Journal of Asian Business and Economic Studies
Volume 26, Special Issue 01 (2019), 04–28

JABES
Journal of Asian Business and Economic Studies

www.jabes.ueh.edu.vn

Journal of Asian Business and Economic Studies

Determinants of firms' total factor productivity in manufacturing industry in Vietnam: An approach of a cross-classified model

NGUYEN THI HOANG OANH [*]

School of Economics, University of Economics Ho Chi Minh City

ARTICLE INFO

Received: 3 Oct, 2018
Revised: 27 Dec, 2018
Accepted: 15 Jan, 2019

Available online

JEL classification:

O1, O25, O4

Keywords:

Total factor productivity;
Cross-classified multilevel models;
Localized industry.

ABSTRACT

This study investigates the determinants of total factor productivity in manufacturing firms in Vietnam using the cross-classified multilevel model. This model enables the study to provide a more proper estimation and to make clear distinctions between firms, region-specific effects, and sector-specific effects. This study combined a data set of Vietnamese manufacturing firms and sectoral variables gathered from the annual data of the Vietnam Enterprises Survey, Technology Competitiveness Survey, and some regional variables from the General Statistics Office's Province Competitive Index during the period from 2011 to 2014. The study found that the main source of firm total factor productivity heterogeneity mostly originates at the firm level. In addition, the interaction between regional (provincial) and sectoral factors also contribute considerably to total factor productivity heterogeneity among firms. At the firm level, both firm size and expenses on technology have a significant positive effect on firm total factor productivity. In addition, firms with exporting activities seem to have a higher total factor productivity. At the regional level, the provinces with a high ratio of well-trained employees may have a positive impact on firm total factor productivity in that province. At the sectoral level, the concentration of sectors in a province may benefit firms belonging to that sector in that province. More interestingly, the study also indicates that the

* Corresponding author.
Email: hoangoanh@ueh.edu.vn (Nguyen Thi Hoang Oanh).

Please cite this article as: Nguyen, T. H. O. (2019). Determinants of firms' total factor productivity in manufacturing industry in Vietnam: An approach of a cross-classified model. *Journal of Asian Business and Economic Studies, 26*(S01), 04–26.

concentration of sectors in a province may benefit firms located in the provinces with a ratio of better trained employees. These findings could lead to policies not only at the firm level but also at the regional level and sectoral level to enhance total factor productivity.

1. Introduction

Total factor productivity (TFP) plays a key role in sustainable development in Vietnam. TFP is the residual of output that is not explained in the amount of capital and labor in the production process. Theoretically, the economy hardly reaches sustainable development if it only depends on capital and labor. In the Solow model (1956), the residual is a black box representing technical change that leads to sustainable development. Practically, the development of some countries, such as South Korea or Singapore, came along with TFP growth. As stipulated by Central Institute for Economic Management (CIEM) (2010), TFP growth in South Korea in the period from 1970 to 1980 was only at 8.3%, but in the next period from 1980 to 1990, this figure reached 31.5%. Therefore, ways to enhance TFP is a principal issue in development policy.

In order to find valuable policy implications on development, it is necessary to study the determinants of TFP. The most important factor in TFP is technological progress that leads to sustained growth (Solow, 1956). However, according to Acemoglu (2009), the heterogeneity in TFP is not necessarily due to technology in the narrow sense. For instance, two firms may have adopted the same technology, but they make use of these techniques in different ways with different degrees of efficiency. These differences are considered in TFP heterogeneity. Cepeda and Ramos (2015) suggested that TFP is a composite of several different elements, such as economies of scale and improved ways of combining resources not only at the level of machines or processes but also through minor adjustments at the level of the factory. However, even if these firms adopt similar technology, they will still have differences in TFP. These differences may originate from the characteristics of their sectors or their locations.

Obviously, the heterogeneity in firm TFP originates primarily from the differences in their characteristics. Differences in size, production technology, human capital, and other firm characteristics may lead to differences in performance. However, firms' performance may be affected by the external economics of scale of their sector. According to Krugman and Obstfeld (2009), the concentration of sectors in a region may lead to positive externalities, such as specialized suppliers, labor market pooling, and knowledge spillovers. Firms are also affected by the environment where they are located. Regions may offer a quantity and quality of various endowments that are alternatively beneficial for firms. Krugman and Obstfeld (2009) notes that external economies support sectors that are localized. Firms tend to be located within a short distance of each other to reap the benefits of external economies. This is confirmed by the existence of several localized industry such as California's Silicon Valley, New York financial sectors.

It is necessary to investigate the determinants of firm TFP with a multilevel cross-classified model. This model enables us to disentangle the effects exerted by firm-specific factors, location, and sectors. Using multilevel equations, the model makes use of the data structures and properly addresses the issue of error correlation across firms that operate in the same region and in the same sector. In addition, by distinguishing between sample sizes at the different levels, the model limits the high risk of type I errors. In single-equation models, the variance is usually underestimated as these models use the entire sample size without differentiating the levels, whereas the variance in any level is correctly estimated in the multilevel cross-classified model. Another advantage of this model is to allow a wide variety of correlation patterns (or variance-covariance structures) by letting each group of observations have its own intercept (possibly slope) randomly deviating from the mean intercept or slope of each group (Seltman, 2015). However, most studies on firm TFP have focused on the determinants as the firms' characteristics (Sjoholm, 1999; Blalock &Veloso, 2007; Waldkirch & Ofosu, 2010; Lopez, 2008; Baptist & Teal, 2014; Fernandes, 2008; Seker & Saliola, 2018). In Vietnam, studies on TFP are still very limited (CIEM, 2010), although TFP is perceived as a key role of development quality (Tran Tho Dat & Do Tuyet Nhung, 2013).

This study makes a contribution as a new approach in investigating TFP in Vietnam by applying the multilevel cross-classified model. With this model, the study explores the separate contributions of firms' characteristics, regional factors, and sectoral factors to heterogeneity to determine firm TFP. The study found that the main source of firm TFP heterogeneity mostly originates at the firm level. More interestingly, the interaction between regional (provincial) and sectoral factors also contribute considerably to TFP heterogeneity among firms. Specifically, 55.95% of the variance in firm TFP is due to firms' characteristics, while the proportion of TFP heterogeneity by the interaction between region and sector is 36.62%.

Additionally, the study may lead to policy suggestions not only for firms but also for regions and sectors. In theory, the concentration of firms in the same sector may bring benefits to the firms. However, the findings on this theory have been inconsistent. This study found that the sector with the highest percentage of firms belonging to industrial parks in a province may have higher TFP. Moreover, the concentration of sectors in a province may be more beneficial to firms located in the provinces that have better trained employees. This finding could be considered as evidence of the Marshall Theory (a part of Marshall - Arrow - Romer), as cited in Beaudry and Schiffauerova (2009). This finding may also lead to policies that encourage firms in manufacturing to join industrial parks. Industrial parks should be established in provinces with the highest percentage of well-trained employees.

This paper has the following structure: Section 2 summarizes a literature review on TFP. Section 3 illustrates TFP measurement and a specific multilevel cross-classified model. Section 4 provides the model's results, and Section 5 makes conclusions.

2. Literature Review

2.1. Total factor productivity (TFP)

Total factor productivity (TFP) identifies the portion of output not explained by the traditionally measured inputs of capital and labor. It is widely known that output is a function of the inputs used by a firm and its productivity (Katayama et al., 2009). This productivity plays an important role in sustainable development as resources become scarcer. Therefore, the measure of TFP as the residual in the production function is necessary for policy implications. Basically, the following Cobb-Douglas production function is used to measure TFP:

$$Y_{it} = A_{it} K_{it}^{\alpha} L_{it}^{\beta} \tag{2.1}$$

where Y_{it} is output of firm I at period t, and A_{it} , K_{it}^{α}, L_{it}^{β} are respectively TFP, capital stock, and labor. Taking a logarithm of equation (2.1), we have:

$$lnY_{it} = lnA_{it} + \beta_k lnK_{it} + \beta_l lnL_{it} \tag{2.2}$$

Supposing $lnA_{it} = \beta_0 + \varepsilon_{it}$ *and* $y_{it} = lnY_{it}, k_{it} = lnK_{it}, l_{it} = lnL_{it}$, then we have:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \varepsilon_{it} \tag{2.3}$$

where β_0 is the mean efficiency level across firms and overtime, and ε_{it} is the time and firm specific deviation from the mean. This can be decomposed into an observable with at least a predictable and unobservable component as follows:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + v_{it} + u_{it}^{q} \tag{2.4}$$

where $\omega_{it} = \beta_0 + v_{it}$ represents firm-level productivity, and u_{it}^{q} represents unexpected deviations from the mean due to measurement error, unexpected delays, or other external circumstances.

Typically, firm productivity could be estimated as follows:

$$\widehat{\omega}_{it} = \hat{v}_{it} + \hat{\beta}_0 = y_{it} - \hat{\beta}_k k_{it} - \hat{\beta}_l l_{it} \tag{2.5}$$

Productivity in levels will be the exponential of $\widehat{\omega}_{it}$, and the productivity measurement depends on equation (2.3).

2.2. Literature review and empirical studies

At the firm level, the most important factor affecting TFP is the firm's technology. In decomposing the components of firm TFP in the United States, Solow (1956) found that a firm's TFP is mostly affected by its technology compared with its capital and labor. Technology is understood as the production method that originates from the firm's knowledge, which may be found by the firm itself through investment in research and development (R&D). The firm may also absorb knowledge from the outside due to knowledge externalities. According to Cohen and Levinthal (1989), the knowledge a firm can absorb depends on its absorptive capacity, which can be enhanced from R&D activities.

In addition, technology can be transferred directly to firms through commercial channels. Most firms in developing countries access technology by purchasing it rather than through R&D investment. Another determinant of TFP is firm size based on the theory of internal economies of scale. This theory confirms that the more goods the firm produces, the less marginal cost the firm bears. The firm can also gain more experience by increasing production, which improves efficiency. This motivates the study to investigate the impact of expenses on technology and firm size on firm productivity.

Another level of concern in this study is the regional level, which refers to factors that are imposed on firms as part of the environment in which firms are located. According to Acemoglu (2009), regional factors can have an impact on firms through a variety of proximate causes. It is widely known that human capital is important both for increasing productivity and adopting technology. Human capital depends on the quality of firms' employees in their labor market. It may be argued that the quality of the training of employees in the region will have an impact on the human capital of firms in that region.

In addition, it is important to consider geographical boundaries in most analyses of sectoral systems (Malerba, 2002). According to him, a sectoral system is highly localized and it frequently defines the specialization of the whole local area, such as in the case of machinery, some traditional industries, and even information technology. Whether this specialization enhances firm productivity is a valuable question for policymakers. The Marshall-Arrow-Romer model, as cited in Beaudry and Schiffauerova (2009), is a well-known model that describes the role of a concentration of firms in similar industries on economic growth. Based on this model, this study tests the hypothesis that firms located in a highly localized sector may have higher TFP than firms in non-localized sectors.

Moreover, economic institutions and the business environment of the region may also impact firms. Consider that institutions are the rules of the game in a society that shapes human interaction. More specifically, economic institutions comprise such things as the structure of property rights, the presence and functioning of markets, and contractual opportunities available to individuals and firms. Acemoglu (2009) devised a theoretical analysis to determine the good economic institutions that encourage factor accumulation and the development and adoption of better technologies.

Obviously, firm TFP, considered as the black box in Solow's model, is very complicated to investigate. It is necessary to explore the determinants of TFP with a multilevel model since TFP is concerned with not only firms' characteristics but also regional and sectoral factors. According to Aiello et al. (2015), it is better to understand the heterogeneity in firm productivity by considering at least three key levels of analysis—firm-specific, location, and sectors—with a cross-classified model in their recent research. Also applying the cross-classified model, this study is more specific than their model due to the interaction between regional and sectoral factors. The interaction between some regional characteristics and sectoral innovation explores what conditions in the region may enhance the spillover effects of sectoral innovation on firm productivity. This new factor enables the study to be innovative and have valuable implications for regional policy.

Several studies investigated TFP at the firm level, sector, and country level. Most studies that investigated TFP at the firm level focused on determinants as firm characteristics. The determinants in these studies are usually exports, imports (Sjoholm 1999; Blalock & Veloso, 2007); foreign presence or foreign technology license (Waldkirch & Ofosu, 2010; Lopez, 2008); technology (Paptist & Teal, 2014), and firm size (Fernandes, 2008). Seker and Saliola (2018) recently found that the determinants of heterogeneity in TFP across countries includes firms' exporting, innovation, access to finance, foreign ownership, and countries' regulations.

While there are several studies on the determinants of TFP, most have approaches based on firm-level characteristics. TFP considers the contribution of factors besides capital and labor on firm output. These factors are so complex that they should be investigated not only at the firm level but also at the regional and sectoral levels. Until now, the number of studies on TFP at a multilevel approach is still limited.

Aiello et al. (2015) is the most recent prominent study on TFP using the cross-classified model. That study combined a data-set of Italian firms with some regional and sectoral variables. By applying the cross-classified model, this study found a clear distinction between firm, region- specific effects, and sector-specific effects. It also found that firms in more innovative sectors may have higher TFP. Moreover, sectors with a high proportion of firms using R&D, public support, and a high propensity to collaborate in innovative projects shall provide benefits for firm productivity.

3. Research Methodology

3.1. The choice of TFP measurement method

This study makes use of the Levinsohn and Petrin (LP) (2003) estimation (see the Appendix). This estimation uses the proxy of intermediate inputs rather than investment to solve the issue of simultaneity bias due to several advantages of intermediate input. First, the data in this study have mostly zero-investment observations that could not satisfy the monotonicity condition. Second, no Olley and Pakes (OP) (1996) proxy is available for these observations. Finally, according to Levinsohn and Petrin (2003), intermediate inputs are easier to verify whether or not the monotonicity condition is consistent with some common technologies used by economics. This means that the sign of the change in intermediate input used for a small change in ω is always positive, as follows:

Sign $(\frac{\partial \iota}{\partial \omega})$ = sign $(f_{\iota l} f_{l\omega} - f_{ll}\, f_{\iota\omega})$, where f_{ll} is the second derivative of f (.) with respect to l.

Under the monotonicity condition, optimizing behavior implies that the marginal product declines as labor increases, so $f_{ll} < 0$ for chosen input bundles. If increases in productivity always make the marginal product of inputs go up weakly, then $f_{\iota\omega} \geq 0$ and $f_{l\omega} \geq 0$, so $- f_{ll}\, f_{\iota\omega} \geq 0$.

Among intermediate inputs, this study uses electricity as the proxy variable. According to Levinsohn and Petrin (2003), counting zero values of a proxy is one natural way to start evaluating its potential usefulness. In our data, electricity has the highest fraction of non-zero observations. Moreover, the price of electricity is mostly stable, which enables it to be a more exact element in estimating the movement of electricity consumption during the period. Finally, electricity could be measured with less error due to the inability of storing electricity for long periods. This could avoid measurement problems if inputs are stored period-to-period and changes in inventories of inputs are not directly observed.

Particularly, this study applies the Wooldridge (2009) estimation, which has more advantages due to using a generalized method of moments (GMM) approach. Unlike two-step procedures in the Levinsohn and Petrin (2003) or Olley and Pakes (1996), GMM estimation does not ignore the potential correlation between the errors in the two steps. In addition, GMM estimation enables testing of the identification assumptions underlying the model and allows for calculating standard errors of input coefficients. Finally, the Wooldridge estimation can account for the first-stage identification problem in the Levinsohn and Petrin (2003) and Olley and Pakes (1996) methods. For robustness, this study applies the Wooldridge estimation.

3.2. The Research Model of Cross-classified Model on Firms' TFP

Data on firm, industry, and location are a prominent example of cross-classified multilevel data. The unit of analysis herein is the firm (level 1), and the higher hierarchy is the industry or the location. In this case, firms can be nested into their industries, but firms are also able to be grouped into their locations (provinces). This data cannot be considered as three-level data because industry is hardly nested into locations or otherwise. A province could have firms in several industries; similarly, firms in any industry could be located in any province. This relationship is demonstrated in Figure 1.

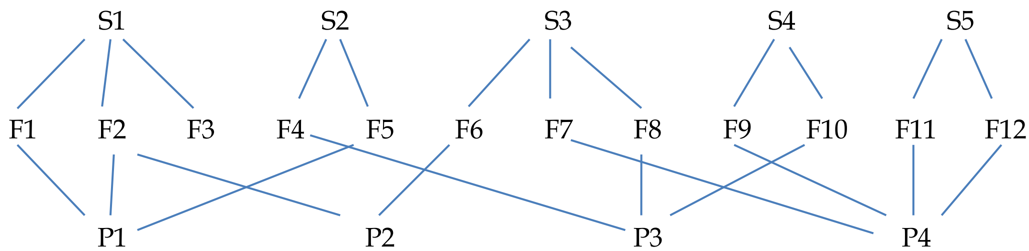

Figure 1. Example on 12 firms at Level 1 nested within a Sector and Province Cross-Classified Multilevel Model at Level 2

Notes: S indicates Sectors, P indicates Provinces, F indicates firms

3.2.1. Model Specification

In order to investigate the separate effects of firm characteristics, region, and sector, this simple model without any independent variables is estimated first as follows:

$$y_{i(sj)} = \gamma_{000} + u_s + u_j + u_{sj} + e_{i(sj)} \tag{3.1}$$

where $y_{i(sj)}$ is the TFP of the firm in sector s and located in region j, γ_{000} is the mean TFP across all sectors and all regions, u_s is the effect of firm i's sector, u_j is the effect of firm i's region, and $e_{i(sj)}$ is the firm-level residual error term. This model includes a random interaction effect between region and sector, u_{sj} . This interaction confirms the different effects that a sector has on firms in different regions, even after accounting for a region's main effects. All of the above effects are assumed independent and normally distributed with zero means and constant variances.

The estimation result of equation (3.1) provides how much firm characteristics, region, and sector are attributed to the heterogeneity of firm TFP. The cross-classified model enables interpretation of the relative magnitude of the variance components by computing variance partition coefficients (VPCs). VPC statistics report the proportion of the response variance that lies at each level of the model's hierarchy. Therefore, the regional VPC_j is calculated as the ratio of the regional variance to the total variance as follows:

$$\mathrm{VPC_j} = \frac{\sigma_j^2}{\sigma_j^2 + \sigma_s^2 + \sigma_{js}^2 + \sigma_e^2}, \tag{3.2}$$

similarly, the sectoral $VPC_s = \frac{\sigma_s^2}{\sigma_j^2 + \sigma_s^2 + \sigma_{js}^2 + \sigma_e^2}$, (3.3)

the interaction between regional and sectoral $VPC_{js} = \frac{\sigma_{js}^2}{\sigma_j^2 + \sigma_s^2 + \sigma_{js}^2 + \sigma_e^2}$, and (3.4)

the firm $VPC_i = \frac{\sigma_e^2}{\sigma_j^2 + \sigma_s^2 + \sigma_{js}^2 + \sigma_e^2}$. (3.5)

Based on this result, this study considers adding a number of variables onto a firm's characteristics, region, and sector specifically into the following model:

$$y_{i(sj)} = \gamma_{000} + \sum_{f=1}^{m} \gamma_f X_{fi(sj)} + \sum_{h=1}^{n} \beta_h . Z_{hij} + \sum_{p=1}^{k} \lambda_p . S_{ps} + \alpha Z_{ij} . S_{is} + u_s + u_j + u_{sj} + e_{i(sj)}, \tag{3.6}$$

where y is the TFP of the ith firm (in logs) operating in sector s and located in region j, X is a vector of m firm-level variables that are considered to be important drivers of TFP, Z presents the variables at the regional level, and S are the variables at the sectoral level. This study considers the result estimation of equation (3.1) to include the number of variables in each investigated level.

This study investigates the effects of expenses on technology transfer in firms and the role of sectoral concentration on firm productivity. It also adds firm size as a characteristic to investigate the effects of economies of scale on TFP. Regional-level variable Z includes the quality of training of employees in the province. The model (3.6) has basic factors as in the other studies on firm productivity that use a cross-classified model (Aiello et al., 2015). This model also enables this study to be exceptional and to potentially provide valuable policy implications using the interaction factor between regional-level variables Z_{ij} and sectoral-level variables S_{is}. According to Castellacci (2008), a lot of attention is paid to the context of innovation, especially regional factors, in evolution theory. In addition, Malerba et al. (2013) suggested that joint analysis of both the regional level and sectoral level is very necessary.

This study combined a data set of Vietnamese manufacturing firms and sectoral variables gathered from the annual data of the Vietnam Enterprises Survey (VES), Technology Competitiveness Survey (TCS), and some regional variables from the General Statistics Office's (GSO) Province Competitive Index during the period from 2011 to 2014. The data description is shown in Table 1

Table 1.

Description of variables in the model of the second objective

Variable	Definitions	Calculations	Source
Firm level	*All in 2015*		
lnTFP	Productivity	Wooldridge method	VES
LnK	Capital stock of firm	Logarithm of total fixed assets	VES
lnL	Labor of firm	Logarithm of total labor	VES
Techexpense	Expenses on technology	Total costs of all machinery/ technology in billion dong	TCS
Size	Firms' size	Logarithm of total assets	VES
Ownership	The firm's ownership	State companies Wholly foreign-owned companies Other foreign-owned companies Other companies	VES
Export	The firm's export status	= 1 if firms have export activities; otherwise = 0	VES
Regional level			
Training	The percentage of trained employees over the untrained employees in the province	Measured by GSO	GSO
Sectoral level			
Concent_Province_IO	Sectoral concentration in a province	The proportion of firms in the same sector belonging to an industrial park in their province	VES

As seen in Table 1, the dependence of the model is lnTFP measured by LP and the Wooldridge estimation. The technical expenses are calculated by the amount of money a firm spends purchasing technology or equipment. The firm's size is the logarithm of the total assets. The ownership variable classifies firms into four types: firms with State equity, firms with 100% foreign equity, firms with partial foreign equity, and other firms. The export variable divides firms into firms with export activities and firms without export activities. The variable Concent_Province_IO defines the percentage of firms in a sector belonging to an industrial park in a province to the total number of firms in that sector and in that province. Regarding regional-level variables, the study collects annual data from the GSO's PCI. The quality of training of employees is the one component in measuring PCI.

Hypothesis testing

As suggested by Solow (1956), technology change was considered to be a main component of firm TFP. This study investigates the role of firm's expenses on technology as a factor of technology change on firm TFP using the following hypothesis:

H_{21}: A firm's expenses on technology has a positive effect on firm productivity.

Originating from the MAR model that confirms the advantages of a cluster of firms in the same sector, this study tests the following hypothesis:

H_{22}: Firms may have higher productivity when located in sectors with a higher percentage of firms belonging to industrial parks in a province.

As proposed in the theoretical analysis of Acemoglu (2009), good economic institutions may encourage factor accumulation and the development and adoption of better technologies. The study investigates the role of well-trained employees in the province on firm productivity by using the following hypothesis:

H_{23}: The percentage of well-trained employees in the province may have a positive effect on firm productivity in that province.

This study also tries to discover what conditions a province should have to enhance the effect of the sector's concentration level on firm productivity using the following hypothesis:

H_{24}: The high percentage of well-trained employees in a province could enhance the effect of the sector's concentration level in that province on firm productivity.

3.2.2. Data

The investigation period of the second objective in this study covers 2011 to 2014. In addition to the VES data, this study also makes use of the data on provinces from the GSO to determine the quality of training in a province. The combination of balance panel data left the number of observations at 1,648 enterprises per year. Over four years, the data show a total number of observations of 6,592. These data cover 38 sectors (see Table A3 in the Appendix) of firms located in 52 provinces (see Table A4 in the Appendix).

4. Model Results and Analysis

4.1. Overview of TFP in Vietnamese manufacturing firms

4.1.1. The sample

The sample in this study comes from the annual survey on Vietnamese Enterprises from 2011 to 2014. This sample only covers firms in manufacturing industries, classified into 38 sectors and located in 52 provinces. The number of provinces was reduced by merging the data through the years in the study period. Table 2 shows the distribution of firms by area, economic activity, and size.

Table 2.

Descriptive statistics of firms in the sample (2011–2014)

			lnTFP			
	Number of firms	% of firms	Mean	Standard Deviation	Min	Max
By territorial area						
Northwest	20	0.3	4.64	0.60	3.22	5.46
Northeast	92	1.4	4.78	0.84	2.12	6.63
Red River Delta	1,720	26.1	4.92	0.69	2.44	7.85
North Central	252	3.8	4.67	0.64	2.64	6.81
South Central	296	4.5	4.70	0.58	2.69	6.38
Central Highlands	32	0.5	4.72	0.68	3.55	6.78
Southeast	3,816	57.9	5.34	0.66	0.14	8.89
Southwest	364	5.5	4.84	1.00	1.85	9.04
By Pavitt Sector[1]						
Supplier dominated	2,188	33.19	5.00	0.62	2.13	9.04
Scale-intensive producer	3,176	48.18	5.14	0.75	0.66	7.91
Science-based	772	11.71	5.53	0.86	0.14	8.88
Specialized equipment supplier	456	6.92	5.12	0.67	2.71	6.86
By Size (Labor)						
Small (<50)	2,353	35.69	4.90	0.75	0.14	9.04
Medium (50-250)	2,858	43.36	5.22	0.70	0.66	8.85
Large (250)	1,381	20.95	5.36	0.65	2.41	8.88
Total	6,592	100	5.14	0.73	0.14	9.04

[1] Following the classification of Pavitt (1984)

As shown in Table 2, firm distribution by area during this period reveals a predominance of enterprises located in the East South region of Vietnam with more than two-thirds of the sample. This study separates sectors into four groups as in Pavitt (1984). Among the groups of sectors (see Appendix A), the distribution shows a concentration of firms in scale-intensive sectors and supplier dominated sectors. Table 2 also indicates that firms in the sample were mainly small (35.69%) and medium (40.36%) enterprises.

The average TFP logarithm in the sample is 5.14, with a high deviation from 0.73 to 9.04. By territorial area, the average TFP logarithm in the South East region is the highest at 5.34. However, the difference in TFP logarithm among all the regions is relatively small. Meanwhile, concerning economic activities, the science-based sector has the highest average TFP logarithm at 5.27. Firm size confirms the role of economies of scale as the large firms have the highest average TFP logarithm.

Table 2 also shows that the high differences in TFP are confirmed at the sectoral level. On average, the best performing firms are in the science-based sector with 5.53, while the lowest performing firms are in the supplier-dominated sectors with only 5.00.

Table 3.

The number of provinces in which a sector is located

Number of Provinces	No. of sectors	Percent
1	1	2.63
2	1	2.63
3	3	7.89
4	3	7.89
6	4	10.53
7	4	10.53
8	5	13.16
10	3	7.89
11	2	5.26
12	2	5.26
13	1	2.63
14	2	5.26
15	3	7.89
16	2	5.26
19	1	2.63
22	1	2.63
Total	38	100

Table 3 presents the distribution of sectors throughout the provinces. Most sectors are located in eight provinces with 13.16% of the total number of sectors. There are also sectors located in several provinces; one sector is located in 22 provinces, and one sector is located in 19 provinces.

Table 4.

The number of sectors in a province

Number of sectors	No. of Provinces	Percent
1	14	26.92
2	8	15.38
3	6	11.54
4	3	5.77
5	1	1.92
6	3	5.77
7	2	3.85
8	1	1.92
9	1	1.92
10	3	5.77
11	2	3.85
13	1	1.92
14	3	5.77
25	1	1.92
33	1	1.92
34	1	1.92
38	1	1.92
Total	52	100

As shown in the Table 4, most of the provinces have a low number of sectors. Twelve provinces have only one sector, and eight provinces have two sectors. Only one province had firms in 38 sectors.

4.1.2. TFP by region and sector

This section verifies whether firm TFP differs across geographical areas and economic sectors in the period from 2011 to 2014 (Figure 2 and Figure 3).

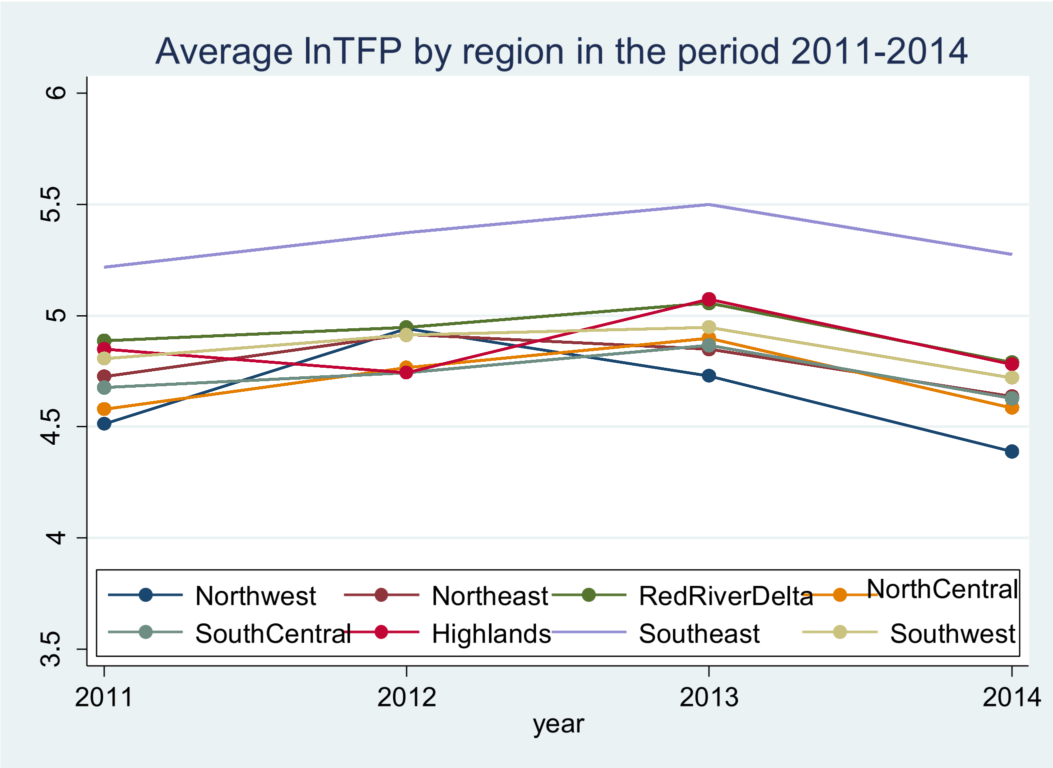

Figure 2. TFP by region from 2011 to 2014

Figure 2 presents the changes in TFP from 2011 to 2014 among the regions. As shown in the figure, TFP in the Southeast region steadily increased while TFP in other regions fluctuated. Among these regions, the Southwest region has the most stable TFP. On average, TFP in the Southeast was always the highest during the period. This reflects the region as being the most developed in the country.

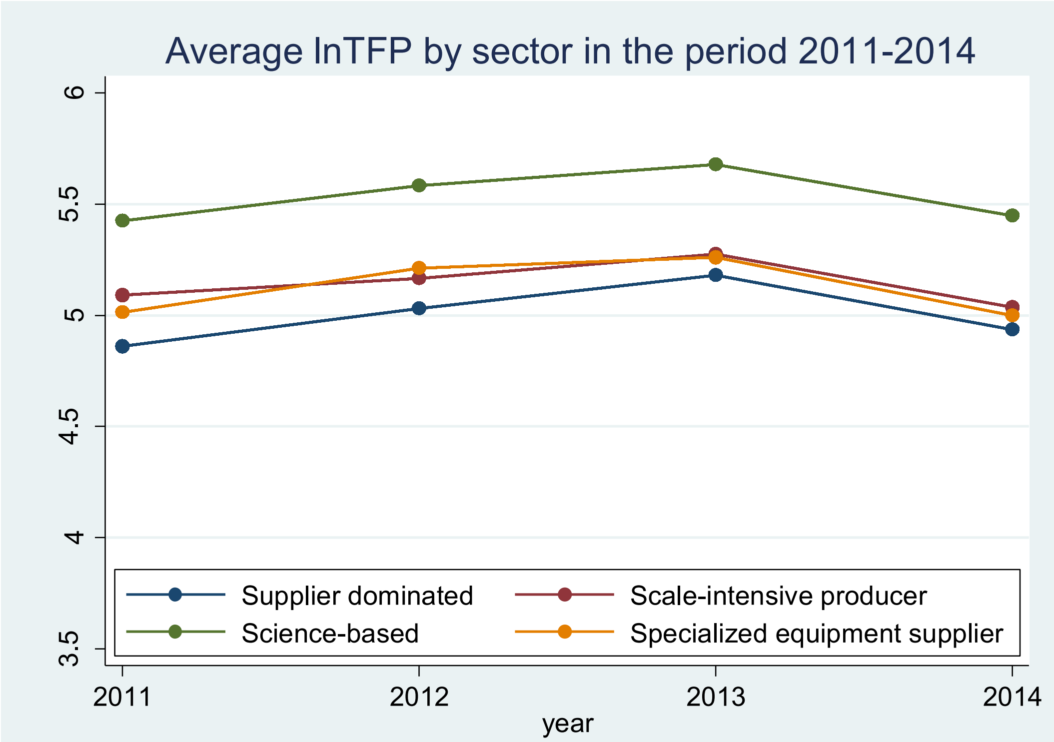

Figure 3. TFP by Pavitt Sector in the 2011-2014

Figure 3 describes TFP heterogeneity among Pavitt sectors. TFP across the Pavitt sectors fluctuated less than across the other regions. The average TFP in the science-based sectors was always higher than the other sectors. The supplier-dominated sectors had the lowest average TFP at the beginning of the period. However, TFP in these sectors was escaping the scale-intensive producer and specialized equipment supplier sectors at the end of the period.

4.1.3. Model results

This section provides the results of the empty cross-classified model from equation (3.1) and the full cross-classified model from equation (3.6) in testing the stipulated hypotheses.

Table 5.

TFP heterogeneity at the firm level: Estimation from the empty model

Model		(1)	(2)	(3)
Constant		4.82***	5.15***	4.86***
		(0.05)	(0.04)	(0.05)
Variance				
	Regions	0.13		0.07
		(0.03)		
	Sectors		0.06	$3.02 * 10^{-24}$
			(0.01)	
	Region * IO			0.2
	Firms	0.45	0.47	0.343
		(0.007)	(0.008)	
Variance Partition Coefficient (VPC)				
	Regions	21.05%		11.41%
	Sectors		11.11%	0.00
	Region*Sector			32.62%
	Firms	78.95%	88.88%	55.95%
LR Test		1,074.16	706.6	2,322.88
Log likelihood		–7,138.7	–7,322.5	–6,514.4
No. region sector				351
No. of groups		52	38	52
No. of firms		6,592	6,592	6,592
Min		4	8	
Max		1,728	756	
Average		126.8	173.5	

Table 5 represents the VPC values attributed to the different sources of variability. The calculations show that 11.41% of the unexplained variation in TFP lies at the regional level and 32.62% with the interaction of sector and region. Meanwhile, the sectoral level primarily did not explain any variability in firm TFP. The remaining variability (55.95%) is explained using firm characteristics, as seen in column (3). With region alone incorporated, seen in column (1), this factor explains a relatively higher percentage of variance (21.05%). A similar thing happens when only sector is considered, as seen in column (2). This result indicates that when only one of the two levels of analysis is investigated in the model as a random effect, it will draw to itself part of the other random effect (Aiello et al., 2015). These improper random disturbance structures may lead to the conclusion that explanatory variable effects are statistically significant, and they will now exist in a correct, cross-classified model (Fielding et al., 2004). LR tests that comparing the cross-classified model (column 3) to a simpler two-level model of firms within-regions (column 1) and two-level firms within-sectors (column 2) (see Table A1 and Table A2 in the Appendix) have confirmed that the cross-classified model offers a significantly better fit to the data. The results indicate that firm characteristics have the most influence on firm TFP and the second most influence is the interaction between regional and sectoral factors.

Table 6.

TFP of Vietnamese manufacturing firms from 2011 to 2014: Multilevel regressions

	(1)	(2)	(3)
Fixed effects			
Constant	2.87***	2.907***	2.98***
	(0.07)	(0.07)	(0.068)
Firm characteristics			
Size	0.177***	0.172***	0.172***
	(0.005)	(0.005)	(0.005)
Expense on technology	0.001***	0.001***	0.001***
	(0.0003)	(0.0003)	(0.0003)
Exports	0.18***	0.16***	0.16***
State		0.018	0.018
Wholly foreign-owned		0.1***	0.1***
Other foreign-owned		0.385***	0.384***
Sectoral characteristics			
Concentration of sector in a province	0.006***	0.006***	
	(0.0008)	(0.0008)	
Regional factor			
The percentage of trained employees over the	0.007*	0.007***	
	(0.004)	(0.004)	
Training*Concentration			0.0008***
			(0.000)

	(1)	(2)	(3)
Random effects			
Variance			
Regions	0.028	0.025	0.0026
Sectors	$2.51*10^{-22}$	$2.28*10^{-22}$	$8.67*10^{-23}$
Region*Sector	0.122	0.123	0.119
Firms	0.266	0.263	0.263
Log Likelihood	–4,911.87	–4,878.45	–4,874.81

Table 6 represents the estimation results of factors at the firm level, regional level, and sectoral level in equation (6). The residuals of this regression have been demonstrated in Figure A1 in the Appendix to reflect the non-existence of heteroscedasticity. Regarding firm characteristics as explanatory factors of TFP, both firm size and expenses on technology have a significant positive effect on TFP in all different model strategies. The larger firms may have higher productivity due to economies of scale. It may also be argued that large firms could have more efficient production by better accessing technology, learning, and dealing with uncertainty and selection processes. This finding is in line with Seker and Saliola (2018), Aiello et al. (2016). Expenses on technology also have a significant positive impact on TFP. This confirms the role of technology in Solow's model as well as in endogenous growth theory. Baptist and Teal (2014) or Mastromarco and Zago (2012) also found the similar findings. However, the effect of this technology channel is sustainably small. The firms should consider the other channels on enhancing firms' technical capacity.

In addition, by controlling the export status of firms, the study found a significantly higher TFP in exporting enterprises than in non-exporting enterprises. This finding is similar to the one from Blalock (2007). This confirms that exporting activities may generate positive externalities on firms to enhance their TFP. In comparison with other firms, state-owned, wholly foreign-owned, or other foreign-owned firms all have higher TFP. Firms with foreign capital may have significantly higher TFP. In contrast, Takii (2004) found that wholly foreign-owned plants tend to have higher productivity than other foreign-owned plants. He confirmed that the relationship between foreign ownership and productivity may be different among industries.

Regarding sectoral characteristics, the concentration of sectors in a province may have a significantly positive effect on firm TFP in that sector. This finding proves the MAR of industrial concentration. Firms may benefit from this concentration due to the advantages of the labor market pool, the input market, and knowledge spillover. This finding is also in line with Beaudry and Schiffauerove (2009) and Lee (2018). With respect to regional characteristics, the percentage of well-trained employees in a province may have a significant impact on firm TFP in that region. This finding confirms the role of the environment in which firms do business. Moreover, the study found that firms in a highly-localized sector located in a province with more well-trained employees may have a significantly higher TFP than others firms.

5. Conclusion

This study found valuable findings for policy implication on enhancing firms' TFP by the multileveled cross-classified model. Firstly, the study confirms the positive effect of firm size and expenses on technology on firms' TFP. The large firms shall have more advantages in reaching higher TFP due to economies of scales. Expenses on technology also have positive but small impact on firms' TFP. This implies that besides purchasing technology or equipment, firms should make advantages of some others channel of technology spillovers. Secondly, the concentration of sectors in a province may also have positive effect of firms' TFP. The policy makers should consider the policies to enhance the industrial cluster. Finally, firms' TFP also are positively affected by the higher percentage of trained employees in their provinces. In addition, these regional conditions may enhance the positive effect of industrial cluster on firms' TFP. It draws the attention of policymakers in building those conditions in provinces for TFP growth■

References

Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. Princeton and Oxford: Princeton University Press.

Aiello, F., Pupo, V., & Ricotta, F. (2015). Firm heterogeneity in TFP, sectoral innovation and location. Evidence from Italy. *International Review of Applied Economics*, *29*(5), 579–607.

Baptist, S., & Teal, F. (2014). Technology and productivity in African manufacturing firms. *World Development*, *64*, 713–725.

Beaudry, C., & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, *38*(2), 318–337.

Blalock, G., & Veloso, F. M. (2007). Imports, productivity growth, and supply chain learning. *World Development*, *35*(7), 1134–1151.

Castellacci, F. (2008). Innovation and the competitiveness of industries: Comparing the mainstream and the evolutionary approaches. *Technological Forecasting and Social Change*, *75*(7), 984–1006.

Cepeda, T. L. E., & Ramos, L. F. C. (2015). Patterns of TFP growth in Mexico: 1991-2011. *North American Journal of Economics and Finance*, *34*, 398–420.

Central Institute for Economic Management (CIEM). (2010). Nâng cao tỷ trọng và tác dụng của năng suất nhân tố tổng hợp. Truy cập từ www.beton6.com/Data/Sites/1/media/Nang%20suat.pdf

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: the two faces of R & D. *The Economic Journal*, *99*(397), 569–596.

Fernandes, A. M. (2008). Firm Productivity in Bangladesh Manufacturing Industries. *World Development*, *10*(36), 1725–1744.

Greene, W. H. (2008). *Econometric Analysis* (6th ed.). Upper Saddle Rive, New Jersey: Prentice-Hall.

Katayama, H., Lu, S., & Tybout, J. R. (2009). Firm-level productivity studies: illusions and a solution. *International Journal of Industrial Organization*, *27*(3), 403–413.

Krugman, P. R., & Obstfeld, M. (2009). *International Economics: Theory and Policy*. Boston: Pearson Addison-Wesley.

Lee, C. Y. (2018). Geographical clustering and firm growth: Differential growth performance among clustered firms. *Research Policy*, *47*(6), 1173–1184.

Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, *70*(2), 317–341.

Lopez, R. A. (2008). Foreign technology licensing, productivity, and spillovers. *World Development*, *36*(4), 560–574.

Malerba, F. (2002). Sectoral systems of innovation and production. *Research Policy*, *31*(2), 247–264.

Malerba, F., Mancusi, M. L., & Montobbio, F. (2013). Innovation, international R&D spillovers and the sectoral heterogeneity of knowledge flows. *Review of World Economics*, *149*(4), 697–722.

Mastromarco, C., & Zago, A. (2012). On modeling the determinants of TFP growth. *Structural Change and Economic Dynamics*, *23*(4), 373–382.

Olley, G.S., & Pakes. A. (1996). The dynamics of productivity in the Telecommunications equipment industry. *Econometrica*, 64(6), 1263-1297.

Pavitt, K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy*, *13*(6), 343–373.

Seker, M., & Saliola, F. (2018). A cross-country analysis of total factor productivity using micro-level data. *Central Bank Review*, *18*(1), 13–27.

Seltman, H.J., (2015). *Experimental Design and Analysis*. Pittsburgh: Carnegie Mellon University. Retrieved from http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf

Sjoholm, F. (1999). Exports, imports and productivity: Results from Indonesian establishment data. *World Development, 27*(4), 705–715.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics, 70*(1), 65–94.

Takii, S. (2004). Productivity differentials between local and foreign plants in Indonesian manufacturing, 1995. *World Development, 32*(11), 1957–1969.

Trần Thọ Đạt & Đỗ Tuyết Nhung. (2013). Vai trò của TFP trong chất lượng tăng trưởng kinh tế ở Việt Nam. *Tạp chí Khoa học Công nghệ Việt Nam, 13*, 12–14.

Waldkirch, A., & Ofosu, A. (2010). Foreign presence, spillovers, and productivity: Evidence from Ghana. *World Development, 38*(8), 1114–1126.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservable. *Economics Letters, 104*(3), 112–114.

Appendix

A1. TFP measurement

The choice of measurement methods on this study based on the comparison of four principal methods including fixed effects, Instrumental variables and GMM, the semi parametric estimation algorithm developed by Olley and Pakes (1996) and the semi parametric estimation algorithm developed by Levinsohn and Petrin (2003).

Fixed effects estimation could be applied when assuming that ω_{it} is firm specific and time-invariant. The estimation (4) becomes:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_i + u_{it}^q \tag{a1}$$

The above equation can be estimated by Least Square Dummy Variable Estimator. This estimation may generate consistent coefficients on labor and capital provided that unobserved productivity ω_{it} does not change over time. This assumption is too strict (Wooldridge, 2009) and this usually result in unreasonably low estimates of the capital coefficient.

The alternative method to estimate the consistent coefficients in the production function is using instrumental variables for the endogenous variables. These instrument variables must satisfy three following requirements. First, instrumental variables must be high correlated with the endogenous covariates. Second, they do not come into the production function directly. Finally, the instrumental variables do not allow to be correlated with the error term (Greene, 2008).

The more popular methods to measure TFP are OP estimation and LP estimation. Both of these methods solves the simultaneity issue by using the proxy for unobserved productivity shocks. In OP estimation, investment decisions depend on capital and productivity. It means that

$$i_t = i_t\ (k_t, \omega_t) \tag{a2}$$

This relation enables to express unobserved productivity as a function of observables:

$$\omega_t = \omega_t\ (k_t, i_t) \tag{a3}$$

where $\Phi_t(.) = i_t^{-1}$ (.). Applying this into the equation (a4), we have:

$$y_t = \beta_0 + \beta_k k_t + \beta_l l_t + \omega_t\ (k_t, i_t) + \eta_t \tag{a4}$$

Both OP and LP estimations proceed in two steps. The first stage is the estimation of the following equation:

$$y_t = \beta_0 + \beta_l l_t + \Phi_t\ (k_t, i_t) + \eta_t \tag{a5}$$

where $\Phi_t\ (k_t, i_t)\ = \beta_0 + \ \beta_k k_t + \omega_t\ (k_t, i_t)$

This estimation results in a consistent estimate of the coefficients on labor.

Taking the expectation of equation (8) conditional on i_t and k_t as follow:

$$\mathrm{E}[y_t \mid i_t, k_t] = \beta_l\ (\mathrm{E}[l_t \mid i_t, k_t\ ]) + \Phi_t\ (k_t, i_t) \tag{a6}$$

Subtracting equation (9) from equation (8), we have:

$$y_t - \mathrm{E}[y_t \mid i_t, k_t] = \beta_l(l_t - \mathrm{E}[l_t \mid i_t, k_t]) + \eta_t \tag{a7}$$

As capital enters $\Phi(.)twice$, OP assumes that ω_t follows a first-order Markov process and that capital does not immediately respond to ξ_t , the innovation in productivity over last's period expectation as follows:

$$\xi_t = \ \omega_t - \ E[\omega_t | \omega_{t-1}] \tag{a8}$$

Then $y_t^* = y_t - \ \beta_l l_t = \beta_0 + \ \ \beta_k k_t\ \ + E[\omega_t|\omega_{t-1}] + \eta_t^*$ (a9)

where $\eta_t^*\ = \ \xi_t + \ \eta_t$. Under these assumptions, a consistent estimate of β_k may be obtained by regressing y_t^* on k_t and a consistent estimate of $E[\omega_t|\omega_{t-1}]$

Table A1.

LR test to compare the Model 1 and the Model 3 in the Table 5

Likelihood-ratio test	LR chi2(3) = 1240.89
(Assumption: Model 1 nested in Model 3)	Prob > chi2 = 0.0000

Table A2.

LR test to compare the Model 2 and the Model 3 in the Table 5

Likelihood-ratio test	LR chi2(3) = 1293.79
(Assumption: Model 2 nested in Model 3)	Prob > chi2 = 0.0000

Table A3.

Description of Sectors

Sector	Description of Sectors	Classification by Pavitt (1984)
S_35	Processing and preserving of meat	2
S_36	Fishery and processing and preserving of fishery product	2
S_37	Processing and preserving of fruit and vegetables	2
S_38	Manufacture of vegetable and animal oils and fats	2
S_40	Manufacture of grain mill products, starches and starch products and bakery products	2
S_41	Manufacture of sugar	2
S_43	Manufacture of coffee and tea	2

Sector	Description of Sectors	Classification by Pavitt (1984)
S_45	Manufacture of macaroni, noodles, couscous and similar farinaceous products; prepared meals and dishes and other food products	2
S_46	Manufacture of prepared animal, fish, poultry feeds	2
S_47	Manufacture of wines	2
S_48	Manufacture of beers	2
S_51	Spinning, weaving and finishing of textiles	1
S_52	Manufacture of other textiles	1
S_53	Manufacture of wearing apparel	1
S_54	Manufacture of leather and related products	1
S_55	Manufacture of footwear	1
S_56	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	1
S_57	Manufacture of paper and paper products	1
S_58	Printing and reproduction of recorded media	1
S_62	Manufacture of basic chemicals, fertilizer and nitrogen compounds, plastics and synthetic rubber in primary forms	3
S_65	Manufacture of other chemical products	3
S_67	Manufacture of pharmaceuticals, medicinal chemical and botanical products	3
S_68	Manufacture of rubber products	2
S_69	Manufacture of plastics products	2
S_70	Manufacture of glass and glass products	2
S_71	Manufacture of non-metallic mineral products	2
S_72	Manufacture of cement	2
S_74	Manufacture of basic iron and steel	2
S_75	Manufacture of basic precious and other non- ferrous metals and Casting of metals	2
S_76	Manufacture of fabricated metal products, except machinery and equipment	2
S_77	Manufacture of computer, electronic and optical products	3
S_81	Manufacture of electric motor, generators, transformers and electricity distribution and control apparatus; batteries and accumulators; wiring and wiring devices	2
S_84	Manufacture of electric lighting equipment; domestic appliances and other electrical equipment	3
S_87	Manufacture of general purpose machinery and special-purpose machinery	4

Sector	Description of Sectors	Classification by Pavitt (1984)
S_89	Manufacture of motor vehicles; trailers and semi- trailers and other transport equipment	2
S_94	Manufacture of furniture	1
S_95	Manufacture of jewelry, bijouterie and related articles; musical instruments; sports goods and games and toys	4
S_98	Repair and installation of machinery and equipment	4

Table A4.

The distribution of provinces by regions

Region	Province	Region	Province	Region	Province
Northwest		*North Central*		*Southeast*	
	Điện Biên		Thanh Hóa		Bình Phước
	Yên Bái		Nghệ An		Tây Ninh
	Hòa Bình		Hà Tĩnh		Bình Dương
Northeast			Quảng Bình		Đồng Nai
	Thái Nguyên		Quảng Trị		Tp.HCM
	Lạng Sơn		TT-Huế		
	Tuyên Quang	*South Central*		*Southwest*	
	Quảng Ninh		Đà Nẵng		Long An
	Bắc Giang		Quảng Nam		Tiền Giang
Red River Delta			Quảng Ngãi		Bến Tre
			Bình Định		Trà Vinh
	Hà Nội		Khánh Hòa		Vĩnh Long
	Phú Thọ		Ninh Thuận		Đồng Tháp
	Vĩnh Phúc		Bình Thuận		An Giang
	Bắc Ninh				Kiên Giang
	Hải Dương	*Central Highlands*			Cần Thơ
	Hưng Yên		Đắk Nông		Hậu Giang
	Thái Bình		Lâm Đồng		Sóc Trăng
	Hà Nam				Bạc Liêu
	Nam Định				Cà Mau

Note: This list includes the provinces in the data. The number of provinces were reduced from 63 provinces to 52 provinces by merging the data in the period.

Figure A1. Standardize values of residuals in the Model 3