Do More Productive Firms Locate New Factories in More Productive Locations? An Empirical Analysis Based on Panel Data of Japan's *Census of Manufactures*

Kyoji Fukao
(Hitotsubashi University, RIETI)

YoungGak Kim
(JSPS, Hitotsubashi University)

HyeogUg Kwon
(Nihon University, RIETI)

Paper prepared for SESSION 5b: Location Choice and Firm Performance (WG2)
Comparative Analysis of Enterprise Data Conference
Tokyo
October 2009

Abstract

Using a Melitz-style model of heterogeneous firms, Baldwin and Okubo (2006) recently presented a theoretical model in which self-sorting occurs and more productive factories choose to locate in more productive areas. The model suggests that firm-specific factors and regional factors affect each other through the endogeneity of location decisions. However, to date there are few studies empirically testing this issue. Against this background, our aim is to examine the relationship between firm and location-specific factors in location decisions using factory-level panel data from Japan's *Census of Manufactures*. We begin by estimating how much of the differences in factories' TFP levels can be explained by firm and by location effects. The estimation results show that both effects have a significant impact on the productivity level of a factory, and that the firm effect is more important than the location effect. We also find a statistically significant negative correlation between firm effect and location effect, and investigate what causes this relationship. One potential explanation is that more productive firms may tend to set up new factories in less productive locations such as rural areas, where factor prices such as land prices and wage rates are usually low, in order to benefit from low factor prices. To examine this, we estimate a mixed logit model of location choice. The results indicate that more productive firms indeed tend to set up new factories in low-productivity locations, which is consistent with our hypothesis.

1. Introduction

The literature on productivity shows that there are large differences in the productivity of factories even in a narrowly defined, highly homogeneous industry (see Bartelsman and Doms, 2000). Many researchers on productivity have looked for firm- or factory-specific factors which may be responsible for such productivity differences, such as human capital, capital vintage, and the characteristics of the firm (in terms of R&D, IT, FDI, exports, etc.) to which the factory belongs. On the other hand, many researchers on regional economics have looked for regional factors which may explain the differences in productivity among different areas, such as agglomeration effects due to local industry-specific knowledge spillovers and natural cost advantages (e.g., Ciccone and Hall, 1996, and Henderson, 2003). Both groups of researchers, however, usually fail to incorporate the perspective of the other group. The former does not take into account location factors, while the latter usually does not fully take account of the characteristics of the firm to which a factory belongs.

Using a Melitz-style model of heterogeneous firms, Baldwin and Okubo (2006) recently presented a theoretical model in which self-sorting occurs and more productive factories choose to locate in more productive areas. Their result suggests that firm-specific factors and regional factors affect each other through the endogeneity of location decisions.

Despite the importance of this issue, there are few empirical studies on this topic, probably because of a lack of appropriate data.[1] Against this background, our aim is to examine this issue using factory-level panel data of Japan's *Census of Manufactures*.[2]

First, we decompose factories' TFP levels into firm effects, location effects, and factory-specific characteristics, such as size and age. Next, we investigate the characteristics of our estimated firm effects and location effects by calculating the coefficients of correlation between these effects and several firm- and location-specific characteristics. Based on the estimated firm and location effects, we examine how much of the total variation in TFP levels across factories can be explained by firm effects and by location effects and test whether more productive firms tend to have factories in more productive locations. We also estimate location choice models and test whether more productive firms tend to open up new factories in more productive locations.

The structure of the paper is as follows. The next section introduces our methodology for estimating firm and location effects on TFP, explains data sources and the construction of variables, and presents our estimation results of firm and location effects. In Section 3, we examine the characteristics of the estimated firm and location effects by calculating the coefficients of correlation

[1] Using start-up data of foreign-owned factories in the United States, Shaver and Flyer (2000) have shown that productive foreign-owned firms tend to locate their activities in less productive regions. Shaver and Flyer suggest that the reason is that they receive fewer net benefits from agglomeration and technology spillovers from other firms.

[2] We were able to gain access to the micro data of the *Census of Manufactures* and the *Basic Survey of Japanese Business Structure and Activities* as part of our research project on the "Firm and Industry Level Analysis of Productivity" at the Research Institute of Economy, Trade and Industry (RIETI).

between these effects and various firm and location characteristics. In addition, we conduct an analysis of variance in order to examine the relative importance of the two effects. In Section 4, we then test whether firm and location effects are positively correlated. We also estimate models of firms' location choices. Section 5 summarizes our results and discusses remaining issues to be investigated in future research.

2. Estimation of Firm and Location Effects on Factories' TFP

We begin our examination of the role of location and firm effects on factory-level TFP by presenting our methodology for measuring TFP and isolating the two effects from other, factory-specific effects. This is followed by an explanation of our data sources and the variables used. Finally, this section presents our estimation results of firm and location effects.

Let us start with our methodology for estimating how much of the differences in factories' TFP levels can be explained by firm effects and how much by location effects. The TFP level in year t of factory i, which belongs to firm f and is located in location l, is assumed to be determined by the following factors: factory age and size, firm effects (measured by a dummy for the firm to which the factory belongs), location effects (measured by a location dummy at the city/town/village level), and industry specific effects (measured by a dummy for the industry to which the factory belongs):[3]

$$TFP_{i,t} = G_t\left(Age_{i,t}, Scale_{i,t}, f_{i,t}, l_i, j_{i,t}\right) \tag{1}$$

Specifically, we estimate the following econometric model:

$$\begin{aligned}\ln TFP_{i,t} = {} & \beta_{a1,t} Age_{i,t} + \beta_{a2,t} Age_{i,t}^2 + \beta_{s1,t} Scale_{i,t} + \beta_{s2,t} Scale_{i,t}^2 \\ & + \sum_{l'} \beta_{l',t} DL_{l'}\left(l_i\right) + \sum_{f'} \beta_{f',t} DF_{f'}\left(f_{i,t}\right) + \sum_{j'} \beta_{j',t} DI_{j'}\left(j_{i,t}\right) + R_{i,t}\end{aligned} \tag{2}$$

where $Age_{i,t}$ is the age of factory i in year t, $Scale_{i,t}$ is the number of employees of this factory in year t, l_i is the location of factory i, and $f_{i,t}$ and $j_{i,t}$ denote the firm and the industry to which factory i belongs in year t. $DL_{l'}$ is a dummy variable for location l' which takes value one if the location of factory i, l_i, is equal to l', otherwise this dummy variable takes value zero. Similarly, $DF_{f'}$ is a dummy variable for firm f' which takes value one if the firm to which factory i belongs, $f_{i,t}$, is equal

[3] We use the industry classification of the Japan Industrial Productivity Database (JIP), which subdivides the manufacturing sector into 52 industries.

to *f'*, otherwise this dummy variable takes value zero, and $DI_{j'}$ is a dummy variable for industry *j'* which takes value one if the industry to which factory *i* belongs, $j_{i,t}$ is equal to *j'*, otherwise this dummy variable takes value zero. $R_{i,t}$ is the residual term. The coefficient $\beta_{l',t}$ shows the location effect of location *l'* on factories' TFP level, while the coefficient $\beta_{f',t}$ shows the firm effect of firm *f'* on factories' TFP level.

We classify factories into six subsectors based on factories' main output and calculate the relative TFP level of each factory vis-à-vis the industry average TFP level. Following Good, Nadiri, and Sickles (1997), we measure the TFP level of factory *i* in year *t* in a certain industry in comparison with the TFP level of a hypothetical representative factory in year *t* in that industry using the following equation:

$$\ln TFP_{i,t} = (\ln Q_{i,t} - \overline{\ln Q_t}) - \sum_{n=1}^{N} \frac{1}{2}(s_{n,i,t} + \overline{s_{n,t}})(\ln X_{n,i,t} - \overline{\ln X_{n,t}}) \qquad (3)$$

where $Q_{i,t}$ is the gross output of factory *i* in year *t*, $s_{n,i,t}$ is the cost share of the *n*-th input, and $X_{n,i,t}$ is the amount of the *n*-th input at factory *i* in year *t*. Variables with upper bars denote the arithmetic mean of each variable over all factories in that industry in year *t*. Three inputs, labor, capital, and intermediate input, are taken into account in our analysis.

The main data source for this paper is the longitudinal data of the *Census of Manufactures* conducted by the Ministry of Economy, Trade and Industry (METI) since 1909. This census covers all manufacturing factories with four or more employees.[4] Since 1996, the data include information on factories' affiliation, so that we can group factories according to their parent firms, although the data do not include detailed information on parent firms. We used data for the period of 1996-2003. Input and output deflators and rates of capital depreciation are taken from the Japan Industrial Productivity Database (JIP).

Typically, studies using micro data of the *Census of Manufactures* measure labor input in terms of man-hours, calculated as the product of the number of employees and the industry average of annual working hours. The underlying assumption is that the labor service provided by one hour of work is the same for all workers in a particular industry. However, if the labor service provided by one hour of work differs across regions, our estimates of location effects will be biased. For example, factories in a certain ward of the Tokyo metropolitan area might employ more skilled workers than factories in other areas. If we do not take account of this difference and measure labor input by

[4] Factories with three or fewer employees are included in specific years, starting with the 1981 survey and then in years ending with 0, 3, 5, and 8.

man-hours, we will overestimate the TFP level of factories in this ward.

We deal with this problem by measuring labor input using the total wage bill of a factory and adjusting this for regional wage premiums. Assuming that wage differentials between workers with different characteristics reflect differences in their marginal productivity, the wage bill should provide a good measure of labor input. At the same time, however, given that labor mobility in Japan is not perfect and there exist substantial regional wage premiums, it is important to take such regional wage premiums into account to avoid underestimating labor input in regions where the wage premium is high. Therefore, we adjust the wage bill data using estimation results for regional wage premiums obtained by Daiji Kawaguchi and Ryo Kambayashi based on micro data of the *Basic Survey on Wage Structure* as part of their background analysis for a recent study (Kawaguchi and Kambayashi, 2009).[5] Thus, we measure labor input as the total wage bill divided by the wage premium for the city where the factory is located.[6]

Capital input is measured as real capital stock, which is defined as the product of the nominal book value of tangible fixed assets (taken from the *Census of Manufactures*) and the book-to-market value ratio for each industry, which is calculated using industry level investment data and the book value of industry-level capital stock from the *Census of Manufactures*. Intermediate input is calculated as the sum of the cost of manufacturing and selling, and general and administrative expenses, minus wages and capital depreciation.

We assume that firm- and location-specific effects on productivity are persistent, and use a three-year rolling panel estimation to estimate equation (2). We will refer to estimated coefficients based on observations from 1996 to 1998 as estimates for 1998, those based on observations from 1997 to 1999 as estimates for 1999, etc. Overall, we have sets of estimates for 1998, 1999, 2000, 2001, 2002, and 2003.

In the case of independent factories, we cannot decompose the TFP level into the firm effect and the location effect. Similarly, in the case of locations with just one factory, we cannot decompose the TFP of this factory into the firm effect and the location effect. In the case of locations with only a small number of factories, or in the case of firms that only have a small number of factories, our estimates of firm and location effects are likely to be unreliable. To obtain reliable results, we excluded observations of factories belonging to firms with fewer than two factories or firms without multiple locations (that is, firms whose factories all fall within the same location) in each rolling panel. We also excluded observations of factories falling in locations with fewer than two factories

[5] Kawaguchi and Kambayashi estimated regional wage premiums by estimating a Mincer-type wage function with each worker's educational attainment, work experience (defined as age minus years of education minus 6), tenure, quadratic terms of work experience and tenure, factory size, city dummies, and industry dummies as explanatory variables, using all the survey data for full-time male workers across all industries and for all of Japan.

[6] To check the robustness of our results, we also estimate factories' TFP level measuring labor input using man hours. However, the results remain largely unchanged.

in each rolling panel.[7] We should note that our approach is not free from sample selection bias problems.

We estimate the TFP level of each factory for each year using equation (3). The original data set for this paper consists of 2,590,769 (factory times year) observations. We can calculate the TFP level for 736,947 observations. We cannot derive the TFP level for the other 1,853,822 observations (most of them are for small factories) mainly because of the absence of information on capital stocks.[8] Table 1 shows summary statistics for the measured TFP levels of factories.

Insert Table 1

Next, we estimate firm and location effects using equation (2). In order to take account of the possibility that firm and location effects may differ across industries, we estimated equation (2) separately for the following six manufacturing subsectors: materials, chemicals, general machinery, electric machinery, transportation machinery, and miscellaneous products. Table 2 shows the number of observations used for the estimation and the number of locations and firms for which factories are included in the observations for each year. Because of the data-screening processes explained above, the number of observations used for the estimation is substantially smaller than the total number of factories for which it was possible to calculate TFP levels. Table 3 shows the summary statistics of this estimation. 59,357 firm effects and 36,874 location effects were estimated, and the standard deviations of firm effects and of location effects are almost of the same size.

Insert Tables 2 and 3

3. The Characteristics of Firm and Location Effects and Their Relative Importance

In order to examine the characteristics of the estimated firm and location effects, we calculated the coefficients of correlation between these effects and several firm- and location-specific characteristics. The results are shown in Tables 4 and 5. In Table 4, correlation coefficients are calculated across firms, while in Table 5, correlation coefficients are calculated across locations.

Insert Tables 4 and 5

Starting with the correlation between firm effects and other firm characteristics (Table 4), we

[7] We looked at each rolling panel separately and then determined which observations to eliminate.
[8] This problem is more serious for years after 2000. The reason is that from 2000, the *Census of Manufactures* stopped collecting capital stock data for factories with 29 or fewer employees in non-benchmark years.

find a positive correlation between firm effects and the aggregated TFP level of the firm.[9] Firm effects are also positively correlated with factories' gross output as well as the number of factories a firm owns. That is, larger firms tend to generate larger positive firm effects. These results are consistent with Adams and Jaffe (1996) which found a strong correlation between plant-level productivity and the number of plants of a firm.

One caveat with regard to these results, however, is that by relying on the *Census of Manufactures*, which only provides information on manufacturing establishments, the calculation of firms' TFP level does not cover firms' headquarter activities and non-manufacturing activities. Thus, in order to examine the correlation with the TFP level of firms' total activities, we also calculated firms' TFP level using micro-data from the *Basic Survey of Japanese Business Structure and Activities*. Again we find a positive correlation between these two variables.

Turning to the correlations between location-specific effects and other location characteristics (Table 5), we find that location-specific effect are positively correlated with the average TFP level of all factories in the same location.[10] In addition, using micro-data from the *Census of Manufactures* to calculate average wage rates and taking land prices from the *Chiiki-keizai-deta CD-ROM* (*Regional Economy Data CD-ROM*) published by Toyo Keizai, we find that location effects are also positively correlated with the average wage, the regional wage premium, and the average land price of that location. That is, location effects tend to be greater in areas with higher wage rates and higher land prices.

Next, in order to examine the relationship between location effects and the agglomeration of economic activities, we calculate the "manufacturing gravity" of each location, where this is defined as the distance-discounted aggregated sum of manufacturing activities for each location. To obtain this, we first sum up the manufacturing gross output for each location and year using the micro data of the *Census of Manufactures*, and then calculate the manufacturing gravity of location *l* by aggregating these values of all the locations weighted by the inverses of squares of the distances between location *l* and all the other locations.[11] That is, the manufacturing gravity of location *l* at

---

[9] We calculate the aggregated log value of the TFP of firm *f'* in year *t*, $lnTFP_{f',t}$, as a weighted average of the log value of the TFP level of all the establishments which belong to this firm:

$$\ln TFP_{f',t} = \sum_{f(i)=f'} \frac{sales_{i,t}}{\sum_{f(i')=f'} sales_{i',t}} \ln TFP_{i,t}$$

[10] We calculate the aggregated TFP of location *l'*, $lnTFP_{l',t}$, as a weighted average of the log values of the TFP level of all the factories located in this location:

$$\ln TFP_{l',t} = \sum_{l(i)=l'} \frac{sales_{i,t}}{\sum_{l(i')=l'} sales_{i',t}} \ln TFP_{i,t}$$

where $sales_{i,t}$ denotes the total sales of factory *i* in year *t*.

[11] Probably it would be more appropriate to measure agglomeration effects taking into account all economic activities, not just manufacturing activities. However, unfortunately, information on the output

time *t*, $GDP\ gravity_{l,t}$, is calculated by the following equation:

$$GDP\ gravity_{l,t} = \sum_{l'} \frac{1}{Dist_{l',l}^{\ 2}} \sum_{i} Y_{i,l',t} \tag{4}$$

where $Y_{i,l',t}$ is the gross output of factory *i* in location *l'* at time *t*, and $Dist_{l',l}$ is the distance between locations *l'* and *l*. In the case where *l'* is equal to *l*, the distance is calculated as two thirds of the radian of the hypothetical circular area whose square measure is the same as that of location *l*. We find a positive correlation between location effects and manufacturing gravity in many industries. It seems that the agglomeration of manufacturing activity has a positive effect on factories' productivity.

To sum up our results, both the estimated firm effects and the location effects have plausible characteristics. For example, larger firms tend to generate larger positive firm effects, and firm effects are positively correlated with the average TFP level of all the factories of this firm. Location effects are positively correlated with the average TFP level of all factories in a particular location, and location effects tend to be greater for locations with more manufacturing activity (manufacturing gravity).

The next question we address is how much of the total variation in TFP levels across factories can be explained by firm effects and how much by location effects. In order to answer this question, we conduct an analysis of variance (ANOVA). ANOVA provides a measure of the fit of the regression of the contribution of each variable by measuring how well the variation in each independent variable predicts the variation in the dependent variable.

The ANOVA results are shown in Table 6. In the table, "partial sum of squared deviations" denotes how much the variation of each variable contributes to the total variation (total sum of squared deviations) of the dependent variable, that is, each factory's TFP level. The results show that, in all six manufacturing subsectors, both location and firm effects are important in explaining factories' productivity level. About 40-50 percent of the total variation can be explained by these two effects. The table also shows that in all the manufacturing subsectors, the partial sum of squared deviations of the firm effects is greater than the partial sum of squared deviations of the location effects. Thus, to which firm a factory belongs is a more important determinant of this factory's TFP level than in which location this factory is located.

Insert Table 6

---

of all economic activities at the local level is not available.

4. Do More Productive Firms Set Up Their Factories in More Productive Locations?

In this section, we examine the relationship between firm and location effects. Above, it was suggested that firm effects on factory-level productivity have a positive correlation with firm size and the average TFP level of all the factories of this firm. Similarly, location effects have a positive correlation with the average TFP level of all the factories in a particular location and the location's manufacturing gravity. This raises the question whether "self-sorting" occurs in that more productive firms set up their factories in more productive locations.

However, before conducting our analysis of location decisions, we examine the static correlation between firm and location effects across factories. The results are shown in Table 7 and indicate that there is a statistically significant negative correlation between firm and location effects for all six manufacturing subsectors.

Insert Table 7

What causes this negative relationship? One potential explanation is that more productive firms may tend to set up new factories in less productive locations such as rural areas, where factor prices such as land prices and wage rates are usually low, in order to benefit from low factor costs. On the other hand, less productive firms may be unable to locate new factories in rural areas because of their inability to solve logistical problems, which are common in rural areas.

To determine the cause of the negative relationship, we calculate correlation coefficients across factories between firm effects and location characteristics. We use three variables for location characteristics: manufacturing gravity, wage premiums, and land prices.[12] Table 8 shows the results, which indicate that there is no clear pattern suggesting that more productive firms tend to place their factories in locations with low wage premiums or low land prices. While a negative and significant correlation between firm effects and wage premiums can be observed in the materials industry, the correlations between firm effects and wage premiums or land prices in all other subsectors (as well as that between firm effects and land prices in the materials industry) are either insignificant or actually significantly positive.

Insert Table 8

[12] Examining the correlation between these variables across locations, we find a high positive correlation between wage premiums and land prices, a low but statistically significant positive correlations between manufacturing gravity and land prices, and a low but statistically significant negative correlation between manufacturing gravity and wage premiums. The last result may be due to the fact that manufacturing activities in general tend to be located in regions with low wage premiums.

There is also no clear evidence that more productive firms tend to place their factories in locations with a low manufacturing gravity in order to escape some kind of negative congestion effect, such as spillovers of their technology to rival firms. In the case of the transportation machinery and miscellaneous products industries, the correlations between firm effects and manufacturing gravity are negative and significant, although the correlation coefficients are not large. Moreover, in the other four subsectors, the correlation is either positive and significant (in the case of electric machinery) or very close to zero and not significant.

So far, our analysis has focused on factories of all ages. However, many of the factories in our sample were set up a long time ago and the regional characteristics on which location decisions were originally based may have changed since the establishment of these factories. In order to take this into account, we also examine the correlation between firm effects for newly opened factories and regional characteristics in the year they were established.

To do so, we estimate the following mixed logit model of location choice:

$$\Pr(y_{f,t} = l) = \frac{\exp(\beta' z_{f,l,t})}{\sum_{l'=1}^{L} \exp(\beta' z_{f,l',t})} \qquad for \;\; l = 1...L \; and \; t = 1....T$$

where the left-hand side denotes the probability of observing the establishment of a factory in year t by firm f in location l, $z_{f,l,t}$ denotes a vector including characteristics of location l and cross terms between characteristics of firm f and characteristics of location l as its elements, and β is the coefficient vector.

The number of newly opened factories in our data set and their distribution is shown in Table 9, with panel (a) showing the number of observations in the *Census of Manufactures* overall and panel (b) showing the number for which it is possible to calculate firm and location effects. The numbers are not very largely, partly because Japan's manufacturing sector has shrunk rapidly during this period and, as can be seen in panel (a), the number of newly opened factories has declined in tandem; and partly because, as indicated in panel (b) we can calculate both firm and location effects only for a fraction of newly established factories.

Insert Table 9

The results of our estimation of the mixed logit model of location decisions are presented in Table 10.[13] Specification 1 represents the baseline estimation and as independent variables only

[13] The number of potential locations, that is, the number of choices in location decisions is greater than

includes the location effect and the cross-term of location and firm effects. The estimated coefficients on the location effect are positive in most cases and significantly so in two subsectors (chemicals and transportation machinery). This result provides weak evidence suggesting that firms tend to prefer more productive locations. The estimated coefficients on the cross term of firm and location effects are negative and significant in five of the six subsectors. This result implies that more productive firms tend to be less attracted to more productive locations. This finding is consistent with the results of our analysis of the correlation between firm and location effects across factories.

Insert Table 10

As already mentioned, one possible explanation of the negative correlation between firm and location effects is that more productive firms tend to set up their factories in less productive locations, such as rural areas, so as to exploit the lower factor prices there. Moreover, they can do so because they are able to overcome the logistical problems that may be associated with locating production in rural areas – something that typically smaller less productive firms cannot do. If this hypothesis is correct, high factor prices should have a negative effect on firms' location decision. To test this, we added factor price variables such as the regional wage premium and land prices to the baseline specification. The results are reported as specifications 2 and 3. Against expectation, we do not find any statistically significant negative effects of factor prices on location decisions.[14]

Another question of considerable interest is what causes location effects. As many studies in the field of economic geography have argued, one potential factor is the agglomeration of economic activities. Above, we suggested that there was a positive correlation between our location effect and "manufacturing gravity (Table 5). To test the importance of agglomeration effects for location decisions, we replaced the location effect variable with manufacturing gravity in our analysis of location choice. If more productive firms are less attracted by manufacturing gravity, a negative coefficient for the cross term of firm effects and manufacturing gravity is expected. The results are reported as specifications 4 and 5 in Table 10. We find that manufacturing gravity has a positive and statistically significant effect on location decision in all industries. But we do not find negative and significant coefficient for the cross term of firm effects and manufacturing gravity.

As we already pointed out, the number of newly established factories during our observation period is not very large and in some sectors, such as chemicals and transportation machinery,

1,500 in each subsector.

[14] In some industries, the estimated coefficient for the local wage premium was positive and significant, against our conjecture. One possible explanation for this result is that, because the equation does not include sufficient variables to control for the positive effects of agglomeration, the estimate for the local wage premium picks up these effects.

available observations are less than one hundred (Panel b of Table 9). Therefore, in order to check the robustness of our results, we also estimated our location choice model at a more aggregated industry level, dividing manufacturing industry into only two large sectors, machinery (general machinery, electric machinery, and transportation machinery) and non-machinery (materials, chemicals, and miscellaneous products). Estimating our mixed logit model for these two sectors, we find that the results remain largely unchanged (Table 11). As another robustness check of our results, we calculate factories' TFP level using man-hour data to represent labor input instead of the wage bill divided by the local wage premium. The results of our estimation of location choices based on these TFP values are shown in Table 12 and are similar to the estimation results in Table 10.

Insert Table 12

5. Conclusion

Using micro data of Japan's *Census of Manufactures*, we decomposed the TFP level of each factory into firm effects, location effects, and other, factory specific factors such as size and age. Both the estimated firm effects and the estimated location effects have plausible characteristics. That is, larger firms tend to generate larger positive firm effects. and firm effects are positively correlated with the average TFP level of all the factories of this firm. Location effects are positively correlated with the average TFP level of all factories in a particular location, and location effects tend to be positively correlated with manufacturing gravity.

Based on the estimated firm and location effects, we also conducted an analysis of variance and found that both location and firm effects play a role in explaining factories' TFP levels. In addition, comparing the contribution of the two effects, it was found that both location and firm effects have a statistically significant and large influence on factories' productivity, and firm effects were the more important of the two.

Next, our attention turned to the correlation between firm and location effects, which was negative. That is, more productive firms tended to have factories in cities with weaker location effects. We hypothesized that the reason for this may be that more productive firms set up their factories in less productive locations such as rural areas to benefit from low factor prices there. To test this hypothesis, we estimated a location choice model. Consistent with our hypothesis, the estimation results showed that location effects had a positive effect on location decisions and the cross term of firm and location effects a significant negative effect. However, against our expectation, we did not find a negative and significant effect of factor prices, such as wage premiums and land prices, on location choices.

These findings mean that we need to examine our estimates of firm and location effects further.

We found a positive and significant correlation between location effects and manufacturing gravity, but the correlation coefficient is rather small. When we replaced the location effect variable in our location choice analysis with manufacturing gravity, we found a positive and statistically significant effect on location decisions. However, we did not find a negative and significant coefficient on the cross term, that is, firm effects times manufacturing gravity. These results indicate that we still do not understand well what local characteristics cause positive location effects.

In the case of firm effects, we found a positive and significant correlation between firm size and firm effects, but we were unable to examine how other firm specific factors, such as research and development expenditures, information technology investment, or the accumulation of intangible assets, affect firm effects, because of a lack of such data. We hope that we can investigate these issues in the future by matching our micro data of the *Census of Manufactures* with other firm level data,.

There are also several additional issues that still need to be tackled. First, as we have already highlighted, our approach is not free from sample selection bias problems. And second, it is likely that firms make decisions on the location of new factories in the context of their networks already in place. These are factors that a more refined location choice model should address and we hope to develop and test such a model in the future.

References

Adam, J. D. and A. B. Jaffe (1996) “Bounding the Effects of R&D: An Investigation Using Matched Establish-Firm Data,” *RAND Journal of Economics*, 27, pp.700-721.

Baldwin, R. E. and T. Okubo (2006) “Heterogeneous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting,” *Journal of Economic Geography*, 6(3), pp. 323-346.

Bartelsman, E. J. and M. Doms (2000) “Understanding Productivity: Lessons from Longitudinal Microdata,” *Journal of Economic Literature*, 38, pp.569-594.

Ciccone, A. and R. E. Hall (1996) “Productivity and the Density of Economic Activity,” *American Economic Review*, 86, pp.54-70.

Henderson, J. V. (2003) “Marshall’s Scale Economies,” *Journal of Urban Economics*, 53, pp.1-28.

Good, D. H., M. I. Nadiri and R. C. Sickles (1997) “Index Number and Factor Demand Approaches to the Estimation of Productivity,” in M.H. Pesaran and P. Schmidt (eds.), *Handbook of Applied Econometrics: Vol. 2. Microeconometrics*, Oxford, England: Basil Blackwell, pp. 14-80.

Kawaguchi, D. and R. Kambayashi (2009) “Seifu Tokei no Setsugo Data no Sakusei to Riyo: Kogyo Tokei Chosa to Chingin Kozo Kihon Chosa no Rei [Compilation and Use of Merged Micro-Data of Government Surveys: The Case of the *Census of Manufactures* and the *Basic Survey on Wage Structure*],” forthcoming in Yukinobu Kitamura (ed.), *Ohyo Mikuro Keiryo Keizai-gaku* [*Applied Micro-Econometrics*], Tokyo, Japan: Nihon Hyoron-sha.

Shaver, J. M. and F. Flyer (2000) “Agglomeration Economies, Firm Heterogeneity, and Foreign Direct Investment in the United States,” *Strategic Management Journal*, 21, pp.1175-1193.

Table 1. Summary statistics of the productivity of factories

	Obs.	Mean	SD	Min.	Median	Max.
Materials	218,509	-0.120	0.256	-1.078	-0.142	0.843
Chemicals	24,891	-0.135	0.355	-1.459	-0.173	1.261
General machinery	96,965	-0.094	0.240	-0.990	-0.111	0.790
Electric machinery	83,199	-0.162	0.252	-1.146	-0.181	0.833
Transportation machinery	37,591	-0.098	0.227	-1.015	-0.118	0.827
Miscellaneous products	275,792	-0.145	0.263	-1.130	-0.167	0.837
Total	736,947	-0.130	0.259	-1.459	-0.151	1.261

Table 2. Number of observations used for the estimation and number of locations and firms for which factories are included in the observations

Year		Materials	Chemicals	General machinery	Electric machinery	Transportation machinery	Miscellaneous products
1996	Observations	18,483	3,333	4,818	6,805	2,984	18,256
	Locations	1,416	365	706	910	387	1,615
	Firms	2,975	497	1,050	1,273	521	3,037
1997	Observations	21,830	3,700	6,004	8,182	3,441	22,130
	Locations	1,463	387	748	950	403	1,668
	Firms	3,251	527	1,175	1,378	560	3,357
1998	Observations	22,601	3,789	6,308	8,423	3,596	23,108
	Locations	1,489	404	765	960	420	1,686
	Firms	3,390	545	1,218	1,427	595	3,498
1999	Observations	22,571	3,803	6,306	8,275	3,647	23,059
	Locations	1,484	408	762	971	424	1,693
	Firms	3,465	558	1,252	1,447	605	3,559
2000	Observations	18,607	3,401	5,352	7,137	3,316	19,794
	Locations	1,467	402	747	946	421	1,677
	Firms	3,342	551	1,204	1,388	584	3,457
2001	Observations	14,543	2,978	4,367	5,961	2,950	16,427
	Locations	1,435	387	709	903	407	1,598
	Firms	3,164	540	1,111	1,326	566	3,248
2002	Observations	9,810	2,388	3,139	4,545	2,490	12,521
	Locations	1,155	324	577	774	363	1,364
	Firms	1,921	423	743	976	466	2,365
2003	Observations	8,549	2,160	2,695	3,921	2,338	11,222
	Locations	1,110	308	548	732	341	1,304
	Firms	1,777	414	697	886	446	2,203

Table 3. Summary statistics of the estimation of firm and location effects

(1) Firm effects

	Obs.	Mean	SD	Min.	Median	Max.
Materials	17,946	-0.041	0.247	-1.494	-0.049	1.469
Chemicals	3,436	0.248	0.419	-2.363	0.251	2.269
General machinery	6,504	-0.154	0.418	-2.825	-0.146	1.890
Electric machinery	7,992	-0.118	0.315	-1.900	-0.115	1.591
Transportation machinery	3,665	0.091	0.242	-1.460	0.073	1.229
Miscellaneous products	19,814	-0.219	0.288	-2.315	-0.238	1.224
Total	59,357	-0.098	0.329	-2.825	-0.101	2.269

(2) Location effects

	Obs.	Mean	SD	Min.	Median	Max.
Materials	9,568	-0.032	0.238	-1.674	-0.014	1.248
Chemicals	2,602	0.092	0.381	-1.620	0.089	2.172
General machinery	4,668	0.195	0.677	-2.092	0.000	3.243
Electric machinery	6,104	-0.047	0.314	-1.961	-0.045	2.418
Transportation machinery	2,790	0.044	0.259	-0.969	0.036	1.455
Miscellaneous products	11,142	-0.010	0.257	-1.884	0.000	1.704
Total	36,874	0.016	0.361	-2.092	0.000	3.243

Table 4. Correlation between estimated firm effects and characteristics of firms

Manufacturing

		1		2		3		4		5
1	Estimated firm effects	1								
2	Average TFP	0.48	***	1						
3	Total output	0.05	***	0.03	***	1				
4	Number of factories	0.04	***	0.02	***	0.25	***	1		
5	TFP (Basic Survey)	0.10	***	0.16	***	0.12	***	0.09	***	1

Materials

		1		2		3		4		5
1	Estimated firm effects	1								
2	Average TFP	0.65	***	1						
3	Total output	0.02	***	0.02	***	1				
4	Number of factories	0.07	***	0.04	***	0.32	***	1		
5	TFP (Basic Survey)	0.10	***	0.12	***	0.18	***	0.11	***	1

Chemicals

		1		2		3		4		5
1	Estimated firm effects	1								
2	Average TFP	0.43	***	1						
3	Total output	0.09	***	0.16	***	1				
4	Number of factories	0.01		0.02		0.12	***	1		
5	TFP (Basic Survey)	0.12	***	0.19	***	0.17	***	0.05	***	1

General machinery

		1		2		3		4		5
1	Estimated firm effects	1								
2	Average TFP	0.40	***	1						
3	Total output	0.02		0.02	***	1				
4	Number of factories	-0.01		0.04	***	0.37	***	1		
5	TFP (Basic Survey)	0.03	**	0.14	***	0.17	***	0.13	***	1

Electric machinery

		1		2		3		4		5
1	Estimated firm effects	1								
2	Average TFP	0.39	***	1						
3	Total output	0.03	***	0.04	***	1				
4	Number of factories	0.01		0.01		0.59	***	1		
5	TFP (Basic Survey)	0.06	***	0.15	***	0.17	***	0.07	***	1

Transportation machinery

		1		2		3		4		5
1	Estimated firm effects	1								
2	Average TFP	0.42	***	1						
3	Total output	-0.02		0.02	**	1				
4	Number of factories	-0.01		-0.01		0.39	***	1		
5	TFP (Basic Survey)	0.10	***	0.09	***	0.17	***	0.23	***	1

Miscellaneous products

		1		2		3		4		5
1	Estimated firm effects	1								
2	Average TFP	0.54	***	1						
3	Total output	0.04	***	0.04	***	1				
4	Number of factories	0.01		0.02	***	0.46	***	1		
5	TFP (Basic Survey)	0.10	***	0.21	***	0.13	***	0.08	***	1

Note: "Average TFP" is the average TFP level of all factories of a particular firm.

Table 5. Correlation between estimated location effects and characteristics of locations

Manufacturing

	1	2	3	4	5	6
1 Estimated location effects	1					
2 Average TFP	0.29 ***	1.00				
3 Average wage	0.09 ***	0.06 ***	1.00			
4 Wage premium	0.15 ***	0.23 ***	0.56 ***	1.00		
5 Manufacturing gravity	0.08 ***	0.08 ***	0.55 ***	0.46 ***	1.00	
6 Land prices	0.02 ***	0.08 ***	0.58 ***	0.53 ***	0.69 ***	1.00

Materials

	1	2	3	4	5	6
1 Estimated location effects	1					
2 Average TFP	0.28 ***	1.00				
3 Average wage	0.09 ***	0.02 ***	1.00			
4 Wage premium	0.17 ***	0.24 ***	0.54 ***	1.00		
5 Manufacturing gravity	0.06 ***	0.05 ***	0.52 ***	0.44 ***	1.00	
6 Land prices	0.08 ***	0.04 ***	0.58 ***	0.52 ***	0.69 ***	1.00

Chemicals

	1	2	3	4	5	6
1 Estimated location effects	1					
2 Average TFP	0.31 ***	1.00				
3 Average wage	0.00	0.04 ***	1.00			
4 Wage premium	0.05 **	0.09 ***	0.56 ***	1.00		
5 Manufacturing gravity	0.04	0.04 ***	0.50 ***	0.42 ***	1.00	
6 Land prices	-0.11 ***	0.03 **	0.52 ***	0.54 ***	0.65 ***	1.00

General machinery

	1	2	3	4	5	6
1 Estimated location effects	1					
2 Average TFP	0.18 ***	1.00				
3 Average wage	0.02	0.05 ***	1.00			
4 Wage premium	0.01	0.23 ***	0.56 ***	1.00		
5 Manufacturing gravity	0.01	0.07 ***	0.55 ***	0.46 ***	1.00	
6 Land prices	-0.08 ***	0.09 ***	0.56 ***	0.54 ***	0.68 ***	1.00

Electric machinery

	1	2	3	4	5	6
1 Estimated location effects	1					
2 Average TFP	0.32 ***	1.00				
3 Average wage	0.08 ***	0.09 ***	1.00			
4 Wage premium	0.11 ***	0.28 ***	0.57 ***	1.00		
5 Manufacturing gravity	0.05 ***	0.14 ***	0.59 ***	0.50 ***	1.00	
6 Land prices	0.02	0.14 ***	0.58 ***	0.54 ***	0.69 ***	1.00

Transportation machinery

	1	2	3	4	5	6
1 Estimated location effects	1					
2 Average TFP	0.31 ***	1.00				
3 Average wage	-0.11 ***	-0.02 **	1.00			
4 Wage premium	-0.01	0.13 ***	0.58 ***	1.00		
5 Manufacturing gravity	-0.02	0.02 **	0.58 ***	0.46 ***	1.00	
6 Land prices	-0.15 ***	0.04 ***	0.55 ***	0.51 ***	0.66 ***	1.00

Miscellaneous products

	1	2	3	4	5	6
1 Estimated location effects	1					
2 Average TFP	0.35 ***	1.00				
3 Average wage	0.20 ***	0.13 ***	1.00			
4 Wage premium	0.31 ***	0.35 ***	0.53 ***	1.00		
5 Manufacturing gravity	0.19 ***	0.13 ***	0.50 ***	0.44 ***	1.00	
6 Land prices	0.15 ***	0.15 ***	0.57 ***	0.52 ***	0.69 ***	1.00

Note: "Average TFP" is the average TFP level of all factories in a particular location.

Table 6. Results of the analysis of variance

	Materials			Chemicals			General machinery			Electric machinery			Transportation machinery			Miscellaneous products		
	Partial sum of squared deviations	#	F	Partial sum of squared deviations	#	F	Partial sum of squared deviations	#	F	Partial sum of squared deviations	#	F	Partial sum of squared deviations	#	F	Partial sum of squared deviations	#	F
Location	402.8 (11%)	1750	5.8	342.3 (23%)	641	9.5	231.1 (20%)	1052	5.8	281.9 (18%)	1337	5.0	119.2 (22%)	628	6.8	548.0 (12%)	1941	7.0
Firm	1144.4 (32%)	3553	8.1	413.5 (28%)	601	12.2	363.2 (31%)	1329	7.3	394.6 (25%)	1569	5.9	133.1 (24%)	644	7.4	1476.3 (33%)	3740	9.8
Age	2.4 (0%)	1	58.8	1.2 (0%)	1	22.1	0.9 (0%)	1	22.9	0.6 (0%)	1	14.7	0.0 (0%)	1	1.7	2.0 (0%)	1	49.7
Age^2	2.0 (0%)	1	50.9	0.9 (0%)	1	16.3	0.7 (0%)	1	19.2	0.8 (0%)	1	18.1	0.0 (0%)	1	1.5	1.5 (0%)	1	36.2
Size	0.1 (0%)	1	2.5	1.0 (0%)	1	18.1	0.3 (0%)	1	8.8	0.1 (0%)	1	2.0	0.1 (0%)	1	4.4	0.3 (0%)	1	7.6
$Size^2$	0.0 (0%)	1	0.0	1.2 (0%)	1	20.9	0.5 (0%)	1	13.0	0.0 (0%)	1	1.0	0.0 (0%)	1	1.6	0.6 (0%)	1	14.3
Industry	6.2 (0%)	15	10.3	6.3 (0%)	8	13.9	0.2 (0%)	4	1.1	3.3 (0%)	7	11.0	1.5 (0%)	2	26.3	6.5 (0%)	10	16.1
Total	3537.6			1462.2			1170.4			1590.8			550.5			4426.3		
Observations	48,088			10,447			15,984			21,289			10,218			52,932		
R^2	0.516			0.646			0.563			0.509			0.544			0.571		
Adj. R^2	0.456			0.597			0.486			0.431			0.478			0.520		

Note: The industry classification is based on the JIP (Japan Industrial Productivity) Database.

Table 7. Correlation between firm effects and location effects

Industry	Corr. Coef.
Materials	-0.341
Chemicals	-0.548
General machinery	-0.685
Electric machinery	-0.603
Transportation machinery	-0.554
Miscellaneous products	-0.361
Total	-0.355

Note: All coefficients are statistically significant at the 0.1% level.

Table 8. Correlation between estimated firm effects and characteristics of locations

Manufacturing

		1		2		3		4
1	Estimated firm effects	1						
2	Manufacturing gravity	0.00		1				
3	Wage premium	0.07	***	-0.02	***	1		
4	Land prices	0.00		0.03	***	0.41	***	1

Materials

		1		2		3		4
1	Estimated firm effects	1						
2	Manufacturing gravity	0.00		1				
3	Wage premium	-0.04	***	-0.03	***	1		
4	Land prices	-0.01		0.04	***	0.40	***	1

Chemicals

		1		2		3		4
1	Estimated firm effects	1						
2	Manufacturing gravity	0.00		1				
3	Wage premium	0.05	***	-0.08	***	1		
4	Land prices	0.11	***	0.01		0.36	***	1

General machinery

		1		2		3		4
1	Estimated firm effects	1						
2	Manufacturing gravity	0.01		1				
3	Wage premium	0.06	***	-0.05	***	1		
4	Land prices	0.05	***	0.02	**	0.46	***	1

Electric machinery

		1		2		3		4
1	Estimated firm effects	1						
2	Manufacturing gravity	0.04	***	1				
3	Wage premium	0.06	***	0.05	***	1		
4	Land prices	-0.01		0.10	***	0.48	***	1

Transportation machinery

		1		2		3		4
1	Estimated firm effects	1						
2	Manufacturing gravity	-0.05	***	1				
3	Wage premium	0.02	**	-0.04	***	1		
4	Land prices	0.03	**	0.50	***	0.32	***	1

Miscellaneous products

		1		2		3		4
1	Estimated firm effects	1						
2	Manufacturing gravity	-0.03	***	1				
3	Wage premium	0.00		-0.01		1		
4	Land prices	0.02	***	0.02	**	0.44	***	1

Table 9. Number of newly established factories in our data

a. Number of new factories in the micro data of the *Census of Manufactures*

	1997	1998	1999	2000	2001	2002	2003	Total
Materials	1,300	445	467	533	383	385	379	3,892
Chemicals	163	66	58	79	61	80	56	563
General machinery	473	172	198	210	164	189	165	1,571
Electric machinery	457	197	210	221	201	178	148	1,612
Transportation machinery	194	91	80	103	102	93	105	768
Miscellaneous products	1,631	506	543	574	453	460	383	4,550
Total	4,218	1,477	1,556	1,720	1,364	1,385	1,236	12,956

b. Number of new factories for which both firm and location effects can be calculated

	1997	1998	1999	2000	2001	2002	2003	Total
Materials	169	89	87	118	43	49	44	599
Chemicals	27	18	5	20	10	9	5	94
General machinery	29	15	20	15	7	15	7	108
Electric machinery	66	24	18	32	19	22	23	204
Transportation machinery	18	6	8	14	11	9	21	87
Miscellaneous products	234	67	87	106	73	78	40	685
Total	543	219	225	305	163	182	140	1,777

Table 10. Location choice: Estimation results of mixed logit models

	Materials	Chemicals	General machinery	Electric machinery	Transportation machinery	Miscellaneous products
Specification 1						
Location effect	0.143	0.579 *	0.032	-0.161	1.36 ***	0.214
	(0.186)	(0.334)	(0.262)	(0.244)	(0.478)	(0.175)
Firm effect × Location effect	-2.07 ***	-0.935	-2.11 ***	-3.78 ***	-7.38 ***	-1.55 ***
	(0.781)	(0.793)	(0.611)	(0.770)	(1.731)	(0.401)
R-squared	0.001	0.003	0.009	0.009	0.017	0.003
Log-likelihood	-4.30E+03	-5.50E+02	-6.92E+02	-1.35E+03	-5.00E+02	-5.00E+03
Observations	798,635	33,372	69,991	167,322	30,450	1,036,334
Specification 2						
Location effect	-0.259	0.643 *	0.0438	-0.245	1.4 ***	-0.211
	(0.201)	(0.346)	(0.269)	(0.251)	(0.480)	(0.189)
Firm effect × Location effect	-2.13 ***	-1.2	-2.14 ***	-3.84 ***	-7.63 ***	-1.85 ***
	(0.809)	(0.815)	(0.627)	(0.790)	(1.790)	(0.427)
Regional wage premium	2.360 ***	-0.190	1.090	1.590 ***	2.130 **	2.470 ***
	(0.293)	(1.037)	(0.769)	(0.487)	(0.952)	(0.260)
R-squared	0.009	0.003	0.01	0.013	0.022	0.012
Log-likelihood	-4.06E+03	-5.06E+02	-6.47E+02	-1.29E+03	-4.83E+02	-4.85E+03
Observations	736,467	30,020	62,820	154,471	28,664	982,522
Specification 3						
Location effect	-0.0196	0.744 *	0.0524	-0.371	1.71 **	0.141
	(0.267)	(0.415)	(0.343)	(0.363)	(0.677)	(0.276)
Firm effect × Location effect	-2.04 *	-0.913	-3.14 ***	-4.65 ***	-8.89 ***	-1.49 *
	(1.160)	(0.913)	(0.851)	(1.414)	(2.326)	(0.777)
Land prices	0.218 ***	0.019	0.074	0.374 ***	0.038	0.395 ***
	(0.055)	(0.164)	(0.131)	(0.103)	(0.186)	(0.046)
R-squared	0.004	0.005	0.018	0.021	0.027	0.016
Log-likelihood	-2.12E+03	-2.67E+02	-3.87E+02	-5.51E+02	-2.55E+02	-2.58E+03
Observations	208,054	10,800	24,389	36,061	9,669	268,647
Specification 4						
Manufacturing gravity	0.159 ***	0.16 ***	0.166 ***	0.155 ***	0.188 ***	0.165 ***
	(0.010)	(0.028)	(0.023)	(0.018)	(0.025)	(0.012)
Firm effect × Manufacturing gravity	-0.00375	-0.0138	0.00903	0.031	0.0125	0.0161
	(0.044)	(0.074)	(0.064)	(0.063)	(0.087)	(0.035)
R-squared	0.013	0.012	0.015	0.011	0.027	0.014
Log-likelihood	-4.62E+03	-1.03E+03	-9.71E+02	-1.95E+03	-7.96E+02	-5.27E+03
Observations	1,670,404	371,516	351,664	703,328	292,108	1,905,792
Specification 5						
Manufacturing gravity	0.131 ***	0.129 ***	0.144 ***	0.134 ***	0.166 ***	0.143 ***
	(0.014)	(0.038)	(0.029)	(0.022)	(0.032)	(0.015)
Firm effect × Manufacturing gravity	0.0000144	0.00262	0.00948	0.0356	0.0198	0.0251
	(0.058)	(0.097)	(0.080)	(0.076)	(0.109)	(0.044)
Regional wage premium	2.360 ***	2.590 ***	1.970 ***	1.400 ***	3.810 ***	2.080 ***
	(0.247)	(0.526)	(0.536)	(0.377)	(0.577)	(0.232)
R-squared	0.022	0.024	0.02	0.013	0.052	0.02
Log-likelihood	-4.41E+03	-9.74E+02	-9.32E+02	-1.88E+03	-7.61E+02	-4.99E+03
Observations	1,401,650	310,400	295,850	591,700	249,775	1,585,950

Notes: 1. Figures in parentheses are standard errors.

2. * p<.1, ** p<.05, *** and p<.01

3. Labor input is adjusted by the regional wage premium when calculating the TFP index of each establishment

Table 11. Location choice: Estimation results of mixed logit models

	General machinery, electric machinery, and transportation machinery	Materials, chemicals, and miscellaneous products
Specification 6		
Location effect	0.0796	0.884 ***
	(0.219)	(0.135)
Firm effect × Location effect	-3.3 ***	-1.38 **
	(0.933)	(0.559)
R-squared	0.002	0.002
Log-likelihood	-4.20E+03	-1.25E+04
Observations	763,367	3,226,893
Specification 7		
Location effect	-0.15	0.263 *
	(0.232)	(0.151)
Firm effect × Location effect	-3.45 ***	-1.57 ***
	(0.996)	(0.593)
Regional wage premium	2.120 ***	2.650 ***
	(0.294)	(0.160)
R-squared	0.009	0.013
Log-likelihood	-4.00E+03	-1.19E+04
Observations	708,127	3,017,627
Specification 8		
Location effect	0.154	0.633 ***
	(0.328)	(0.220)
Firm effect × Location effect	-4.38 ***	-1.08
	(1.586)	(0.902)
Land prices	0.284 ***	0.359 ***
	(0.056)	(0.030)
R-squared	0.009	0.013
Log-likelihood	-2.04E+03	-6.10E+03
Observations	198,050	761,979
Specification 9		
Manufacturing gravity	0.169 ***	0.159 ***
	(0.012)	(0.007)
Firm effect × Manufacturing gravity	0.0304	0.0146
	(0.051)	(0.029)
R-squared	0.015	0.014
Log-likelihood	-4.50E+03	-1.18E+04
Observations	1,630,700	4,279,524
Specification 10		
Manufacturing gravity	0.148 ***	0.132 ***
	(0.014)	(0.010)
Firm effects × Manufacturing gravity	0.0413	0.0222
	(0.064)	(0.038)
Regional wage premium	2.080 ***	2.320 ***
	(0.249)	(0.155)
R-squared	0.021	0.022
Log-likelihood	-4.33E+03	-1.12E+04
Observations	1,377,400	3,574,450

Notes: 1. Figures in parentheses are standard errors.
2. * p<.1, ** p<.05, *** and p<.01

Table 12. Location choice: Estimation results of mixed logit models, TFP is based on man-hour labor input data

	Materials	Chemicals	General machinery	Electric machinery	Transportation machinery	Miscellaneous products
Specification 11						
Location effect	0.0464	0.285	-0.197	-0.157	0.306	0.381 ***
	(0.197)	(0.295)	(0.250)	(0.279)	(0.411)	(0.146)
Firm effect × Location effect	-1.46 **	-0.736	-1.92 ***	-2.09 ***	-5.35 ***	-0.883 **
	(0.697)	(0.778)	(0.526)	(0.556)	(1.660)	(0.442)
R-squared	0.001	0.001	0.008	0.007	0.009	0.001
Log-likelihood	-4.48E+03	-5.93E+02	-7.35E+02	-1.46E+03	-5.01E+02	-5.27E+03
Observations	867,274	36,353	76,993	188,016	31,311	1,146,733
Specification 12						
Location effect	-0.137	0.273	-0.153	-0.215	0.406	0.259
	(0.213)	(0.313)	(0.261)	(0.291)	(0.423)	(0.163)
Firm effect × Location effect	-1.36 *	-0.887	-1.86 ***	-2.17 ***	-5.68 ***	-1.05 **
	(0.747)	(0.819)	(0.542)	(0.595)	(1.703)	(0.489)
Regional wage premium	2.220 ***	0.013	1.070	1.510 ***	1.990 **	2.450 ***
	(0.283)	(1.012)	(0.748)	(0.469)	(0.960)	(0.250)
R-squared	0.008	0.001	0.009	0.011	0.015	0.011
Log-likelihood	-4.16E+03	-5.36E+02	-6.75E+02	-1.36E+03	-4.82E+02	-5.02E+03
Observations	756,434	31,822	66,707	165,193	28,778	1,024,024
Specification 13						
Location effect	0.0644	0.358	-0.0522	-0.157	0.753	0.257
	(0.290)	(0.386)	(0.337)	(0.422)	(0.600)	(0.237)
Firm effect × Location effect	-1.2	-0.357	-2.31 ***	-2.45 **	-6.7 ***	-0.327
	(1.045)	(0.940)	(0.733)	(1.020)	(2.142)	(0.718)
Land prices	0.213 ***	-0.024	0.058	0.339 ***	0.113	0.408 ***
	(0.054)	(0.157)	(0.127)	(0.101)	(0.190)	(0.045)
R-squared	0.004	0.002	0.012	0.017	0.018	0.016
Log-likelihood	-2.17E+03	-2.89E+02	-4.08E+02	-5.72E+02	-2.54E+02	-2.67E+03
Observations	218,175	11,652	26,128	38,088	9,649	285,892
Specification 14						
Manufacturing gravity	0.157 ***	0.157 ***	0.167 ***	0.155 ***	0.189 ***	0.159 ***
	(0.011)	(0.022)	(0.022)	(0.022)	(0.020)	(0.010)
Firm effect × Manufacturing gravity	-0.0119	-0.012	0.0239	0.0175	0.0195	0.0318
	(0.040)	(0.066)	(0.058)	(0.049)	(0.083)	(0.030)
R-squared	0.013	0.012	0.014	0.01	0.027	0.014
Log-likelihood	-4.74E+03	-1.08E+03	-1.01E+03	-2.01E+03	-7.89E+02	-5.46E+03
Observations	1,712,944	388,532	365,844	726,016	289,272	1,976,692
Specification 15						
Manufacturing gravity	0.129 ***	0.128 ***	0.144 ***	0.134 ***	0.166 ***	0.135 ***
	(0.015)	(0.030)	(0.028)	(0.027)	(0.026)	(0.013)
Firm effect × Manufacturing gravity	-0.00887	-0.0163	0.0394	0.016	0.0305	0.0477
	(0.053)	(0.086)	(0.072)	(0.060)	(0.102)	(0.037)
Regional wage premium	2.360 ***	2.620 ***	1.960 ***	1.390 ***	3.770 ***	1.990 ***
	(0.245)	(0.514)	(0.526)	(0.371)	(0.581)	(0.228)
R-squared	0.022	0.024	0.02	0.013	0.051	0.02
Log-likelihood	-4.50E+03	-1.02E+03	-9.70E+02	-1.94E+03	-7.54E+02	-5.18E+03
Observations	1,430,750	324,950	307,975	611,100	247,350	1,644,150

Notes: 1. Figures in parentheses are standard errors.

2. * p<.1, ** p<.05, *** and p<.01

3. The standard Tornqvist index is used for calculating the TFP index of each factory.