Humanities and Social Sciences Letters
2024 Vol. 12, No. 2, pp. 248-274
ISSN(e): 2312-4318
ISSN(p): 2312-5659
DOI: 10.18488/73.v12i2.3701
© 2024 Conscientia Beam. All Rights Reserved.

Human capital spillovers and plant productivity in ASEAN: Evidence from the plant panel data

Muhammad Arditya Rindimajangga

Independent Scholar, Semarang, Indonesia.
Email: muhammadarditya@gmail.com

Article History
Received: 7 November 2023
Revised: 30 January 2024
Accepted: 12 March 2024
Published: 2 April 2024

Keywords
ASEAN
Cobb-Douglas
Human capital spillover
Manufacturing sector
Plant productivity
Worker.

ABSTRACT

This paper examines whether local human capital spillovers can affect plant productivity in three Association of Southeast Asian Nations (ASEAN) countries using panel data from the World Bank enterprise survey between 2009 and 2015. According to the literature, spillovers in the form of human capital are critical for economic growth. The spillovers exist if plants can produce more in the region with the abundant number of skilled workers given their inputs. This study uses the augmented Cobb-Douglas production function to evaluate the influence of the non-production worker proportion outside the plant which serves as a proxy for skilled workers. The findings indicate that plant production is impacted by spillovers of human capital. This study finds that a 1 percentage point increase in the proportion of nonproduction workers within a region will enhance the productivity of plants in all regions by roughly 7.1–8.9 percent. Moreover, productivity rises are found to be higher for low-tech plants. According to the economic significance of these spillover effects, the value-added per plant will rise by US$ 330,000 or US$ 866.75 million for the manufacturing sector as a whole. According to the findings, policymakers should encourage plants to provide workers with additional training and raise the level of human capital through education.

Contribution/Originality: This paper fills the research gap on how to measure the impact of human capital spillovers and their policy implications for firm productivity in ASEAN. The study also makes a contribution within multiple countries or regional contexts which is often scarce in the firm-level productivity literature.

1. INTRODUCTION

Economists have long identified that knowledge spillovers in the form of human capital are crucial to determine the long-run economic growth (Lucas, 1988). Education or experience-based learning is two ways to accumulate human capital. In the context of the latter, Marshall (1920) emphasised that social interaction among workers in the same location could generate learning opportunities to implement productivity-enhancing ideas that increase firm productivity. The quality of the pool of workers will determine the success rate of such collaboration. For example, organizations find it easier to hire employees who can use their company's assets efficiently if their workplace is in a region with a higher level of education. One thing workers may do to accomplish this aim is to learn new skills and information from other knowledgeable employees at the location. The possibility of a face-to-face relationship increases with close proximity. Therefore, it will be simpler for them to engage and exchange such items. Rosenthal and Strange (2008) explained that there is a negative relationship between the influence of this activity and spatial closeness. In other words, the exchange of ideas will sharply attenuate when the distance among skilled workers

© 2024 Conscientia Beam. All Rights Reserved.

increases. The aforementioned fact suggests that the location of a plant is significant since the presence of educated workers in the area plays a crucial role in their ability to get new insights and ultimately impacts their output.

This study examines the presence of local human capital spillovers on plant productivity for manufacturing industries in three ASEAN countries. Plants are expected to produce a larger output if they are located in a particular region with a high level of human capital for any given input rather than similar plants located in an area where the human capital level is low. The findings of this study can inform policymakers about whether they have to invest in human capital upgrading in various ways. This study primarily uses a more direct method to verify this hypothesis by estimating the augmented Cobb-Douglas production function at the plant level using data from the World Bank Enterprise Survey for the years 2009 and 2015.

This study examines the degree of human capital within the region as the proportion of non-production workers among all workers in the region. However, previous empirical works mainly used educational attainment (i.e., the fraction of workers with a college degree or above) to capture regional human capital levels. This study uses indirect indicators such as the share of non-production workers to capture labour skill intensity because of the lack of comprehensive data on workers' education in the survey. According to Berman, Bound, and Griliches (1994) this measure will also not be significant.

> Both conceptually and empirically, the production and nonproduction worker distinction closely mirrors the distinction between blue- and white-collar occupations. The blue-collar and white-collar classification closely reflects an educational classification of high school and college (pp. 371-372).

This statement implies that the use of nonproduction workers as the degree of labour skills is acceptable because it reflects the separation between blue and white-collar workers. In turn, blue and white-collar worker categories mirror the educational distinction. In other words, non-production workers tend to be more educated than production workers and as a result, the former type is more likely to absorb new ideas and knowledge.

Empirical literature indicates that human capital spillovers are an important aspect of determining plant productivity especially in the manufacturing sector (Chang, Wang, & Liu, 2016; Liu, 2014; Moretti, 2004b). A small amount of systematic research has been done to determine the extent of spillovers in developing countries despite their importance for policy consequences. All of the earlier research was conducted in developed countries. Hence, this paper will fill the gap by estimating the magnitude of human capital spillovers in three ASEAN countries.

This study contributes significantly to the corpus of knowledge about human capital externalities despite the fact that the empirical methodology used in it is quite comparable to that of Moretti (2004b). The three ways in which this study differs from the others are as follows: First, this study is the first attempt to employ panel-plant level data for multiple countries. Previous studies have emphasised the magnitude of local skilled workers on plant productivity for a given country. Another difference comes from the use of the dataset. This research uses information from enterprise surveys to internally assess regional human capital levels rather than creating matched employer-employee data like Moretti did. Such computation is relevant and reliable since the survey is sufficient to exploit the overall variation in plant conditions within regions. Lastly, this study employs a different instrument to eliminate the endogeneity issue compared to the previous studies. The proportion of skilled workers at the regional level is correlated with the lagged number of higher education institutions.

The remainder of the paper is organised as follows: Section 2 describes the related theoretical and empirical literature on human capital spillovers. The augmented production function used is explained in section 3. This section also presents the estimation issues that I need to address. Section 4 discusses the data source and variables I employ in this paper. The estimated effect of human capital spillovers on plant productivity is described in section 5. This section also provides sensitivity exercises for the main specification. Finally, section 6 is the concluding remark.

© 2024 Conscientia Beam. All Rights Reserved.

2. LITERATURE REVIEW

Marshall (1920) was the first to analyse the theoretical literature about human capital spillovers and their relationship to business productivity. He believed that the localization of skilled employees in a limited area was the main aspect of productivity enhancement. This view is then supported by Kuznets (1962) who elaborates that

> creative effort flourishes in a dense intellectual atmosphere. The possibility of more intensive intellectual contact afforded by greater numbers may be an important factor in stepping up the rate of additions to new knowledge (pp. 328-329).

Both views focused on the average quality of the local workforce. However, it is not the sole determinant of economic productivity but also depends on the role of skill diversity and cultures. Jacobs (1969) highlights that the interactions among people with different knowledge and perspectives within a city tend to generate new ideas and innovations. These arguments are then formalised with the learning model proposed by Lucas (1988) who proposes that the benefit associated with urban areas comes from the firm acquiring ideas from the citizens where they operate. Other types of learning model have been richly presented in the literature (Glaeser, 1999; Jovanovic & Rob, 1989). On the other hand, Acemoglu (1996) builds a model where aggregate human capital can affect plant productivity in the absence of learning mechanisms. In his model, the human capital of workers and the physical investment of firms are complementary. For instance, if workers are expected to increase their educational attainment in the future, firms will decide to invest in more advanced machinery and equipment now to match the expected worker's skill requirements and thus raise productivity. More recently, Duranton and Puga (2004) analyse that the spatial concentration and structure of the local workforce can enhance firm productivity through three mechanisms such as sharing, matching and learning. The first type includes sharing narrower specialization and risk by establishing labour pooling. The matching process focuses on improving the quality of the match between workers and employers. Finally, learning mechanisms are based on knowledge accumulation and diffusion.[1]

Several papers have examined the empirical extent of local human capital spillover either at the regional or firm level. The first wave of studies focused on its impact on individual wage based on the Mincerian model. Rauch's (1993) fundamental study was the first to identify the impact of the variance in human capital between cities on wages. He discovered that a one-year increase in the average number of years spent in school in any particular city would result in a three to five percent rise in the average hourly wage based on the 1980 US Census of population. Nevertheless, the main limitation of Rauch's paper is failing to deal with the endogeneity problem associated with city-level human capital. Acemoglu and Angrist (2000) use the state variation of compulsory education turn out and child labour laws as instruments for average schooling across cities to overcome such issue. They also address the endogeneity concern in individual schooling by instrumenting it with a quarter of birth. They only find little evidence of such spillovers in practice. Similarly, Moretti (2004a) implemented two instrumental variables at the city level which are age structure and land grant institutions. He also uses the share of college graduates as the measure of city human capital level. He highlighted that one percentage point increase in the proportion of college graduates in a particular city increased the average wage by 0.6-1.2 percent in 1980. The main difference in findings between Acemoglu and Angrist (2000) and Moretti (2004a) lies in the instrument they use. The latter uses the share of college graduates which reflect the upper bound estimate of human capital spillovers whereas the former employs compulsory schooling laws that mostly affect lower education.

However, another strand of literature on human capital spillovers has shifted away from a wage model to a firm productivity model by estimating production functions. De La Fuente and Domenech (2001) found that there is a positive relationship between the growth rate of average years of schooling and productivity using a cross-country dataset. However, employing cross-country data is not particularly useful as the human capital level is possibly correlated with unobserved shocks that can affect productivity. Hence, the subsequent studies tried to explore

[1] It is critical to notice that in this paper, I do not test hypothesis in any of these theoretical aspects. Thus, testing them is well beyond the scope of this study.

© 2024 Conscientia Beam. All Rights Reserved.

exogenous variation at the city or regional level to tackle this issue. Adams and Jaffe (1996) investigate the effect of transfer knowledge composition within and across firms on plant productivity for the chemical industry in the US over the period 1974-1988. They use the number of research and development (R&D) performed in labs which can be conducted informally at plants, formally at parent firm or by other plants in the same city or industry as the explanatory variable. They document that R&D spillover are essential both within and across plants. The most intriguing result was that the amount of spillovers decreased with increasing geographic separation between labs and plants. They suggest that such a condition reflects the rising cost of communication. Similarly, growth and the spatial distribution of scientists are the main factors determining the birth and location of the biotechnology industry in the US (Lynne, Michael, & Marilynn, 1998). A number of outstanding scientists in cities will influence the location decisions of firms in such sectors.

Both studies emphasise high-tech industries. Moretti (2004b) then systematically tries to estimate the impact of regional human capital spillovers on plant productivity for all US manufacturing industries between 1982 and 1992. Moretti indicated that the impact of inter-industry human capital spillovers is quite low, only responsible for an average of 0.1% increase in firm output per year during the period. Estimated spillovers tend to come from high-tech industries and they are virtually zero for low-tech industries. More importantly, the scale of spillovers between plants within a city depends on their interaction level. Spillovers are found to be significant for plants that regularly interact with each other.

Recent studies on human capital spillovers using the productivity model particularly follow Moretti's (2004b) study. The only difference essentially comes from the country sample and the instruments they apply when controlling the endogeneity problem. More specifically in Asia, Liu (2014) and Chang et al. (2016) report that a high level of city human capital is positively associated with larger plant productivity for the manufacturing sector in China and Taiwan respectively while Park (2016) does not find a similar result for Korean manufacturing plants. Lu (2022) suggests that a high level of educated workers in the Special Economic Zones (SEZs) area of the Yangtze Delta of China improve the productivity of the firm within the SEZ and its neighboring area. Current research in Europe indicates that the long-term benefits of human capital spillovers will decrease in contrast to Asia. Eklund and Pettersson (2019) argue that an increase of 1 percent in educated worker share will improve firm productivity by around 0.4 to 1.0 percent but the effect diminishes over time. Similarly, according to Kijek and Kijek (2020) the relationship between human capital and productivity is non-linear indicating that positive effects continue to a certain point at which they become negative. In addition, there is conflicting empirical data about local human capital spillovers based on both the wage and productivity approaches.

3. EMPIRICAL FRAMEWORK

The hypothesis of this paper is whether plants located in the region with a higher quality of human capital will generate more output with any given input. Thus, production function will be estimated to evaluate the effect of human capital spillovers on plant productivity. However, such a strategy is not straightforward to identify. Unobservable factors can affect the productivity of plants and are correlated with the overall human capital across regions. More productive plants may decide to produce in areas with an ample amount of highly skilled workers for any reasons that are not associated with human capital spillovers.

This study begins by explaining how production function can fit to capture such externalities. The remaining part of this section then discusses the endogeneity problems, the panel data model options and the best econometric strategy to address them.

3.1. Model Specifications

First, consider the following standard Cobb-Douglas production function:

© 2024 Conscientia Beam. All Rights Reserved.

$$Y_{ijrt} = A_{ijrt} LNP_{ijrt}^{\alpha} LP_{ijrt}^{\beta} K_{ijrt}^{\delta} \quad (1)$$

Where Y_{ijrt} indicates the value added of plant i belonging to industry j, in region r, and at time t; LNP_{ijrt} and LP_{ijrt} denote the number of nonproduction and production workers in the plant respectively and K_{ijrt} is capital stock. If this study assumes that plant total factor productivity, A_{ijrt}, reflects the regional human capital level, then A_{ijrt} can be expressed as follows:

$$\ln A_{ijrt} = \gamma HS_{rt} + \theta X'_{ijrt} + \varepsilon_i + \varepsilon_j + \varepsilon_r + \varepsilon_t + \varepsilon_{rt} + \varepsilon_{ijrt} \quad (2)$$

HS_{rt} is the share of non-production workers in a particular region r outside the firm and coefficient of the interest γ represents the external effect of human capital on productivity. In Equation 2, this study emphasises the human capital spillovers within a region. This study does not distinguish the spillovers within and between industries. Hence, estimates of γ can be interpreted as the upper bound on the total spillover magnitude. Statistically, the main challenge from Equation 2 is how to establish that the coefficient γ consistent and not correlated with time-varying shocks which are embodied in ε_{rt} and ε_{ijrt}. If γ=0, the model goes back to the standard production function without spillovers. ε_i, ε_j, and ε_r are time-invariant unobserved productivity shocks at the plant, industry and region respectively. Similarly, specific time-varying factors and region-specific time-varying shocks are represented by ε_t and ε_{rt}. However, this study does not include any industry-specific time-varying shocks that are shared by all plants in the same industries because it will cause perfect multicollinearity with permanent heterogeneity at the industry level (ε_j). This study also adds control for individual plant characteristics, X_{ijrt}, as years of operation and single or multi-plant categories.

Equation 1 can be rewritten into a natural log and plugged back into $\ln A_{ijrt}$, and then the production function.

$$\ln y_{ijcrt} = \delta + \alpha \ln LNP_{ijrt} + \beta \ln LP_{ijrt} + \delta \ln K_{ijrt} + \gamma HS_{rt} + \theta X'_{ijrt} + f_i + f_t + \varepsilon_{rt} + \varepsilon_{ijrt} \quad (3)$$

It is important to notice that c indexes the country where the region r is located. All the regressions are estimated using clustered standard errors on plants.

The main advantage of using a panel dataset is that this study can control the permanent and time-varying factors that affect productivity and the overall share of nonproduction workers. Plant-fixed effect f_i absorbs any unobserved time-invariant variations at the plant, regional and industry level (ε_i, ε_j, and ε_r). Similarly, the year-fixed effect f_t deal with the time-varying productivity shock common to all plants. Nevertheless, this study leaves the time-varying, region-specific factors ε_{rt} as they are . If region-time is fixed effect, f_{rt}, is added as a covariate, it will cause perfect multicollinearity as the human capital spillovers variable HS_{rt} is at the region-year level. As a result, Equation 3 cannot be estimated properly. When plant specific characteristics can be controlled, this study is not able to do the same thing at the regional level due to data limitations across these countries. Hence, this study provides an alternative measure of those changes over time within an area to overcome this problem. One can expect that the regional human capital level will have a positive impact on productivity and hence the coefficient γ is predicted to be positive. However, unobserved factors can affect either region-human capital or plant productivity. This will be the source of the identification threat that needs to be tackled.

The endogeneity problem from Equation 3 is a major problem. It arises if the time-varying factors at the regional level ε_{rt} are positively correlated with the share of college graduates within the region H_{rt}. When such a problem is not solved, the magnitude of coefficient γ will be biased upward.

The common solution to this concern is employing the two-stage least square estimation technique. It will be crucial for selecting the right instrument while employing this strategy. It has to be correlated with the share of nonproduction workers (relevant condition) but has no relevance to the productivity shocks of individual plant (exogeneity condition). Additionally, the instrument has to vary at the regional-year level to extract a portion of the region-specific change in the share of nonproduction workers. Then, this study proposes the logarithm of total higher education institutions for the instrument. This study uses two-stage least squares to examine the causal

© 2024 Conscientia Beam. All Rights Reserved.

relationship and the degree of human capital spillovers using an instrument that allegedly fits these two criteria. More formally, the first stage equation is described as follows:

$$HS_{rt} = \tau \ln HEI_{r,t-1} + \sigma \ln LNP_{ijrt} + \rho \ln LP_{ijrt} + \pi \ln K_{ijrt} + \vartheta X'_{ijrt} + f_i + f_t + \varepsilon_{rt} \quad (4)$$

Here, HS_{rt} denotes the share of nonproduction workers as defined earlier in this section. $HEI_{r,t-1}$ is the number of universities and colleges in the previous year. In other words, this study examines the proportion of nonproduction workers employed, supposing that there would be a greater exogenous relationship between higher education institutions and the quantity of highly competent people. The second stage regression of the IV approach uses the predicted value of the share of nonproduction workers in the first stage, $\widehat{HS_{rt}}$, and estimates it with the following equation:

$$\ln y_{ijcrt} = \delta + \alpha \ln LNP_{ijrt} + \beta \ln LP_{ijrt} + \delta \ln K_{ijrt} + \gamma \widehat{HS_{rt}} + \theta X'_{ijrt} + f_i + f_t + \varepsilon_{rt} + \varepsilon_{ijrt} \quad (5)$$

Again, the coefficient of interest is γ. If the instrument is valid, it will tackle the positive biased attributed to simple OLS estimates and γ will further reveal regional human capital spillovers.

This study makes reference to those two conditions in order to verify that the instrument satisfies its validity requirements. The relevant condition can be tested by looking at the significance of the estimated result of the first stage in the two-stage least square analysis. In terms of the exogeneity criterion, an instrument is deemed invalid if unobservable productivity shocks positively influence plant production and likely increase the number of universities in the area. However, it is unclear whether it is the result of unobservable productivity shocks in the current year because the instrument is based on data from the prior year. Hence, it is assumed that the exogeneity condition is satisfied in the analysis.

The next subsection will describe the details of econometric concerns to obtain unbiased estimates. However, this research will provide an overview of the panel data models that are frequently used in empirical papers before entering into the topic of estimation difficulties.

3.2. Linear Static Panel Data Models

This study begins by estimating the presence of human capital spillovers as in baseline. Equation 3 uses linear static panel data to estimate approaches (i.e., pooled OLS, random effects and fixed effects). The method of estimating starts with the use of pooled ordinary least squares and random effect incorporating cluster-robust standard errors to accommodate heteroscedasticity. It is assumed that the unobserved heterogeneity of each individual plant persists from year to year and clustering is performed at the company level to enable serial correlation of the error term (both the idiosyncratic and the unobserved errors) across time.

A typical example of a linear static panel data model will be denoted as follows:

$$y_{it} = \beta_0 + \beta_1 x_{it} + u_i + \varepsilon_{it} \quad (6)$$

Where i=1,…,N and t = 1,….,T. The main characteristic of Equation 6 is that coefficient β_1 and β_2 do not vary across individual or over time which does not allow for individual heterogeneity. This criterion leads to the concept of the pooled least square model. As a result, it is assumed that there is no contemporaneous correlation of error terms with the independent variables x_{it}.[2] Thus, the estimators are consistent and can be estimated using standard Ordinary Least Square technique. However, the restricted assumption tends to be violated in the case of a panel dataset due to its extreme impossibility. Heterogeneity across individuals tends to happen as the unobserved characteristics that are not captured in the regressors will be included in error terms which leads to incorrect standard errors.

An alternate model is then proposed known as the random effect model. The main rationale behind the random effect model is that the variations across individuals are assumed to be random and not correlated with regressors.

[2] $E(x_{it}, u_i) = 0$ and $E(x_{it}, \varepsilon_{it}) = 0$ for t=1….,T.

© 2024 Conscientia Beam. All Rights Reserved.

If we change Equation 6 a little bit where the individual specific effects u_i are assumed to be uncorrelated with the regressor x_{it};

$$y_{it} = \beta_0 + \beta_1 x_{it} + u_i + \varepsilon_{it}; i = 1, \dots, N; t = 1, \dots, T. \quad (7)$$

$$\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \text{ and } u_i \sim N(0, \sigma_u^2)$$

We use Breusch and Pagan's (1980) Lagrangian Multiplier (LM) test for random unobserved individual effects after the pooled ordinary least squares and random effect techniques for the baseline model. If the p-value of Breusch and Pagan LM is less than 5 percent significant level, the null hypothesis of homoscedastic errors is rejected and then the random effects model is selected over the pooled ordinary least square estimator. According to convention in relevant empirical work, the random effect estimator has to be compared to the fixed effects model in order to be determined to be the best model option.

This study then proceeds to estimate the model using the fixed effect estimator which transforms the data through a demeaning procedure that removes the unobserved heterogeneity from the model and is then estimated using the standard ordinary least square estimation technique. The fixed effects model is the most commonly used between the two models in this study. This is because the random effect model makes the restrictive assumption that unobservable factors are not random and uncorrelated with regressors which does not reflect the data structure of most datasets.

Suppose we have the following generic static panel data linear regression model stated in Equation 8 where a_i and u_{it} are the unobserved time-invariant plant-fixed effects and the idiosyncratic error term respectively which include all other time-varying unobservable.

$$y_{it} = \alpha_i + \beta' x_{it} + u_{it} \quad (8)$$

The ordinary least squares estimator will be biased in the setting of ordinary least squares if the unobserved fixed time-invariant effect fi is associated with the regressor xit which is typically due to simultaneity bias or missing variables. This is because the tight exogeneity requirement is then violated. One way of overcoming this issue is the inclusion of dummy variables for each cross-section where for every observation of the cross-sectional variable one dummy is assigned 1 and for all other cross-sectional variables 0 is assigned. This method is known as the Least Squares Dummy Variable model or LSDV model. It should be noted that the LSDV is theoretically the same as the fixed effects estimator but a major weakness is the issue of degree of freedom losses. In light of this, the LSDV model is not preferred.

Then, the preferred option is what is known as the time demeaning (within) transformation process. Now, each variable in Equation 9 is averaged to obtain the following expression in Equation 9:

$$\bar{y}_{it} = \bar{\alpha}_i + \beta'_1 \bar{x}_i + \bar{u}_i \quad (9)$$

Where $\bar{y}_{it}$, $\bar{x}_i$, $\bar{\alpha}_i$ and $\bar{u}$ are the time averages of the dependent variable, independent variable, the time-invariant unobserved factor and idiosyncratic error term. The core of this process is that it subtracts Equation 9 from Equation 8 to obtain a fixed effect estimator.

$$(y_{it} - \bar{y}_{it}) = (\alpha_i - \bar{\alpha}_i) + \beta'(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i) \quad (10)$$

A transform equation with time-demeaned variables is obtained which is as follows:

$$\ddot{y}_{it} = \beta'_1 \ddot{x}_{it} + \ddot{u}_{it} \quad (11)$$

Where $\bar{\alpha}_i = \frac{1}{N}\sum_{t=1}^{N} \alpha_{it}$ and $\bar{X}_i = \frac{1}{N}\sum_{t=1}^{N} X_{it}$. Equation 11 makes it clear that the fixed effect that is time invariant and constant has been eliminated ($\bar{\alpha}_i = \alpha_i$). Equation 11 may be evaluated using ordinary least squares once the unobserved time invariant fixed effects have been eliminated.

The Hausman test is used to choose between the fixed and random effect models. It is used to verify the orthogonality condition of the random effect and the regressors (Greene, 2003). However, the standard Hausman test does not function on non-independent and identically distributed errors (non-i.i.d) as we assumed from the onset in this paper. Hence, a robust version of the Hausman test will be used in this article.

© 2024 Conscientia Beam. All Rights Reserved.

3.3. Estimation Issues

The econometric problems described in this section are related to the potential unobservable heterogeneity that may happen across plants and over time under a fixed effect model. As stated before, the main assumption of suitable identification after controlling plant characteristics is that there is no correlation between the fraction of nonproduction workers in region-observable variation HS_{rt} and the time-varying regional unobserved factor ε_{rt}. If this assumption is violated, the coefficient of interest γ will be inconsistent. There are two sources of threats to identification regarding such a relationship. The first problem comes from the existence of reverse causality. It is a condition where plants with better management practices, for instance, choose to locate their production in regions with a high proportion of nonproduction workers because they can attract them with a higher wage. Therefore, the local density of such employees increases and provides an overestimated coefficient of regional human capital. This issue is unlikely to be a significant one because the model incorporates the plant fixed effect which also incorporates region-specific variability and permanent industries. Secondly, the variation in local amenities that is not included in HS_{rt} but is correlated with regional human capital level and plant productivity at the same time will also pose an omitted variable problem. The exclusion of such feature will be captured in the region error term ε_{rt}. An obvious example of this is local infrastructure such as highways and airports. Both transport means will certainly raise productivity for all plants in any region and at the same time attract workers to move to that area. If this happens, the estimates of γ will lead to a positive bias. However, there are no plants that move to another region in the dataset between two subsequent years. This study can include plant fixed effect f_i to eliminate heterogeneity across regions. Additionally, I also employ an instrumental variable approach to adjust to this endogeneity. More explicitly, this study uses the one-year lagged number of higher education institutions as the instrument.

My estimates may also be invalid if any positive relationship between plant productivity and human capital spillovers are found to be caused by time-varying factors other than the proportion of nonproduction workers in the region. Obviously, if this possibility is ignored, the estimates tend to exhibit an upward bias. Thus, the inclusion of the year-fixed effect is essential to be able to avoid the failure to recognise these unobservable factors.

Another potential econometric problem is the relationship between firm-level human capital and the random error term. The latter may contain plant-specific unobserved factors that affect productivity. Such a potential issue will lead to inconsistent estimates. This study addresses this issue by using time-varying plant-specific control factors including multi-plant status and plant operation longevity.

This study additionally investigates the regional human capital endowment using a different metric in order to obtain reliable estimates of the regional human capital variable. The measure involves the log of the average monthly real wage within the regions. This indicator can capture the local quality of workers besides schooling. Plants located in areas with abundant nonproduction workers will pay a higher wage because these employees are anticipated to be more productive and hence increase plant output. Therefore, it is an appropriate inference that nonproduction employees receive higher salaries than production employees since they are more likely to be competent.

A final concern is related to theory when input factors such as labour and capital cannot be treated as exogenous. Endogenous inputs are the only issue to the extent that they will result in biased estimates of coefficient γ. Therefore, this study assumes that this problem is solved after controlling for the plant -fixed effect. It directly estimates Total Factor Productivity (TFP) as the function of HS_{rt} using Levinsohn and Petrin's (2003) approach in order to evaluate the reliability of my findings to this assumption. The findings mostly align with the primary model indicating that labour and capital are entirely exogenous in this case. In addition, it is apparent that several possible causes of endogeneity have been considered. However, I cannot completely rule out the possibility that some of the estimated impacts are a reflection of regional productivity shocks. I also carry out a set of specification tests to prove that the positive impact on plant productivity in this study is solely due to human capital spillovers.

© 2024 Conscientia Beam. All Rights Reserved.

4. DATA AND VARIABLES

The data used in this study is taken from World Bank Enterprise Surveys for 3 ASEAN countries, namely Indonesia, the Philippines and Vietnam for the period 2009 and 2015. This study excludes other ASEAN countries due to differences in the year surveyed since the World Bank just began to survey them at the beginning of the 2000s. This study only focuses on plants in the manufacturing sector which is consistent with previous studies. It contains approximately 700 manufacturing plants for each year within a country but some plants leave and join the survey in 2015 thus creates unbalanced panel format. However, this study only includes plants that appear in both years to construct a balanced panel dataset. Finally, there are 198 plants from all countries combined in my dataset. Although the number of plants seems small, enterprise surveys can sufficiently represent the whole manufacturing sector within the country because they apply stratified random sampling.[3] The dataset also covers detailed information about individual plants such as the net value of capital, locations, industrial classification, cost of production, sales and others.

The dataset also shows the number of employees, both production and nonproduction employees and average schooling years. However, the information on the latter is restricted to production workers only. That is the reason this study does not use schooling as the variable of interest. Hence, it only uses the proportion of white-collar (nonproduction) workers as a proxy for college graduates share. The enterprise survey defines nonproduction worker as "workers employed in sales (including driver-salespersons), sales delivery (highway truck drivers and their helpers), janitorial and guard services, advertising, credit, collection, installation and servicing of own products, clerical and routine office functions, executive, purchasing, financing, legal, personnel (including cafeteria, medical, etc.)" (World Bank, 2017). In addition, this study also constructs an alternative measure of regional human capital to check the validity of the core results by employing the log of the average yearly wage outside the plant within a region calculated as the yearly labour cost divided by the total number of workers.

This study uses value-added as an indicator of plant output. Value added is defined as sales less the cost of production which includes the cost of electricity and materials. Labour will be measured and separated using the number of production and nonproduction employees. Such an indicator is used instead of a number of work hours because of data constraint. The net book value of machinery, vehicles and equipment is assigned as a proxy for plant capital stock. However, the value of each variable is presented in local currency. The research also includes the overall cost of goods as an input element. Thus, these local currencies are converted into US dollars using the official exchange rate from the World Development Indicator (WDI-WB) database.[4]

This research also uses the statistical yearbook of each country to gather data on the number of higher education institutions one year behind schedule. The regional distribution of all universities and institutions is included in this paper. The regional classification between the Statistical Yearbook and the enterprise survey is the same. Therefore, this study determines how many higher education institutions are based on the latter and it applies those numbers to every sample plant in the dataset without making any additional adjustments.

This study will prioritize the industrial sectors in its analysis to align with value-added terms and ultimately boost productivity. The industrial classification is based on four-digit level manufacturing industries. They are regrouped into 17 sectors at a two-digit level to capture more plants within industries when examining the spillovers effect based on the sectoral analysis. Nevertheless, the number of plants for each industry is not perfectly distributed because some industries have a larger contribution to the national economy than others. Large plants have a greater possibility of being oversampled as a result. The plant site is another area where the same technique

---

[3] Enterprise Survey use stratified random sampling methodology to collect the key variables from plants (See Data Appendix for detailed explanation). The strata used are based on plant size, sector, and region.

[4] 2009: US$1=IDR 10389.94; US$1= PHP 47.68; US$1= VND 17065.08.

2015: US$1=IDR 13389.41; US$1= PHP 45.50; US$1= VND 21697.57.

© 2024 Conscientia Beam. All Rights Reserved.

is applicable. Not all regions are included in the survey. Nine provinces are selected for Indonesia, five districts for the Philippines and only four regions for Vietnam. There will be a probability that more skilled workers prefer to work in one of those regions.[5]

It is also important to keep in mind the caveat of my results presented here due to data availability. The most matched panel dataset for ASEAN countries is only for the period 2009 and 2015. Plants included in the survey are the only small part of plant population in each country and therefore not essentially representative to describe the whole effect in ASEAN. The World Bank acknowledged that one important outcome is the oversampling of large plants because of their higher economic production. The present study estimates the individual regressions based on three general plant size categories: small, medium and large in the robustness analysis in order to manage them (see section 4.2.4 for more details).

Table 1 provides summary statistics for the main variables used in this paper. Columns in Table 1 contain the mean and standard deviation for each year and country.[6] One can verify that both sales and value-added increased significantly between 2009 and 2015. The average sales of each plant increased greatly from around 7 million U.S. dollars in 2009 to 78 million U.S dollars in 2015, this figure increased nearly 11 times in six years. During this period, the average value-added also rose rapidly by more than 600%. On the other hand, capital stock is the variable that slightly fell during the term of interest. Each plant's net value of capital, on average, declined approximately 6% from a base of 1.5 million U.S dollars in 2009. The cost of purchasing materials and fuel (energy) increased at a rate that was comparable to that of the inputs needed for production. The cost of materials increased roughly by more than 346% and there was a 485% increase in energy expenses respectively. As a result, energy costs are as important as the cost of materials.

Table 1. Means of summary statistics.

Variables	All countries		Indonesia		Vietnam		Philippines	
	2009	2015	2009	2015	2009	2015	2009	2015
Sales (*1000)	7658.1	77980.1	5988.6	74532.3	2749.4	5073.6	13914.7	131938.6
Value added (*1000)	3876.0	40005.9	2690.5	55590.6	590.9	1458.3	8310.2	34832.9
Capital (*1000)	1542.8	1447.7	1780.5	1014.8	478.9	678.4	1770.0	2788.1
Materials (*1000)	2815.1	12574.8	2411.5	16080.0	1946.1	2615.5	4163.3	12214.0
Energy (*1000)	159.1	932.0	145.5	1480.5	30.8	68.7	268.8	282.0
Number of production workers	145.8	398.7	109.5	180.7	92.4	80.9	109.5	87.4
Number of nonproduction workers	35.9	124.0	41.4	54.0	13.0	12.8	41.4	32.5
Nonproduction workers share in the region	0.200	0.213	0.193	0.207	0.128	0.130	0.258	0.280
Average yearly wage in region	2234.3	7099.4	1478.1	2852.8	1699.6	3560.1	4051.3	18271.3
Age of plants	17.3	22.5	19.7	24.5	10.1	15.2	17.5	23.2
A proportion of plants belong to multi-unit firms	0.15	0.13	0.22	0.09	0.09	0.06	0.07	0.26
Number of plants	194		105		35		54	

Note: (*1000) indicates that all monetary values are reported in US dollar. Nonproduction worker share and average yearly wage are calculated outside the plant.

Moreover, there is no discernible difference between the two periods in the average percentage of white-collar workers among the three countries in ASEAN. In particular, there was only one percentage point increase in the fraction of nonproduction workers within a region outside the plant in the space of six years. This can be inferred

[5] The fraction of nonproduction workers for each region is presented in Data Appendix.

[6] Summary statistics for each country in details are presented in Data Appendix.

© 2024 Conscientia Beam. All Rights Reserved.

by a slight increase in total nonproduction workers from an average of 36 to 41 people employed while there is a small decline of around nine nonproduction workers during the period of interest. On the other hand, the average age of individual plants slightly increased while the fraction of multi-unit plants achieved a minor decline during the period. Overall, summary statistics show that plant level characteristics increased significantly whereas the regional human capital level remained stable between these two years.

It is now possible to illustrate the pattern for each country as well. During the period when plants in the Philippines had the most substantial rise in every aspect, productivity, input variables (except the capital stock) and aggregate human capital expanded significantly. For example, we can look at value-added as a productivity measure. It is interesting that per plant value-added in Indonesia is approximately 2.5 times larger than in Vietnam whereas Indonesia's and Vietnam's value-added is just 10 and 5 percent of each plant's value-added in the Philippine respectively over a six-year period. This fact implies that there are huge variations in productivity across plants within the country.

This condition also applies to worker endowments as well. Although the number of skilled workers is the highest for plants in Indonesia, one cannot say that manufacturing plants in Indonesia have the best worker's skill intensity if not compared to total production workers. Moreover, plants in the Philippine have the highest employment share of nonproduction workers where every three production workers consists of one nonproduction workers (3:1). On the other hand, in Indonesia, one skilled worker is equivalent to four unskilled worker (4:1) while the comparison between production and nonproduction workers in Vietnamese plants is around seven to one (7:1). As a result, the share of highly-skilled workers within the region among countries is also the largest for Philippine manufacturing plants.

This disaggregated means also reveal the similar features for years of operation and multi-unit status as plant-specific controls. Filipino plants have been operating longer, nearly 20 years, than the similar plants in Indonesia and Vietnam. Similarly, 17 percent of total manufacturing plants in the Philippines have more than one establishment compared to 15 and 7 percent for Indonesian and Vietnamese plants respectively. Based on the summary statistics, this study can show that there may be a significant and positive correlation between human capital spillovers and productivity, as indicated by the figures for plants in the Philippines. The predicted outcome aligns with the results presented in section 5.

5. RESULT

This study initially conducts a series of tests in order to choose the optimal panel data models for the analysis. It then continues to estimate the production function in Equation 3 and use the logarithm of value-added as a dependent variable. This study also employs instrumental variable estimation as estimated using Equations 4 and 5 to eliminate the endogeneity issue. Then, a series of robustness checks are carried out to strengthen my findings. It is critical to note that this study does not separate estimation results for each country since it takes them as an integrated analysis.

5.1. Basic Results

5.1.1. Ordinary Least Square Estimation

This study starts the analysis by choosing the best static panel data models to be used in this paper. Thus, it estimates the baseline Equation 3 under different models and the results are reported in Table 2. Examining the statistical significance of the Breusch-Pagan test allows us to select between the random effect model and the pooled least square model. The p-value of this test is 0.2382, more than 5 percent level of significance. It means that the null hypothesis cannot be rejected and thus the pooled least squares model is more appropriate. However, one may assume that there is unobservable heterogeneity across plants in the dataset that has an impact on productivity. This study must determine whether the fixed effect model can be used with the Hausman test in order to validate

© 2024 Conscientia Beam. All Rights Reserved.

this claim. This test is used to choose between a random effect and a fixed effect model. According to the p-value of the Hausman test, the null hypothesis is rejected. This verdict suggests that the variation across plants is correlated with the independent variables and then a fixed effect estimation is selected over a random effect. Therefore, this study continues to use the fixed effect model for analysis throughout the paper.

Table 2. Estimates of human capital spillovers: model choices.

Variables	Pooled OLS	Random effect	Fixed effect
Nonproduction worker share in region	4.061** (1.700)	4.380*** (1.684)	8.822*** (3.027)
ln (Production workers)	0.566*** (0.116)	0.575*** (0.115)	0.334 (0.214)
ln (Non-production workers)	0.731*** (0.149)	0.712*** (0.123)	0.289 (0.216)
ln (Capital)	0.191** (0.075)	0.183*** (0.047)	0.053 (0.066)
Plant specific controls	+	+	+
R^2	0.53	0.53	0.43
Breusch-Pagan test		0.2382	
Hausman test		0.0012	

Note: All specifications include plant characteristics such as years of operation and multi-unit status but without a dummy year variable. The sample size is 388 observations. *** p<0.01, ** p<0.05.

Table 3 reports the regression results under different specifications of Equation 3. This study begins by documenting the relationship between the share of nonproduction workers in the region and plant productivity. It also includes a dummy equal to one if the plant belongs to the multi-establishment unit and a logarithm of plant longevity to control for plant-specific characteristics that affect productivity.

This study starts with a model in column (1) where only factor endowments and plant-fixed effect are included. It eliminates estimates of unobserved variability in the permanent plant, industry and region. Identification of the human capital spillovers comes from changes in plant productivity and non-production worker share between 2009 and 2015. The coefficient on nonproduction worker share in column (1) is 8.857.

I modify the model specification in column (2) by maintaining the plant-fixed impact and include more plant controls like plant age and multi-plant status. Assessing spillovers provides an estimated impact of 8.822 which is quite similar to the number in column (1). This implies that plant characteristics do not completely affect the impact of aggregate human capital at the regional level.

Similarly, this study modifies the definition slightly by eliminating several plant controls and including the year-fixed impact. The spillover impact becomes smaller to 7.320 as documented in column (3). This figure indicates that time-invariant unobserved characteristics and plant-specific controls may confound the external human capital faced by the plants and lead to underestimating the effect of human capital spillovers. Bloom, Schankerman, and Van Reenen (2013) pointed out that spillovers will result in two contradictive effects on firm performance: a positive effect from knowledge spillovers and a negative impact from product market rivalries. The plant-fixed impact increases after correcting for plant competition within areas which causes downward biased results if the calculation does not clearly take this into consideration.

The most comprehensive requirements, found in column (4) are that a one percentage point increase in the proportion of nonproduction workers is linked to a 7.1 percent increase in productivity. The average annual rise in non-production workers between 2009 and 2015 was almost 0.2 percentage points which should be taken into consideration when interpreting this statistic. Therefore, an increase in nonproduction workers of 0.2 percentage points is estimated to raise value-added by 1.42 percent annually. This amounts to almost $55,000 annually for the usual plant across three ASEAN countries.

© 2024 Conscientia Beam. All Rights Reserved.

Table 3. Fixed effect estimates of human capital spillovers.

Variables	(1)	(2)	(3)	(4)
Non-production worker share in region	8.857*** (3.364)	8.822*** (3.393)	7.320** (3.331)	7.152** (3.450)
ln (Production workers)	0.345 (0.220)	0.334 (0.220)	0.297 (0.221)	0.302 (0.220)
ln (Non-production workers)	0.281 (0.231)	0.289 (0.231)	0.300 (0.232)	0.295 (0.233)
ln (Capital)	0.047 (0.089)	0.053 (0.090)	0.080 (0.090)	0.079 (0.090)
Plant specific controls		+		+
Plant-fixed effect	+	+	+	+
Year-fixed effect			+	+
R^2	0.44	0.43	0.47	0.48

Note: Standard errors are adjusted for clustering on the plant. Specifications (2) and (4) include plant characteristics such as years of operation and multi-unit status. The sample size is 388 observations. *** p<0.01, ** p<0.05.

One may wonder why the coefficients of all input variables are not statistically different from zero despite having positive signs. It means that there may be a relationship between the inputs and the unobservable productivity shocks. However, this condition is not a major issue because the study's focus is on the investigation of the human capital spillovers variable. In any case, this research attempts to address this issue by using the TFP specification which is discussed in section 5.2.1 to demonstrate that there is no need for concern.

The empirical findings of this study support the existence of a positive human capital spillover effect within the literature debate. Lu (2022) in their empirical study shows that firm and subsequently regional productivity benefits from a high level of human capital in Chinese SEZs. Similarly, Eklund and Pettersson (2019) using a firm-level dataset in Sweden also find that the impact of human capital spillover on firm productivity is positive albeit to a smaller magnitude with a diminishing return in the long-term.

In a nutshell, Table 3 results provide strong evidence that a region's level of human capital contributes significantly to an improvement in plant production. The value-added of a plant over a five-year period will increase by 7.1 to 8.8 percent for every percentage point increase in the nonproduction worker share of that plant in any given region as previously mentioned in the preceding paragraphs. The impacts are nearly three times larger than those observed by Moretti (2004b) if these estimated statistics are annualized. A primary reason for this is that the growth effect is expected to be stronger for countries in the catching-up process of the economic development stage (Benhabib & Spiegel, 1994). A common situation in developing countries is that their level of human capital is still low whereas human capital in advanced economies has reached a high level. The opportunity to improve such conditions is more open for countries in the "take-off" phase. As a result, the effect on economic performance which is measured using plant productivity in this paper will be greater than for those identified as developed economies.

5.1.2. Instrumental Variable Estimation

There is some possible endogeneity bias in the estimators despite the fact that several unobserved variables have been specifically taken into account in section 5.1.1. The inclusion of plant and year-fixed effect only deals with omitted variable bias but not the reverse causality problem. For instance, higher-producing facilities are more likely to be found in areas with a better-educated labour population which leads to the employment of more non-production workers. Therefore, this study tries to solve this type of issue using instrumental variable approach. This study suggests using the number of higher education institutions within regions at one year lag as the variable

© 2024 Conscientia Beam. All Rights Reserved.

of the instrument following Moretti (2004a). However, in this case, this study will restrict my sample to plants located in Indonesia and the Philippines only due to the lack of information about them in Vietnam. [7]

The presence of higher education institutions will be a valid instrument if it is correlated with the regional human capital level and uncorrelated with the error term. This study can test for the former requirement but not for the latter if it has only one instrument. They are often thought to be a suitable instrument for the local human capital level because they are the best place to produce highly-skilled workers. This study uses the lagged number of universities and colleges, so they will not affect current productivity. However, it still could be the case that higher education institutions improve plant productivity beyond the effect that they have on creating a more skilled population. For example, higher education institutions may also bring external dollars into the local area through grants to conduct research. Such a dollar stream might increase local plant production since plants could employ research results for their own benefit through contracts with private companies or public access. Thus, the presence of a higher education institution may be a questionable instrument.

According to Moulton (1990), when micro unit observations such as those of firms, contain aggregate variables, the standard errors will be biased downward if the error term of individual units that have the same value of the aggregate variables exhibits a correlation. In this case, this study assigns the logarithm of total higher education institutions in the regions to each plant based on its location. Hence, test statistics for this approach are sensitive to the type of clustered standard error used. Clustering standard errors at the regional level will lower first-stage F-statistics which have a tendency to view instruments as weak but at the same time, it will lower Hansen-J statistics, increasing the likelihood of interpreting the model as correctly specified. Since the theory is not clear, standard errors will be grouped at the regional level in this study because the spillovers variable fluctuates over time inside the region [8] as well.

Table 4. Estimates of human capital spillovers: IV estimates.

Variables	(1)	(2)
First stage:		
ln (Number of universities)	0.064*** (0.024)	0.060**(0.026)
F statistic	6.73	4.87
Second stage:		
Non-production worker share in region	59.131*** (17.719)	77.528*** (25.418)
ln (Production workers)	0.577* (0.307)	0.734**(0.363)
ln (Non-production workers)	0.252 (0.265)	0.257 (0.322)
ln (Capital)	0.031 (0.076)	0.011 (0.090)
Plant specific control		+

Note: Figures in parentheses are clustered standard errors at the regional level. All regressions includes the plant-fixed effect. Specification (2) include plant characteristics such as years of operation and multi-unit status. The sample size is 318 observations. *** p<0.01, ** p<0.05, * p<0.1.

The two-stage least square estimation results are highlighted in Table 4. The first-stage regressions in both models clearly show that the employment percentage of nonproduction workers and the number of universities are positively correlated. The first-stage coefficients are between 0.060 and 0.064. The instrument's mean (in logarithmic value) is 5.470 which could help in the interpretation of the first-stage regression. The number of higher education institutions accounts for 0.044-0.047 percentage-point increase in the aggregate nonproduction workers share outside the plant or about 1.5-1.7 percent of the typical increase in skilled workers over a six-year period. In addition to that, the coefficients are statistically significant at 1 percent level. This result implies that the proposed instrumental variable satisfies the relevant condition. However, the F-statistics in the first stage

[7] The regional distribution of higher education institutions is only available for Philippines and Indonesia. As a consequence, This study limits the sample to plants located in both countries only.

[8] This study does not apply clustered standard errors at regional level when using OLS estimation because it does not match between two different data sources.

© 2024 Conscientia Beam. All Rights Reserved.

regressions are reasonably small, less than 10. According to Stock and Yogo (2005), this study cannot reject the null hypothesis of the weak instrument at the 5 percent significance level. This indicates that the instrument used in this study cannot predict endogenous variables very well.

The second stage estimation specifies that when the aggregate human capital within the region (as explained by the instrumental variable) rises by 1 percentage point, plant productivity under models (1) and (2) will increase by 59.1 and 77.5 percent over the five-year time span. If we compare the coefficients between the ordinary least square and instrumental variable models, the magnitude of the former is lower than the latter. Thus, one may underestimate the spillovers effect of highly skilled workers when the endogeneity problem is not dealt with. Findings under two-stage least square estimation also show us that our estimates of human capital spillovers are robust and can be stated as causal inference. In a nutshell, the use of the instrumental variable technique yielded results that indicate the presence of external human capital that plants in the region experience positively influences plant productivity.

5.2. Robustness Checks

5.2.1. Estimation using TFP

TFP is the dependent variable in this subsection of this study that serves as a plant productivity indicator. A beneficial application of this metric is to examine the potential relationship between factor inputs and productivity shocks that go unnoticed. If the spillover effect based on TFP is contradictory to the findings in Section 5.1, the previous results may be regarded as unstable ones. When using TFP estimation, this study argues that a correlation may exist between factors of production and productivity shocks. Positive productivity shocks such as new technology incorporated into the production process will encourage plants to raise their inputs because they will also be expected to produce more in the future. Hence, input endowments become endogenous in this case.

If this problem is not addressed properly, results based on the OLS technique may be biased. Thus, Olley and Pakes (1996) suggest a method to address such issue by using the investment decision of the plants to capture the correlation between unobserved productivity shocks and factor inputs. They argued that changes in investment level is a useful instrument to reflect the change in unobservable productivity shocks and these are found to be monotonically increasing.

However, the main drawback of this method is that the level of investment in some plants is usually zero since plants may not invest every year. Hence, Levinsohn and Petrin (2003) point out this limitation and provide a more appropriate instrument using intermediate inputs such as material or electricity costs. This study uses Levinsohn and Petrin's (2003) approach (L&P TFP) as the plant productivity indices due to the data indicating that investment value for several plants is lacking. In this case, this study uses the logarithm of material cost as a proxy to capture the unobserved productivity shocks at the plant level.[9]

Estimation results in Table 4 show that the estimated human capital spillover effects vary from 1.987 to 3.857. An increase of one percentage point in the number of nonproduction workers will result in a 2.4% rise in the TFP of the plant as shown in the third row which represents the most robust TFP specification. There is no statistically significant difference between Tables 2 and 5 despite the fact that the former's coefficients are significantly bigger than the latter.

This outcome indicates that the human capital level in the region has a positive effect on plant productivity under different productivity measures.

[9] In the dataset, nearly 70% of observations have zero investment value.

© 2024 Conscientia Beam. All Rights Reserved.

Table 5. Estimates of human capital spillovers based on TFP.

Variables	(1)	(2)	(3)
Non-production worker share in region	3.857** (2.020)	1.987** (0.970)	2.404* (1.470)
Plant-fixed effect	+		+
Year-fixed effect		+	+

Note: Standard errors are adjusted for clustering on the plant. The dependent variable is Levinsohn and Petrin (2003)'s TFP (L&P TFP). The proxy used is the logarithm of material cost. All specifications include plant characteristics such as years of operation and multi-unit status. The sample size is 388 observations. ** $p<0.05$, * $p<0.1$.

5.2.2. Alternative Measures for Region-Human Capital

This study now examines the robustness of the projected spillover effect in the case of a different region-human capital is employed. Following Liu (2014) the new indicator is the logarithm of the average yearly wage within a region, excluding the plant itself. The identification mainly comes from the variations over time within the region. The underlying assumption is that the wage is a fairly good proxy to determine the level of human capital endowment under conventional wisdom. Firms pay their workers equal to their marginal product and the marginal product is positively correlated with the human capital of workers. Based on this theory, if the proportion of non-production workers is high in any region, one may expect that the productivity in that area will be large as well.

However, firms may not pay their workers marginal product and wages is determined by the profitability of the firms. To deal with this problem, all regressions include the plant-fixed effect and as a result, the spillover effect is identified by within-firm variations in the wage bill. Therefore, as long as firms pay proportionally more non-production workers than production workers and there is no huge variation in profitability across years, this proxy is acceptable.

Table 6. Estimates of human capital spillovers: an alternative measure of region-human capital.

Variables	(1)	(2)	(3)	(4)
Log of average yearly wage in the region	0.381** (0.174)	0.393** (0.183)	0.020 (0.243)	0.034 (0.246)
ln (Production workers)	0.316 (0.227)	0.311 (0.228)	0.274 (0.227)	0.280 (0.226)
ln (Non-production workers)	0.307 (0.240)	0.307 (0.244)	0.288 (0.246)	0.281 (0.247)
ln (Capital)	0.060 (0.097)	0.063 (0.097)	0.087 (0.097)	0.085 (0.097)
Plant specific controls		+		+
Plant-fixed effect	+	+	+	+
Year-fixed effect			+	+
R^2	0.52	0.51	0.52	0.51

Note: Standard errors are adjusted for clustering on the plant. Specification (2) and (4) include plant characteristics such as years of operation and multi-unit status. The sample size is 385 observations. ** $p<0.05$.

This study entirely duplicates the regressions in Table 2 using this new measure. Table 6 shows that the spillover estimates are both positive and statistically significant. Column (1) can provide an example; one percent increase in wage outside the plant within a region is positively associated with overall plant productivity by 0.38%. Additionally, it is also possible to combine both plant-specific characteristics and control for time-varying shocks by adding a year-fixed effect to the model. Surprisingly, column (3) shows that although the projected spillover impact is statistically insignificant, it still has a positive sign. Furthermore, the magnitude is far lower than in the two previous rows. Despite such imperfection, this measure is partially a good proxy for region-human capital.

The potential explanation for this finding is that the variation of plant profitability over the five-year period seems high and regional wage only slightly change during the period. This condition is often called the sticky-wage theory. This theory states that the pay of employed workers tends to respond more slowly to changes in the firm or

© 2024 Conscientia Beam. All Rights Reserved.

economic performance. This study presents an obvious relationship between the average annual pay level in the region outside the plant and the value-added of the plant to further clarify this idea (both are in logarithmic values).

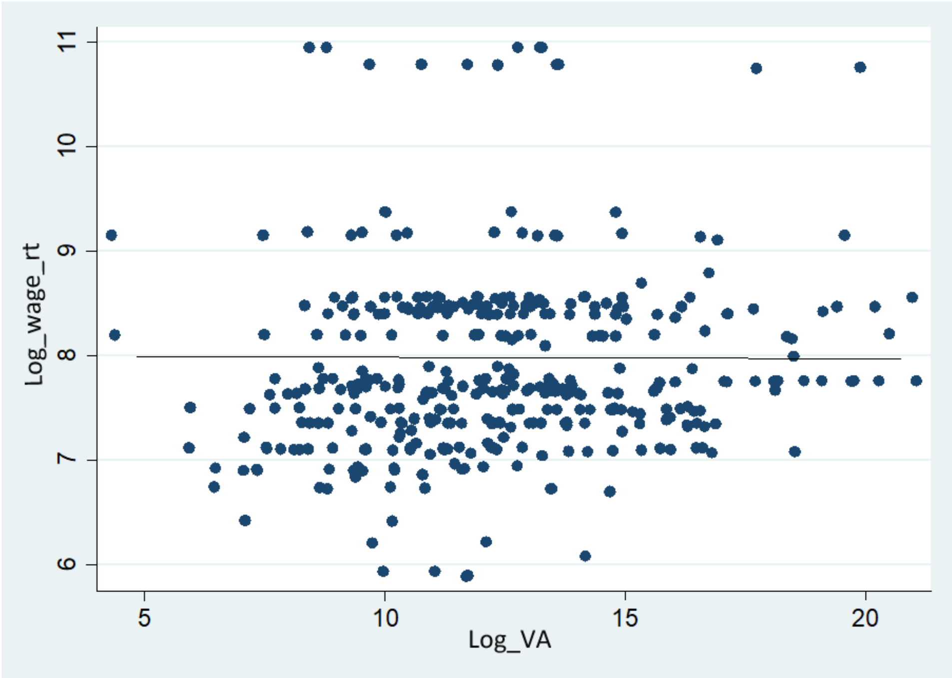

Figure 1. Scatter plot of wage rate and value-added.

Note: All dots included are from both 2009 and 2015.

Figure 1 illustrates that the regional wage levels are not positively attributed to higher plant productivity. Many of the region's productive plants are more likely to pay their employees at a rate comparable to that of less productive plants. Therefore, it cannot be inferred whether the regional wage rate can reflect the aggregate level human capital in the area. According to Solow (1979), employers will provide wage stability regardless of output level in order to avoid costs particularly in the short term. Furthermore, businesses will not lower salary rates even in the event of a decline in output as doing so will negatively impact employee morale and hence lower output. Therefore, it is better for plants to pay a steady salary as it ensures the continuity of production during a crisis when their output declines.

5.2.3. Human Capital Spillovers in High-Tech Industries

Since the impact of spillovers may vary depending on the industry, we now focus our research on the division between high-tech and low-tech sectors. In general, human capital spillovers within high-tech industries are expected to be larger than those in low-tech industries because the importance of human capital in the production process is more substantial for the former type of industry. The underlying argument is that plants in high-tech industries which face rapid change in technological progress are more sensitive to human capital spillovers than plants in low-tech sectors. This study uses an approach that separates each plant within a region into high-tech and low-tech sectors based on the Organisation for Economic Co-operation and Development (OECD) categorization in order to confirm this prediction.[10] This study then goes on calculating the share of non-production workers faced by each plant in high-tech and low-tech industries separately.

[10] The definition of high-tech industries is calculated based on R&D intensities. They consist of high-tech and medium high-tech industries. These include chemicals and pharmaceuticals; aircraft and spacecraft; ordinary machinery; transportation equipment; electrical machinery; electronics and telecommunications; instruments and office machinery; motor vehicles and trailers.

© 2024 Conscientia Beam. All Rights Reserved.

In the next step, this study tests whether human capital externalities are more important in high-tech than low-tech industries. More importantly, this study also examines whether the spillovers from human capital in high-tech industries of the region as a whole are more important for high-tech plants than human capital from low-tech industries of the region and vice versa. Orlando (2004) claims that the size of spillovers between firms depends on sectorial proximity where spillovers among firms in similar industries are usually larger than among firms with different types of industries. Thus, it is expected that the medical industry will benefit more from white-collar workers in the chemical industry than from the food industry. If the outcomes do not meet these expectations, a new pattern in the spillover relationship between high-tech and low-tech industries develops.

Table 7. Plants and worker distribution by industrial technological intensity.

Plant category	Number of plants			Number of nonproduction workers		
	Indonesia	Vietnam	Philippines	Indonesia	Vietnam	Philippines
Low-tech	87	32	30	45.6	13.4	20.7
High-tech	18	3	24	55.6	8.1	57.3

Note: These figures above are averaged between 2009 and 2015.

Table 7 shows the details of the distribution of plants based on technological intensity along with the average numbers of workers employed by each type of industry. We can clearly see that high-tech sectors hire more skilled workers than low-tech plants, although the number of plants in high-tech industries is far lower than their counterparts. However, this pattern does not occur in Vietnam. Each country indicates a similar feature which is consistent with our analysis of summary statistics. The manufacturing sector in the Philippines has the highest proportion of plants in high-tech industries. 40% of total plants in the Philippines are classified as high-tech plants. This coincides with the fact that on average, the number of nonproduction workers per plant hired in the Philippines is 57, the biggest among other countries. The reverse condition does happen in the Vietnamese manufacturing sector. From these figures, we can expect that the size of spillovers will be higher for high-tech rather than low-tech industries.

This study uses the dummy variable for each industry type that interacts with human capital outside the plant making use of it for both high-tech and low-tech industries' estimation approaches. The estimation results are presented in Table 8 with the logarithm of value-added as the dependent variable. The first row focuses on human capital spillovers from high-tech industries on high-tech plants and low-tech plants. Coefficients in the second row similarly show the extent of human capital spillovers from low-tech sectors to high-tech and low-tech plants. All entries in each panel are from the same regression. For instance, the coefficient on the percentage of non-production workers in high-tech industries interacts with a dummy variable in column (1) which equals one if the relevant plant is high-tech. Similarly, in row 1 column (2), the dummy variable equal to one if the relevant plant is low-tech interacted with the coefficient on the fraction of white-collar workers in high-tech sectors.

Results in Table 8 show that low-tech plants reap the benefit of more skilled workers in both high-tech and low-tech plants, with the effect of low-tech industries (12.564) being approximately four times greater than that of high-tech industries (2.834) while there is no significant impact of workers on high-tech plants. The implication for low-tech plants is higher in high-tech industries although the effect is larger in low-tech industries. These findings show that the presence of highly skilled workers will generate spillovers only for low-tech plants. The conclusion drawn from this subsection somehow is not in line with our expectations of human capital spillovers, (i.e., the spillovers between relevant industries have a larger effect and exhibit greater magnitude in high-tech industries). We can infer that the composition of the manufacturing sector in ASEAN countries is dominated by low-tech industries such as the food and textile industries. The importance of such low-skilled industries is enormous for the economies of those countries. Thus, the presence of skilled workers within the region will mainly spill over to them

© 2024 Conscientia Beam. All Rights Reserved.

since the proportion of plants in high-tech industries is not substantial and less influential to have an impact on other plants even for other high-tech plants.

Table 8. Human capital spillovers between high- and low-tech industries.

Variables	High-tech plants (1)	Low-tech plants (2)
Non-production worker share of high-tech industries in the region.	0.301 (2.584)	2.834** (1.153)
Non-production worker share of low-tech industries in the region.	8.530 (7.295)	12.564** (6.508)
Plant-fixed effect	+	+
Time-fixed effect	+	+

Note: Standard errors are adjusted for clustering on the plant. All specifications include plant characteristics such as years of operation and multi-unit status. The sample size is 388 observations. ** p<0.05.

5.2.4. Other Tests for Robustness Checks

This study concludes this subsection by performing a series of alternative specifications to review the results reported in Table 9. Firstly, this study uses the logarithm of sales instead of value-added as the dependent variable, and the result is shown in the first row.[11] The coefficient decreases to 5.212 compared to the value of 7.152 in Table 3 of row (4). This study also divides the plants into size groups of 5-19 employees, 20-99 employees and equal to or more than 100 employees. Then, it runs separate regressions to estimate human capital spillovers within these categories. The results which are documented in rows 2-4 show that no clear pattern arises. Spillovers only occur for small and large plants only while there is an insignificant and small effect for medium plants. Large plants can generate higher value-added from the presence of skilled workers than small plants. One percentage point increase in the proportion of non-production workers will induce plants to produce 14.3% higher value-added for small plants whereas this figure is nearly 19.2% for plants with 100 workers or more.

Table 9. Other robustness checks.

Specifications	The coefficient of white-collar workers share outside plants in the region
(1) Sales	5.212** (2.650)
(2) Small plants (5-19workers)	14.255* (8.377)
(3) Medium plants (20-99 workers)	0.567 (4.950)
(4) Large plants (Workers+100)	19.209*** (6.796)
(5) Single-unit plants	8.285** (3.588)
(6) Multi-unit plants	63.245*** (0.068)
(7) Translog production function	7.063** (3.502)

Note: (2) The number of observations is 89; (3) The number of observations is 159; (4) The number of observations is 140; (5) The number of observations is 333; (6) The number of observations is 55. Multi-unit status is excluded for estimating specifications in row (5) and (6). All regressions include the plant and year-fixed effect. Standard errors are adjusted for clustering on the plant. *** p<0.01, ** p<0.05, * p<0.1.

The next iteration is to separate the plants based on their multi-unit status. Single-plant firms should be more responsive to changes in their local conditions than firms that have plants in several locations (Henderson, 2003). According to this argument, multi-plant firms rely more on their internal-firm connections and hence, these firms

[11] This study uses log of sales because my dataset does not provide information on output.

© 2024 Conscientia Beam. All Rights Reserved.

are more protected from the local external environment than single-plant firms. Thus, it is expected that the magnitude of human capital spillovers is stronger for single-unit plants than multi-establishment plants. Rows 3-4 of Table 8 report the corresponding results for single-unit and multi-unit plants. The findings are surprisingly not consistent with our prediction. The impact of human capital spillovers for multi-unit firms is 63.245 and only 8.460 for single-unit firms. This condition may happen when multi-establishment plants apply a dual human capital networking system, internal and external. These firms take advantage of being located in dispersed locations to absorb regional spillovers better which complements their ability to exploit internal-firm human capital.

This work completes a final specification check using a translog specification, a more versatile type of production function. The square of each log input and the interaction term between each log input are then included in the estimation of Equation 3. The coefficient of interest is somewhat sizeable compared to the main specifications in Table 3. One percentage point increase in non-production workers within the region will increase the value-added of the plants by approximately 7 to 9 percent under the Cobb-Douglas functional form whereas the impact of non-production workers on the plant's value added is expected to be around seven percentage points using the translog production function. This finding implies that the results are robust with regard to the different types of production functions.

6. CONCLUSION

This paper investigates the existence of regional human capital spillovers in ASEAN. In principle, it tests whether plants located in a region with a high level of human capital can produce more output with a similar amount of input. This study also provides an instrumental variable approach to solve the endogeneity issue caused by the correlation between regional human capital level and plant productivity. Additionally, this study also performs some robustness checks to test the validity of the findings under alternative measures or specifications.

The findings clearly show that the percentage of nonproduction workers in the region rises by 1 percentage point and roughly increases plant productivity by 7.15 percent after adjusting for time-varying heterogeneity, unobservable plant characteristics, and the sample plant itself.

The study uses the number of higher education institutions in the preceding year as a tool to assist in a two-stage estimate in order to address the endogeneity problem resulting from unobservable time-varying factors in regional human capital and plant productivity. The results also show that the causal effect of human capital spillovers obviously happens and is significant since we may underestimate the impact when using OLS estimation. However, this instrument is considered a weak instrument.

This study additionally investigates the spillover impact using Levinsohn and Petrin (2003) as the plant productivity indicator, serving as a check on the primary specification. There is no statistical difference between Ordinary Least Square and TFP estimation although the magnitude of spillovers obtained is substantially lower when using the latter. This suggests that the factor endowments are exogenous and uncorrelated with regional human capital spillovers.

In addition, this study also employs the average yearly wage within the region as another independent variable to measure the aggregate level of human capital in the area. The results suggest that the effect of human capital spillovers is still associated with higher productivity, although it cannot confirm the spillover effects when time-varying unobservable factors exist.

This study separates the human capital in a specific region into low-tech and high-tech sectors to investigate the effects of spillover from each kind of business highlighting the spillovers effect across industries within a region. The results show that low-tech and high-tech industries only spill over the benefits of their highly skilled workers to plants in low-tech industries only while the effect from the former is found to be stronger. This kind of situation reflects that industries with highly skilled workers are far less influential in affecting other plants' productivity in developing countries particularly in ASEAN.

© 2024 Conscientia Beam. All Rights Reserved.

6.1. Policy Implications

This study presents some microeconomic evidence that human capital spillovers are robust and significantly benefit productivity in the ASEAN manufacturing industry. The monetary benefit of this spillover effect is large which yields a productivity gain with a rising per plant value-added of US$ 55,000 and a total gain of US$ 866.75 million over the next six years.[12] There could be several explanations for the relatively large magnitude of regional human capital spillovers in ASEAN. Rosenthal and Strange (2008) explain that the benefit of human capital spillovers is inversely related to distance and depends on the proximity to highly skilled workers. Thus, it is more likely that plants in ASEAN are concentrated in the clustered area (e.g., industrial district) and the reason behind it is to have closer access to skilled workers. When this does exist, workers will have more opportunity to interact more with the workers outside their plants as they work in close collaboration and therefore facilitate the flow of ideas. Then, such new knowledge will be translated into greater plant productivity.

Two policy implications can be drawn from our study. First, there is clear evidence that firms are benefiting from locating in areas with abundance highly skilled workers. Policies that want to reduce the regional inequality of human capital levels need to consider this fact carefully and alter the incentives firms have to locate in regions with a higher concentration of human capital density. The same explanation also applies to the labour force. Skilled workers are more likely to move to areas where more industrious jobs are accessible and again trying to adjust the incentive arrangement in this respect is a vital (although challenging) aspect of these policies. Therefore, incentives driven by policies will determine how firms and workers are willing to move to areas with lower quality of human capital spreading the impact on plant productivity around the country. Second, while workers' skills are closely associated with workers' education encouraging the local population to pursue more education is an effective way in the context of regional development. It will raise the stock of skilled labour, and at the same time, it will increase the supply of better human capital which helps plants hire skilled workers at a lower wage. My findings suggest that the establishment of new universities or colleges by the public or private sectors can help accomplish this goal. Policies that promote plants to enhance their proficiency level can also serve as a valuable tool for regional development.

6.2. Limitations

This paper will conclude by providing some limitations to this work regarding any interpretation of the findings. First, this paper does not observe the effect of spillovers separately for each industry with different characteristics such as technology intensity. When technological progress is not similar across industries, some industries may benefit more from the human capital level of industries with similar characteristics. Moreover, since plants in the service sector are not included here, this study cannot investigate whether the spillover effect coming from the service sector can affect manufacturing productivity. Second, this work does not decompose the spillovers within and between industries as this study emphasises regional boundaries. When the connection between two or more industries is strong, the effect of human capital is expected to be large as well. For example, if plants in industries A and B have vertical linkage, the level of human capital in industry A will have a greater impact on industry B and vice versa.

6.3. Future Research Suggestions

Since the presence of spillovers involving human capital has been demonstrated, the continuing goal is to identify the precise mechanism through which these spillovers occur. One of the possible sources is knowledge exchange through foreign technology licensing. When plants acquire foreign technology, it is assumed that only

[12] The number of plants in each year is 2580 (2009) and 2673 (2015). With an average of 2627 plants, the pecuniary benefit for the manufacturing sector is estimated at around US$ 866.75 million.

© 2024 Conscientia Beam. All Rights Reserved.

skilled workers can operate it most effectively. According to this study's analysis of the dataset, about 22% of ASEAN manufacturing sector facilities purchased such foreign technology during both periods. Hence, an investigation into the relationship between foreign technology and the human capital spillover effect will be an interesting topic for further research.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The author states that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The author declares that there are no conflicts of interests regarding the publication of this paper.

REFERENCES

Acemoglu, D. (1996). A microfoundation for social increasing returns in human capital accumulation. *The Quarterly Journal of Economics, 111*(3), 779-804. https://doi.org/10.2307/2946672

Acemoglu, D., & Angrist, J. (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER Macroeconomics Annual, 15*, 9-59. https://doi.org/10.1086/654403

Adams, J. D., & Jaffe, A. B. (1996). Bounding the effect of R&D: An investigation using matched establishment-firm data. *RAND Journal of Economics, 27*(4), 700-721. https://doi.org/10.3386/w5544

Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary Economics, 34*(2), 143-173. https://doi.org/10.1016/0304-3932(94)90047-7

Berman, E., Bound, J., & Griliches, Z. (1994). Changes in the demand for skilled labor within US manufacturing: Evidence from the annual survey of manufactures. *The Quarterly Journal of Economics, 109*(2), 367-397. https://doi.org/10.3386/w4255

Bloom, N., Schankerman, M., & Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica, 81*(4), 1347–1393. https://doi.org/10.3982/ECTA9466

Breusch, T. S., & Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies, 47*(1), 239-253. https://doi.org/10.2307/2297111

Chang, C.-F., Wang, P., & Liu, J.-T. (2016). Knowledge spillovers, human capital and productivity. *Journal of Macroeconomics, 47*, 214-232. https://doi.org/10.1016/j.jmacro.2015.11.003

De La Fuente, A., & Domenech, R. (2001). Schooling data, technological diffusion, and the neoclassical model. *American Economic Review, 91*(2), 323-327.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics, 4*, 2063-2117. https://doi.org/10.1016/s1574-0080(04)80005-1

Eklund, J., & Pettersson, L. (2019). Education, Human capital spillovers, and productivity: Evidence Swedish firm level production functions. In E. Lehmann, & K. Keilbach (Eds.), From Industrial Organization to Entrepreneurship: A Tribute to David B. Audretsch. In (pp. 203-223). Cham: Springer.

Glaeser, E. L. (1999). Learning in cities. *Journal of Urban Economics, 46*(2), 254-277.

Greene, W. (2003). *Econometric analysis* (5th ed.). Upper Saddle River: Prentice Hall.

Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics, 53*(1), 1-28.

Jacobs, J. (1969). *The economy of cities*. New York: Random House.

Jovanovic, B., & Rob, R. (1989). The growth and diffusion of knowledge. *The Review of Economic Studies, 56*(4), 569-582.

Kijek, A., & Kijek, T. (2020). Nonlinear effects of human capital and R&D on TFP: Evidence from European regions. *Sustainability, 12*(5), 1808. https://doi.org/10.3390/su12051808

Kuznets, S. (1962). *Population change and aggregate output. In Demographic and Economic Change in Developed Countries*. Princeton: Princeton University Press.

© 2024 Conscientia Beam. All Rights Reserved.

Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies, 70*(2), 317-341. https://doi.org/10.1111/1467-937x.00246

Liu, Z. (2014). Human capital externalities in cities: Evidence from Chinese manufacturing firms. *Journal of Economic Geography, 14*(3), 621-649. https://doi.org/10.1093/jeg/lbt024

Lu, Z. (2022). Human capital spillovers from special economic zones: Evidence from Yangtze Delta in China. *Economies, 10*(5), 99. https://doi.org/10.3390/economies10050099

Lucas, J. R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22*(1), 3-42.

Lynne, G., Michael, R. D., & Marilynn, B. B. (1998). Intellectual of human capital and the birth of US biotechnology enterprises. *The American Economic Review, 88*(1), 290-306. https://doi.org/10.1093/oso/9780199256938.003.0008

Marshall, A. (1920). *Principles of economics* (8th ed.). London: McMillan.

Moretti, E. (2004a). Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics, 121*(1-2), 175-212. https://doi.org/10.1016/j.jeconom.2003.10.015

Moretti, E. (2004b). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *American Economic Review, 94*(3), 656-690. https://doi.org/10.1257/0002828041464623

Moulton, B. R. (1990). An illustration of a pitfall in estimating the effect of aggregate variables on micro unit. *The Review of Economics and Statistics, 72*(2), 334-338. https://doi.org/10.2307/2109724

Olley, S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica, 64*(6), 1263-1297. https://doi.org/10.2307/2171831

Orlando, M. (2004). Measuring spillovers from industrial R&D: On the importance of geographic and technological proximity. *RAND Journal of Economics, 35*(4), 777-786. https://doi.org/10.2307/1593773

Park, W. (2016). How large are human capital spillover? Evidence from Korea. *KDI Journal of Economic Policy, 38*(4), 77-94. https://doi.org/10.23895/kdijep.2016.38.4.77

Rauch, J. E. (1993). Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics, 34*(3), 380-400. https://doi.org/10.1006/juec.1993.1042

Rosenthal, S. S., & Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics, 64*(2), 373-389. https://doi.org/10.1016/j.jue.2008.02.006

Solow, R. M. (1979). Another possible source of wage stickiness. *Journal of Macroeconomics, 1*(1), 79-82. https://doi.org/10.1016/0164-0704(79)90022-3

Stock, J., & Yogo, M. (2005). Testing for weak instrument in linear IV regression. In Identification and Inference for Econometric Models. In (pp. 80-108). New York: Cambridge University Press.

World Bank. (2017). *Survey methodology for enterprise surveys*. Retrieved from http://www.enterprisesurveys.org/methodology

APPENDIX

Appendix A. Supplementary results.

World Bank Enterprise Survey covers 2580 plants in 2009 and 2673 plants in 2015 for three ASEAN countries combined. First of all, this study calculates the share of nonproduction workers and average yearly wage in the region, outside the plant, from whole manufacturing plants in the dataset. Then, to obtain a balanced sample used in this paper, this study excludes all plants that do not appear in both years. It also deletes some of the relevant variables that have missing values for at least one year. The resulting balanced sample has 194 plants in each year of our analysis. This dataset provides information on sales, capital, production and nonproduction workers, industry, and the location.

Instead of using matched plant-labour survey dataset, this study constructs the share of nonproduction workers in particular region internally from Enterprise Survey. This measure will not lead to multicollinearity with the amount of production and nonproduction workers in each plant because the reference plant is omitted from the calculation.

© 2024 Conscientia Beam. All Rights Reserved.

However, two datasets are matched using the lagged number of higher education institutions when in IV method. The figures for them are taken from Statistical Yearbook, more specifically in the education section. Then, each of them is assigned to the dataset and take logarithmic value.

To assess the quality of the regional human capital variable, this study chooses average monthly wage in the region as the alternative indicator. Data on wages, from Enterprise Survey, are plant averages by dividing the total labour cost by total workers times 12, excluding the relevant plant. It is imputed by assuming that nonproduction workers are paid more than production workers. Therefore, when nonproduction workers rise, wage bill for the plants will increase as well.

Enterprise Surveys uses stratified random sampling methodology to collect the key variables from plants. It implies that they are pooled into specific strata or attributes and then having equal chance to be selected in the survey to form a random sample. The main advantage of this methodology is it can capture plants that are proportional to the overall plant population with various characteristics. The strata used in this survey are based on firm size, business sector, and region. Size of the plants are separated into small (between 5-19 employees), medium (between 20-99 employees), and large (more than or equal to 100). Plants are also categorised into either manufacturing, service, or other service sector. Finally, the survey also groups firms depending on their location.

Table A1 presents the total plant in each region within each country. Table A2, Table A3, Table A4, and Table A5 exhibit the summary statistics of variables for all combined countries and each country of Indonesia, Philippines, and Vietnam respectively. Table A6 shows the share of nonproduction workers in each region in all countries. Table A7 describes the regression result when it is estimated using Equation 3 for each data period. Lastly, Table A8 and A9 reports the regression results of Equation 3 using Pooled OLS and Random Effect estimation techniques.

Table A1. Number of plants, by region.

Country	Region	Number of plant	Percentage to total plant
Indonesia	Bali	14	7.1%
	Banten	26	13.3%
	DKI Jakarta	18	9.2%
	Jawa Barat	29	14.8%
	Jawa Tengah	8	4.1%
	Jawa Timur	4	2.0%
	Lampung	2	1.0%
	Sulawesi Selatan	1	0.5%
	Sumatera Utara	3	1.5%
Philippines	Calabarzon	15	7.7%
	Central Luzon	8	4.1%
	Metro Cebu	4	2.0%
	Metro Manila	3	1.5%
	NCR excluding Manila	24	12.2%
Vietnam	Central North	3	1.5%
	Mekong River Delta	2	1.0%
	North Central and Central Coastal	4	2.0%
	Red River Delta	8	4.1%
	South East	16	8.2%
	Southern Central Coastal	2	1.0%
Total		196	100%

Note: Plants are identified as which appear in both year (2009 and 2015).

© 2024 Conscientia Beam. All Rights Reserved.

Table A2. Summary statistics: Overall.

Variables	Overall		2009		2015	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Sales (*1000)	42795.3	307764.8	7658.1	26384.0	77980.1	432147.5
Value added (*1000)	21941.7	121024.7	3876.0	14576.6	40005.9	168823.5
Capital (*1000)	1495.2	5605.6	1542.8	6217.0	1447.7	4935.0
Materials (*1000)	7659.7	35575.3	2815.1	11108.1	12574.8	48828.7
Energy (*1000)	534.3	7720.9	159.1	968.5	932.0	11025.2
Number of production workers	141.3	398.7	145.8	472.4	136.7	309.0
Number of nonproduction workers	38.2	124.0	35.9	84.6	40.6	153.9
Nonproduction workers share in region	0.206	0.066	0.200	0.055	0.213	0.075
Average yearly wage in region	4647.9	9041.7	2234.3	1245.7	7099.4	12315.0
Age of plants	19.9	12.0	17.3	12.0	22.5	11.4
Proportion of plants belong to multi-unit firm	0.14	0.35	0.15	0.36	0.13	0.34

Note: (*1000) indicates that all monetary values are reported in thousand US Dollar. Nonproduction worker share and average yearly wage are calculated outside the plant.

Table A3. Summary statistics: Indonesia.

Variables	Overall		2009		2015	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Sales (*1000)	40260.5	154724.9	5988.6	14704808	74532.3	213358.5
Value added (*1000)	29140.6	134821.9	2690.5	8258.7	55590.6	187210.6
Capital (*1000)	1397.6	6010.1	1780.5	7190.2	1014.8	4538.3
Materials (*1000)	9245.7	42876.4	2411.5	4606.1	16080.0	59824.0
Energy (*1000)	813.0	10298.3	145.5	963.2	1480.5	14536.3
Number of production workers	181.5	510.2	182.3	611.8	180.7	385.8
Number of nonproduction workers	47.4	156.2	40.7	80.9	54.0	205.9
Nonproduction workers share in region	0.200	0.045	0.193	0.037	0.207	0.051
Average yearly wage in region	2165.4	1742.2	1478.1	500.0	2852.8	2212.6
Age of plants	22.1	12.2	19.7	12.6	24.5	11.2
Proportion of plants belong to multi-unit firm	0.15	0.36	0.22	0.42	0.09	0.28

Note: (*1000) indicates that all monetary values are reported in thousand US Dollar. Nonproduction worker share and average yearly wage are calculated outside the plant.

Table A4. Summary statistics: Philippines.

Variables	Overall		2009		2015	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Sales (*1000)	72926.6	542089.6	13914.7	44509.7	131938.6	764322.9
Value added (*1000)	21571.5	131364.3	8310.2	24646.9	34832.9	184046.0
Capital (*1000)	2279.1	6398.5	1770.0	6173.2	2788.1	6634.5
Materials (*1000)	8073.6	30628.8	4163.3	19480.2	12214.0	38923.7
Energy (*1000)	275.0	1246.6	268.8	1252.7	282.0	1253.0
Number of production workers	98.5	209.6	109.5	242.5	87.4	172.1
Number of nonproduction workers	37.0	85.4	41.4	112.6	32.5	44.8
Nonproduction workers share in region	0.267	0.056	0.258	0.019	0.280	0.076
Average yearly wage in region	10958.2	15468.3	4051.3	618.7	18271.3	19778.7
Age of plants	20.3	11.2	17.5	11.0	23.2	10.8
Proportion of plants belong to multi-unit firm	0.17	0.37	0.07	0.26	0.26	0.44

Note: (*1000) indicates that all monetary values are reported in thousand US Dollar. Nonproduction worker share and average yearly wage are calculated outside the plant.

© 2024 Conscientia Beam. All Rights Reserved.

Table A5. Summary statistics: Vietnam.

Variables	Overall		2009		2015	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Sales (*1000)	3911.5	8096.5	2749.4	8060.9	5073.6	8079.3
Value added (*1000)	1024.6	2463.5	590.9	1447.1	1458.3	3136.0
Capital (*1000)	578.7	1224.0	478.9	657.4	678.4	1608.7
Materials (*1000)	2280.8	5852.3	1946.1	6566.0	2615.5	5115.0
Energy (*1000)	48.6	174.1	30.8	60.3	68.7	246.6
Production workers	86.7	155.0	92.4	145.3	80.9	166.0
Nonproduction workers	12.9	14.7	13.0	12.1	12.8	17.1
Nonproduction workers share in region	0.129	0.033	0.128	0.039	0.130	0.027
Average yearly wage in region	2629.9	1286.0	1699.6	377.8	3560.1	1196.4
Age of plants	12.7	9.9	10.1	9.0	15.2	10.4
Proportion of plants belong to multi-unit firm	0.07	0.26	0.09	0.28	0.06	0.24

Note: (*1000) indicates that all monetary values are reported in thousand US Dollar. Nonproduction worker share and average yearly wage are calculated outside the plant.

Table A6. Proportion of nonproduction workers, by region.

Country	Region	Proportion of nonproduction workers		
		Overall	2009	2015
Indonesia	Bali	0.201	0.206	0.197
	Banten	0.218	0.176	0.261
	DKI Jakarta	0.209	0.175	0.243
	Jawa Barat	0.196	0.230	0.162
	Jawa Tengah	0.104	0.103	0.105
	Jawa Timur	0.216	0.242	0.190
	Lampung	0.206	0.177	0.235
	Sulawesi Selatan	0.210	0.204	0.216
	Sumatera Utara	0.240	0.200	0.280
Philippines	Calabarzon	0.250	0.279	0.202
	Central Luzon	0.333	0.267	0.379
	Metro Cebu	0.161	0.224	0.099
	Metro Manila	0.300	0.262	0.307
	NCR excluding Manila	0.271	0.245	0.298
Vietnam	Central North	0.128	0.128	0
	Mekong River Delta	0.197	0.239	0.155
	North Central-Central Coastal	0.143	0	0.143
	Red River Delta	0.155	0.145	0.165
	South East	0.100	0.096	0.104
	Southern Central Costal	0.167	0.167	0

Note: Figures are computed excluding the plants itself. The sample of plants used each year is balanced panel dataset. The figures above include the excluded plants due to certain circumstances. The sample size is 388 observations.

Table A7. Cross-sectional estimates of human capital spillovers.

Variables	2009	2015
Nonproduction worker share in region	4.252* (2.341)	3.270 (2.358)
ln (Production workers)	0.289** (0.134)	0.776*** (0.191)
ln (Nonproduction workers)	0.721*** (0.157)	0.677*** (0.219)
ln (Capital)	0.367*** (0.065)	0.121 (0.097)
Plant-specific controls	+	+
R^2	0.63	0.51

Note: Robust standard errors are in parentheses and each column is a separate regression and has 194 observations. *** p<0.01, ** p<0.05, * p<0.1.

© 2024 Conscientia Beam. All Rights Reserved.

Table A8. Pooled OLS estimates of human capital spillovers.

Variables	**(1)**	**(2)**
Nonproduction worker share in region	3.887** (1.655)	4.061** (1.700)
ln (Production workers)	0.564*** (0.122)	0.566*** (0.116)
ln (Nonproduction workers)	0.728*** (0.137)	0.731*** (0.149)
ln (Capital)	0.201*** (0.074)	0.191** (0.075)
Plant-specific controls		+
R^2	0.53	0.53

Note: Robust standard errors are in parentheses. The sample size is 388 observations. *** p<0.01, ** p<0.05.

Table A9. Random effect estimates of human capital spillovers.

Variables	(1)	(2)
Nonproduction worker share in region	4.260** (1.762)	4.380*** (1.684)
ln (Production workers)	0.573*** (0.119)	0.575*** (0.115)
ln (Nonproduction workers)	0.711*** (0.140)	0.712*** (0.123)
ln (Capital)	0.191** (0.076)	0.183*** (0.047)
Plant-specific controls		+
R^2	0.53	0.53

Note: Robust standard errors are in parentheses. The sample size is 388 observations. *** p<0.01, ** p<0.05.

Views and opinions expressed in this article are the views and opinions of the author(s), Humanities and Social Sciences Letters shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.

© 2024 Conscientia Beam. All Rights Reserved.