For citation: Hoang H. T., Huynh L. T. D. & Chen G. S. (2019). How New Economic Geography Explain Provincial Wage Disparities: Generalised Methods of Moments Approach. Ekonomika Regiona [Economy of Region], 15(1), 205-215

**doi 10.17059/2019-1-16**
**JEL: F12, F15, J31, R11, R12**

***H. T. Hoang*** [a], ***L. T. D. Huynh*** [b], ***G. S. Chen*** [c]
[a] Duy Tan University (Da Nang, Vietnam)
[b] The University of Da Nang (Da Nang, Vietnam; e-mail: linhhtd@due.edu.vn)
[c] University of New England (Armidale, NSW, Australia)

# HOW NEW ECONOMIC GEOGRAPHY EXPLAINS PROVINCIAL WAGE DISPARITIES: GENERALISED METHODS OF MOMENTS APPROACH [1]

*In recent years, new and promising developments have made new economic geography (NEG) a popular framework for examining the spatial distribution of economic activity around the world. A major NEG prediction is that wages are higher in regions with a large market and easy access to suppliers of intermediate inputs. Based on this principle, we examine this hypothesis by using provincial data in Vietnam. Since the Doi Moi reform in 1986, the Vietnamese economy has successfully transformed from a centrally planned to a market-based economy. The Vietnamese experience has been seen as a valuable case study for other economies in transition. This paper estimates a structural model of NEG using Vietnamese provinces data for the period 2000 — 2012. Using instrumental variables based on the principle of generalised method of moments (GMM), we take account of potential endogeneity problem between regressors. We provide evidence that the industrial linkages and trade costs are statistically significant and quantitatively important in explaining variation in provincial wages. This finding is robust to controlling for a wide range of considerations. Moreover, local governments need to strengthen human capital as the first step towards improving average wages. Indeed, a better education system is not only instrumental to raising average earnings per worker, but also crucial for minimizing income inequality in the long run.*

**Keywords:** New economic geography, agglomeration economies, industrial linkages, market access, trade costs, wage equation, living costs, human capital, provincial panel data, generalised method of moments, Vietnam

## 1. Introduction

Deriving from Krugman's [1] seminal work on trade theory, the use of new economic geography (NEG) has emerged as a trend in the study of the spatial distribution of economic activity. The central premise of NEG rests on the notion that economic agglomeration is a self-reinforcing phenomenon and responsible for creating regional inequality. As summarised in the seminal books by Fujita, et al. [2] and Baldwin, et al. [3] the NEG literature provides rigorous theoretical expositions on issues ranging from patterns of economic agglomeration and the origin of regional inequalities, to spillovers in the industrial cluster. Since the NEG-based models often offer clear a priori expectation on a particular effect, Head and Mayer [4] argued strongly for these models to be tested empirically. However, despite a rich theoretical contribution, NEG-based models have been criticised for lacking empirical support. In the comprehensive literature review of Redding [5], a recurring theme is that empirical research using NEG remains sparse, relative to its rich theoretical predictions to date. In this light, our study aims to test the empirical validity of NEG in the context of Vietnamese economy.

One of the key results from NEG is that the interaction between agglomeration and dispersion forces determines the nominal wage in the region [4]. In general, the existing literature has used NEG to explain international income inequalities and inter-regional wage gaps. The findings from these lines of research suggest that nominal wages can be explained by market access, cost-of-living, industrial linkages and trade cost [5]. For example, after examining the 1994 bilateral trade data among 101 countries, Redding and Venables [6] concluded that both market access and supplier access have positive effects on national income. Investigating the nominal wage variations across 3,075 US counties, Hanson [7] found that local wage rates increase with proximity to large market. Similarly, Steven, et al. [8] confirmed this positive market access–local wage nexus for German districts.

Turning to developing countries, Ma [9] examined the wage gap between the coastal and in-

[1] © Hoang H. T., Huynh L. T. D., Chen G. S. Text. 2019.

land areas in China.[1] Using provincial data for the period 1990–2000, the results showed that better market access and supplier access accounted for the higher wage in the coastal provinces.[2] In Indonesia, Amiti and Cameron [11] found that, on average, firms located in the best market-access and supplier-access region were willing to pay a wage premium up to 20 %. Using the 1999 survey data, Fally et al. [12] also found a positive association between local wage, market access and supplier access in Brazilian states and industries. In this paper we apply NEG model to examine the wage determinations using Vietnamese provincial data for the period of 2000–2012. The contribution of this paper is threefold.

Firstly, we want to establish whether the Krugman [1] model can be verified, that is to say we want to find if the key model parameters are significant and offer a meaningful description of the features of Vietnamese economy. The main equation to be estimated will be a wage equation which indicates that wages will be higher in those regions that have easy access to market but falls with trade costs. The main geographical unit of analysis in this study is the Vietnam's provinces.

Secondly, following the works of Helpman [13] and Hanson [7], we extend Krugman (1991) model by introducing a non-tradeable good into the model.[3] Specifically, we use the cost of living in province as it may not only include the price of manufacturing varieties but also the price of the non-tradeable good.

Finally, to the best of our knowledge, this is the first attempt to systematically analyse wage determination in Vietnamese provinces through the lens of NEG. Therefore, our paper's results will contribute to not only the existing empirical literature of economic geography but also policy makers in Vietnam and other transitional economies.

The rest of this paper is organised as follows. Section 2 discusses the theoretical background that motivates the empirical strategy. Section 3 provides a description of the datasets and empirical methodology. The empirical results are discussed in section 4, which is then followed by concluding remark section.

## 2. Theoretical framework

### *2.1. The logic*

We begin by assuming two symmetric regions, namely the North (*N*) and the South (*S*). If, for some reason, workers migrate from region *N* to region *S* and spend their income locally we would expect an expansion of the market size in region *S*. This expansion, according to Krugman [1], entices firms to follow these migrant workers and relocate in the enlarged region. Such a decision is easily justifiable for market-seeking firms as proximity to final consumers provides a greater scope for realising increasing returns to scale in the production process. Meanwhile, it is also in the best interest of efficiency-seeking firms to be located in the enlarged region where labour supply is plentiful. Apart from these increasing-returns and labour-availability considerations, the clustering of firms in a single location helps to foster deeper and wider industrial linkages, strengthening the extent of external economies that benefit all firms in the vicinity [14]. In the words of Fujita and Thisse [15], these aforementioned developments are typical examples of pecuniary and technological externalities brought about by economic agglomeration.

However, given these aforementioned agglomeration forces it is tempting to conclude that economic activity would eventually concentrate in a single location only, or as Baldwin, et al. [16] put it, catastrophic agglomeration. Clearly, such a conclusion is premature and neglects the fact that there are countervailing factors working against the benefits of economic agglomeration. In our earlier example, one such factor can arise from a temporary reduction in prices and a sudden surge in wages brought about by fierce competition in the goods and labour markets in region *S*. Furthermore, the influx of migrant workers increases the demand for the limited supply of tradeable and non-tradeable goods, pushing up the cost of living in region *S*. This cost-of-living pressure forces firms in region S to remunerate their workers better; firms that do not pay well stand to lose highly productive workers to their competitors [3]. This rising marginal cost on the one hand, and falling marginal revenue on the other, gradually erode any economic profit to be made from the decision to relocate. It is worth noting that during this transition those firms choos-

[1] Although less prosperous regions had experienced increased economic growth during the last decade as a result of preferential policies and investment, regional disparities between coastal and interior regions continued to be a general phenomenon in China [10].

[2] Ma [9] separated his sample into three types of ownership, namely FIEs, state-owned enterprises and collective-owned enterprises.

[3] Krugman's model, namely the core-periphery model, has been criticised for lacking analytical tractable solutions [2]. Therefore, apart from reflecting reality, the introduction of the non-tradeable consumption goods is critical for obtaining tractable analytical solutions.

ing to stay in region *N* receive higher markups.[1] Similarly, those workers who remained in region *N* enjoy higher wages because of the fall in labour supply. Even if the fall in labour demand dominates the fall in labour supply in region *N*, possibly caused by fewer firms hiring workers in the region, workers there can still benefit from the falling cost of living, preserving the purchasing power of their wages. Taken together, these countervailing factors weaken the incentives for firm relocation and partly explain why economic activity remains dispersed across regions [13].

Another factor working against catastrophic agglomeration is the effect of distance in shipping tradeable goods between regions. Clearly, location choice is inconsequential if firms can ship goods to and from the markets free of charge. In reality, however, transporting goods from the place of manufacturing to the place of consumption can be costly. In order to reduce the cost of trade, firms have every incentive to locate in proximity to the final market [17]. It is not difficult to see how catastrophic agglomeration can emerge as a result of firms' attempting to be nearer their consumers, especially when trade cost is prohibitive. By the same token, we can infer that economic activity will spread across regions as trade cost starts to fall. In fact, Puga [18] used numerical simulations to show that agglomeration and dispersion forces respond differently to the same change in trade cost. Baldwin, et al. [16] elaborated this asymmetric response by demonstrating that falling trade cost erodes the relative strength of dispersion forces faster than agglomeration forces, helping to explain why economic activity remains concentrated in certain regions rather than in one single location when trade cost is excessive.

In the next section, we describe the modified core-periphery (*CP*) model on two grounds [1, 7, 13]. Firstly, the modified *CP* model is able to account for the huge wave of internal labour migration in Vietnam. And secondly, this model places an explicit emphasis on the role of non-tradeable consumption goods in shaping the regional distribution of economic activity in Vietnam [7, 18].

### *2.2. The model*

We begin by assuming that there is a finite set of ex ante homogenous provinces (*J*) in the country, and that manufacturing firms there operate under the Dixit-Stiglitz monopolistic competition. Consistent with the convention, the markup pricing rule for a profit-maximising manufacturing firm is given by:

$$p_{ij} = \left(\frac{\sigma}{\sigma-1}\right)\beta W_j,\ j = 1, ..., J, \qquad (1)$$

where $p_{ij}$ denotes the price of manufacturing variety $m_j$ produced in province *j*, but consumed in province *i*. Recall that one of the main dispersion forces in the model is trade cost. Here, we assume the trade cost for transporting $m_j$ from province *j* to province *i* takes the "iceberg" form so that $p_{ij} = p_j T_{ij}$, with $T_{ij}$ capturing the notion of a given portion of $m_j$ melt en route.[2] $W_j$ denotes the wage rate in province *j*, and $\beta w_j$ the marginal cost. In equation (1) both $\beta$ and $\sigma$ are constant parameters, with $\sigma > 1$ as the elasticity of substitution between manufacturing varieties and $\beta$ as the unit input coefficient for labour.

The total demand in province *i* for manufacturing variety $m_j$ produced region *j* is given by:

$$m_{ij} = p_{ij}^{-\sigma}\mu Y_i G_i^{\sigma-1}, \qquad (2)$$

$$G_i = \left[\sum_{j=1}^{J} n_j \left(p_j T_{ij}\right)^{(1-\sigma)}\right]^{1/(1-\sigma)}, \qquad (3)$$

where $Y_i$ is the income in province *i* and $\mu$ the constant share of province *i*'s income spent on $m_j$. Let us digress to (3) for the moment. Now if we define $n_j$ as the number of manufacturing varieties in province *j*, then we can interpret $G_i$ as the price index in province *i* for all manufacturing varieties produced in province *j*. Intuitively, (3) describes the market equilibrium for manufacturing varieties in province *i* whose price index increases with the proportion of manufacturing varieties that must be imported from distant provinces.

According to the markup pricing rule from equation (1), the zero-profit condition under the Dixit-Stiglitz monopolistic competition suggests the break-even supply of manufacturing variety mj to be given by:

$$m_j = \frac{\alpha(\sigma-1)}{\beta}, \qquad (4)$$

where $\alpha$ is fixed cost of labour. We know that the market is in equilibrium when demand equals supply, and in our model, this point occurs when (2) is equal to (4):

[1] Following our earlier discussion the reduction of firms in region *N* implies less competition in the goods and labour markets. Unlike firms in region *S*, firms in region *N* receive higher prices for their outputs and pay lower wages to their workers. As a result, these firms in region *N* enjoy higher markups since their marginal revenue exceeds marginal cost.

[2] Given that our focus is on shipping $m_j$ from province *j* to province *i* we will assume that any shipping of $m_j$ in side province *j* incurs zero trade cost, namely $T_{ij} = 1\ \forall\ i = j$. Meanwhile, we ignore export costs and assume that all goods are consumed domestically.

$$\mu p_{ij}^{-\sigma} Y_i G_i^{\sigma-1} = \frac{\alpha(\sigma-1)}{\beta}. \quad (5)$$

Since we already have $p_{ij}$ from (1), we can substitute it into (5) and solve for $W_j$. After manipulations, the equilibrium nominal wage rate for manufacturing workers in province *j* is given by:

$$W_j = \theta\left(\sum_{j=1}^{J} Y_j T_{ij}^{1-\sigma} G_j^{\sigma-1}\right)^{1/\sigma}. \quad (6)$$

In the literature, (6) is commonly referred to as the wage equation by Redding and Venables [6] or the market-potential function by Hanson [7]. Intuitively, given the elasticity of substitution σ, this equation shows that the wage rate rises with provincial income $Y_j$ and manufacturing price index $G_j$[1], but falls with trade costs $T_{ij}$.

Despite shedding qualitative insights into the determination of provincial wage rate, (6) has been criticised for lacking analytical tractable solutions [2]. To counter this criticism, Helpman [13] and Hanson [7] introduced a non-tradeable good into the model. Specifically, they assumed that the supply of this non-tradable good $H_j$ is fixed in province *j* and the consumers there exhibit identical preferences defined by the Cobb-Douglas utility function:

$$U_j = M_j^{\mu} H_j^{1-\mu}, \quad (7)$$

where $M_j$ represents the consumption of all manufactured varieties in province *j*, $H_j$ the consumption of the non-tradeable good in province *j*, $(1 - \mu)$ the share of income spent on the non-tradable good and μ the share of income spent on all manufacturing varieties.

For tractability, we will assume that the expenditure on this non-tradeable good in province *j* takes the following form:

$$P_j H_j = (1-\mu) Y_j, \quad (8)$$

where $P_j$ is the price of the non-tradable good in province *j*, Intuitively, (8) states the equilibrium condition between the supply and demand of the non-tradeable good.

Finally, recall that workers migrate from one province to another in response to the difference in provincial wage rates. We can infer from this point that inter-regional labour migration ceases when the real wage rate in provinces *i* and *j* is equalised, that is,

$$W_i \Big/ \left(P_i^{1-\mu} G_i^{\mu}\right) = W_j \Big/ \left(P_j^{1-\mu} G_j^{\mu}\right) = \bar{\omega}. \quad (9)$$

[1] This is also a measure of the cost-of-living for province *j*.

In effect, (9) implies that the real wage rate $\bar{\omega}$ must be identical in provinces *i* and *j* in the long run. According to Hanson [7], such equilibrium is brought about by that fact that workers are free to migrate from the province with high cost of non-tradeable good to a low cost one. Given that the stock of non-tradeable good is fixed in each province this migration of workers eventually equalises the real wage rate across all provinces. To state this differently, the term $P_i^{1-\mu} G_i^{\mu}$ in (9) is a broadened concept of the cost of living in province *i* as it not only includes the price of manufacturing varieties $G_i$, but also the price of the non-tradeable good $P_i$.

Based on (6) and (9), we arrive at the benchmark wage equation in our empirical analysis by expressing the wage in province *i* at time *t* as a log-linear function of agglomeration and dispersion force:

$$\log W_{it} = \alpha_0 + \beta_1 \log Y_{it} + \beta_2 \log P_{it} + \\ + \beta_3 \log N_{it} - \beta_4 T_{it} + \varepsilon_{it}, \quad (10)$$

where subscript *i* refers to the province and *t* the time period. From the left-hand side of (10), *W* denotes the provincial wage rate, *Y* the extent of market access, $P_{it}$ the level of cost-of-living, $N_{it}$ the scope of industrial linkages, $T_{it}$ the range of trade cost and $\varepsilon_{it}$ is the stochastic error term. Intuitively, (10) suggests that the provincial wage rate increases with market access, but decreases with trade cost. This interpretation is consistent with the firms' preference to produce in the large market and serve foreign ones from there. We expect the provincial wage rate to rise with the firms' ability to access markets, both at home and abroad, because it increases the demand for labour, creating a tight labour market in the short run. Apart from a tight labour market, higher cost-of-living in the province erodes the real purchasing power of wages that prompts firms to pay more, in order to retain the existing workers and attract new ones. Similar to market access, industrial linkages have a positive effect on the provincial wage rate, due to the intense competition among firms for the limited supply of qualified workers in the province.

## 3. Measurements, data and econometric issues

### *3.1. Measurements*

*3.1.1. Provincial wage rate*

We calculate the dependent variable, provincial wage rate (*AW*) by dividing total wage bill by total head count workers in the province [19]. Compared to alternative wage series based on

household or enterprise surveys, the provincial wage rate in this study offers advantages in that it covers all dependent employees, both full-time and part-time workers and includes earnings from all wage and salary jobs.

*3.1.2. Market access*

The most frequently used measure of market access in related literature is regional output [7, 20, 21]. In this study, we selected provincial product per capita, GPPC, as the proxy for market potential, on the basis that it is positively related to local market size and controls the effects of population size [20, 22, 23]. In general, we expect a positive effect on provincial wage rate. [1]

*3.1.3. Living costs*

In previous studies housing stock was commonly used as the proxy for cost-of-living [7, 8].[2] However, as similar data was unavailable at the provincial level in Vietnam during our sample period, we relied on the survey data taken from the Vietnam Household Living Standard Survey (VHLSS). This data provided the cost-of-living (LC) for each province in 2002, 2004, 2006, 2008, 2010 and 2012. We expect cost-of-living to have positive effect on provincial wage rate.

The VHLSS is a household survey conducted by the General Statistical Office (GSO) of Viet Nam every two years. The survey collected information through face-to-face interviews conducted by interviewers with household heads and key commune officials in communes containing sample enumeration areas. Sample of the VHLSS is selected in the way to represent the entire country, regions and provinces/cities.[3]

*3.1.4. Industrial linkages*

Following Aitken, et al. [26], we captured the notion of industrial linkages by the level of aggregate fixed assets in the province (FA), on the basis that it covers the investment status of firms in the province. We expect this measure will have positive effect on provincial wages.

*3.1.5. Trade cost*

Since there is no publicly available measure for trade cost in Vietnam, we used the volume of freight traffic (FT), which is the product of the volume of goods freight and the actual distance freighted on the road system in the province.[4] Logically, a high volume of freight traffic reflects the long distance that goods have to travel on the road in order to enter and exit the province.[5] Furthermore, high freight traffic can be an indicator of severe congestion in the road system, as heavy reliance on this mode of transportation possibly reflects the lack of alternative transportation modes such as rail and air. Taken together, a high volume of freight traffic is correlated with high trade costs within the province and this proxy is expected to have a negative effect on the provincial wage rate.[6]

*3.1.6. Human capital*

According to Hanson [7] and Head and Mayer [19], the extent of regional wage gap is partly determined by the stock of human capital in the region. In order to take this into account, we also introduced the human capital variable (HS) as an additional regressor in (10). Specifically, we measured HS by tertiary enrolment, as a share of the population in the province. This data is collected from Statistical Year Book of the GSO. Additionally, Head and Mayer [19] suggested that market access could have an indirect effect on wages via the channel of encouraging higher levels of education. Therefore, we constructed the interaction term (INT) as the product of GPPC and HS, and introduced it to (10), in order to capture the interaction of the host province's level of human capital and market access, which could raise provincial wages.

### *3.2. Data*

Our analysis is based on the dataset consisting of 60 Vietnamese provinces for selected years during the 2000–2012 period. The variables for wages, market access, industrial linkages and trade costs were obtained from the Statistical Year Book published by Vietnam's General Statistics Office (GSO). Data for living costs was obtained from the 2012 VHLSS conducted by GSO. All nominal variables were then converted into 1999 constant price using the provincial consumer price index (CPI).[7]

[1] We also consider alternative measures of market access including provincial GDP (GPP) and provincial population size (POP) and use them to check the robustness of our results.

[2] Hanson [7] used housing units from the U.S. Census of Population and Housing, whereas Steven, et al. [8] used the number of rooms in residential dwellings per German district.

[3] Details of sample design and sampling procedures had been presented in Phung and Phong [24] and Glewwe [25].

[4] The unit of measurement is million tons km. Data is taken from the annual Statistical Yearbook of Vietnam issued by the GSO from 2000 to 2012.

[5] Please note that we ignore export costs and assume that all goods are consumed domestically.

[6] It may be argued that high volume of freight traffic may reflect high efficiency in road transportation. However, in the case of Vietnam, the volume of freight traffic is highly correlated with the severity of congestion on the roads, especially in major cities such as Ho Chi Minh City and Hanoi.

[7] The reason for using CPI is that the provincial GDP deflator is not available in Vietnam.

### *3.3. Econometric issues*

Several econometric issues needed to be addressed for estimating (10). First and foremost, there is potential endogeneity among the regressors in (10); for example, an increase in either *AW* or *FA* can lead to an increase in *GPPC*.[1] Wooldridge [28] pointed out that potential endogeneity among the regressors, if left unaddressed, can result in biased OLS estimates. In order to control for this complication we estimate (10) by instrumental variable (*IV*), based on the principle of generalised method of moments (*GMM*), with lagged regressors serving as the internal instruments. Specifically, we instrumented all regressors with a two-year lag to coincide with the biannual VHLSS used for cost of living (*LC*). We also adding lagged first differences as instrumental variables for market access (*GPPC*), industrial linkages (*FA*) and trade costs (*FT*) [29]. The GMM was introduced in Hansen [30], and later refined by Wooldridge [28]. In general, *IV/GMM* estimation works in a similar manner to a standard IV estimator such as 2SLS. The implementation of *IV/GMM* estimation includes three steps: (i) Estimate the equation using the instrument set, (ii) Estimate the residuals, then estimate the optimal weighting matrix based on the set of moment conditions, (iii) Calculate the efficient estimator coefficients and their variance-covariance matrix using the estimated optimal weighting matrix [31]. To test for the validity of instruments, we performed the Hansen *J*-statistic test for the appropriateness of our model specifications and any over-identification problems. [2]

Geographically, provinces in Vietnam can be separated in to six broad regions according to climate, culture and language.[3] In order to control for these region-specific effects we introduced a regional dummy, *REG*, to (10). Furthermore, we included a time dummy variable, $\eta_t$, to reduce the possibility of correlation across provinces in the error terms, $\varepsilon_{it}$ [33].

Finally, given that our sample included some Vietnamese provinces that share similar attributes with each other it is possible for the variance of the stochastic error term to vary systematically across these provinces. In order to account for heteroskedasticity of an unknown form in the error term, we applied the robust option in Stata to correct for this issue [28, 32]. In addition, we also correct all standard errors for clustering within each province using a generalisation of the White method.[4] We then tested the statistical significance of the estimated coefficients using their standard errors based on this robust variance matrix.

### *3.4. Empirical model*

The basic specification of our empirical model is given by:

$$AW_{it} = \alpha_0 + \beta_1 GPPC_{it} + \beta_2 LC_{it} + \\ + \beta_3 FA_{it} - \beta_4 FT_{it} + REG + \eta_t + \varepsilon_{it}, \qquad (11)$$

where subscripts i and t denote the province and the year respectively. *AW* denotes provincial average wage, *GPPC* market access, *LC* the cost-of-living, *FA* the industrial linkages, *FT* the trade cost and $\varepsilon_{it}$ the stochastic error term. *REG* and $\eta_t$ are regional dummy and time dummy variables, respectively.

Since the variable LC is based from sample estimates, there may be concerns about sampling variance and sampling errors for small places [34, 35].[5] Indeed, the data in this study is aggregated at province level to generate consistent and coherent estimates. According to Phung and Phong [24], the sampling errors at provincial data level was determined at an optimal level and the number of households per commune was larger than 1,500 persons.

Tables 1 and 2 present summary statistics and correlation matrix of the variables used in the estimation, respectively.

## 4. Results

### *4.1. Baseline results*

Data limitation forced us to use an IV set that consisted of two-year lagged regressors and restrict our analysis to only estimations corresponding to the year 2004, 2006, 2008, 2010 and 2012. Table 3 presents the OLS and IV/GMM results for Equation (11). The adjusted R-square statistics in Table 3 mean that our model can explain over 85 % of the variation in the provincial average wage rate for above years. In addition, the F-statistics show that we can reject the null hypothesis and conclude that estimated coefficients are jointly significantly different from zero. Since we used lag2

[1] This potential endogeneity is also called reverse causality which arises when firms choose their location according to local productivity, directly linked to the local wages [27].

[2] The IV/GMM estimator is performed by applying ivreg2 in Stata12. Baum and Schaffer [31] and Baum, et al. [32] provided details of ivreg2 and routines for estimation.

[3] These regions include including the Red River Delta, the Northern Midlands and Mountain Areas, the South East Areas, the Central Coast, the Central Highlands and the Mekong River Delta.

[4] See Baum and Schaffer [31] and Baum, et al. [32].

[5] We would like to thank anonymous referees for these helpful comments.

Table 1

**Summary statistics**

| Variables | | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| *AW* | Average wages | 780 | 9.28 | 0.50 | 6.90 | 10.63 |
| *GPPC* | GDP per capita | 780 | 1.97 | 0.59 | 0.67 | 4.06 |
| *LC* | Living costs | 360 | 1.29 | 0.55 | -1.14 | 2.42 |
| *FA* | Fixed assets | 780 | 8.09 | 1.45 | 3.92 | 13.38 |
| *FT* | Freight traffic | 780 | 5.77 | 1.36 | 1.18 | 9.83 |
| *REG* | Regional dummies | 780 | 3.28 | 1.77 | 1.00 | 6.00 |
| *HS* | Human capital | 780 | 3.50 | 0.32 | 2.51 | 4.59 |
| *INT* | Interaction term between *GPPC* and *HS* | 780 | 6.88 | 2.08 | 1.73 | 14.64 |
| *GPP* | Provincial *GDP* | 780 | 9.03 | 0.93 | 6.40 | 12.48 |
| *POP* | Provincial population size | 780 | 13.97 | 0.55 | 12.54 | 15.85 |

Notes: all variables are in logarithmic transformation of their value, except dummies.
Source: Compiled from the data sources including the Statistic Year Book, the VHLSS and unpublished resources.

Table 2

**Correlation matrix**

| | *AW* | *GPPC* | *GPP* | *POP* | *LC* | *FT* | *FA* | *REG* | *HS* | *INT* |
|---|---|---|---|---|---|---|---|---|---|---|
| ***AW*** | 1.00 | | | | | | | | | |
| ***GPPC*** | 0.49 | 1.00 | | | | | | | | |
| ***GPP*** | 0.43 | 0.83 | 1.00 | | | | | | | |
| ***POP*** | 0.20 | 0.34 | 0.80 | 1.00 | | | | | | |
| ***LC*** | 0.90 | 0.58 | 0.57 | 0.34 | 1.00 | | | | | |
| ***FT*** | 0.22 | 0.59 | 0.80 | 0.72 | 0.42 | 1.00 | | | | |
| ***FA*** | 0.61 | 0.70 | 0.84 | 0.67 | 0.70 | 0.74 | 1.00 | | | |
| ***REG*** | 0.21 | 0.28 | 0.27 | 0.16 | 0.22 | 0.00 | 0.03 | 1.00 | | |
| ***HS*** | –0.15 | –0.05 | 0.00 | 0.06 | –0.24 | 0.16 | 0.04 | –0.40 | 1.00 | |
| ***INT*** | 0.43 | 0.95 | 0.80 | 0.34 | 0.50 | 0.62 | 0.69 | 0.15 | 0.24 | 1.00 |

Source: Compiled from the data sources described in the text

of the independent variables to construct the instrument set, the number of observations reduces from 360 in column (1) to 300 in column (2). In order to save space, we have not listed the estimation results of time dummies and regional dummies. However, it is worth noting that the inclusion of regional dummies increased the fit of the estimated model, since these dummies were significant at conventional levels. [1] Therefore, we continued to employ regional dummies in the robustness check section.

In column (1), the pooled OLS estimates are consistent with our expectations, except for GPPC. In particular, the pooled OLS results show that living costs (*LC*) and industrial linkages (*FA*) have positive effects on provincial wages. In contrast, trade costs (*FT*) reduce the provincial wages. However, these mixed results may be due to the bias of *OLS* estimates as discussed in above section. Therefore, we switched to the *IV/GMM* estimates in order to obtain robust and unbiased results. Results of *IV/GMM* estimates in column (2) of Table 3 show that, with the exception of the coefficient on market access (*GPPC*), all coefficients have the correct sign and are statistically significant at the 1 % level. Specifically, we found a positive coefficient on cost-of-living (*LC*), supporting a positive cost-of-living effect on provincial wage rate. To put this finding into perspective, we estimated that a 1 % increase in cost-of-living is likely to raise provincial wage rate roughly 0.9 %. Further, we estimated that a 1 % increase in fixed assets stock (*FA*) brings about 0.1 % rise in provincial wage rate. In part, this finding is consistent with the notion that the clustering of firms increases the demand for labour, pushing up the provincial wage rate in the short term. In terms of trade cost (*FT*), our result shows that a 1 % increase in *FT* leads to around 0.14 % reduction in provincial wage, all things constant. Meanwhile, the Hansen *J*-statistic at the bottom of column (2) suggests that there is no over-identification in the internal instrument set. Taken together, *IV/GMM* estimates appear to perform well and the proposed instrument set is appropriate.

[1] Coefficients of time dummies are insignificant at all conventional level of confidence. Hence, we ignore time dummies in the next section.

Table 3

**Determinants of provincial wages: OLS and IV/GMM estimators**

| | (1) *OLS* | (2) *IV/GMM* |
|---|---|---|
| | *AW* | *AW* |
| *GPPC* | 0.030 | –0.027 |
| | (0.056) | (0.089) |
| *LC* | 0.870*** | 0.910*** |
| | (0.033) | (0.062) |
| *FT* | –0.128*** | –0.140*** |
| | (0.014) | (0.025) |
| *FA* | 0.093*** | 0.102*** |
| | (0.017) | (0.030) |
| Constant | 8.082*** | 8.113*** |
| | (0.104) | (0.181) |
| Observations | 360 | 300 |
| Regional dummies | Yes | Yes |
| Year dummies | Yes | Yes |
| $R^2$ | 0.874 | 0.878 |
| adj. $R^2$ | 0.869 | 0.872 |
| *F* test | 139.135 | 61.070 |
| *p*-value | | 0.000 |
| Hanson *J*-statistic | | 1.479 |
| *p*-value | | 0.687 |

Notes:
Standard errors in parentheses and *, **, *** indicate significance at 10 %, 5 % and 1 % levels, respectively.
Standard errors adjusted for arbitrary heteroscedasticity and arbitrary intra-cluster correlation.
Instrument set in column (2) includes lag2 of GPPC, LC, FA and FT and lagged first differences of GPPC, FA and FT.
For brevity, the coefficients of regional dummies and year dummies are not reported here.
Sources: Author's estimation.

Since some correlation coefficients exceed 0.6 in Table 2, we also performed the variance inflation factor (*VIF*) procedure to detect possible multicollinearity in the regressors. According to Wooldridge [36], if *VIF* exceeds the value of 10, one may conclude that there could be potential multicollinearity in the estimated regression. From the *VIF* results in Table 4, there is not enough evidence to conclude that multicollinearity presents a problem in our regression.

### *4.2. Robustness checks*

*4.2.1. Alternative measures of market access*

It is possible that results obtained thus far, could have been driven by our selection of the market access measure. As such, we tested the preceding hypothesis by examining the effect that changing the market access measure from the provincial GDP per capita (*GPPC*), to the provincial *GDP* (*GPP*) and provincial population size (*POP*). Table 5 provides the results for this change to an alternative measure, holding constant all other specifications as per Table 3. As can be seen from Table 5, the results remain largely unchanged, with the magnitude of the coefficient on cost-of-living (*LC*), trade cost (*FT*) and industrial linkages (*FA*) being affected moderately. Notably, coefficients of the market access variables are statistically insignificant at all conventional levels across all columns in Table 5. The Hansen *J*-statistics at the bottom of all columns suggest that the *IV/GMM* estimator is appropriate and free from instrument proliferation. The adjusted *R*-square results and *F*-test statistics are similar to the baseline model (column (1) of Table 5), confirming the goodness-of-fit of our models. Taken together, Table 5 shows that our results are not sensitive to the way in which market access is measured.

### *4.3. The effect of human capital*

In this section, we examine the effect of human capital on provincial wage rate. We also considered the effect of, if any, provincial human capital has had on the market access-wages nexus in Vietnam. In constructing this term we demeaned both *GPPC* and *HS* before multiplying them by each other. Since the pair-wise correlation between *GPPC* and the interaction term, *INT*, is extremely high (0.95), we have to omit *GPPC* from Equation (11) to avoid multicollinearity problem [37]. Column (2) in Table 6 presents the results with the inclusion of human capital, *HS*, and column (3) show the results for adding interaction term, *INT*.

Clearly, the inclusion of human capital (*HS*) and the interaction term (*INT*) does not signifi-

Table 4

**VIF calculation for selected years**

| Variable | *VIF* | *1/VIF* | Variable *REG* | *VIF* | *1/VIF* | Variable Year | *VIF* | *1/VIF* |
|---|---|---|---|---|---|---|---|---|
| *GPPC* | 4.27 | 0.23 | 2 | 2.9 | 0.35 | 2004 | 1.73 | 0.58 |
| *LC* | 2.63 | 0.38 | 3 | 2.02 | 0.49 | 2006 | 1.94 | 0.52 |
| *FA* | 5.89 | 0.17 | 4 | 1.38 | 0.73 | 2008 | 2.25 | 0.44 |
| *FT* | 3.15 | 0.32 | 5 | 1.98 | 0.50 | 2010 | 2.69 | 0.37 |
| | | | 6 | 2.33 | 0.43 | 2012 | 3.14 | 0.32 |
| Mean *VIF* | 2.74 | | | | | | | |

Source: Author's estimation.

Table 5

**Alternative measures for market access: *IV/GMM* estimator**

| | (1) *IV/GMM* | (2) *IV/GMM* | (3) *IV/GMM* |
|---|---|---|---|
| | *AW* | *AW* | *AW* |
| *GPPC* | −0.018 | | |
| | (0.057) | | |
| *GPP* | | −0.037 | |
| | | (0.055) | |
| *POP* | | | −0.025 |
| | | | (0.065) |
| *LC* | 0.917*** | 0.907*** | 0.906*** |
| | (0.059) | (0.061) | (0.063) |
| *FT* | −0.136*** | −0.122*** | −0.128*** |
| | (0.025) | (0.027) | (0.025) |
| *FA* | 0.091*** | 0.096** | 0.086** |
| | (0.029) | (0.038) | (0.035) |
| Constant | 8.179*** | 8.363*** | 8.492*** |
| | (0.161) | (0.262) | (0.756) |
| Observations | 300 | 300 | 300 |
| Regional dummies | Yes | Yes | Yes |
| Year dummies | No | No | No |
| $R^2$ | 0.875 | 0.876 | 0.875 |
| adj. $R^2$ | 0.871 | 0.872 | 0.872 |
| *F* test | 64.313 | 63.659 | 57.423 |
| *p*–value | 0.000 | 0.000 | 0.000 |
| Hanson *J*–statistic | 2.645 | 2.596 | 2.598 |
| *p*–value | 0.450 | 0.458 | 0.458 |

Notes:

Standard errors in parentheses and *, **, *** indicate significance at 10 %, 5 % and 1 % levels, respectively.

Standard errors adjusted for arbitrary heteroscedasticity and arbitrary intra-cluster correlation.

Instrument set in column (1) includes lag2 of GPPC, LC, FA and FT and lagged first differences of GPPC, FA and FT.

Instrument set in column (2) includes lag2 of GPP, LC, FA and FT and lagged first differences of GPP, FA and FT.

Instrument set in column (3) includes lag2 of POP, LC, FA and FT and lagged first differences of POP, FA and FT.

For brevity, the coefficients of regional dummies are not reported here.

Sources: Author's estimation.

cantly alter the results obtained thus far. In line with our a priori expectations, wages are generally higher in provinces with higher cost-of-living (*LC*), stronger industrial linkages (*FA*) and lower trade cost (*FT*). The results in columns (2) and (3) show that education also has an important influence on wages. The coefficients on human capital (*HS*) suggest that a 10 % increase in human capital raises provincial wages by 0.9 %. Notably, coefficients of market access variable in column (2) and the interaction term (*INT*) in column (3) are both statistically insignificant, indicating the absence of the market access effect on provincial wage for the period of this study.

## 5. Conclusion

This study has examined determinants of average wages across provinces in Vietnam over the period 2000–2012. We find that cost of living, industrial linkages and trade costs are significant determinants of provincial wage. However, we could not find any evidence of market access impacts wage in specific provinces. We also provided

Table 6

**Inclusion of human capital and interaction term: *IV/GMM* estimator**

| | (1) *IV/GMM* | (2) *IV/GMM* | (3) *IV/GMM* |
|---|---|---|---|
| | *AW* | *AW* | *AW* |
| *GPPC* | −0.018 | −0.002 | |
| | (0.057) | (0.040) | |
| *HS* | | 0.092* | 0.094* |
| | | (0.052) | (0.052) |
| *INT* | | | 0.002 |
| | | | (0.012) |
| *LC* | 0.917*** | 0.909*** | 0.907*** |
| | (0.059) | (0.036) | (0.036) |
| *FT* | −0.136*** | −0.136*** | −0.138*** |
| | (0.025) | (0.015) | (0.016) |
| *FA* | 0.091*** | 0.088*** | 0.089*** |
| | (0.029) | (0.018) | (0.019) |
| Constant | 8.179*** | 7.842*** | 7.822*** |
| | (0.161) | (0.214) | (0.211) |
| Observations | 300 | 300 | 300 |
| Regional dummies | Yes | Yes | Yes |
| Year dummies | No | No | No |
| $R^2$ | 0.875 | 0.877 | 0.877 |
| adj. $R^2$ | 0.871 | 0.873 | 0.873 |
| *F* test | 64.313 | 158.043 | 156.100 |
| *p*–value | 0.000 | 0.000 | 0.000 |
| Hanson *J*-statistic | 2.645 | 2.468 | 3.685 |
| *p*–value | 0.450 | 0.481 | 0.298 |

Notes:

Standard errors in parentheses and *, **, *** indicate significance at 10 %, 5 % and 1 % levels, respectively.

Standard errors adjusted for arbitrary heteroscedasticity and arbitrary intra-cluster correlation.

Instrument set in columns (1) and (2) includes lag2 of GPPC, LC, FA and FT and lagged first differences of GPPC, FA and FT.

Instrument set in column (3) includes lag2 of INT, LC, FA and FT and lagged first differences of INT, FA and FT.

For brevity, the coefficients of regional dummies are not reported here.

Sources: Author's estimation.

robust evidences for our baseline results based on several checks including control for the extent of endogeneity of independent variables, different measures of market access and the inclusion of human capital and its interaction term in *NEG* wage equation.

Our findings suggest several policy implications for various levels of government in Vietnam to consider. The benefits arising from industrial linkages combined with the effect of trade costs put forward another strong case for upgrading local infrastructure and establishing business support schemes. In addition, local governments need to emphasise their comparative advantage as part of the overall developmental policy. Provinces with an abundant supply of low-cost labour should promote labour-intensive industries such as textiles and clothing, agriculture and food industries. In doing so, these provinces may attract more local or foreign investment, thereby, generating job opportunities for locals.

Meanwhile, local governments need to strengthen human capital as the first step towards improving average wages. Indeed, a better education system is not only instrumental to raising average earnings per worker, but also crucial for minimising income inequality in the long run. According to GSO [38], the education and vocational training systems in Vietnam remain outdated and fail to meet the current economic demands. In 2001 only 11.8 % of total employed labour was officially classified as 'trained'. Eleven years later, this ratio has increased by only 4.6 %, to a total of 16.4 %. Therefore, local government needs to encourage private and foreign investors to participate in the vocational and tertiary education sector, through tax concessions and other incentives. For instance, local governments could partner with private providers of vocational schools, or a university, to design curricula that cater for the needs of local labour markets. Moreover, local governments should reward those employers offering on-the-job training and those who work closely with vocational and tertiary institutions, in developing specific programmes in order to meet recruiters' expectations.

## References

1. Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy, 99(3),* 483–499.

2. Fujita, M., P. Krugman, and A. J. Venables (1999). *The Spatial Economy: Cities, Regions, and International Trade.* The United States of America: The MIT Press, Cambridge, Massachussets, 70.

3. Baldwin, R., et al. (2003). *Economic Geography and Public Policy.* Princeton, New Jersey, United States of America: Princeton University Press, 59.

4. Head, K. and T. Mayer (2004). The Empirics of Agglomeration and Trade, *Handbook of Regional and Urban Economics,* 2609–2669.

5. Redding, S. J. (2010). The Empirics of New Economic Geography. *Journal of Regional Science, 50(1),* 297–311.

6. Redding, S. and A. J. Venables (2004). Economic Geography and International Inequality. *Journal of International Economics, 62,* 53–82.

7. Hanson, G. H. (2005). Market Potential, Increasing Return, and Geographic Concentration. *Journal of International Economics, 67,* 1–24.

8. Steven, B., G. Harry, and S. Marc (2004). The Spatial Distribution of Wages: Estimating the Helpman-Hanson Model for Germany. *Journal of Regional Science, 44(3),* 437–466.

9. Ma, A. C. (2006). Geographical location of foreign direct investment and wage inequality in China. *World Economy, 29(8),* 1031–1055.

10. Golley, J. (2007). China's western development strategy and nature versus nurture. *Journal of Chinese Economic and Business Studies, 5(2),* 115–129.

11. Amiti, M. and L. Cameron (2007). Economic Geography and Wages. *The Review of Economics and Statistics, 89(1),* 15–29.

12. Fally, T., R. Paillacar, and C. Terra (2010). Economic geography and wages in Brazil: Evidence from micro-data. *Journal of Development Economics, 91(1),* 155–168.

13. Helpman, E. (1998). The Size of Regions, In D. Pines, E. Sadka, and I.Zilcha (eds.), *Topics in Public Economics: Theoretical and Applied Analysis,* Cambridge: Cambridge University Press, 21.

14. Krugman, P. and A.J. Venables (1995). Globalization and the Inequality of Nations. *The Quarterly Journal of Economics, 110(4),* 857–880.

15. Fujita, M. and J.-F. Thisse (1996). Economics of Agglomeration. *Journal of the Japanese and International Economies, 10,* 339–378.

16. Baldwin, R., et al. (2003). The core-periphery model: Key features and effects. *The Monopolistic Competition Revolution in Retrospect,* 213–235.

17. Ottaviano, G.I., T. Tabuchi, and J.-F. Thisse (2002). Agglomeration and trade revisited. *International Economic Review, 43,* 409–436.

18. Puga, D. (1999). The rise and fall of regional inequalities. *European Economic Review, 43(2),* 303–334.

19. Head, K. and T. Mayer (2006). Regional wage and employment responses to market potential in the EU. *Regional Science and Urban Economics, 36(5),* 573–594.

20. Cheng, L. K. and Y. K. Kwan (2000). What are the determinants of location of foreign direct investment? The Chinese experience. *Journal of International Economics, 51(2),* 379–400.

21. Sun, Q., W. Tong, and Q. Yu (2002). Determinants of foreign direct investment across China. *Journal of International Money and Finance, 21(1),* 79–113.

22. Agiomirgianakis, G., D. Asteriou, and K. Papathoma (2006). *The Determinants of foreign Direct Investment: A Panel Data Study for the OECD Countries.* Department of Economics: School of Social Sciences, City University London [Electronic source] Retrieved from http://openaccess.city.ac.uk/id/eprint/1420 (Date of access: 11.01.2018).

23. Hilber, C. A. L. and I. Voicu (2010). Agglomeration Economies and the Location of Foreign Direct Investment: Empirical Evidence from Romania. *Regional Studies, 44(3),* 355–371.

24. Phung, D. T. and N. Phong (2006). *Vietnam household living standard survey (VHLSS), 2002 and 2004: Basic information.* Hanoi: General Statistics Office, [Electronic source]. Retrieved from http://microdata.worldbank.org/index.php/catalog/2306 (Date of access: 28/07/2018).

25. Glewwe, P. (2005). *Procedure for calculating nominal and real expenditures, and poverty indicators, for the 2002 Viet Nam Household Living Standards Survey (VHLSS).* University of Minnesota, Twin Cities, June, 2003. 13, 17–25.

26. Aitken, B., A. Harrison, and R.E. Lipsey (1996). Wages and foreign ownership — A comparative study of Mexico, Venezuela, and the United States. *Journal of International Economics, 40(3–4),* 345–371.

27. Combes, P.-P. (2011). The empirics of economic geography: how to draw policy implications? *Review of World Economics, 147(3),* 567–592.

28. Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, Massachusetts, London, england: The MIT Press, 29.

29. Anderson, T. W. and C. Hsiao (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics, 18(1),* 47–82.

30. Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econemetrica, 50(4),* 1029–1054.

31. Baum, C. F. and M. E. Schaffer (2003). Instrumental variables and GMM: Estimation and testing. The Stata Journal, 3(1), p. 1–31.

32. Baum, C. F., M. E. Schaffer, and S. Stillman (2007). Enhanced routines for instrumental variables / GMM estimation and testing. *The Stata Journal, 7(4),* 465–506.

33. Roodman, D. (2009). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics, 71(1),* 135–158.

34. Fay III, R. E. and R. A. Herriot (1979). Estimates of income for small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association, 74(366a),* 269–277.

35. Betti, G. and A. Lemmi, *Advances on income inequality and concentration measures.* 2008, Routledge, 368.

36. Wooldridge, J. M. (2009). *Introductory Econometrics: A Modern Approach.* the United States of America: South-Western Cengage Learning. 49.

37. Gujarati, D. N. (2003). *Basic Econometrics.* New York: McGraw-Hill Higher Education, 15.

38. GSO Report on the 2011 *Vietnam Labour Force Survey.* (2012) [Electronic source]. Retrieved from https://www.gso.gov.vn/default_en.aspx?tabid=515&idmid=5&ItemID=12541 (Date of access: 08/09/2016).

## Authors

**Hien Thanh Hoang** — PhD in Economics, Institute of Socio-Economics Research, Duy Tan University; ORCID: 0000–0002–1182–078X; Scopus Author ID: 57205611543 (03 Quang Trung street, Hai Chau district, Da Nang, Vietnam; e-mail: hoangthanhhien@dtu.edu.vn).

**Linh T. D. Huynh** — PhD in Economics, Chair of International Business Division, University of Economics, The University of Da Nang; ORCID: 0000–0001–9714–522X; Scopus Author ID: 57205610900 (71 Ngu Hanh Son street, Ngu Hanh Son district, Da Nang city, Vietnam; e-mail: linhhtd@due.edu.vn).

**George S. Chen** — PhD in Economics, UNE Business School, University of New England; ORCID: 0000–0002–8519–9340; SCOPUS ID: 56013242400 (Room 816, W40, UNE Business school, University of New England, Armidale, NSW, 2351, Australia; e-mail: gchen2@une.edu.au).