterogeneous, and whether the employment effect of minimum wage differs across markets, depending on employers' market power or frictions in the labor market. Specifically, I estimate the surplus between the value of the marginal product of labor and the wage rate from standard production function estimations. I then tested the minimum wage impact on employment growth across the extent of surplus that plants face in the local labor markets. By applying the estimation framework to a Japanese manufacturing census, I first observed that plants significantly reduced their employment growth in response to increases in minimum wage. However, the estimated negative impact masks the heterogeneity in plants' behavioral response: an increase in minimum wage affected plants in the sample in rather different ways, depending on the surplus that plants face. I found that in response to an increase in minimum wage, plants that initially experienced a large surplus did not significantly reduce their employment growth. Interestingly, albeit insignificant, the estimates become larger when plants have larger surplus. While the main data does not contain the wages and hours of work for individual employees, computation from another source of administrative wage records confirms that the minimum wage effects are concentrated in plants with larger proportions of minimum wage workers. Although a lack of individual hourly wage information prevents us from obtaining precise estimates, the results found in this chapter largely support the view that the local labor market is diverse and plants respond to the minimum wage shock differently, depending on the extent of surplus employers face in the labor market. The results in this study speak to previous studies' observation of mixed employment evidence of minimum wage. Recent studies suggest a high level of labor market concentrations across regions in the U.S. (Azar et al., 2018; Benmelech et al., 2018). Thus, the relatively small or zero effect of minimum wage observed in previous studies may be attributable to the fact that employers faced a surplus in the local labor market and did not reduce their employment in response to the increase in minimum wage. The results also provide an important implication for policymakers. The fact that the local labor markets operate in different market mechanisms requires attention when revising minimum wages, because an increase in minimum wage can slow down the employment growth in specific industries or regions. Policymakers often take local economic conditions into account in revising the minimum wage; however, they pay little attention to the extent of frictions or market power that employers can exercise in local labor markets. Increasing the minimum wage is not a simple redistribution policy. The results in this study suggest that policy makers should account for variant extent of competitions across heterogeneous markets.

4.6 Appendix

4.6.1 Target amounts (instrumental variables)

This Appendix lays out further details on the instrumental variables, baseline target and benefit target. The baseline target is determined by the prefecture's rank. The Central Council classified each prefecture into one of four ranks, based on its economic condition. Figure 4:5 shows the evolution of the baseline target over time for each rank. The graph shows the baseline target for years when hourly baseline target was available. The Central Council began issuing the hourly baseline target in 2002. Prior to 2002, they issued the baseline target at the daily rate. The lines indicate the baseline target proposed each year in terms of the percentage of existing minimum wage, as an average over prefectures in the same rank. Ranks A and B include urban or populated prefectures (e.g., Tokyo, Osaka), whereas C and D include rural or less populated prefectures. The graphs suggest that prefectures received a proposal of raising their minimum wages by 0 to 2.6 %, depending on their rank and the year. Among all four ranks, the baseline target has equally increased after 2007, except for 2009 and 2011, when the financial crisis and the Great Eastern Japan Earthquake hit Japan. Important to the identification strategy, the baseline target was proposed at the rank level, not at the prefecture level. Thus, the minimum wage variations here do not reflect any decisions made at each prefectures' local council.

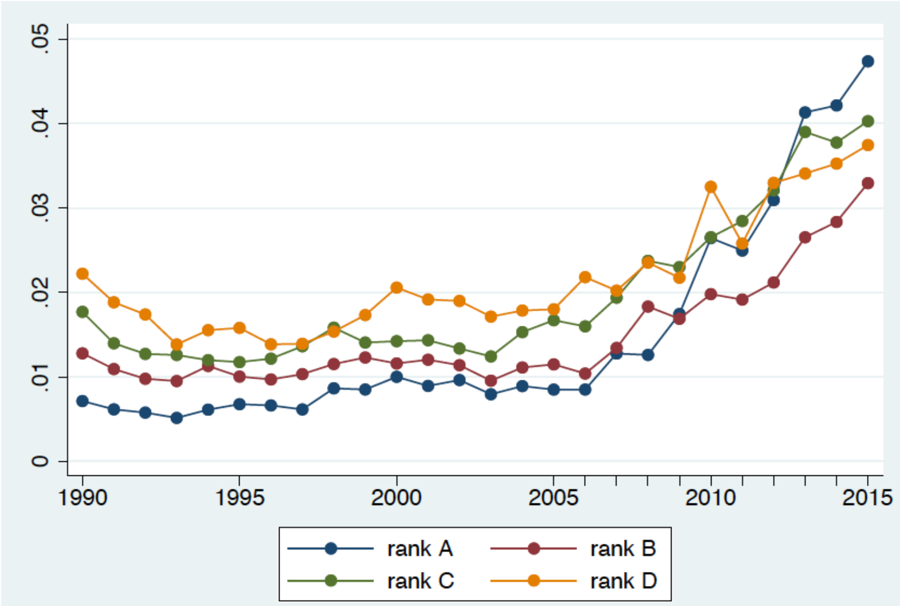
Figure 4:5 Baseline target to raise minimum wage



Note: Lines indicate the average amount of baseline target in terms of the percentage of the existing regional minimum wage. The average is taken over prefectures within the same rank and weighted by prefecture population. The baseline target is the amount that the Central Minimum Wage Council proposes to local authorities at prefectures to determine how much they should raise their regional minimum wage.

Despite that the baseline target shown in Figure 4:5 was not very large on the annual basis, the continuous increases in the target amount have substantially raised the proportion of workers affected by it. Figure 4:6 presents the trend of the proportion of minimum wage workers for each rank. The graph suggests that prefectures faced sharp increases in the proportion of minimum wage workers in the late 2000s, even among prefectures in rank A. An increasing number of plants have been exposed to the minimum wage shock due to the increases in the baseline target.

Figure 4:6 Proportion of minimum wage workers by prefecture rank



Source: Basic Surveys on Wage Structure (BSWS), Minimum Wage Determination Directory

Note: The lines indicate average proportions of minimum wage workers for each rank. The rank consists of 5 to 17 prefectures. The Central Minimum Wage Council proposes the same target amount to prefectures in the same rank.

The benefit target was introduced by the Central Minimum Wage Council in 2007 to only those prefectures that initially had higher benefit levels than minimum wage earnings. Unlike the baseline target, the benefit target was proposed in terms of the amount necessary to close the gap between the minimum wage and the benefit level at each prefecture. The benefit target, in addition to the baseline target was proposed for 12 prefectures, which I define as exposed prefectures. The 12 prefectures are Hokkaido, Aomori, Miyagi, Akita, Saitama, Chiba, Tokyo, Kanagawa, Kyoto, Osaka, Hyogo, and Hiroshima. The Central Council requested the exposed prefectures to close the gap in the following couple of years, albeit not all at once in that year.The Central Council initially proposed that prefectures close the gap within the next two years, in principle; however, they later allowed options to extend the deadline, according to the Minimum Wage Determination Directory (Labour Investigation Bureau) issued every year. The local minimum wage councils determined the extent of closing the gap in that year, based on the proposed amount of the benefit target, and replaced their baseline target with this projected amount in case the latter was higher. Figure 4:7 shows the evolution of the benefit target over time for exposed and non-exposed prefectures. The graph is shown until 2013 because the baseline target in 2014 was sufficient to close the remaining gap, and thus, the Central Council issued no benefit target in 2013. The lines indicate the benefit target proposed each year in terms of the %

of the existing minimum wage. In 2008, the exposed prefectures had at most 11.6% gap. The benefit target becomes smaller as the exposed prefectures follow the proposed target and increase the minimum wage over the years.

Figure 4:7 Benefit target to raise minimum wage

Note: Lines indicate the average benefit target in terms of percentage of the existing minimum wage in that prefecture. The average is taken over the exposed prefectures as defined in text and weighted by prefecture population. The benefit amount is the extra target amount introduced only to those prefectures that initially had higher benefit levels than minimum wage earnings.

Figure 4:1 in the main text presents the proportion of minimum wage workers by the extent of exposure to the amendment shock. The exposed prefectures experienced disproportionately higher increases in the proportion of minimum wage workers after the amendment than the other prefectures. Important to the identification strategy, the benefit target in the current year does not reflect a specific economic trend in the same year. Exposed prefectures are located in both urban (e.g., Tokyo and Osaka) and rural areas (e.g., Hokkaido and Akita). The welfare benefits were initially high in urban regions because of their relatively high living costs. Regions with cold weather have also had relatively high benefits because of their high heating costs. Thus, local economic conditions in the current period are unlikely to be the primary factor that explains this differential shock. Again, the amendment has provided the Central Council with significant initiative in raising the target level, attenuating the concern on local economic conditions, which can potentially confound the identification variation. I consider that the 2007 amendment initiated sizable and exogenous increases in those affected by the minimum wage and exploit this variation to identify the employment effect of the minimum wage.

4.6.2 Estimating production function and labor market surplus

In this Appendix, I explain the method of estimating production functions and labor market surplus in detail. As laid out in Section 4.3.2, I define the translog production function as follows:

$$\begin{aligned} \ln Q_{it} = {} & \beta_K \ln K_{it} + \beta_L \ln L_{it} + \beta_M \ln M_{it} + \beta_{KK} (\ln K_{it})^2 \\ & + \beta_{LL} (\ln L_{it})^2 + \beta_{MM} (\ln M_{it})^2 + \beta_{KL} \ln K_{it} \ln L_{it} \\ & + \beta_{KM} \ln K_{it} \ln M_{it} + \beta_{LM} \ln L_{it} \ln M_{it} + u_{it}. \end{aligned} \tag{14}$$

OLS estimates are not consistent because the inputs are positively correlated with unobserved productivity, $E(X_{it} u_{it}) \neq 0$, where $X_{it} \in \{K_{it}, L_{it}, M_{it}\}$. I therefore adopt the system GMM approach proposed in Blundell and Bond (1998, 2000) for the baseline results. In this method, the unobserved productivity is decomposed into three terms:

$$u_{it} = \delta_i + \omega_{it} + m_{it} \tag{15}$$

where δ_i denotes average productivity of plant i and is captured by plant fixed effects. ω_{it} denotes a productivity shock unobserved by the econometrician. The shock is observed by the managers before determining inputs. This term is therefore the main source of endogeneity. m_{it} denotes a measurement error or a productivity shock after the amounts of inputs are determined. The average productivity can be correlated with the levels of inputs but must be independent from the changes in the inputs, $E(\delta_i \mid \Delta \ln X_{it}) = 0$ for $t \geq 2$. The measurement error can be correlated with the contemporaneous levels of inputs, but must be independent from the inputs in the previous periods, $E(m_{it} \mid \ln X_{it-s}) = 0$ for $s \geq 1$.

The dynamic process of the productivity shock is specified as

$$\omega_{it} = \rho \omega_{it-1} + \xi_{it} \tag{16}$$

where ρ is a parameter of the productivity process and ξ_{it} is an innovation term. ξ_{it} is the deviation from the expected productivity shock; it is therefore independent from all of the inputs in the previous periods, $E(\xi_{it} \mid \ln X_{it-s}) = 0$ for s ≥ 1.

Substituting the process into the production function, the following expression is derived as

$$\begin{aligned}\ln Q_{it} = & \sum_{X}[\beta_X \ln X_{it} - \rho\beta_X \ln X_{it-1} + \beta_{XX}(\ln X_{it})^2 \\ & - \rho\beta_X(\ln X_{it-1})^2] \\ & + \sum_{X,X'}[\beta_{XX'} \ln X_{it} \ln X'_{it} - \rho\beta_{XX'} \ln X_{it-1} \ln X'_{it-1}] \\ & + \rho \ln Q_{it-1} + (1-\rho)\delta_i + \xi_{it} + m_{it} - \rho m_{it-1}\end{aligned} \tag{17}$$

To obtain parameter estimates in this model, I first estimate a vector of parameters γ in the following estimating equation:

$$\begin{aligned}\ln Q_{it} = & \sum_{X}[\gamma_X \ln X_{it} + \gamma'_X \ln X_{it-1} + \gamma_{XX}(\ln X_{it})^2 \\ & + \gamma'_{XX}(\ln X_{it-1})^2] \\ & + \sum_{X,X'}\left[\gamma_{XX'} \ln X_{it} \ln X'_{it} + \gamma'_{XX'} \ln X_{it-1} \ln X'_{it-1}\right] \\ & + \gamma_Q \ln Q_{it-1} + d_i + v_{it}\end{aligned} \tag{18}$$

The moment conditions for the first-differenced equations are written as

$$E\left[\begin{pmatrix} \ln Q_{it-s} \\ \ln X_{it-s} \\ (\ln X_{it-s})^2 \\ \ln X_{it-s} \ln X'_{it-s} \end{pmatrix} \Delta v_{it}\right] = \mathbf{0}, \text{ for s} \geq 3. \tag{19}$$

On the other hand, the moment conditions for the levels equations are written as

$$E\left[\begin{pmatrix} \Delta \ln Q_{it-2} \\ \Delta \ln X_{it-2} \\ \Delta(\ln X_{it-2})^2 \\ \Delta(\ln X_{it-2} \ln X'_{it-2}) \end{pmatrix} (d_i + v_{it})\right] = \mathbf{0}. \tag{20}$$

Using consistent estimates of the unrestricted parameters and the variance-covariance matrix, I impose the restrictions $\gamma_X = -\gamma'_X\gamma_Q$ by minimum distance to obtain the restricted parameter vector. Hempell (2005) describes the procedure in detail.

To measure the labor market surplus, I then calculate the output elasticities for each input. In the calculation, I use the median values for the inputs in each industry-prefecture group:

$$\hat{\varepsilon}_K = \hat{\beta}_K + 2\hat{\beta}_{KK} \ln \overline{K} + \hat{\beta}_{KL} \ln \overline{L} + \hat{\beta}_{KM} \ln \overline{M} \tag{21}$$

$$\hat{\varepsilon}_L = \hat{\beta}_L + 2\hat{\beta}_{LL} \ln \overline{L} + \hat{\beta}_{KL} \ln \overline{K} + \hat{\beta}_{LM} \ln \overline{M} \tag{22}$$

$$\hat{\varepsilon}_M = \hat{\beta}_M + 2\hat{\beta}_{MM} \ln \overline{M} + \hat{\beta}_{KM} \ln \overline{K} + \hat{\beta}_{LM} \ln \overline{L} \tag{23}$$

where $\overline{X}$ shows the median value for input X. The median values are taken across industry-prefecture groups. The cost shares are also aggregated into industry-prefecture groups by taking median values. Finally, the markup and the labor market surplus are measured by taking the ratios of these elasticities and cost shares.

Table 4:8 Estimated η by industry

	mean	p25	p50	p75
livestock products	0.35	0.3	0.36	.4
food and fisheries	0.17	0.14	0.16	.19
fine grain milling	1.03	0.66	1.05	1.35
organic fertilizers feed	0.36	0.21	0.28	.51
beverage	0.69	0.53	0.65	.84
tobacco	0.04	0.03	0.03	.07
textile goods	0.64	0.57	0.64	.74
sawing lumbering wooden products	0.43	0.37	0.43	.48
furniture equipment	0.62	0.58	0.62	.64
coated paper pulp,paper and paperboard	0.77	0.66	0.75	.78
pre press binging	0.67	0.59	0.69	.75
furs leather and leather products	0.39	0.32	0.37	.38
rubber products	1.34	1.22	1.42	1.55
chemical fertilizer	0.89	0.53	0.8	1.16
basic inorganic chemical products	0.46	0.38	0.42	.52
basic organic chemical products	0.01	0	0	.02
organic chemical products	0.74	0.5	0.67	.9
final chemical products	0.64	0.56	0.66	.7
medical and pharmaceutical products	0.69	0.53	0.66	.71
glass and glass products	0.44	0.37	0.4	.47
cement and cement products	1.48	1.46	1.51	1.63
ceramics and porcelain	0.86	0.67	0.8	.97
other ceramic and clay products	1.45	1.32	1.56	1.59
pi giron crude steel	0.55	0.31	0.59	.66
other iron and steel	0.59	0.54	0.59	.63
refining non-ferrous metal smelting	1.45	1.45	1.45	1.45
non-ferrous metal products	1.42	1.4	1.47	1.51
metal products for building and construction	0.97	0.81	0.91	1.01
other metal products	0.52	0.48	0.5	.52
general industrial machinery	0.83	0.75	0.82	.88
special industrial machinery	0.58	0.53	0.58	.61
other general machinery	0.5	0.45	0.5	.54
equipment for office and service	0.76	0.7	0.71	.81
heavy electrical machinery	0.26	0.23	0.25	.28
electronics-applied equipment,electronic measuring instrument	0.65	0.57	0.61	.77
semiconductor element, integrated circuit device	1.12	0.75	1.18	1.36
electronic components	0.94	0.84	0.89	1.07
other electrical equipment	0.9	0.81	0.88	.95
motorcar	1.24	1.24	1.31	1.38
automotive parts automobile accessories	0.44	0.39	0.44	.49
other transportation equipment	1.13	0.86	1.06	1.4
precision machine	0.65	0.6	0.68	.69
plastic products	0.44	0.41	0.43	.46
other manufactured products	1.3	1.26	1.3	1.45
Total	0.64	0.45	0.58	0.76

Note: This table summarises statistics from plant-level observations. Estimates are obtained from translog production function estimations. The translog production functions are estimated separately for each industry group. The estimation procedure follows System GMM. The estimated production function estimates are then used to calculate η by using median values for other input variables within each prefecture-industry group (see Section 4.3.2).

Source: The data comes from Census of Manufactures (METI).

Table 4:9 Estimated η by region

	mean	p25	p50	p75
Hokkaido	.46	.14	.48	.64
Aomori	.51	.15	.38	.86
Iwate	.63	.41	.54	.78
Miyagi	.62	.36	.52	.88
Akita	.7	.52	.6	.89
Yamagata	.63	.52	.57	.75
Fukushima	.66	.45	.56	.83
Ibaragi	.61	.41	.47	.81
Tochigi	.62	.44	.51	.79
Gunma	.55	.4	.49	.72
Saitama	.68	.46	.59	.75
Chiba	.6	.4	.51	.71
Tokyo	.75	.68	.79	.79
Kanagawa	.67	.53	.6	.77
Niigata	.64	.48	.63	.81
Toyama	.81	.51	.61	1.07
Ishikawa	.65	.51	.67	.75
Fukui	.73	.52	.74	.74
Yamanashi	.67	.46	.66	.83
Nagano	.64	.49	.6	.77
Gifu	.64	.46	.54	.67
Shizuoka	.64	.49	.5	.78
Aichi	.61	.39	.52	.65
Mie	.63	.39	.54	.75
Shiga	.62	.43	.5	.83
Kyoto	.7	.51	.67	.82
Osaka	.65	.49	.6	.7
Hyogo	.67	.48	.56	.77
Nara	.59	.4	.49	.71
Wakayama	.58	.44	.54	.68
Tottori	.68	.42	.61	.78
Shimane	.69	.49	.71	.78
Okayama	.6	.45	.56	.66
Hiroshima	.59	.45	.55	.65
Yamaguchi	.58	.32	.61	.71
Tokushima	.68	.46	.59	.89
Kagawa	.61	.33	.54	.76
Ehime	.63	.49	.62	.74
Kochi	.59	.31	.52	.67
Fukuoka	.61	.4	.61	.69
Saga	.61	.33	.49	.69
Nagasaki	.58	.31	.63	.76
Kumamoto	.71	.43	.51	1.22
Ohita	.62	.44	.59	.67
Miyazaki	.62	.46	.54	.66
Kagoshima	.51	.27	.51	.7
Okinawa	.64	.41	.52	.74
Total	.64	.45	.58	.76

Note: This table summarises statistics from plant-level observations. Estimates are obtained from translog production function estimations. The translog production functions are estimated separately for each industry group. The estimation procedure follows System GMM. The estimated production function estimates are then used to calculate η by using median values for other input variables within each prefecture-industry group (see Section 4.3.2).

Source: The data comes from Census of Manufactures (METI).

Table 4:10 Predicting proportions of minimum wage workers with administrative wage records

	(1) $proportion(wage_t < 1.2mw_{t+1})$	(2) $proportion(wage_t < 1.2mw_{t+1})$
$(wagebill/worker)$	-0.00733***	-0.00887***
	(0.000106)	(0.000111)
$(wagebill/worker)^2/10^3$	0.0154***	0.0181***
	(0.000313)	(0.000349)
$(wagebill/worker)^3/10^6$	-0.0134***	-0.0154***
	(0.000384)	(0.000445)
$(wagebill/worker)^4/10^{11}$	0.399***	0.456***
	(0.0166)	(0.0198)
$plantsize/10^3$	-0.0514***	
	(0.00286)	
$plantsize^2/10^6$	0.0240***	
	(0.00186)	
$plantsize^3/10^{11}$	-0.330***	
	(0.0350)	
$plantsize^4/10^{13}$	0.00133***	
	(0.000184)	
Ratio of regular workers	0.258***	
	(0.00401)	
Constant	1.201***	1.557***
	(0.0132)	(0.0122)
N	82,509	82,509
Adjusted R-squared	0.658	0.652

Note: Robust standard errors clustered at the prefecture level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Each column represents the estimates from a plant-level linear regression to predict a proportion of minimum wage workers at each plant. wagebill/worker indicates average annual wage bill per employee. plantsize indicates a number of employees at each plant.

Source: The data come from administrative wage records, Basic Survey on Wage Structures (BSWS). See Section 4.3.3 for details.

4.6.3 Variable construction for production function estimation

1. Gross output

Gross output is measured as the sum of shipments, revenues from repairing and fixing services, and revenues from performing subcontracted work. Gross output is deflated by the output deflator taken from the Japan Industrial Productivity (JIP) Database 2011 and converted to values in constant prices of 2000.

2. Intermediate input

Intermediate input is defined as the sum of raw materials, fuel, electricity and sub-contracting expenses for consigned production used by the plant. Using the corporate goods price index (CGPI) published by Bank of Japan, intermediate input is converted to values in constant prices of 2000.

3. Capital input

Capital input (K_{it}) is measured as real capital stock, defined as follows:

$$K_{it} = BV_{it} \frac{INK_{jt}}{IBV_{jt}} \qquad (24)$$

where BV_{it} is the initial net book value of plant i, INK_{jt} represents the initial net capital stock of the whole industry in constant 2000 price, and IBV_{jt} is the initial net book value of the whole industry. That is, INK_{jt}/IBV_{jt} stands for the ratio of real value in constant 2000 price to book value of capital stock of the whole industry in year t. INK_{jt} is calculated as follows. First, as a benchmark, I took the data on the book value of tangible fixed assets in 1975 from the Financial Statements Statistics of Corporations published by Ministry of Finance. I then converted the book value of year 1975 into the real value in constant 2000 prices using the investment deflator provided in the JIP 2011. Second, the net capital stock of industry j, INK_{jt}, for succeeding years is calculated using the perpetual inventory method.

$$INK_{jt} = (1 - \delta_{jt})INK_{jt-1} + I_{jt} \qquad (25)$$

I_{jt} stands for the real investment in industry j and in year t.We used the investment deflator in the JIP 2011. The sectoral depreciation rate (δ_{jt}) used is taken from the JIP 2011.

4. Labor input

For labor input, I use the total number of workers at each plants.

5. Value added

Value added is defined as the difference between gross output and intermediate input.

4.6.4 Estimation results on employment levels

In this Appendix, I provide the explanation on the regression for employment levels instead of the employment growth. I estimate the following equation:

$$\ln L_{it} = \alpha + \gamma \ln mw_{p,t-1} + \delta_t I_j + g_i + x_{pt}\beta + \nu_{it} \tag{26}$$

where g_i is a plant fixed effect. I also include the same set of vairables as in equation 12. The following table shows the estimation results. In the first column, the level of employment is regressed on the level of minimum wage with the plant fixed effect and year fixed effect. I take the first lag for the levels of minimum wage as in the main text. The coefficient for the minimum wage is negative and statistically significant. In column 2, time-varying control variables are included as explanatory variables. Then, the year fixed effect is replaced with an industry-year fixed effect to capture the different business cycles by industries in column 3. The estimates are significantly negative across specifications, although the magnitude gets larger when time-varying controls are included. While the main focus is on the employment growth, I found consistent estimates even when I regress the employment level on the minimum wage.

Table 4:11 Estimation results on employment levels

	(1)	(2)	(3)
$ln(mw_{p,t-1})$	-0.386***	-0.632***	-0.616***
	(0.0522)	(0.0666)	(0.0670)
N	297,577	297,577	297,577
Time FE	Country	Country	Industry
Time-varying controls	No	Yes	Yes

Note: Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Prefecture control variables include log-population and the share of those aged 15-65.

Source: The data comes from Census of Manufactures (METI).

4.6.5 Robustness against alternative production function specifications

This section conducts further robustness tests against alternative specifications of $\hat{\eta}$, namely, alternative procedures to estimate production functions. One important concern on the results in the main analyses is the potential endogeneity in labor market surplus. For instance, if the local labor market becomes competitive and $\hat{\eta}$ becomes closer to 1, the harsh competitive environment may induce firms to invest more in technologies that can substitute labor inputs away from production, and at the same time improve the efficiency in the production process in the long run. A reduction in production costs can expand the firm's production. If this is the case, it may invite further increases in minimum wage, since economic expansion often accelerates upward revisions of the minimum wage. Because I construct the labor market parameter, $\hat{\eta}$, from production function estimates based on observations in 2001–2014, it is possible that the estimates using the level of $\hat{\eta}$ in Table 4:6 disproportionately selects plants in specific prefecture-industry groups. Furthermore, $\hat{\eta}$ also includes the contemporaneous changes in labor input, L_{it}, since I use industry-prefecture-level median values of input factors and cost shares to calculate the output elasticities for each input.

In order to address this endogeneity concern, Panels A and B in Table 4:15 conduct robustness estimations using the information prior to the main sample period. In particular, panel A constructs $\hat{\eta}$ from the same production function estimates in Table 4:7 (i.e., system GMM estimations separately for each industry, 2001–2014), but with median input values and median cost shares from the pre-sample period (2001–2007) only. In Panel B, $\hat{\eta}$ is constructed from production function estimates from the pre-sample period only (i.e., system GMM estimations separately for each industry, 2001–2007), and median input values and median cost shares are taken from observations in pre-sample period, 2001–2007. Observations in the analyses here are limited to those plants with $\hat{S}_{i,t=2008} > 0.1$. I apply the same IV estimation framework as before.

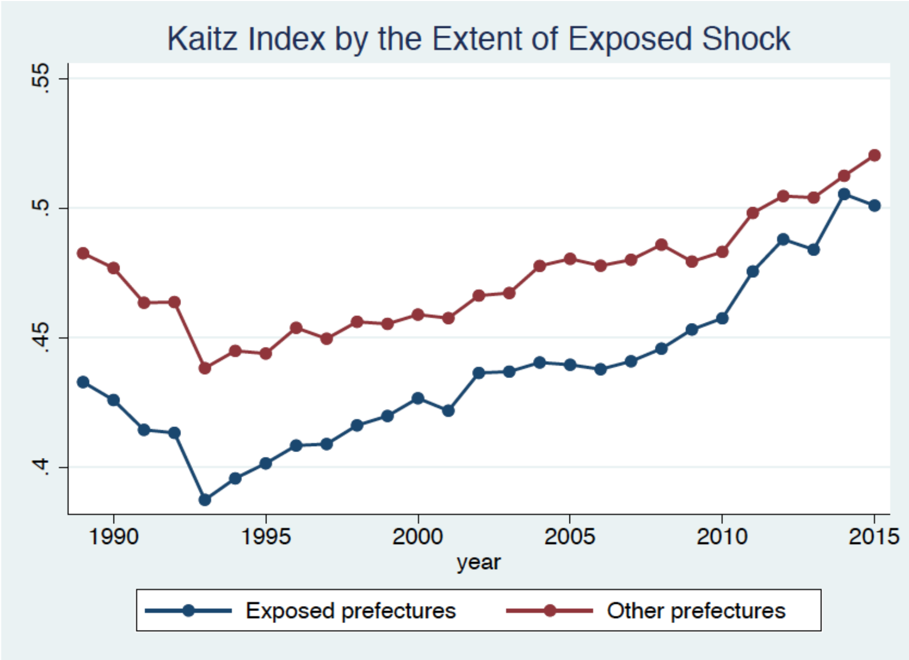
The results in Panels A and B show that the negative impacts of minimum wage are again concentrated among those plants with $\hat{\eta}$ closer to one. In both panels, the elasticities of employment with respect to minimum wage take significant and negative values of about −0.8 and −1 in columns (5) and (6). On the other hand, the estimates for the smaller $\hat{\eta}$ are smaller in magnitude and not significant, except for cases in column (3) where the magnitude of the estimates are slightly larger in both panels.

It should be noted that a non-trivial proportion of plants are dropped from these specifications due to the lack of a parameter estimate for η. For instance, in Panel A, I drop plants in those industryprefecture groups where the output elasticities of input factors take unrealistic values such as negative. In Panel B, those industry-prefecture groups with a relatively fewer plants are dropped from the sample due to the insufficient sample size in production function estimations. Moreover, while these estimations can mitigate the endogeneity concerns on the surplus, the production function estimations ignore technological advancements after the Lehman shock since they only use information from prior to the event. If the market structure has been substantially changed over time, limiting the production function estimations prior to the 2007 amendment can also be misleading. Because of these limitations, I consider the preferred estimates as those presented in Tables 4:6 and 4:7. However, the estimation results shown in Panels A and B in Table 4:15 still point to a similar direction as the previous estimations.

So far, I estimated $\hat{\eta}$ using system GMM suggested by Blundell and Bond (2000). As explained in section 4.3.2, I adopt system GMM since it provides consistent coefficient estimates in the presence of plant fixed effects. However, in order to examine the flexibility of the framework over different estimation procedures, I also estimate production functions by following Wooldridge (2009) to construct $\hat{\eta}$. Wooldridge (2009) extends the two-step estimation framework in Levinsohn and Petrin (2003) to a two-equation system which provides more efficient estimators by allowing the error component to be correlated with current labor input but not the labor input at the previous period. Similar to the previous case, I estimate production functions separately for each industry with observations between 2001 to 2014. The median values are obtained from observations between 2001 to 2007. Panel C in Table 4:15 presents the estimation results. Although a pattern in the magnitude of the estimate is not as clear as previous cases, I again obtained significant and negative estimates for a case with little surplus or $\hat{\eta}$ close to 1. Unfortunately, in many industry-prefecture groups, this procedure also provides us with negative output elasticity values and I dropped these industry-prefecture groups from the sample. Since the final sample size in Table 4:15 is disproportionately reduced, I consider that the preferred estimates are those based on system GMM estimations, although the results with Wooldridge (2009)'s method also point to similar implications.

4.6.6 Other figures and tables

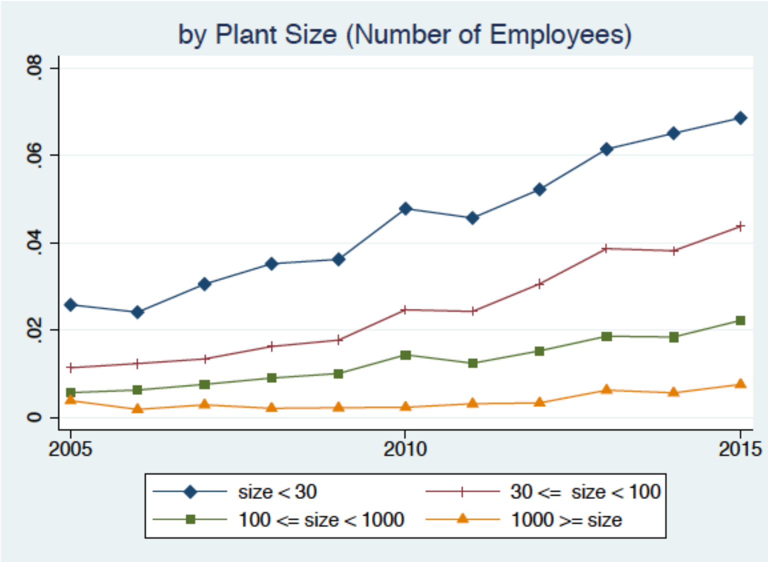
Figure 4:8 Kaitz Index by extent of exposed shock



Note: The exposed prefectures are those that initially had relatively lower benefit levels than minimum wage earnings; therefore, were exposed to intense increases in minimum wage after the revision of Minimum Wage Act, which was approved in 2007. The Kaitz index is constructed by dividing average minimum wage in the prefectures by average market wage in the prefectures.

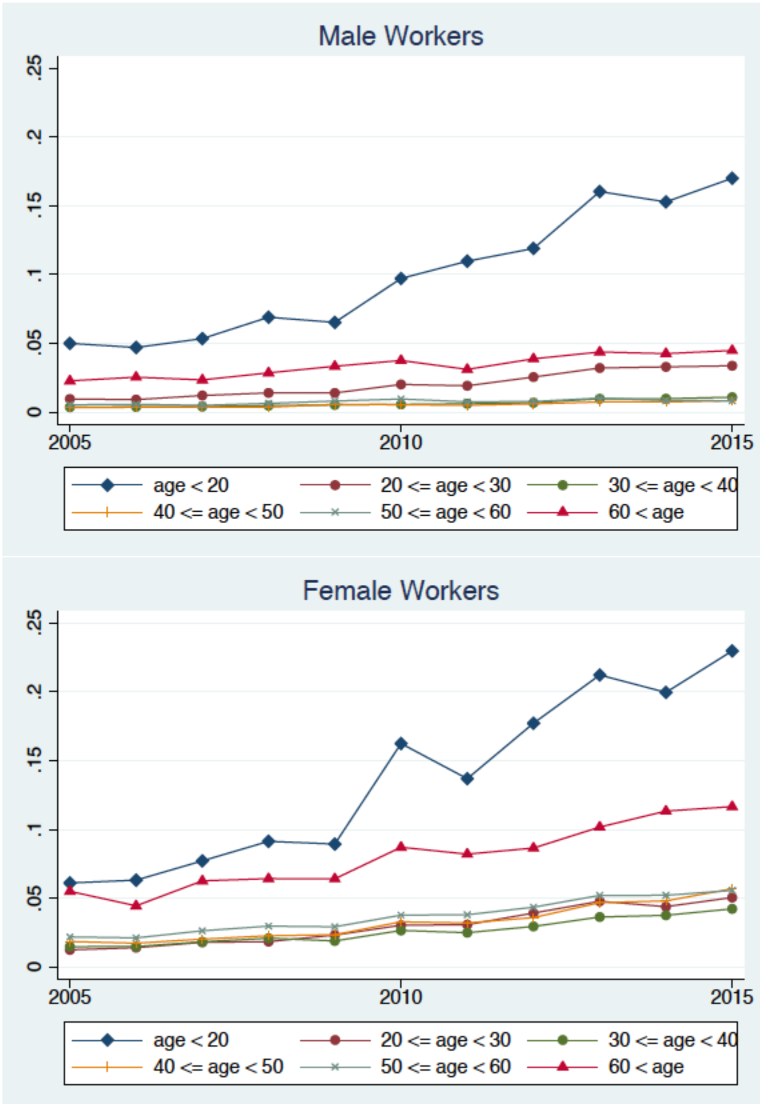
Source: Based on authors' calculation from Basic Survey of Wage Structures (Japanese Ministry of Health, Labour and Welfare).

Figure 4:9 Shares of minimum wage workers by plant size



Source: Based on authors' calculation from Basic Survey of Wage Structures (Japanese Ministry of Health, Labour and Welfare).

Figure 4:10 Shares of minimum wage workers by gender and age group



Source: Based on authors' calculation from Basic Survey of Wage Structures (Japanese Ministry of Health, Labour and Welfare).

Table 4:12 First-stage estimation results

Panel A. all plants

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
$base_{rt-1}/10^3$	0.804***	0.852***	0.936***	0.877***	0.807***	0.804***
	(0.00278)	(0.0205)	(0.00800)	(0.00421)	(0.00358)	(0.00333
$benefit_{pt-1}/10^3$	0.298***	0.214***	0.252***	0.286***	0.300***	0.301***
	(0.000779)	(0.00395)	(0.00263)	(0.00128)	(0.00101)	(0.000945
N	281,388	6,966	34,491	120,173	173,345	199,693
first-stage F statistic	106.2	71.3	113.7	99.2	94.5	94.3
partial R-squared	0.532	0.384	0.500	0.525	0.526	0.527

Panel B. plants with $\hat{S}_{i,t=2008} > 0.1$

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
$base_{rt-1}/10^3$	0.864***	0.886***	0.940***	0.938***	0.875***	0.874***
	(0.00435)	(0.0216)	(0.0109)	(0.00683)	(0.00599)	(0.00561)
$benefit_{pt-1}/10^3$	0.281***	0.205***	0.237***	0.263***	0.278***	0.280***
	(0.00135)	(0.00425)	(0.00341)	(0.00225)	(0.00189)	(0.00180)
N	111,398	5,504	17,423	44,936	61,272	70,272
first-stage F statistic	107.7	80.1	91.2	80.3	86.8	91.7
partial R-squared	0.528	0.395	0.485	0.510	0.514	0.515

Panel C. Plants with $\hat{S}_{i,t=2008} \leq 0$ (placebo test)

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
$base_{rt-1}/10^3$	0.785***	0.578***	1.061***	0.846***	0.795***	0.790***
	(0.00900)	(0.184)	(0.0291)	(0.0127)	(0.0110)	(0.0102)
$benefit_{pt-1}/10^3$	0.304***	0.275***	0.258***	0.303***	0.309***	0.307***
	(0.00231)	(0.0400)	(0.0122)	(0.00350)	(0.00294)	(0.00266)
N	25,526	49	2,662	12,241	17,473	19,896
first-stage F statistic	105.1	105.1	155.8	123.1	100.9	101.4
partial R-squared	0.564	0.747	0.576	0.585	0.567	0.567

Note: This table presents the first-stage estimation results for Table 4:7. Dependent variable is $\Delta \ln mw_{p,t-1}$. Robust standard errors clustered at the prefecture level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Each column controls for industry-specific linear trends and prefecture control variables. Prefecture control variables (x_{pt}) include log-population and the share aged 15-65.

Source: The data comes from Census of Manufactures (METI).

Table 4:13 Reduced-form estimates

Panel A. all plants

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
$benefit_{pt-1}/10^3$	-0.174***	-0.169	-0.256*	-0.157**	-0.155**	-0.167***
	(0.0356)	(0.216)	(0.149)	(0.0615)	(0.0647)	(0.0581)
$base_{rt-1}/10^3$	-0.248	1.192	0.327	0.148	0.0592	-0.186
	(0.242)	(1.597)	(0.895)	(0.388)	(0.343)	(0.295)
N	281,388	6,966	34,491	120,173	173,345	199,693

Panel B. plants with $\hat{S}_{i,t=2008} > 0.1$

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
$benefit_{pt-1}/10^3$	-0.192***	-0.331	-0.172	-0.254**	-0.260***	-0.242***
	(0.0540)	(0.282)	(0.165)	(0.0993)	(0.0939)	(0.0849)
$base_{rt-1}/10^3$	-0.622	1.106	0.0499	-0.475	-0.642	-0.728
	(0.401)	(1.623)	(0.816)	(0.524)	(0.450)	(0.465)
N	111,398	5,504	17,423	44,936	61,272	70,272

Panel C. Plants with $\hat{S}_{i,t=2008} \leq 0$ (placebo test)

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
$benefit_{pt-1}/10^3$	-0.146	0.405	-0.470	-0.0284	0.0203	0.0264
	(0.122)	(0.692)	(0.582)	(0.156)	(0.149)	(0.168)
$base_{rt-1}/10^3$	-1.187*	-0.229	0.325	0.150	-0.368	-0.641
	(0.671)	(2.103)	(1.932)	(1.169)	(0.876)	(0.674)
N	25,526	49	2,662	12,241	17,473	19,896

Note: Table presents reduced-form estimation results, where I regress $base_{r,t-1}$ and $benefit_{p,t-1}$ on $\Delta \ln L_{it}$. Robust standard errors clustered at the prefecture level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Each column controls for industry-specific linear trends and prefecture control variables. Prefecture control variables (x_{pt}) include log-population and the share aged 15-65.

Source: The data comes from Census of Manufactures (METI).

Table 4:14 Minimum wage effects by number of rival plants (IV estimation)

Panel A. all plants

	(1) all	(2) $n(rival) \leq q_{0.2}$	(3) $n(rival) \leq q_{0.4}$	(4) $n(rival) \leq q_{0.6}$	(5) $n(rival) \leq q_{0.8}$
$\Delta ln(mw_{p,t-1})$	-0.517***	-3.377**	-0.443	-0.527	-0.453**
	(0.124)	(1.528)	(0.698)	(0.359)	(0.207)
N	281,388	5,659	17,946	43,973	98,279
mean(number of rival plants)	83.7	1.7	4.2	7.7	16.4
% of first-stage variation	100 (base)	87.3	86.3	87.3	92.5
first-stage F statistics	106.2	146.2	177.0	178.4	232.4
p-value (overidentification)	0.42	0.10	0.51	0.89	0.88

Panel B. plants with $\hat{S}_{i,t=2008} > 0.1$

	(1) all	(2) $n(rival) \leq q_{0.2}$	(3) $n(rival) \leq q_{0.4}$	(4) $n(rival) \leq q_{0.6}$	(5) $n(rival) \leq q_{0.8}$
$\Delta ln(mw_{p,t-1})$	-0.695***	-4.109	-1.070	-0.206	-0.315
	(0.237)	(2.870)	(1.000)	(0.537)	(0.461)
N	111,398	1,605	5,819	15,190	35,715
mean(number of rival plants)	87.3	1.6	4.4	7.9	17.0
% of first-stage variation	100 (base)	81.2	85.2	87.3	92.8
first-stage F statistics	107.8	53.2	103.8	156.6	230.3
p-value (overidentification)	0.95	0.43	0.44	0.74	0.30

Panel C. Plants with $\hat{S}_{i,t=2008} \leq 0$ (placebo test)

	(1) all	(2) $n(rival) \leq q_{0.2}$	(3) $n(rival) \leq q_{0.4}$	(4) $n(rival) \leq q_{0.6}$	(5) $n(rival) \leq q_{0.8}$
$\Delta ln(mw_{p,t-1})$	-0.684	-1.204	-3.018**	-1.770**	-0.868
	(0.444)	(1.381)	(1.175)	(0.843)	(0.634)
N	25,526	843	2,285	4,928	10,231
mean(number of rival plants)	81.9	1.5	3.7	7.1	14.5
% of first-stage variation	100 (base)	90.1	87.2	87.8	91.9
first-stage F statistics	105.1	234.5	261.0	140.0	187.0
p-value (overidentification)	0.35	0.62	0.83	0.89	0.44

Note: q_p equals to the number of rival plants within the same prefecture-industry cell at the (p*100) th percentile. The percentiles are constructed over prefecture-industry groups. Table presents second-stage estimation results, where I use basert−1 and benefitpt−1 as instruments to $\Delta \ln mw_{p,t-1}$. Robust standard errors clustered at the prefecture level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Each column controls for industry-specific linear trends and prefecture control variables. Prefecture control variables (x_{pt}) include log-population and the share aged 15-65. % of first-stage variation is calculated by dividing standard deviation in first-stage prediction by the same standard deviation in column (1) or baseline case.

Source: The data comes from Census of Manufactures (METI).

Table 4:15 Robustness against alternative parameter constructions (IV estimation)

Panel A. Parameters constructed with median values in pre-sample period (2001-2007)

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
$\Delta ln(mw_{p,t-1})$	-0.902***	-0.104	-0.570	-0.528	-0.806***	-0.775***
	(0.304)	(1.180)	(0.739)	(0.362)	(0.284)	(0.275)
N	77,563	4,979	14,248	42,623	60,610	69,127
% of first-stage variation	100 (base)	86.8	93.3	96.4	100.2	102.3
first-stage F statistics	98.0	74.3	78.7	69.6	82.9	92.5
p-value (overidentification)	0.71	0.22	0.20	0.86	0.99	0.98

Panel B. Production functions estimated with pre-sample observations only (2001-2007)

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
$\Delta ln(mw_{p,t-1})$	-0.703**	-0.487	-0.793**	-0.863***	-0.893***	-1.054***
	(0.285)	(0.456)	(0.386)	(0.286)	(0.299)	(0.311)
N	93,244	15,028	32,414	42,504	52,474	55,294
% of first-stage variation	100 (base)	97.5	99.3	100.0	100.3	100.2
first-stage F statistics	100.4	83.7	74.5	97.9	94.3	99.8
p-value (overidentification)	0.96	0.88	0.52	0.99	0.95	0.87

Panel C. Production functions estimated by Wooldridge (2009)'s method

	(1) all	(2) $\hat{\eta} < 0.2$	(3) $\hat{\eta} < 0.4$	(4) $\hat{\eta} < 0.6$	(5) $\hat{\eta} < 0.8$	(6) $\hat{\eta} < 1$
$\Delta ln(mw_{p,t-1})$	-0.539**	0.329	-0.966	-0.956*	-0.543	-0.646*
	(0.231)	(1.497)	(0.634)	(0.522)	(0.442)	(0.333)
N	60,641	2,265	8,670	12,293	20,277	33,922
% of first-stage variation	100 (base)	93.9	95.4	96.9	93.0	102.5
first-stage F statistics	91.7	94.7	96.6	102.2	59.6	150.4
p-value (overidentification)	0.39	0.21	0.89	0.68	0.95	0.95

Note: Table presents second-stage estimation results, where I use $base_{r,t-1}$ and $benefit_{p,t-1}$ as instruments to $\Delta \ln mw_{p,t-1}$. Robust standard errors clustered at the prefecture level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In panel A, $\hat{\eta}$ is constructed from the same production function estimates (System GMM, industry-level, 2001-2014) in Table 4:6, but with median input values and median cost shares from pre-sample period (2001-2007) only. See Section 4.3.2 for the exact procedure to construct the market parameter using these median values. In panel B, $\hat{\eta}$ is constructed from production function estimates with pre-sample period only (System GMM, industry-level, 2001-2007). Median input values and median cost shares are taken from those from the pre-sample period, 2001-2007. In panel C, $\hat{\eta}$ is constructed from production function estimates with Wooldridge (2009)'s method (industry-level, 2001-2014). Median input values and median cost shares are taken from observations from 2001-2007. Each column controls for industry-specific linear trends and prefecture control variables. Prefecture control variables (x_{pt}) include log-population and the share aged 15-65. % of first-stage variation is calculated by dividing standard deviation in first-stage prediction by the same standard deviation in column (1) or baseline case.

Source: The data comes from Census of Manufactures (METI).

Chapter 5 Heterogeneous Impacts of Free Trade Agreements: The Case of Japan[1]

This chapter investigates the trade creation effects of Japan's free trade agreements (FTAs) using aggregate trade data for the years 1996–2015. I estimate various specifications of a gravity model. The main finding is that the effects of Japan's FTAs are not clearly observed when the gravity model is specified with three types of fixed effects, i.e. exporter-year fixed effects, importer-year fixed effects, and country-pair fixed effects. In fact, the effects of FTAs vary substantially among trade partners and around half of the FTAs increase Japan's trade values. The results also suggest that FTAs with small trade partners tend to have large effects on Japan as well as other countries. Recently enforced FTAs, however, increase Japan's import values more rapidly.

5.1 Introduction

Free trade agreements (FTAs) are currently the dominant form of commercial policy. The pattern of trade policies in the last two decades has been characterized mainly by the proliferation of FTAs. According to the Regional Trade Agreements Information System (RTA-IS) in the World Trade Organization (WTO), the cumulative number of physical RTAs increased from 24 in 1992 to 286 in 2017.[2]

Japan started to seek trade liberalization through FTAs around 2000 and established the first FTA with Singapore in 2001. Since then, 15 FTAs with 17 countries have been established by August 2018.[3] In addition, Japan recently signed two important multilateral trade agreements, the Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP) and the economic partnership agreement (EPA) with the European Union (EU). While East Asian countries including Japan

---

[1] The revised version of this chapter is in *Asian Econoomic Papers* 18 (2), 1-20.

[2] See the following website: http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx (last accessed 30 August 2018).

[3] See ::1 for the list of partner countries. More detailed information is available in Ando and Urata (2015). While most of Japan's trade deals are EPAs, covering a wide range of issues, I use the term FTA throughout the chapter for brevity.

began forming FTA networks only recently, they are catching up with Western countries and are expected to play a vital role against protectionism.

This chapter investigates the trade creation effects of Japan's FTAs. I focus on Japan for three reasons. First, despite increasing interest among policymakers, the ex post evaluation of FTAs is extremely limited. No papers apply recently developed methods of analysis to Japan's FTAs. The gravity model, which is commonly used in ex post studies, is developing rapidly in the academic literature and is now estimated differently from the way it was ten years ago. Furthermore, some recent studies show that different FTAs have very different effects, and trade creation effects are not clearly observed in some cases. Even under FTAs, Japan has not eliminated trade barriers for agricultural goods, which are heavily protected by most-favored-nation (MFN) tariff rates. In addition, MFN tariff rates on manufacturing goods are already low and there is little scope to reduce them further through FTAs. It is therefore unclear whether Japan's FTAs have increased its own trade values. Trade creation is crucial for demonstrating the positive welfare effects of FTAs; if trade creation does not occur, a reconsideration of Japan's trade and commercial policies would be called for.

The second reason is that Japan provides a suitable case study for exploring the heterogeneous effects of FTAs in the context of the surge in regionalism and the proliferation of bilateral trade agreements since the 1990s. European countries trade mainly among themselves and they established trade agreements with each other soon after WWII. The United States also established trade agreements with its major trading partners, Canada and Mexico, around 1990 by signing the Canada–U.S. Free Trade Agreement (CUSFTA) and the North American Free Trade Agreement (NAFTA). In contrast, Japan, the world's fourth largest exporter and third largest economy, began negotiations of bilateral trade agreements with the countries of the Association of Southeast Asian Nations (ASEAN) after 2000.

Finally, Japan's FTA partners vary substantially in terms of economic development, ranging from Cambodia, Laos, and Myanmar to Australia, Singapore, and Switzerland. The differences in the trade creation effects across different FTAs can be attributed to the characteristics of the partner countries.

The objective of this chapter is to evaluate Japan's FTAs. To this end, I first estimate the effects of Japan's and other countries' FTAs using a state-of-the-art gravity model.[4] I include three types of

[4] In the first stage, I estimate the crude effects of FTAs. Various kinds of related factors such as interactions among other FTAs and the formation of production networks are included.

fixed effects and estimate using Poisson pseudo maximum likelihood (PPML).[5] I estimate the gravity model using various specifications to compare the coefficients for the FTA dummies. In addition, the effects of each of the FTAs are separately identified to take heterogeneity into account. Then the estimated crude effects of individual FTAs are regressed on some variables to explore which FTAs have larger effects. This chapter therefore contributes to the literature by examining the determinants of successful FTAs.

The structure of the chapter is as follows. I briefly review the literature on the impacts of FTAs in the next section. In Section 5.3, I describe the source of the data and provide a descriptive analysis. Section 5.4 discusses the econometric methodology for estimating the gravity model. Results of the estimation are presented in Section 5.5, followed by conclusions in the last section.

5.2 Related literature

FTAs are major instruments for promoting international trade in the 21st century. The ex post effects of FTAs are usually estimated using a gravity model; Baier and Bergstrand (2007) recommend the use of panel data to remove all time-invariant bilateral factors not controlled for in the traditional specification. A similar specification is applied in Magee (2008), although the effects of FTAs weaken when the gravity model is estimated with fixed effects. Cipollina and Salvatici (2010) conduct a meta-analysis and robustly reject the hypothesis that FTAs have no effects. Large effects are also confirmed in Eicher, Henn, and Papageorgiou (2012).

While the effects of FTAs have been established in many papers, Kohl (2014) points out that the trade creation effects are heterogeneous and only about one-quarter of agreements are actually trade promoting.[6] Heterogeneity is also studied in Baier, Yotov, and Zylkin (2019) and Baier, Bergstrand, and Clance (2018). Zylkin (2016) examines the heterogeneous effects of FTAs, using the case of NAFTA. The differences between these papers and ours are threefold. First, I focus on the heterogeneity of directional effects rather than agreement-specific effects or pair-specific effects. This is because Japan's FTAs are bilateral except for the agreements with ASEAN (ASEAN–Japan Comprehensive Economic Partnership, AJCEP), initially enforced in December 2008 between Japan and four countries in ASEAN. Second, I compare the coefficients estimated in various specifications to

[5] The roles of these fixed effects are discussed in Section 5.4.

[6] Among the 166 agreements studied in Kohl (2014), only 44 agreements have a trade-promoting effect.

consider what is important for the evaluation. This analysis provides a good benchmark for future studies because the state-of-the-art specification is computationally burdensome. Finally, the sample period, 1996–2015, includes recent agreements. Because Kohl (2014) finds that FTAs signed after 1990 have smaller effects, extending the sample period is not a trivial point.

Some papers examine FTAs in Japan or East Asia.[7] Ando and Urata (2011) investigate the impact of the Japan–Mexico EPA and find large trade creation effects for some products. Ando and Urata (2015) conduct a similar analysis for three of Japan's FTAs with Malaysia, Thailand, and Indonesia. Yamanouchi (2017) examines the effects of Japan's FTAs by using the Trade Statistics of Japan published by the Ministry of Finance and Japan Customs. While the effects of individual FTAs are estimated in these papers, only Japan's trade data are used. The trade values of Japan's FTA partners with third countries are not considered. Vietnam, for example, is undertaking rapid liberalization, including WTO accession in 2007. It is therefore useful to separately identify the effects of Japan's FTAs with Vietnam from the effects of Vietnam's unilateral trade liberalization. Okabe (2015) explores FTAs formed by ASEAN countries and their trade partners, so AJCEP is studied in the paper. She concludes that the impact of AJCEP is unclear. Furthermore, while the effects of each of Japan's FTAs are estimated, a country-pair dummy is not included in the specification.

In this chapter, I use world trade date to estimate the effects of Japan's FTAs. I estimate various specifications of a gravity model and place emphasis on the importance of estimating the effects of FTAs using the correct specification. In addition, I discuss the characteristics of the partners with which FTAs are working well.

5.3 Data

The trade data used in this chapter are obtained from UN Comtrade. The sample period extends from 1996 to 2015. I first construct a large dataset of 156 countries to interpolate missing trade values. Then the sample to estimate the gravity model is limited to 69 countries.

We include a country in the dataset if its import data are available for more than 11 years during the period 1996–2015. All of Japan's FTA partners are then added regardless of data availability. Many countries have some missing import data. I interpolate the missing import values using the

[7] Some papers study the effects of FTAs in East Asia other than for Japan. Yean and Yi (2014), for example, explore the ASEAN–China FTA. Chia (2013; 2015) discuss the prospect of economic integration in the region.

export data reported by exporters. As import values are reported as cost, insurance, and freight (cif), while export values are reported as free on board (fob), the gap must be estimated.[8] I regress import values on a quartic of export values, a quartic of bilateral distance (population weighted), a quartic of log of importer GDP, a quartic of log of importer population, other gravity variables (a contingency dummy, a common colonizer dummy, and a common language dummy), variables related to trade policies (a FTA dummy, a customs union (CU) dummy, a partial scope agreement (PSA) dummy, a common currency dummy, an importer EU dummy, and an importer WTO dummy), and an exporter-year fixed effect. I obtained these variables, other than some trade policy variables, from the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) website, constructed by Head, Mayer, and Ries (2010) and Head and Mayer (2014).[9] The information on the FTAs, CUs, and PSAs is obtained from the Mario Larch Regional Trade Agreements Database from Egger and Larch (2008).[10]

The actual and predicted values have a high correlation coefficient of 0.92, suggesting that the estimation is valid and the missing import values are well approximated by the corresponding export values.

The sample I use in the main analysis is smaller because of computational difficulty. I then select the countries by ranking trade values averaged over 20 years. The sample includes Japan's FTA partners and the countries with a ranking of export or import values higher than 60. The sample includes 69 countries.[11]

Table 5:1 shows the evolution of Japan's exports to all countries and FTA partners. The trade flows under FTAs are shaded. Japan's exports to all countries were around 410 billion U.S. dollars in 1996. The total export value has increased rapidly over the period 2003–08. While Japan's total exports collapsed during the 2008-09 global financial crisis, they soon recovered before decreasing slightly. From Table 5:1, it is difficult to identify the effects of Japan's FTAs because most of them were enforced during the expansion period, although the export values increased after the

[8] In the dataset, the correlation of trade values reported by importers and by exporters is 0.88, but the mean of the logged trade gap, the difference between log of import values and log of export values, is 0.26 and the median is 0.14.
[9] The dataset is available at the following website: http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=8 (last accessed 2 September 2018).
[10] The dataset is available at the following website: http://www.ewf.uni-bayreuth.de/en/research/RTA-data/index.html (last accessed 2 September 2018)
[11] Mongolia is included in the sample as an FTA partner of Japan; however, the FTA was not enforced in the sample period.

enforcement. In addition, while export values to developing countries such as Myanmar and Vietnam grew most rapidly, I must account for the impacts of the deepening integration of the world economy.

Table 5:1 Evolution of Japan’s export values to FTA partners

FTA partner	Export 1996	Export value in 1996 = 100																	
	(million US dollar)	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Singapore	23,842	97	71	77	97	67	61	64	80	80	83	90	108	78	102	110	99	85	84	77
Mexico	3,837	116	118	132	168	210	243	197	275	340	398	425	424	297	391	429	460	445	457	452
Malaysia	19,225	90	59	70	89	73	74	74	88	86	90	98	101	80	107	111	104	93	87	71
Chile	950	111	104	66	73	57	56	74	104	133	154	208	338	167	357	308	273	260	248	221
Thailand	20,445	78	49	59	74	67	72	89	109	127	126	146	163	122	185	206	242	200	174	152
Indonesia	8,504	97	50	34	63	55	51	49	71	81	64	76	177	115	199	228	267	226	199	155
Brunei	139	169	72	61	68	51	147	89	100	124	154	127	242	192	176	201	192	150	104	171
Laos	49	80	54	72	70	45	62	54	58	78	90	149	263	297	67	89	147	226	185	148
Myanmar	106	90	84	93	273	367	171	213	191	186	221	359	514	549	206	377	835	1605	1544	1447
Vietnam	1,208	113	120	152	190	180	207	246	294	337	389	512	682	618	746	861	960	957	1064	1174
Philippines	7,578	104	84	86	91	92	104	109	106	111	101	95	93	76	96	92	91	73	73	89
Switzerland	2,112	100	100	108	110	95	83	100	116	111	117	138	183	156	167	221	238	190	188	166
Cambodia	81	104	86	109	72	75	78	96	102	123	159	173	140	146	192	305	294	214	324	520
India	2,186	97	112	116	101	81	97	106	135	168	216	266	356	305	378	513	565	479	455	440
Peru	440	109	121	108	110	97	93	83	81	101	128	178	290	210	311	298	341	326	251	244
Australia	7,859	106	106	110	119	106	115	141	163	175	174	202	229	171	223	236	250	228	196	188
World (156 countries)	410,174	102	94	101	117	103	106	119	142	151	163	180	198	149	188	206	206	189	182	164

Note: While trade values are basically reported by importers, missing values are interpolated from corresponding export values.

Source: UN Comtrade.

The evolution of Japan’s import values is like that of its export values. As shown in Table 5:2, total imports almost doubled over the period 2003–08. After the collapse in 2009, import values recovered quickly but then decreased. As for each of the FTA partners, import values have grown more rapidly than export values. For example, imports from Cambodia have increased by 147 times in the last two decades.

Table 5:2 Evolution of Japan's import values to FTA partners

FTA partner	Import 1996 (million US dollar)	Import value in 1996 = 100																		
		1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Singapore	7,323	80	64	74	87	73	68	74	85	91	102	96	107	83	111	118	119	101	107	107
Mexico	1,890	85	65	87	126	106	95	94	114	134	149	167	201	148	184	210	233	223	226	251
Malaysia	11,750	96	73	92	123	109	95	107	120	124	131	148	197	142	193	259	279	253	248	182
Chile	2,763	107	86	91	102	88	77	95	151	185	262	295	286	192	280	355	337	290	294	217
Thailand	10,212	93	80	86	103	101	102	116	138	152	165	179	203	156	205	240	231	215	212	200
Indonesia	15,194	96	71	82	107	97	93	108	123	137	158	174	214	143	186	224	212	190	168	130
Brunei	1,392	100	74	75	118	121	109	131	135	164	167	179	326	239	295	409	430	340	288	168
Laos	23	90	85	58	51	29	28	31	34	34	52	51	77	115	161	415	527	458	491	416
Myanmar	103	96	87	98	116	99	106	135	174	198	239	287	306	331	375	573	653	738	837	840
Vietnam	2,018	108	86	97	130	129	125	153	191	225	262	303	450	344	405	572	747	705	764	750
Philippines	4,522	110	98	117	159	141	144	155	182	170	176	193	186	141	175	197	206	204	225	196
Switzerland	3,563	96	84	94	92	92	92	108	135	141	143	146	180	176	190	220	230	204	202	207
Cambodia	7	200	245	526	795	1005	1141	1362	1519	1606	1831	2117	1843	2172	3165	4691	6158	8882	11761	14755
India	2,843	93	76	79	92	78	73	76	91	112	142	146	184	131	200	240	246	248	245	171
Peru	427	127	67	68	82	99	100	101	159	164	309	523	495	389	510	548	656	619	411	290
Australia	14,229	102	91	90	104	101	98	105	136	172	196	219	334	244	316	398	396	358	338	244
World (156 countries)	327,328	98	81	90	110	102	98	112	133	151	169	182	223	162	203	251	261	245	239	182

Note: While trade values are basically reported by importers, missing values are interpolated from corresponding export values.

Source: UN Comtrade.

Overall, I cannot conclude from this simple analysis that FTAs have had significant effects on Japan's international trade, although trade with some countries increased rapidly after an FTA was enforced. Instead, I turn to explore the contribution of FTAs using a correctly specified gravity model with three types of fixed effects.

5.4 Estimation method

In this chapter, I use the standard gravity framework established by Anderson and van Wincoop (2003), in which the bilateral trade value from country i to country j, $p_{ij}x_{ij}$, takes the form:

$$p_{ij}x_{ij} = \frac{Y_i E_j}{Y}\left(\frac{t_{ij}}{\Pi_i P_j}\right)^{1-\sigma}, \qquad (1)$$

where Y_i, E_j, Y, t_{ij}, P_j, Π_i, and σ denote the total sales of country i, the total expenditure in country j, the sum of the sales all over the world, bilateral iceberg trade costs from country i to country j, an inward multilateral resistance term, an outward multilateral resistance term, and the elasticity of substitution, respectively.

We estimate the following equation, based on Baier and Bergstrand (2007) and Yotov, Piermartini, Monteiro, and Larch (2016), corresponding to equation (1):

$$\begin{aligned} Trade_{ijt} = \exp\big(&\alpha^{W} FTA_{ijt}^{W} + \alpha^{JX} FTA_{ijt}^{JX} + \alpha^{JM} FTA_{ijt}^{JM} + \beta_1 CU_{ijt} \\ &+ \beta_2 PSA_{ijt} + \beta_3 CommonCurrency_{ijt} + \delta_{ij}^{B} + \delta_{it}^{X} + \delta_{jt}^{M}\big) \\ &+ \varepsilon_{ijt}, \end{aligned} \tag{2}$$

where $Trade_{ijt}$ is the aggregate trade value from country i to country j at year t. FTA_{ijt}^{W} is a dummy variable and equal to one if the trading countries are included together in an FTA and both are not Japan. FTA_{ijt}^{JX} and FTA_{ijt}^{JM} are FTA dummies and equal to one if the trading countries are included together in an FTA and the exporter or importer is Japan. The FTA dummies are separated by partners in the estimating equation when I focus on the heterogeneity of FTAs across partners. I interpret positive coefficients for these FTA dummies as evidence of trade creation.[12] CU_{ijt} is equal to one if both countries are included in the same customs union. Similarly, PSA_{ijt} is a partial scope agreement dummy and $CommonCurrency_{ijt}$ is a common currency dummy. δ_{ij}^{B} is a country-pair fixed effect and reflects all time-invariant factors that affect the bilateral trade values, such as distance, language, and the historical relationship between two countries. δ_{it}^{X} is an exporter-year fixed effect and reflects the production capacity of the exporter, outward multilateral resistance, and unilateral trade policies such as WTO accession. Finally, δ_{jt}^{M} is an importer-year fixed effect and reflects the total expenditure of the importer, inward multilateral resistance, and unilateral trade policies such as the reduction of MFN tariff rates. In addition, the combination of these two country-year fixed effects controls for the log of the levels of the bilateral exchange rates. In all estimations, standard errors are clustered by country pair.

Equation (2) is estimated initially by OLS. However, the bias of the OLS estimator has been pointed out recently. Santos Silva and Tenreyro (2006) show that when a log-linearized model such as the gravity model is estimated by OLS, heteroskedasticity affects both consistency and efficiency. They recommend specifying the conditional variance as proportional to the conditional mean and estimating the log-linearized model by PPML. The baseline specification is therefore the estimation with three types of fixed effects by PPML.[13] I can deal with the zero trade flows problem by PPML.

[12] The estimated coefficients for the FTA dummies are considered to be average effects over time. While I do not explicitly consider the phase-in effects in the first stage, it is partially addressed in the second stage.
[13] PPML with high-dimensional fixed effects is computationally demanding. In this chapter, I use the Stata command ppml_panel_sg written by Larch, Wanner, Yotov, and Zylkin (2017). See their paper for the detailed procedure.

In addition, the use of PPML is supported by the need to satisfy the adding up constraint (Arvis and Shepherd 2013; Fally 2015).

To explore the role of the country-pair fixed effects, I also estimate the gravity equation as follows:

$$\begin{aligned} Trade_{ijt} = \exp\big(&\alpha^{W} FTA_{ijt}^{W} + \alpha^{JX} FTA_{ijt}^{JX} + \alpha^{JM} FTA_{ijt}^{JM} + \beta_1 CU_{ijt} \\ &+ \beta_2 PSA_{ijt} + \beta_3 CommonCurrency_{ijt} + \gamma_1 \ln dist_{ij} \\ &+ \gamma_2 Contiguity_{ij} + \gamma_3 CommonLanguage_{ij} \\ &+ \gamma_4 CommonColonizer_{ij} + \delta_{it}^{X} + \delta_{jt}^{M}\big) + \varepsilon_{ijt}. \end{aligned} \tag{3}$$

In this specification, log of distance, a contingency dummy, a common language dummy, and a common colonizer dummy are added instead of the country-pair fixed effect. Compared with equation (2), this specification ignores the effects of unobservable factors related to the level of bilateral trade flows. The endogeneity bias of trade policies becomes severe if those factors are closely related to the determinants of trade policies.

We also estimate a panel version of the naïve gravity equation (Head and Mayer, 2014). In this specification, country-year fixed effects are dropped as follows:

$$\begin{aligned} Trade_{ijt} = \exp\big(&\alpha^{W} FTA_{ijt}^{W} + \alpha^{JX} FTA_{ijt}^{JX} + \alpha^{JM} FTA_{ijt}^{JM} + \beta_1 CU_{ijt} \\ &+ \beta_2 PSA_{ijt} + \beta_3 CommonCurrency_{ijt} + \lambda_1 \ln GDP_{it} \\ &+ \lambda_2 \ln GDP_{jt} + \lambda_3 \ln Population_{it} + \lambda_4 \ln Population_{jt} \\ &+ \lambda_5 \ln Remoteness_{it} + \lambda_6 \ln Remoteness_{jt} + \lambda_7 WTO_{it} \\ &+ \lambda_8 WTO_{jt} + \delta_{ij}^{B} + \delta_{t}^{T}\big) + \varepsilon_{ijt}, \end{aligned} \tag{4}$$

where $Remoteness_{it} = \left(\sum_{j \neq i} \frac{GDP_{jt}}{dist_{ij}}\right)^{-1}$ is the inverse of the market potential function and measures the degree of isolation from the rest of the world. $Remoteness_{jt}$ is defined in the same way and it is the inverse of supply potential. Other additional explanatory variables are log of GDP, log of population, a WTO dummy, and a year fixed effect. In this specification, multilateral resistance terms are not included. While the remoteness indices are used instead, Anderson and van Wincoop (2003) and Head and Mayer (2014) criticize their use because the indices do not have a solid theoretical foundation. Furthermore, the coefficients for the FTA dummies are biased if FTAs enter into force simultaneously with unilateral trade liberalization other than WTO accession.

5.5 Estimation results

In this section, I provide the estimation results of equations (2)–(4). I first estimate the average treatment effects of all FTAs. Japan's FTAs are then separated from the other FTAs. I further decompose the effects of Japan's FTAs by partner countries. Individual FTAs all over the world are also investigated to explore the determinants of successful FTAs.

5.5.1 Trade creation effects of all FTAs

Before estimating the effects of Japan's FTAs, I first estimate the effects of all FTAs. I start with the results of the traditional gravity specification (equation [4]) by OLS. As reported in column (1) of Table 5:3, the coefficients for the FTA dummy and the customs union dummy are both positive and statistically significant. The coefficients for the other gravity variables are also consistent with the standard gravity model estimates in previous studies.

Table 5:3 Estimation results for overall FTAs

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) PPML	(5) PPML	(6) PPML	(7) PPML w/o zero
FTA	0.367***	0.202***	0.118***	0.374***	0.104**	-0.00703	-0.00786
	(5.856)	(6.216)	(3.266)	(5.655)	(2.051)	(-0.127)	(-0.142)
CU	0.238**	0.561***	0.292***	0.656***	0.448***	0.0606	0.0535
	(2.363)	(9.950)	(4.390)	(6.536)	(6.739)	(0.786)	(0.688)
PSA	0.0181	0.0768	0.0663	-0.154	0.0993	-0.000895	-0.00176
	(0.231)	(0.989)	(0.837)	(-1.431)	(1.631)	(-0.0207)	(-0.0407)
Common Currency	-0.584***	-0.139***	-0.00833	-0.122	0.0302	-0.0386	-0.0396
	(-6.059)	(-4.521)	(-0.209)	(-1.590)	(0.907)	(-1.251)	(-1.284)
ln(Distance)	-1.108***			-0.669***			
	(-27.46)			(-17.48)			
Contiguity	0.612***			0.353**			
	(4.019)			(2.257)			
Common Colonizer	0.635***			0.407***			
	(5.358)			(6.170)			
Common Language	0.509***			0.122*			
	(6.936)			(1.938)			
Exporter ln(GDP)		0.0710***			0.268***		
		(7.134)			(3.145)		
Importer ln(GDP)		0.218***			0.740***		
		(16.61)			(23.22)		
Exporter ln(Population)		1.448***			0.698***		
		(12.86)			(5.788)		
Importer ln(Population)		0.555***			0.0680		
		(7.393)			(0.611)		
Exporter ln(Remoteness)		0.233**			0.746***		
		(2.035)			(5.415)		
Importer ln(Remoteness)		-0.0708			-0.119		
		(-0.601)			(-0.898)		
Exporter WTO		0.594***			0.366***		
		(12.65)			(5.632)		
Importer WTO		0.465***			0.226***		
		(10.33)			(4.945)		
Observations	91,267	91,262	91,262	93,840	93,760	93,760	91,262
exporter-year fixed effects	Yes	No	Yes	Yes	No	Yes	Yes
importer-year fixed effects	Yes	No	Yes	Yes	No	Yes	Yes
exporter-importer fixed effects	No	Yes	Yes	No	Yes	Yes	Yes

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculations.

We also estimate equation (3) and report the results in column (2). While the coefficient for the FTA dummy is significantly positive, it has halved in value. This implies that FTAs are more likely to be signed between country pairs with high ex ante trade values, conditional on gravity variables. While this is the opposite result to that of Baier and Bergstrand (2007), it is qualitatively consistent with Magee (2008). The coefficient for the customs union dummy increases. Overall, the coefficients for the other country-year variables are positive and statistically significant as expected.

Column (3) of Table 5:3 shows the estimation results with all three types of fixed effects. While the coefficient for the FTA dummy is positive and statistically significant, it has decreased compared with columns (1) and (2). This result implies that the role of trade policies is overestimated when the gravity model is specified with all three types of fixed effects.

The results of the estimations by PPML are almost the same as those by OLS. As reported in columns (4)–(6), the coefficients for the FTA dummies and the customs union dummies are positive and statistically significant. One notable difference between the results of OLS and PPML is that when three types of fixed effects are included, the coefficients for the FTA dummies and the customs union dummies decrease to almost zero, which means that these trade policies have no trade creation effects on average.

A potential reason for the lack of significant effects is the choice of the sample period. Baier and Bergstrand (2007) and Magee (2008), for example, use the sample periods 1960–2000 and 1980–98, respectively. The coefficient for the FTA dummy in the present chapter reflects the impacts of only recently signed FTAs because the trade creation effects of FTAs enforced before 1996 are absorbed into the country-pair fixed effect. Kohl (2014) points out that FTAs enforced after the 1990s performed poorly.

5.5.2 Trade creation effects of Japan's FTAs

We next turn to the trade creation effects of Japan's FTAs. Table 5:4 shows the results of estimation, in which the effects of Japan's FTAs are separated from those of others. Other covariates are included in the estimation, but not reported in the table to save space. As reported in column (1) of Table 5:4, the coefficients for Japan's FTA dummies are positive and statistically significant for exports and positive and slightly insignificant for imports when the country-year fixed effects are included but the country-pair fixed effects are not. While the results change when the country-pair fixed effects are included in the regression, the point estimates are still large in the estimation with three types of fixed effects. In addition, this result holds when the gravity model is estimated by PPML. As reported in column (6), the point estimates of the coefficients for Japan's FTA dummies are 0.067 for exports and 0.086 for imports, but they are not statistically significant.

Table 5:4 Estimation results for Japan's FTAs and others

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) PPML	(5) PPML	(6) PPML	(7) PPML w/o zero
Japan's FTA (export)	0.672***	-0.191*	0.178	0.504***	-0.00362	0.0672	0.0667
	(3.636)	(-1.758)	(1.583)	(3.245)	(-0.0249)	(0.781)	(0.775)
Japan's FTA (import)	0.760	-0.0525	0.106	0.423*	0.173**	0.0864	0.0855
	(1.615)	(-0.513)	(0.808)	(1.821)	(2.516)	(1.312)	(1.296)
Other FTA	0.397***	0.154***	0.0517	0.296***	0.0738	-0.0439	-0.0460
	(6.164)	(4.789)	(1.437)	(4.267)	(1.430)	(-0.684)	(-0.716)
Observations	91,267	91,262	91,262	93,840	93,760	93,760	91,262
exporter-year fixed effects	Yes	No	Yes	Yes	No	Yes	Yes
importer-year fixed effects	Yes	No	Yes	Yes	No	Yes	Yes
exporter-importer fixed effects	No	Yes	Yes	No	Yes	Yes	Yes

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculations.

Overall, Japan's FTAs do not appear to have positive effects on trade values when the model is correctly specified. While the coefficients are statistically insignificant, the point estimates are not small. These results weakly suggest that the effects of Japan's FTAs are heterogeneous.

5.5.3 Trade creation effects of Japan's individual FTAs

The results of the previous subsection suggest that some of Japan's FTAs are working well, but others are not. To examine this point further, I decompose the effects of Japan's FTAs by partner countries.

Table 5:5 presents the coefficients for Japan's FTA dummies. I regard column (4), which includes three types of fixed effects and estimates by PPML, as the most reliable result.[14] I also add the results from other studies, Ando and Urata (2015) and Yamanouchi (2017), in columns (5) and (6). Based on column (4), the export values from Japan to Australia, Chile, India, Indonesia, Mexico, Myanmar, Thailand, and Vietnam are positively affected by the FTAs with these countries. The FTA with Myanmar (AJCEP) has the largest effect and it increased Japan's exports to Myanmar by $\exp(0.517) - 1 = 67.7$ percent. This result is surprising because Myanmar's tariff rates were not

[14] As robustness checks, I also estimate the model (1) without zero trade values, (2) without interpolated values, (3) limiting the sample to impose one-year intervals, and (4) including pair-specific linear trends. The results do not change substantially.

lowered in the sample period under AJCEP. Therefore, this implies that removing nontariff barriers is crucial for trade creation. In contrast, FTAs with Brunei, Cambodia, Laos, Peru, Philippines, and Switzerland have no significant effects on Japan's exports. The coefficients for FTAs with Malaysia and Singapore are negative and statistically significant.

Table 5:5 Estimation results for Japan's individual FTAs

	(1)		(2)		(3)		(4)		(5) Ando and Urata (2015)		(6) Yamanouchi (2017)	
VARIABLES	OLS		PPML		PPML		PPML		OLS		PPML	
	Export	Import	Export	Import	Export	Import	Export	Import	Export	Import	Export	Import
FTA (Australia)	0.285**	0.350***	0.312**	1.469***	0.0401	0.139***	0.153**	0.190**				
	(2.202)	(3.352)	(2.415)	(3.658)	(0.896)	(3.442)	(2.135)	(2.008)				
FTA (Brunei)	0.132	0.803***	0.188	2.473***	0.0685	0.557***	0.108	0.627***			-0.0429	0.370***
	(1.495)	(3.597)	(0.951)	(6.187)	(1.600)	(16.12)	(1.119)	(3.025)				
FTA (Cambodia)	0.125	-0.245	-0.491**	-0.0673	0.134***	1.245***	0.0570	0.610***			0.114***	0.413***
	(0.988)	(-1.595)	(-2.226)	(-0.192)	(2.809)	(31.42)	(0.470)	(3.252)				
FTA (Chile)	0.744***	0.132	0.147	1.350***	0.462***	0.369***	0.397***	0.155	0.593***	0.0533	0.595***	0.109**
	(5.023)	(1.228)	(1.094)	(4.393)	(10.85)	(9.481)	(2.728)	(1.603)	(4.556)	(0.578)		
FTA (India)	-0.175	-0.252***	-0.631***	-0.639***	0.469***	0.330***	0.170**	-0.0246				
	(-1.156)	(-3.389)	(-3.347)	(-2.680)	(11.24)	(9.168)	(2.511)	(-0.323)				
FTA (Indonesia)	0.271**	-0.0343	0.585***	0.993***	0.302***	-0.000329	0.265**	-0.0129	-0.106	-0.303***	0.0122	-0.201**
	(2.508)	(-0.458)	(4.385)	(5.322)	(6.463)	(-0.00699)	(2.561)	(-0.213)	(-0.710)	(-2.856)		
FTA (Laos)	0.853***	1.041***	-0.634	-0.730**	-0.180***	1.161***	-0.151	0.594***			0.560***	1.108***
	(5.839)	(5.470)	(-1.192)	(-1.994)	(-3.274)	(22.66)	(-1.097)	(2.650)				
FTA (Malaysia)	-0.520***	0.0652	0.796***	0.711***	-0.452***	0.231***	-0.178**	0.304***	-0.148	-0.0262	-0.220***	-0.0515
	(-3.965)	(0.853)	(6.387)	(3.270)	(-8.978)	(6.386)	(-2.490)	(3.599)	(-1.201)	(-0.301)		
FTA (Mexico)	0.192*	-0.223**	0.361***	-0.799***	0.534***	0.214***	0.428***	0.00810	0.628***	0.264***	0.498***	0.141**
	(1.795)	(-2.340)	(3.012)	(-3.572)	(12.84)	(3.680)	(4.567)	(0.128)	(4.800)	(2.844)		
FTA (Myanmar)	0.864***	0.874***	0.515**	0.211	0.0588	0.530***	0.517***	0.679***				
	(6.553)	(4.820)	(2.334)	(0.547)	(1.074)	(5.468)	(3.776)	(2.734)				
FTA (Peru)	0.148	0.238*	0.0522	0.632**	0.185***	0.341***	0.0251	-0.0108				
	(1.263)	(1.929)	(0.356)	(2.182)	(4.430)	(9.035)	(0.225)	(-0.0902)				
FTA (Philippines)	-0.0955	-0.199**	0.267	0.433*	-0.716***	-0.158***	-0.144	0.199*	-0.216	-0.400***	-0.181***	-0.369***
	(-0.871)	(-2.169)	(1.544)	(1.958)	(-15.71)	(-4.556)	(-1.583)	(1.841)	(-1.464)	(-3.822)		
FTA (Singapore)	-0.316***	-0.113	0.397**	-0.348*	-0.486***	-0.202***	-0.307***	-0.138	-0.0863	-0.118	-0.485***	-0.407***
	(-3.012)	(-1.357)	(2.563)	(-1.705)	(-9.341)	(-4.993)	(-4.272)	(-0.884)	(-0.445)	(-0.857)		
FTA (Switzerland)	0.0576	0.184**	-0.0444	0.471*	0.251***	0.137***	0.0990	-0.00131	0.914***	0.0844	0.391***	0.128**
	(0.544)	(2.246)	(-0.325)	(1.945)	(6.393)	(2.615)	(0.919)	(-0.0139)	(4.704)	(0.612)		
FTA (Thailand)	0.343***	-0.281***	1.203***	0.506***	0.208***	0.295***	0.211***	0.0366	0.0366	-0.0761	0.0664	-0.0248
	(3.687)	(-3.308)	(9.873)	(2.762)	(4.124)	(7.980)	(3.393)	(0.444)	(0.281)	(-0.824)		
FTA (Vietnam)	-0.0664	-0.448***	0.575***	0.598**	0.321***	0.551***	0.209*	-0.153	0.239	0.0845	0.182***	0.175***
	(-0.520)	(-3.978)	(3.814)	(2.436)	(5.802)	(10.54)	(1.856)	(-1.177)	(1.599)	(0.799)		
Observations	91,262		93,840		93,760		93,760		360	360	1,908	1,908
Exporter-year fixed effects	Yes		Yes		No		Yes		Yes	Yes	Yes	Yes
Importer-year fixed effects	Yes		Yes		No		Yes		No	No	No	No
Exporter-importer fixed effects	Yes		No		Yes		Yes		Yes	Yes	Yes	Yes

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Columns (1)–(4) are calculated by the author. Column (5) is taken from Ando and Urata (2015, Table 10). Column (6) is taken from Yamanouchi (2017, Tables 43 and 47).

Import values, however, increased significantly because of FTAs with Australia, Brunei, Cambodia, Laos, Malaysia, Myanmar, and Philippines. The coefficients are statistically insignificant for the other FTAs. The largest effect on imports is also observed for the FTA with Myanmar.

We also check the symmetry of the trade creation effects using a joint test of the hypothesis that all pairs of coefficients are equal for exports and imports. The chi-square statistic is 103.1 and the hypothesis is clearly rejected.

Tables 5:6 and 5:7 present the matrices of the root mean square differences (RMSDs) between the coefficients for the FTA dummies across specifications. Table 5:6 shows the RMSDs of the coefficients for Japan's exports to its FTA partners. The coefficients in the baseline specification (6) differ little from those estimated without country-year dummies. The choice of estimator does not matter much. However, the specifications without the country-pair dummies (1) and (4) show very different coefficients. Table 5:7 shows the same matrix for Japan's imports. The results are like the case of exports, and misspecification is problematic if the country-pair fixed effects are excluded from the estimation. If the country-pair fixed effects are not included, the trade creation effects of the FTAs are overestimated because of endogeneity.

Table 5:6 Matrix of RMSDs between coefficients for Japan's FTAs (exports)

export	(1)	(2)	(3)	(4)	(5)	(6)
(1) OLS, No pair dummies	0.00					
(2) OLS, No country-year dummies	1.20	0.00				
(3) OLS, three types	0.68	0.60	0.00			
(4) PPML, No pair dummies	0.89	0.79	0.66	0.00		
(5) PPML, No country-year dummies	0.97	0.29	0.43	0.65	0.00	
(6) PPML, three types	0.85	0.43	0.32	0.51	0.23	0.00

Source: Author's calculations.

Table 5:7 Matrix of RMSDs between coefficients for Japan's FTAs (imports)

import	(1)	(2)	(3)	(4)	(5)	(6)
(1) OLS, No pair dummies	0.00					
(2) OLS, No country-year dummies	1.89	0.00				
(3) OLS, three types	1.62	0.46	0.00			
(4) PPML, No pair dummies	1.29	1.08	0.92	0.00		
(5) PPML, No country-year dummies	1.74	0.44	0.53	0.99	0.00	
(6) PPML, three types	1.69	0.38	0.32	0.89	0.34	0.00

Source: Author's calculations.

In this subsection, the effects of the individual FTAs are explored. I found that each of Japan's FTAs affects trade values in a different way. Therefore, I conclude that one of the reasons for the absence of the trade creation effects of Japan's FTAs is that aggregation across all trade partners obscures the positive effects. I should note that about half of Japan's FTAs have positive and statistically significant impacts.

5.5.4 Trade creation effects of each individual FTA

In this subsection, I consider the determinants of the trade creation effects. To this end, I first obtain the coefficients for each individual FTA. I then regress the estimated coefficients on some variables. This two-step estimation is suggested in Baier, Yotov, and Zylkin (2019).

The first stage is estimated by OLS as in Kohl (2014) because of the computational difficulty. As explained in the previous subsection, the differences between coefficients in the two estimates are small. Among the 725 directional flows within active FTAs, 256 (35 percent) of these flows have positive and statistically significant values at 5 percent levels. Mean and median values are 0.113 and 0.105, respectively.[15] Compared with Baier, Yotov, and Zylkin (2019), these values are small and suggest that recent FTAs have weaker effects.

To examine the effects of individual FTAs, I regress the estimated coefficients on some other variables in the second stage. The estimating equation is as follows:

$$\alpha_{ij} = \mu_1 \overline{\delta^X}_i + \mu_2 \overline{\delta^M}_j + \mu_3 \delta^B_{ij} + \mu_4 FTAyears_{ij} + u_{ij}, \qquad (5)$$

where $\overline{\delta^X}_i$ and $\overline{\delta^M}_j$ are the means of exporter-year and importer-year fixed effects estimated in the first stage, respectively. These variables measure the size of countries in terms of trade values. δ^B_{ij} is the estimated country-pair fixed effect and a proxy for the strength of trade linkages before the FTAs were established. Intuitively, these three variables are expected to have negative coefficients because there is no room for more trade if a country is already open to trade. The last variable, $FTAyears_{ij}$, equals the number of years for which the country pair has an FTA. The coefficient is expected to be

[15] 301 (42%) are insignificant and 168 (23%) are negatively significant.

positive if recent FTAs are not effective. The cohort effects of FTAs are not identified from phase-in effects, so the positive coefficient may be a result of phase-in effects.[16]

In another specification, I replace the estimated fixed effects with some gravity variables as follows:

$$\alpha_{ij} = \nu_1 \overline{\ln GDP}_i + \nu_2 \overline{\ln GDP}_j + \nu_3 \ln Distance_{ij} + \nu_4 FTAyears_{ij} + u_{ij}, \qquad (6)$$

where the expected signs are the same as for equation (5), except ν_3 because the country-pair fixed effects are smaller for longer distances. Following Baier, Yotov, and Zylkin (2019), I simply use heteroskedasticity-robust standard errors. Furthermore, I include country fixed effects again instead of these variables in some estimations of the second stage.

The estimation results of equation (5) are shown in Table 5:8. In the first column, the signs of the coefficients are as expected. The FTAs generally increase trade values if the trading countries are smaller and the relationship between the two countries is weak. In addition, the results imply that relatively old FTAs are more effective.

[16] Yamanouchi (2017) does not find phase-in effects for Japan's FTAs.

Table 5:8 Estimation results for the effects of individual FTAs using estimated fixed effects

	(1)	(2)	(3)	(4)	(5)
Importer-year fixed effect	-0.0529***	-0.0550***	-0.0369		
	(-3.828)	(-3.729)	(-1.519)		
Importer-year fixed effect		-0.0759	-0.0869***		
× Export from Japan		(-1.423)	(-3.583)		
Exporter-year fixed effect	-0.0447*	-0.0437*		-0.0631*	
	(-1.822)	(-1.650)		(-1.868)	
Exporter-year fixed effect		-0.0860		-0.0716**	
× Import to Japan		(-1.331)		(-2.121)	
Country pair fixed effect	-0.128***	-0.139***	-0.160***	-0.163***	-0.178***
	(-4.329)	(-4.531)	(-3.241)	(-4.786)	(-3.366)
Country pair fixed effect		-0.0191	-0.0182	0.137	0.146**
× Export from Japan		(-0.128)	(-0.368)	(1.534)	(2.123)
Country pair fixed effect		0.171**	0.0406	0.161***	0.0637
× Import to Japan		(2.070)	(0.526)	(4.753)	(1.000)
Number of years under FTA	0.0147***	0.0158***	0.0253***	0.0231**	0.0289**
	(2.795)	(2.981)	(3.088)	(2.540)	(2.351)
Number of years under FTA		0.0249	-0.0328***	0.0147	0.00231
× Export from Japan		(1.319)	(-4.009)	(1.095)	(0.153)
Number of years under FTA		0.000187	-0.00323	-0.0536***	-0.0391***
× Import to Japan		(0.0142)	(-0.286)	(-5.884)	(-3.888)
Observations	725	725	723	723	721
Adjusted R-squared	0.070	0.075	0.184	0.155	0.250
Exporter fixed effects	No	No	Yes	No	Yes
Importer fixed effects	No	No	No	Yes	Yes

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculations.

In columns (2)–(5), the interaction terms with Japan's export and import dummies are added to explore the characteristics of Japan's FTAs.[17] While the results are not stable, the effects of Japan's FTAs are larger when the partner countries are smaller. However, the trade linkages before FTAs are signed are not important. More interestingly, recently enforced FTAs increase Japan's import values. In the case of Japan, agricultural goods are excluded from the negotiations in initial agreements. Protection is likely to be weakened and tariff rates for some products such as beef are lowered in the recent agreements with Australia.[18,19]

[17] I also check the results of regressions estimated using Japan's variables only. Of course, the sample size is so small that the results are only indicative, but the main results are unchanged from Tables 5:8 and 5:9.
[18] I cannot see the effects of lowered tariff rates using simple statistics because Japan's import values of agricultural goods from Australia decreased after the FTA entered into force. As explained in Section 5.3, I cannot make conclusions about the effects of FTAs from the descriptive analysis. In fact, Australia's total exports also decreased in 2015.
[19] It is difficult to explain the large effects of Japan's recent FTAs by the differences of contents across Japan's FTAs. Almost all FTAs have chapters on investment and trade in services. Provisions on intellectual property, movement of natural persons, and government procurement are not limited to recent FTAs.

In Table 5:9, I present the estimation results of equation (6). While the main results are unchanged from the previous analyses, I find some notable differences. First, exporter GDP does not explain the effectiveness of an FTA.[20] This implies that some factors related to export values but not to GDP are key determinants of trade creation effects. One potential explanation is the endowment of natural resources. The export values of natural resources are not closely related to tariff rates, so the countries specializing in those resource sectors cannot increase exports via FTAs. Similarly, the coefficients for distance are statistically insignificant in some specifications, and therefore the role of distance is less clear than the country-pair fixed effects. However, distance is important for Japan's import under FTAs.

Table 5:9 Estimation results for the effects of individual FTAs using GDP and distance

	(1)	(2)	(3)	(4)	(5)
Importer ln(GDP)	-0.0345**	-0.0350*	-0.0238		
	(-2.047)	(-1.893)	(-0.878)		
Importer ln(GDP)		-0.101**	-0.118***		
× Export from Japan		(-2.263)	(-4.345)		
Exporter ln(GDP)	-0.0251	-0.0261		-0.0307	
	(-0.957)	(-0.905)		(-0.942)	
Exporter ln(GDP)		-0.124**		-0.148***	
× Import to Japan		(-2.427)		(-4.550)	
ln(Distance)	0.0751*	0.0728*	0.105	0.0904	0.174*
	(1.756)	(1.671)	(1.474)	(1.309)	(1.689)
ln(Distance)		0.347**	0.271***	0.0148	0.161
× Export from Japan		(2.439)	(3.799)	(0.743)	(1.573)
ln(Distance)		0.409***	0.0326**	0.144**	-0.151
× Import to Japan		(2.695)	(2.393)	(2.089)	(-1.370)
Number of years under FTA	0.0143***	0.0150***	0.0210**	0.0210**	0.0268**
	(2.621)	(2.694)	(2.420)	(2.544)	(2.107)
Number of years under FTA		-0.0361	-0.0456***	0.000480	0.0109
× Export from Japan		(-1.444)	(-5.261)	(0.0209)	(0.383)
Number of years under FTA		-0.0357	-0.0416***	-0.0617***	-0.0383***
× Import to Japan		(-1.604)	(-3.720)	(-7.478)	(-3.884)
Observations	725	725	723	723	721
Adjusted R-squared	0.010	0.008	0.119	0.067	0.181
Exporter fixed effects	No	No	Yes	No	Yes
Importer fixed effects	No	No	No	Yes	Yes

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculations.

[20] Baier, Yotov, and Zylkin (2019) report positive coefficients for both exporter and importer GDP.

In this subsection, I discussed which types of countries have effective FTAs. In a nutshell, a small ex ante trade value means substantial scope to increase trade via FTAs. Japan's FTA partners are so far mainly located in the Asia–Pacific region and actively transacting with Japan before signing the agreements. Japan can be integrated with the global market through FTAs along with many developing countries.

5.6 Conclusion

While FTAs are one of the major commercial policies of the 21st century, Japan's FTAs have not been evaluated adequately. This chapter therefore investigates the effects of Japan's FTAs using a recently developed gravity framework and explores the determinants of the effects of FTAs.

The estimation results do not indicate the presence of trade creation effects for Japan's FTAs on average. However, the effects of Japan's FTAs vary substantially across partners and around half of the FTAs increase Japan's trade values. Positive impacts on Japan's exports are more likely to be observed for small partners. Japan's imports from FTA partners tend to increase when the partner countries are small and distant. More importantly, Japan's recent FTAs have larger effects.

The results suggest that there is a little scope to increase trade values with some countries. This implies that political resources should be directed toward negotiations with developing countries if the government is aiming to integrate the economy into the global market through FTAs. However, large-scaled multilateral trade agreements may not be effective in terms of trade creation, even though large amounts of effort are spent on the negotiation. Of course, the impacts on investment and other forms of international cooperation are different and those multilateral agreements may play important roles in regulating world trade systems and supporting new types of globalization. These are issues open for future research.

Bibliography

Aaronson, D., S. Agarwal, and E. French. 2012. The Spending and Debt Responses to Minimum Wage Hikes. *American Economic Review* 102 (7), 3111–3139.

Aaronson, D. and E. French. 2007. Product Market Evidence on the Employment Effects of the Minimum Wage. *Journal of Labor Economics* 25 (1), 167–200.

Abel, A. B. 1983. Optimal Investment under Uncertainty. *American Economic Review* 73 (1), 228-233.

Ackerberg, D. A., K. Caves, and G. Frazer. 2015. Identification Properties of Recent Production Function Estimators. *Econometrica* 83 (6), 2411–2451.

Aghion, P., J. Cai, M. Dewatripont, L. Du, A. Harrison, and P. Legros. 2015. Industrial Policy and Competition. *American Economic Journal: Macroeconomics* 7 (4), 1-32.

Akdoğu, E. and P. MacKay. 2008. Investment and Competition. *Journal of Financial and Quantitative Analysis* 43 (2), 299-330.

Allegretto, S., A. Dube, M. Reich, and B. Zipperer. 2017. Credible Research Designs for Minimum Wage Studies: A Response to Neumark, Salas, and Wascher. *ILR Review* 70 (3), 559-592.

Anderson, J. E. and E. van Wincoop. 2003. Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review* 93 (1):170–192.

Ando, M. and S. Urata. 2011. Impacts of the Japan-Mexico EPA on Bilateral Trade. RIETI Discussion Paper 11-E-020.

Ando, M. and S. Urata. 2015. Impacts of Japan’s FTAs on Trade: The Cases of FTAs with Malaysia, Thailand, and Indonesia. RIETI Discussion Paper 15-E-104.

Andrews, D. and F. Cingano. 2014. Public Policy and Resource Allocation: Evidence from Firms in OECD Countries. *Economic Policy* 29 (78), 253-296.

Arbatli, E. C., S. J. Davis, A. Ito, N. Miake, and I. Saito. 2017. Policy Uncertainty in Japan. NBER Working Paper No. 23411.

Arellano, M. and S. Bond. 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies* 58 (2), 277-297.

Arvis, J.-F. and B. Shepherd. 2013. The Poisson Quasi-Maximum Likelihood Estimator: A Solution to the ‘Adding Up’ Problem in Gravity Models. *Applied Economics Letters* 20 (6), 515–519.

Asker, J., A. Collard-Wexler, and J. De Loecker. 2014. Dynamic Inputs and Resource (Mis) Allocation. *Journal of Political Economy* 122 (5), 1013-1063.

Aw, B.-Y., X. Chen, and M. J. Roberts. 2001. Firm-level Evidence on Productivity Differentials and Turnover in Taiwanese Manufacturing. *Journal of Development Economics* 66 (1), 51–86.

Azar, J. A., I. Marinescu, and M. I. Steinbaum. 2017. Labor Market Concentration. NBER Working Paper No. 24147.

Azar, J. A., I. Marinescu, M. I. Steinbaum, and B. Taska. 2018. Concentration in US Labor Markets: Evidence from Online Vacancy Data. NBER Working Paper No. 24395.

Bach, T. 2014. Misallocation and Growth in an Entry Liberalization Period. Unpublished.

Baier, S. L., Y. V. Yotov, and T. Zylkin. 2019. On the Widely Differing Effects of Free Trade Agreements: Lessons from Twenty Years of Trade Integration. *Journal of International Economics* 116 (1), 206–226.

Baier, S. L. and J. H. Bergstrand. 2007. Do Free Trade Agreements Actually Increase Members' International Trade? *Journal of International Economics* 71 (1), 72–95.

Baier, S. L., J. H. Bergstrand, and M. W. Clance. 2018. Heterogeneous Effects of Economic Integration Agreements. *Journal of Development Economics* 135 (1), 587–608.

Baily, M. N., C. Hulten, and D. Campbell. 1992. Productivity Dynamics in Manufacturing Plants. *Brookings Papers on Economic Activity: Microeconomics* 1992, 187-267.

Baker, M., D. Benjamin, and S. Stanger. 1999. The Highs and Lows of the Minimum Wage Effect: A Time-Series Cross-Section Study of the Canadian Law. *Journal of Labor Economics* 17 (2), 318–350.

Banerjee, A. V. and E. Duflo. 2005. Growth Theory through the Lens of Development Economics. In Handbook of Economic Growth Vol.1A edited by P. Aghion and S. N. Durlauf, 473-552. Elsevier.

Banerjee, A. V. and B. Moll. 2010. Why Does Misallocation Persist? *American Economic Journal: Macroeconomics* 2 (1), 189-206.

Bar-Ilan, A. and W. C. Strange 1996. Investment Lags. *American Economic Review* 86 (3), 610-622.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta. 2013. Cross-Country Differences in Productivity: The Role of Allocation and Selection. *American Economic Review* 103 (1), 305–334.

Bellone, F. and J. Mallen-Pisano. 2013. Is Misallocation Higher in France than in the United States? University of Nice-Sophia-Antipolis. Unpublished.

Benmelech, E., N. Bergman, H. Kim. 2018. Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages? NBER Working Paper No. 24307.

Bentolila, S., G. Bertola. 1990. Firing Costs and Labour Demand: How Bad is Eurosclerosis? *Review of Economic Studies* 57 (3), 381–402.

Bernard, A. B., and J. B. Jensen. 1995. Exporters, Jobs, and Wages in US Manufacturing: 1976-1987. *Brookings Papers on Economic Activity. Microeconomics* 1995, 67-119.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott. 2007. Firms in International Trade. *Journal of Economic Perspectives* 21 (3), 105-130.

Bloom, N. 2009. The Impact of Uncertainty Shocks. *Econometrica* 77 (3), 623-685.

Bloom, N. 2014. Fluctuations in Uncertainty. *Journal of Economic Perspectives* 28 (2), 153-175.

Bloom, N., S. Bond, and J. Van Reenen. 2007. Uncertainty and Investment Dynamics. *Review of Economic Studies* 74 (2), 391-415.

Blundell, R. and S. Bond. 1998. Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics* 87 (1), 115–143.

Blundell, R. and S. Bond. 2000. GMM Estimation with Persistent Panel Data: An Application to Production Functions. *Econometric Reviews* 19 (3), 321-340.

Bontempi, M. E., R. Golinelli, and G. Parigi. 2010. Why Demand Uncertainty Curbs Investment: Evidence from a Panel of Italian Manufacturing Firms. *Journal of Macroeconomics* 32 (1), 218-238.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt. 2002. Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence. *Quarterly Journal of Economics* 117 (1), 339-376.

Broda, C., J. Greenfield, and D. Weinstein. 2006. From Groundnuts to Globalization: A Structural Estimate of Trade and Growth. NBER Working Paper No. 12512.

Bulan, L. T. 2005. Real Options, Irreversible Investment and Firm Uncertainty: New Evidence from US Firms. *Review of Financial Economics* 14 (3), 255-279.

Bulan, L., C. Mayer, and C. T. Somerville 2009. Irreversible Investment, Real Options, and Competition: Evidence from Real Estate Development. *Journal of Urban Economics* 65 (3), 237-251.

Busso, M., L. Madrigal, and C. Pages-Serra. 2012. Productivity and Resource Misallocation in Latin America. *The B.E. Journal of Macroeconomics* 13 (1), 903–932.

Caballero, R. J. 1991. On the Sign of the Investment-Uncertainty Relationship. *American Economic Review* 81 (1), 279-288.

Caballero, R. J., T. Hoshi, and A. K. Kashyap. 2008. Zombie Lending and Depressed Restructuring in Japan. *American Economic Review* 98 (5), 1943-1977.

Calligaris, S. 2015. Misallocation and Total Factor Productivity in Italy: Evidence from Firm-Level Data. *Labour* 29 (4), 367–393.

Camacho, A. and E. Conover. 2010. Misallocation and Productivity in Colombia's Manufacturing.

IDB Working Paper Series 6783.

Card, D., and A. B. Krueger. 1994. Minimum Wages and Employment: A Case Study of the Fast–Food Industry in New Jersey and Pennsylvania. *American Economic Review* 84 (4), 772–793.

Card, D., A. B. Krueger. 2000. Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Reply. *American Economic Review* 90 (5), 1397–1420.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer. 2019. The Effect of Minimum Wages on Low-Wage Jobs. *Quarterly Journal of Economics* 134 (3), 1405-1454.

Chaney, T. 2008. Distorted Gravity: the Intensive and Extensive Margins of International Trade. *American Economic Review*, 98 (4), 1707-1721.

Chia, S. Y. 2013. The Emerging Regional Economic Integration Architecture in East Asia. *Asian Economic Papers* 12 (1), 1–37.

Chia, S. Y. 2015. Emerging Mega-FTAs: Rationale, Challenges, and Implications. *Asian Economic Papers* 14 (1), 1–27.

Cipollina, M., and L. Salvatici. 2010. Reciprocal Trade Agreements in Gravity Models: A Meta-Analysis. *Review of International Economics* 18 (1), 63–80.

Clemens, J. and M. Wither. 2019. The Minimum Wage and the Great Recession: Evidence of Effects on the Employment and Income Trajectories of Low-Skilled Workers. *Journal of Public Economics* 170, 53–67.

Da Rocha, J. M., and P. Pujolas. 2011. Policy Distortions and Aggregate Productivity: The Role of Idiosyncratic Shocks. *The B.E. Journal of Macroeconomics* 11 (1), 1-36.

Davis, S. J., J. C. Haltiwanger and S. Schuh. 1996. Job Creation and Destruction. Cambridge: MIT Press.

De Loecker, J. and F. Warzynski. 2012. Markups and Firm-Level Export Status. *American Economic Review*, 102 (6), 2437-2471.

De Loecker, J. and P. K. Goldberg. 2013. Firm Performance in a Global Market. *Annual Review of Economics* 6, 201–227.

de Vries, G. J. 2014. Productivity in A Distorted Market: The Case of Brazil's Retail Sector. *Review of Income and Wealth* 60 (3), 499–524.

Dheera-Aumpon, S. 2014. Misallocation and Manufacturing TFP in Thailand. *Asia-Pacific Economic Literature* 28 (2), 63–76.

Dixit, A. K. and R. S. Pindyck. 1994. Investment under Uncertainty. Princeton University Press.

Dobbelaere, S., K. Kiyota, and J. Mairesse. 2015. Product and Labor Market Imperfections and

Scale Economies: Micro-Evidence on France, Japan and the Netherlands. *Journal of Comparative Economics* 43 (2), 290–322.

Dobbelaere, S. and J. Mairesse. 2013. Panel Data Estimates of the Production Function and Product and Labor Market Imperfections. *Journal of Applied Econometrics* 28 (1), 1–46.

Draca, M., S. Machin, and J. Van Reenen. 2011. Minimum Wages and Firm Profitability. *American Economic Journal: Applied Economics* 3 (1), 129–151.

Dube, A., T. W. Lester, and M. Reich. 2010. Minimum Wage Effects across State Borders: Estimates Using Contiguous Counties. *Review of Economics and Statistics* 92 (4), 945–964.

Dube, A., T. W. Lester, and M. Reich. 2016. Minimum Wage Shocks, Employment Flows, and Labor Market Frictions. *Journal of Labor Economics* 34 (3), 663–704.

Edmond, C., V. Midrigan, and D. Y. Xu 2015. Competition, Markups, and the Gains from International Trade. *American Economic Review* 105 (10), 3183-3221.

Egger, P. and M. Larch. 2008. Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis. *Journal of International Economics* 76 (2), 384–399.

Eicher, T. S., C. Henn, and C. Papageorgiou. 2012. Trade Creation and Diversion Revisited: Accounting for Model Uncertainty and Natural Trading Partner Effects. *Journal of Applied Econometrics* 27 (2), 296–321.

Eisfeldt, A. L. and D. Papanikolaou. 2013. Organization Capital and the Cross-Section of Expected Returns. *Journal of Finance* 68 (4), 1365-1406.

Ericson, R. and A. Pakes. 1995. Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *Review of Economic Studies* 62 (1), 53-82.

Eslava, M., J. Haltiwanger, A. Kugler, and M. Kugler. 2004. The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia. *Journal of Development Economics* 75 (2), 333–371.

Falch, T. 2010. The Elasticity of Labor Supply at the Establishment Level. *Journal of Labor Economics* 28 (2), 237–266.

Fally, T. 2015. Structural Gravity and Fixed Effects. *Journal of International Economics* 97 (1), 76–85.

Foster, L., J. Haltiwanger, and C. J. Krizan. 2001. Aggregate Productivity Growth: Lessons from Microeconomic Evidence. In New Developments in Productivity Analysis, edited by C. R. Hulten, E. R. Dean, and M. J. Harper, 303-372. University of Chicago Press.

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

Foster, L., J. Haltiwanger, and C. Syverson. 2016. The Slow Growth of New Plants: Learning about Demand? *Economica* 83 (329), 92-129.

Fujii, D., Y. Saito, and T. Senga. 2017. The Dynamics of Inter-firm Networks and Firm Growth. RIETI Discussion Paper 17-E-110.

Gandhi, A., S. Navarro, and D. Rivers. 2013. On the Identification of Production Functions: How Heterogeneous is Productivity? Unpublished.

Ghosal, V. and P. Loungani 2000. The Differential Impact of Uncertainty on Investment in Small and Large Businesses. *Review of Economics and Statistics* 82 (2), 338-343.

Giuliano, L. 2013. Minimum Wage Effects on Employment, Substitution, and the Teenage Labor Supply: Evidence from Personnel Data. *Journal of Labor Economics* 31 (1), 155–194.

Grenadier, S. R. 1996. The Strategic Exercise of Options: Development Cascades and Overbuilding in Real Estate Markets. *Journal of Finance* 51 (5), 1653-1679.

Grenadier, S. R. 2002. Option Exercise Games: An Application to the Equilibrium Investment Strategies of Firms. *Review of Financial Studies* 15 (3), 691-721.

Guiso, L. and G. Parigi. 1999. Investment and Demand Uncertainty. *Quarterly Journal of Economics* 114 (1), 185-227.

Ha, D. T. T. and K. Kiyota. 2014. Firm-level Evidence on Productivity Differentials and Turnover in Vietnamese Manufacturing. *Japanese Economic Review* 65 (2), 193–217.

Ha, D. T. T. and K. Kiyota. 2015. Misallocation, Trade Liberalization, and the Global Financial Crisis: The Case of Vietnamese Manufacturing. Keio University. Unpublished.

Hara, H. 2017. Minimum Wage Effects on Firm-Provided and Worker-Initiated Training. *Labour Economics* 47, 149–162.

Harasztosi, P. and A. Lindner. 2019. Who Pays for the Minimum Wage? *American Economic Review* 109 (8), 2693-2727.

Head, K., and T. Mayer. 2014. Gravity Equations: Workhorse, Toolkit, and Cookbook. In Handbook of International Economics Vol. 4, edited by G. Gopinath, E. Helpman, and K. Rogoff, 131-195. Elsevier.

Head, K., T. Mayer., and J. Ries. 2010. The Erosion of Colonial Trade Linkages After Independence. *Journal of International Economics* 81 (1), 1–14.

Hall, R. E. 1988. The Relation between Price and Marginal Cost in US Industry. *Journal of political Economy* 96 (5), 921-947.

Hopenhayn, H. 1992. Entry, Exit, and Firm Dynamics in Long Run Equilibrium. *Econometrica* 60 (5), 1127-50.

Hopenhayn, H. A. 2014. Firms, Misallocation, and Aggregate Productivity: A Review. *Annual Review of Economics* 6, 735-770.

Hosono, K. and M. Takizawa. 2013. Misallocation to Jigyousho no Dynamics (Misallocation and the Dynamics of Establishment). *Financial Review* 112, 180–209 (in Japanese).

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

Hsieh, C. and P. J. Klenow. 2013. Misallocation and Manufacturing TFP in China and India: Correction Appendix. Stanford University. Unpublished.

Kambayashi, R., D. Kawaguchi, and K. Yamada. 2013. Minimum Wage in a Deflationary Economy: the Japanese Experience 1994-2003. *Labour Economics* 24, 264–276.

Kawaguchi, D., R. Kambayashi, Y. G. Kim, H.U. Kwon, S. Shimizutani, K. Fukao, T. Makino, and I. Yokoyama. 2007. Nenkou Chingin wa Seisansei to Kairi Shiteiruka (Do Seniory Wages Differ from Prodcutivity?). *Keizai Kenkyu* (*Economic Research*) 58 (1), 61–90. (in Japanese)

Kawaguchi, D. and Y. Mori. 2009. Is Minimum Wage an Effective Anti-poverty Policy in Japan? *Pacific Economic Review* 14 (4), 532–554.

Kohl, T. 2014. Do We Really Know That Trade Agreements Increase Trade? *Review of World Economics* 150 (3), 443–469.

Krugman, P. 1980. Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review* 70 (5), 950-959.

Kuhn, P. 2004. Is Monopsony the Right Way to Model Labor Markets? A Review of Alan Manning's Monopsony in Motion. *International Journal of the Economics of Business* 11 (3), 369-378.

Kulatilaka, N. and E. C. Perotti 1998. Strategic Growth Options. *Management Science* 44 (8), 1021-1031.

Larch, M., J. Wanner, Y. V. Yotov, and T. Zylkin. 2019. Currency Unions and Trade: A PPML Reassessment with High-dimensional Fixed Effects. *Oxford Bulletin of Economics and Statistics*, 81 (3), 487-510.

Lee, J., J. McCullough, and R. Town. 2013. The Impact of Health Information Technology on Hospital Productivity. *The RAND Journal of Economics* 44 (3), 545–568.

Leon-Ledesma, M. and D. Christopoulos. 2016. Misallocation, Access to Finance, and Public Credit: Firm-level Evidence. *Asian Development Review* 33 (2), 119-143.

Levinsohn, J. and A. Petrin. 2003. Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies* 70 (2), 317–341.

Lu, Y., Y. Sugita, and L. Zhu. 2017. Wage Markdowns and FDI Liberalization, Hitotsubashi University. Unpublished.

Magee, C. S. P. 2008. New Measures of Trade Creation and Trade Diversion. *Journal of International Economics* 75 (2), 349–362.

Manning, A. 2003. Monopsony in Motion: Imperfect Competition in Labor Markets. Princeton University Press.

McDonald, R. and D. Siegel. 1986. The Value of Waiting to Invest. *Quarterly Journal of Economics* 101 (4), 707-728.

McMillan, M. and D. Rodrik. 2011. Globalization, Structural Change and Productivity Growth. In Making Globalization Socially Sustainable, edited by M. Bacchetta and M. Jansen, 49-84. World Trade Organization and International Labour Office.

Meer, J. and J. West. 2016. Effects of the Minimum Wage on Employment Dynamics. *Journal of Human Resources* 51 (2), 500–522.

Melitz, M. J. 2003. The Impact of Intra-industry Trade Reallocations and Aggregate Industry Productivity. *Econometrica* 71 (6), 1695–1725.

Melitz, M. J. and G. I. Ottaviano 2008. Market Size, Trade, and Productivity. *Review of Economic Studies* 75 (1), 295-316.

Melitz, M. J. and S. Polanec 2015. Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit. *The RAND Journal of Economics* 46 (2), 362-375.

Midrigan, V. and D. Y. Xu. 2014. Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review* 104 (2), 422–458.

Mizobata, H. 2014. What Determines the Japanese Firm Investments: Real or Financial? *Applied Economics* 46 (3), 303-311.

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

Neumark, D., J. M. I. Salas, W. Wascher. 2014. Revisiting the Minimum Wage-Employment Debate: Throwing Out the Baby with the Bathwater? *ILR Review* 67 (3), 608–648.

Neumark, D. and W Wascher. 1992. Employment Effects of Minimum and Subminimum Wages: Panel Data on State Minimum Wage Laws. *ILR Review* 46 (1), 55–81.

Nishida, M., A. Petrin, M. Rotemberg, and T. K. White. 2016. Measuring Cross-country Differences in Misallocation. US Census Bureau Center for Economic Studies Paper No. CES-WP-16-50.

Nishida, M., A. Petrin, M. Rotemberg, and T. K. White. 2017. Are We Undercounting

Reallocation's Contribution to Growth? US Census Bureau Center for Economic Studies Paper No. CES-WP-13-55R.

Nishimura, K., T. Nakajima, and K. Kiyota. 2005. Does the Natural Selection Mechanism Still Work in Severe Recessions? – Examination of the Japanese Economy in the 1990s. *Journal of Economic Behavior and Organization* 58 (1), 53–78.

Novy-Marx, R. 2007. An Equilibrium Model of Investment under Uncertainty. *Review of Financial Studies* 20 (5), 1461-1502.

Ogawa, K. and K. Suzuki 2000. Uncertainty and Investment: Some Evidence from the Panel Data of Japanese Manufacturing Firms. *Japanese Economic Review* 51 (2), 170-192.

Oikawa, K. 2010. Uncertainty-Driven Growth. *Journal of Economic Dynamics and Control* 34 (5), 897-912.

Okabe, M. 2015. Impact of Free Trade Agreements on Trade in East Asia. ERIA Discussion Paper Series 2015-1.

Okudaira, H. 2018. The Economic Costs of Court Decisions Concerning Dismissals in Japan: Identification by Judge Transfers. *International Review of Law and Economics* 53, 60–75.

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

Petrin, A., and J. Levinsohn 2012. Measuring Aggregate Productivity Growth Using Plant-Level Data. *The RAND Journal of Economics* 43 (4), 705-725.

Petrin, A. and J. Sivadasan. 2013. Estimating Lost Output from Allocative Inefficiency, with an Application to Chile and Firing Costs. *Review of Economics and Statistics* 95 (1), 286–301.

Pindyck, R. S. 1988. Irreversible Investment, Capacity Choice, and the Value of the Firm. *American Economic Review* 78 (5), 969-985.

Porter, M. E. and A. M. Spence. 1982. The Capacity Expansion Process in a Growing Oligopoly: the Case of Corn Wet Milling. In The Economics of Information and Uncertainty, edited by J. J. McCall, 259-316. University of Chicago Press.

Ransom, M.R. and D. P. Sims. 2010. Estimating the Firm's Labor Supply Curve in a New Monopsony Framework: Schoolteachers in Missouri. *Journal of Labor Economics* 28 (2), 331–355.

Restuccia, D. and R. Rogerson 2008 Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic dynamics* 11 (4), 707-720.

Restuccia, D. and R. Rogerson. 2013. Misallocation and Productivity. *Review of Economic Dynamics* 16 (1), 1–10.

Santos Silva, J. M. C. and S. Tenreyro. 2006. The Log of Gravity. *Review of Economics and*

Statistics 88 (4), 641–658.

Söderbom, M. and F. Teal. 2004. Size and Efficiency in African Manufacturing Firms: Evidence from Firm-Level Panel Data. *Journal of Development Economics* 73 (1), 369–394.

Staiger, D. O., J. Spetz, and C. S. Phibbs. 2010. Is There Monopsony in the Labor Market? Evidence from a Natural Experiment. *Journal of Labor Economics* 28 (2), 211–236.

Syverson, C. 2011. What Determines Productivity? *Journal of Economic Literature* 49 (2), 326–365.

Thompson, P. 2010. Learning by Doing. In Handbook of the Economics of Innovation Vol.1, edited by B. H. Hall and N. Rosenberg, 429-476. Elsevier.

Tran, T. C., X. S. Le, and K. A. Nguyen. 2008. Vietnam's Small and Medium Sized Enterprises Development: Characteristics, Constraints and Policy Recommendations. In SMEs in Asia and Globalization (ERIA Research Project Report 2007-5), edited by H. Lim, 323-364, Economic Research Institute for ASEAN and East Asia.

Waugh, M. E. 2010. International Trade and Income Differences. *American Economic Review* 100 (5), 2093–2124.

Williams, J. T. 1993. Equilibrium and Options on Real Assets. *Review of Financial Studies* 6 (4), 825-850.

Wolfers, J. 2006. Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results. *American Economic Review* 96 (5), 1802–1820.

Wooldridge, J. M. 2009. On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables. *Economics Letters* 104 (3), 112–114.

World Bank. 2014. World Development Indicators. World Bank.

Yamanouchi, K. 2017. Nihon no Jiyuu Boueki Kyoutei (FTA) no Boueki Soushutsu Kouka (The Trade Creation Effects of Japan's Free Trade Agreements). Mitsubishi Economic Research Institute (in Japanese).

Yean, T. S. and A. K. J. Yi. 2014. Re-examining the Impact of ACFTA on ASEAN's Exports of Manufactured Goods to China. *Asian Economic Papers* 13 (3), 63–82.

Yotov, Y. V., R. Piermartini, J.-A. Monteiro, and M. Larch. 2016. An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model. World Trade Organization.

Ziebarth, N. 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.

Zylkin, Thomas. 2016. Beyond Tariffs: Quantifying Heterogeneity in the Effects of Free Trade Agreements. GPN Working Paper Series GPN2016_006.