# The Effect of Knowledge Accessibility on the International Inequality*

Kentaro Nakajima†

January 7, 2005

Abstract

This paper estimates the structural model of new economic geography by using cross-country data of income per worker, bilateral trade, and bilateral flight passengers in order to explain international income disparity. We provide evidence that the geographical accessibility to foreign markets is statistically significant. It is consistent to early researches. And we provide evidence that the geographical accessibility to foreign knowledge is also statistically significant and this is robust to controlling for a wide range of considerations which includes primary geographical factors, political factors, and institutional factors.

*Keywords:* Economic development; economic geography; International trade

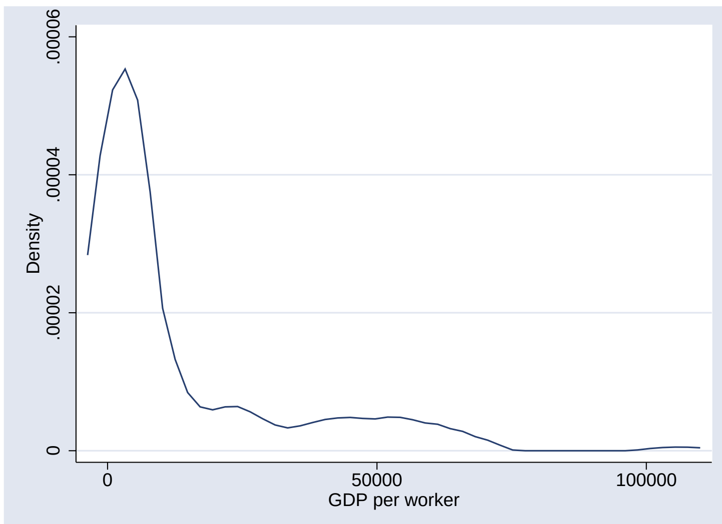
* I am very grateful to my academic adviser Professor Takatoshi Tabuchi for his encouragement and very helpful comments. Of course, all remaining errors are mine.

† Graduate School of Economics, University of Tokyo

# 1 Introduction

International income inequality has been the one of the most serious problem in the world economy. The gap between high income countries and low income countries are too huge. For example, the highest income par capita is more than 500 times to the lowest country. And the income distribution is also unequal.

Figure 1 is a nonparametric kernel density function of GDP per worker on the year of 1999. It is clear that there exist a large number of low income countries, and a few numbers of high




Data from World Development Indicators 2003.
*This kernel densities function is computed by using Epanechnicov kernel function

Figure 1 Kernel densities function of GDP per worker.

income countries from this figure.

About this problem, economic growth theory has been the one of the mainstream of the research. These researches mainly investigate the role of capital accumulation on the economic growth. Especially, endogenous growth theory has focused on the knowledge capital as the source of technical externality. For example, Grossman and Helpman (1991) emphasized that knowledge capital spillover from foreign countries is an important factor on the economic growth. Barro and Sala-i-Martin (1995) is a useful monograph about the works of this growth theory.

On the contrary, the theory called new economic geography (NEG) which stemmed from the epoch-making research by Krugman (1991) gave the analysis from a new viewpoint to such problem. Since 1990's, through the theoretical development of NEG theory, that has become one of the most persuasiveness theory about this wage disparity. The key factor of these studies has been pecuniary externality. Fujita et al. (1999) is a representative monographs of theoretical aspects, and Ottaviano and Thisse (2004) summarized more recent NEG theoretical developments.

And there exist several studies have been made on combining these studies. The studies of Baldwin et al. are regarded as the hybrid of NEG and endogenous growth theory. They have made theoretical NEG models named capital constructing model which is based on Grossman and Helpman (1991)'s endogenous growth model and standard NEG model. The model can treat both capital accumulation as a critical factor of the economic growth theory and pecuniary externality as a critical factor of the NEG theory simultaneously. For example, Baldwin et al. (2001) have applied this model to explaining north-south economic bifurcation in detail. They emphasized the importance of foreign knowledge spillover as well as pecuniary externality for firm's location. And they analyzed circumstantially the relationship between market access and foreign knowledge spillover. Baldwin and Martin (2004) summarized these studies and Baldwin et al (2003) is a useful monograph about these research.

Over the past few years, several empirical studies have been made on explaining this wage disparity by using NEG model. And most of these studies have argued the importance of market accessibility. Head and Miyor (2004) is a useful survey about these empirical works. Especially, Redding and Venables (2004) is an important study which has empirically explained international wage disparity by using NEG model framework. They concluded that market access is important geographical factor on the international wage disparity. According to their analysis, foreign market access alone explains 35% of international income variations.

However, following Baldwin et al.'s theoretical analysis, we should pay much attention to knowledge spillover as well as market and supplier access. The purpose of this paper is to investigate the role of foreign knowledge spillover on the international wage disparity by using NEG framework. The concrete method is the following. We define *knowledge access* as the geographical accessibility to the foreign country's knowledge. And we empirically investigate its role on the international wage disparity by using NEG framework. While this research is extension of Redding and Venables (2004) , it is also the first step of the empirical study about Baldwin et al. 's series of studies.

The structure of this paper is as follows. Section 2 explains a theoretical framework. Section 3 explains an empirical framework and provides the definition of the knowledge access. Section 4 presents the estimation of the trade equation, and section 5 presents the estimation of the wage equation. Section 6 concludes.

## 2 Theoretical framework

The structure of our model is based on that of Redding and Venables (2004). Their model is constructed by standard NEG model with intermediate goods[*1]. We use their model and extend to deal with knowledge access.

The world consists of $i \in \{1, \ldots, R\}$ countries. Our model is focused on only manufacturing sector which produces differentiated goods by using increasing returns to scale technology. Each firm's product is differentiated from that of other firms and it is used both consumption and as an intermediate good.

We start analysis on the demand side first. We assume the consumer's preference for manufacturing goods is a constant elasticity of substitution (CES) with $\sigma$ representing an index of product differentiation. And firm's demand for intermediate goods is same as consumption. These functions are as follows

$$U_j = \left[\sum_i^R \int_{n_i} x_{ij}(z)^{(\sigma-1)/\sigma} dz\right]^{\sigma/(\sigma-1)} = \left[\sum_i^R n_i x_{ij}^{(\sigma-1)/\sigma}\right]^{\sigma/(\sigma-1)}, \quad \sigma > 1. \tag{1}$$

The second equation satisfies from the fact that all products produced each country $j$ and demanded by country $i$ is the same quantity in an equilibrium. A price index $G_j$ which is the minimum cost to get one unit of utility derived from this utility function is

$$G_j = \left[\sum_i^R \int_{n_i} p_{ij}(z)^{1-\sigma} dz\right]^{1/(1-\sigma)} = \left[\sum_i^R n_i p_{ij}^{1-\sigma}\right]^{1/(1-\sigma)}. \tag{2}$$

Its second term is also obtained by symmetry between country $i$'s production on the equilibrium. By using Shepard's lemma on the price index, we derive country $j$'s total demand for each goods

$$x_{ij} = p_{ij}^{-\sigma} E_j G_j^{(\sigma-1)}, \tag{3}$$

[*1] For further theoretical detail, see Fujita et al.(1999), chapter14.

where $E_j$ is the country $j$'s total expenditure on manufactures which used as consumption and intermediate goods.

Next, we analyze the production side. We assume each country $i$'s firm's cost function is as follows

$$C(x_i) = G_i^{\alpha} w_i^{\beta} v_i^{\gamma} c_i (F + x_i), \tag{4}$$

where $x_i \equiv \sum_j x_{ij}$. This technology has increasing returns to scale. There are three types of inputs, combined by Cobb-Douglas technology. At first, $w_i$ is the wage and the share is $\beta$. Second, $v_i$ is the internationally mobile factor with price $v_i$ and the share is $\gamma$. Third, $\alpha$ is the share of intermediate goods and $\alpha + \beta + \gamma = 1$. The use of this cost function obtains the profit function of country $i$'s firm

$$\pi_i = \sum_j^R p_{ij} x_{ij} / T_{ij} - G_i^{\alpha} w_i^{\beta} v_i^{\gamma} c_i (F + x_i). \tag{5}$$

$T_{ij}$ is an iceberg transport cost between country $i$ and $j$. If one unit of the manufacturing good shipped from country $i$ to $j$, only a fraction $1/T_{ij}$ of original unit actually arrives. And if trade is costless, $T_{ij} = 1$.

Maximizing this firms profit function, the f.o.b.price $p_i$ explained by $p_{ij} = p_i T_{ij}$ becomes a constant markup over marginal cost, given by

$$p_i = G_i^{\alpha} w_i^{\beta} v_i^{\gamma} c_i \sigma / (\sigma - 1). \tag{6}$$

By using that, the profit of each country $i$ firm is

$$\pi_i = (p_i / \sigma)[x_i - (\sigma - 1)F]. \tag{7}$$

We suppose that there is free entry and exit in response to profits or losses, so firm's zero-profit conditions satisfied. Thus, firm's production and their sales equals a constant $\bar{x} \equiv (\sigma - 1)F$. From demand function eq(3), we obtain

$$p_i^{\sigma} \bar{x} = \sum_j^R E_j G_j^{\sigma - 1} (T_{ij})^{1 - \sigma}. \tag{8}$$

Substituting eq(5), we obtain wage equation,

$$\bar{x} (G_i^{\alpha} w_i^{\beta} v_i^{\gamma} c_i \sigma / (\sigma - 1))^{\sigma} = \sum_j^R E_j G_j^{\sigma - 1} T_{ij}^{1 - \sigma}. \tag{9}$$

This wage equation is the price of the immobile factor of production. The purpose of this paper is estimating this equation empirically.

To analyze market access and supplier access, we have to rewrite trade equation. Using eq(3), the trade equation is rewritten the form of trade value as

$$n_i p_i x_{ij} = n_i p_i^{1-\sigma} (T_{ij})^{1-\sigma} E_j G_j^{\sigma-1}. \tag{10}$$

The right-hand side of this equation contains both demand and supply variables. The term $E_j G_j^{\sigma-1}$ is country $j$'s market capacity, and $n_i p_i^{1-\sigma}$ is the supply capacity. These will be argued in detail on the next section. And price index specified iceberg transport cost is as follows

$$G_j = \left[ \sum_i^R n_i (p_i T_{ij})^{1-\sigma} \right]^{1/(1-\sigma)}. \tag{11}$$

These three equations are fundamental equations for estimating knowledge and market access. In the full general equilibrium, we can derive $E_j$ and $n_i$ endogenous by using factor market clearing condition. But we set $E_j$ and $n_j$ exogenous from the labor immobile assumption. Thus, this wage equation become expressing the wages which manufacturing firms each location affords to pay.

# 3 Estimation

## 3.1 Market access and Supplier access

At first, we estimate trade equation in order to estimate bilateral transportation costs. By using this, we construct market access and supplier access. We define $m_i \equiv E_i G_i^{\sigma-1}$ and $s_i \equiv n_i p_i^{1-\sigma}$ such that country $i$'s market capacity and supplier capacity. Rewrite trade equation by using $m_i$ and $s_i$,

$$n_i p_i x_{ij} = s_i (T_{ij})^{1-\sigma} m_j. \tag{12}$$

Next we construct market access (MA) and supplier access (SA). Market access is the transport cost weighted sum of the market capacities of all partner countries, and supplier access is that of supplier capacities, so

$$MA_i = \sum_j E_i G_i^{\sigma-1} (T_{ij})^{1-\sigma} = \sum_j (T_{ij})^{1-\sigma} m_j \tag{13}$$

$$SA_j = \sum_i n_i (p_i T_{ij})^{1-\sigma} = \sum_i s_i (T_{ij})^{1-\sigma}. \tag{14}$$

By using predicted values from trade equation estimation, we construct these market access and supplier access. And substituting these equations into eq(5), we can rewrite wage equation as

$$(w_i^{\beta} v_i^{\gamma} c_i)^{\sigma} = AG_i^{-\alpha\sigma} \sum_j^R E_j G_j^{\sigma-1} (T_{ij})^{1-\sigma} = A \left( \sum_j s_j (T_{ij})^{1-\sigma} \right)^{\frac{\alpha\sigma}{\sigma-1}} \left( \sum_j (T_{ij})^{1-\sigma} m_j \right) \tag{15}$$
$$= A(SA_i)^{\frac{\alpha\sigma}{\sigma-1}} (MA_i),$$

where $A$ is the constant term. This equation is log-linear of MA and SA, and it includes $w_i$ in the left hand side. To estimate this estimation function, we need to get MA and SA value.

## 3.2 Knowledge access

Next we construct knowledge access (KA) as the index of the knowledge spillover from foreign countries.

We assume knowledge can be divided two types by the way of transportation. For one thing is internationally mobile freely knowledge and another one is mobile restrictedly knowledge. Mobile freely means it can be used same value on everywhere. For example, knowledge transported on books, papers, internets are included in this group. We assume this type of knowledge are included by $v_i$ in the cost function. On the other hand, mobile restrictedly knowledge is which we cannot use equivalently at everywhere. A representative example is the knowledge which is transported by face to face communication. The importance of this type of knowledge spillover has been argued, especially much of case studies.

We suppose transport cost of mobile restrictedly knowledge is the function of the distance between country $i$ to $j$ which is denoted by $dist_{ij}$. It seems a reasonable approximation. Then, we can set the flow of knowledge by a gravity model specification.

We define $FK_{ij}$ is the flow of knowledge from country $j$ to $i$. And $T_{ij}^k$ [*2]is an iceberg transport cost for knowledge flow from country $i$ to $j$ as $T_{ij}^k = dist_{ij}^{\delta^k}$ . Thus, each country's peculiar factor and distance between countries determine the flow of knowledge as below,

$$FK_{ij} = \frac{country\ i's\ factor \quad country\ j's\ factor}{dist_{ij}^{\delta^k}}. \tag{16}$$

Next, we specify the factors of each country. Knowledge exporting country's factor mainly includes the country's quantity and quality of knowledge. On the contrary, knowledge im-

---

[*2] This $T_{ij}^k$ can be regarded as same as $\lambda$ used by Baldwin et al.(2001) (2003).

porting country's factor mainly includes the capability of receiving knowledge and it includes few geographic factors. Now we define knowledge access as *the distance weighted sum of available foreign knowledge* to investigate fundamental determinants of per capita income without resorting to own country's technology and knowledge. Thus, on making knowledge access, we have to set knowledge importing country's factor is common between every country, because knowledge importing country's factor mainly includes the technology and fundamental knowledge of the country but geographic factors.

Following these reason, the definition of knowledge access is as below.

**Definition [Knowledge access]**

*The country i's knowledge access is the distance weighted sum of available foreign knowledge. It is explained as*

$$KA_i \equiv \sum_{j \neq i} dist_{ij}^{\delta^k} \; country\ j's\ factor. \tag{17}$$

And we assume this KA affects total factor productivity, and decreases production cost as $c_i = KA_i^{-\varphi}$ where $\varphi > 0$. By using this assumption, wage equation eq. (15) is rewritten as below,

$$(w_i^\beta v_i^\gamma KA_i^{-\varphi})^\sigma = A(SA_i)^{\frac{\alpha\sigma}{\sigma-1}}(MA_i). \tag{18}$$

That is the equation which we want to estimate. In order to estimate this equation, we need several preparations, especially estimating market access, supplier access, and knowledge access.

# 4 Trade equation and knowledge access estimation

## 4.1 Data sources and sample size

Data on bilateral trade flows for a cross-section of over 55 countries are obtained from the World Bank's COMTRADE database. And data on bilateral flight passengers for a cross-section of 55 countries are obtained ICAO's On-Flight Origin and Destination. On our estimation, we use these available 55 countries.

## 4.2 Trade equation estimation

At first, we construct market access and supplier access. For this purpose, we estimate trade equation by using gravity model as following reason. The transport cost $T_{ij}$ is not observable, but setting $T_{ij} = dist_{ij}^{\delta}$ where $dist_{ij}$ is a distance between country $i$ and $j$ is regarded as a reasonable approximation. Thus, the trade equation can be rewritten as below

$$\ln(EX_{ij}) = \theta^T + \mu_i^T cty_i + \lambda_j^T ptn_j + \delta_1^T \ln(dist_{ij}) + \delta_2^T bord_{ij} + u_{ij}^T. \tag{19}$$

where $EX_{ij}$ is the value of exportation from country $i$ to country $j$, $cty_i$ and $ptn_j$ is the dummy variable of country and partner[*3], and $bord_{ij}$ is a dummy variable for whether an exporting country and importing partner share a common border.

By using predicted values $\hat{\mu}_i^T, \hat{\lambda}_i^T, \hat{\delta}_1^T, \hat{\delta}_2^T$, we construct $\hat{MA}$ and $\hat{SA}$ as below

$$\begin{aligned} \hat{MA}_i &= D\hat{MA}_i + F\hat{MA}_i \\ &= (\exp(ptn_i))^{\hat{\lambda}_i^T} (T_{ii})^{1-\sigma} + \sum_{j \neq i} (\exp(ptn_j))^{\hat{\lambda}_j^T} dist_{ij}^{\hat{\delta}_1^T} (\exp(bord_{ij}))^{\hat{\delta}_2^T}, \end{aligned} \tag{20}$$

$$\begin{aligned} \hat{SA}_j &= D\hat{SA}_j + F\hat{SA}_j \\ &= (\exp(cty_j))^{\hat{\mu}_j^T} (T_{jj})^{1-\sigma} + \sum_{i \neq j} (\exp(cty_i))^{\hat{\mu}_i^T} dist_{ij}^{\hat{\delta}_1^T} (\exp(bord_{ij}))^{\hat{\delta}_2^T}, \end{aligned} \tag{21}$$

where $D\hat{MA}_i$ and $D\hat{SA}_i$ is country $i$'s domestic effect, and $F\hat{MA}_i$ and $F\hat{SA}_i$ is its foreign countries' effect. On the construction of $D\hat{MA}$ and $D\hat{SA}$, we assume trade cost is same as foreign countries which is 100km away and with a common border.

By using OLS, we estimate this trade equation on a data of 1999. The result is Table 1 . The coefficient of $\ln(dist)$ is negative, and the coefficient of common border is positive. That is consistent to our theory, and these coefficients are statistically significant at the 1% level. The null hypothesis the coefficients on either the country dummies or partner dummies are equal to zero is rejected at the 1% level by a $F$-test. Then, we construct $\hat{MA}$ and $\hat{SA}$ by using these predicted coefficients.

[*3] These dummy variables include the effects of economic variables, for instance GDP, trade openness, and so on. But we cannot observe completely these economic variables, so we use these dummies as a substitute. The consistency of this method is argued Redding and Venables (2004) and it justified the use of this method.

Table 1 Trade equation estimation

| | |
|---|---|
| Dependent | ln($X$) |
| Observations | 2970 |
| Year | 1999 |
| ln($dist$) | $-1.556^{**}$ |
| | (0.079) |
| ln($bord$) | $0.298^{**}$ |
| | (0.266) |
| Country dummy | yes |
| Partner dummy | yes |
| $R^2$ | 0.4705 |
| F(·) | 68.56 |
| Prob>F | 0.0000 |

White heteroskedasticity robust standard errors in parentheses.
*Denotes statistical significance at the 10% level. **Denotes statistical significance at the 5% level.

## 4.3 Knowledge access estimation

Next, we estimate knowledge access equation to construct predicted value of knowledge access. We use the proxy of the flow of knowledge is the number of flight passengers between each country's first airport[*4]. Taking logarithm eq.(16), we obtain the estimation function,

$$\ln(TRAVEL_{ij}) = \theta^K + \mu_i^K cty_i + \lambda_i^K ptn_j + \delta_1^K \ln(dist_{ij}) + \delta_2^K bord_{ij} + u_{ij}^K, \qquad (22)$$

where $TRAVEL_{ij}$ is the number of passengers from country $j$ to country $i$. Because we cannot observe all of the country's effects which determine the flow of knowledge, we also use country and partner dummies as same as trade equation estimation.

By using predicted values $\hat{\mu}_i^K, \hat{\lambda}_i^K, \hat{\delta}_1^K, \hat{\delta}_2^K$ to the definition of knowledge access eq.(17), we construct a fitted value of knowledge access $\hat{KA}_i$ as

$$\begin{aligned} \hat{KA}_i &= \hat{DKA}_i + \hat{FKA}_i \\ &= (\exp(ptn_i))^{\hat{\lambda}_i^K} T_{ii}^k + \sum_{j\neq i} (\exp(ptn_j))^{\hat{\lambda}_i^K} dist_{ij}^{\hat{\delta}_1^K} (\exp(bord_{ij}))^{\hat{\delta}_2^K}. \end{aligned} \qquad (23)$$

As same as $\hat{MA}$ and $\hat{SA}$, on the construction of $\hat{DKA}$, we set trade cost is same as foreign

[*4] We regard the use of first airports considerably removes tourists and makes possible to choose business users.

countries which is 100km away and with a common border. The data on bilateral number of flight passengers is obtained from ICAO's On-Fright Origin Destination. But it has much missing values because it counts only direct flight lines. So we estimate this knowledge access on a data which is removed these missing values. Then, number of observation decreases to 683. Border dummy, country and partner dummies are same as trade equation. Same as trade equation estimation, we regress by OLS. The estimation result is Table 2 .

Table 2 Knowledge access estimation

| | |
|---|---|
| Dependent | $\ln(TRAVEL)$ |
| Observations | 683 |
| Year | 1999 |
| $\ln(dist)$ | -0.655** |
| | (0.056) |
| $\ln(bord)$ | 0.277** |
| | (0.121) |
| Country dummy | yes |
| Partner dummy | yes |
| $R^2$ | 0.6482 |
| F(·) | 12.42 |
| Prob>F | 0.0000 |

White heteroskedasticity robust standard errors in parentheses.
*Denotes statistical significance at the 10% level. **Denotes statistical significance at the 5% level.

Same as the result of trade equation estimation, the coefficient of log distance is negative, that of common border dummy is positive. These are consistent to our theory. And these coefficients are statistically significant at the 1% level. Null hypothesis of country dummies and partner dummies are zero is also rejected by the $F$-test. But in comparison with the trade equation estimation, the coefficient of log distance is lower than that of the trade equation. Thus, we suppose knowledge transportation needs lower transport cost than manufacturing goods transportation.

# 5 Wage equation estimation

## 5.1 Econometric preparation

Next, we estimate wage equation. Solving the wage equation eq(11) by $w$, and taking logarithms, we can derive an estimation function

$$\ln w_i = \frac{-\sigma\gamma \ln v_i + \log A}{\sigma\beta} + \frac{\varphi}{\beta} \ln KA_i + \frac{1}{\sigma\beta} \ln MA_i + \frac{\alpha}{\beta(\sigma - 1)} \ln SA_i. \tag{24}$$

Now, by using the assumption that internationally mobile factors are same value on every countries, we can set, for all country $i$, taking $v_i$ is constant such that $v_i = v$ and rearranging, we derive

$$\ln w_i = \xi + \psi_1 \ln KA_i + \psi_2 \ln SA_i + \psi_3 \ln MA_i + \eta_i. \tag{25}$$

And substituting fitted value of access indicator, we obtain an estimation function,

$$\ln w_i = \zeta + \psi_1 \ln \hat{KA}_i + \psi_2 \ln \hat{SA}_i + \psi_3 \ln \hat{MA}_i + \varepsilon_i. \tag{26}$$

For estimating this equation, we need to settle several difficulties. The first is errors in variables problem. Our estimation uses predicted value instead of unobservable real MA, SA, KA values. Thus, there exist the possibility an error term has a correlation with regressors. In order to resolve this possibility, we estimate by IV as well as OLS. On the IV estimation, we use two-step least square estimation (2SLS) . And we use White-robust standard error to correct heteroskedasticity condition. Instrument variables we use are the distance from the three main markets and sources of supply for manufactures and knowledge (the United States, Japan, and the European Union[*5]), and domestic total flight passengers, and dummies whether the residents of the country can understand English communication or not (English dummy). In order to test the overidentification condition, we use Hansen's $J$ statistics.

Second, there exist high correlations between regressors. A correlation coefficient between $\ln \hat{MA}$ and $\ln \hat{KA}$ is 0.768, moreover that between $\ln \hat{MA}$ and $\ln \hat{SA}$ is 0.899. Thus, we can not estimate simultaneously at least $\ln \hat{MA}$ and $\ln \hat{SA}$. Our concern is mainly knowledge access. For these reason, we remove supplier access, and use market access as manufacturing goods accessibility.

[*5] We use Belgium as the central point of the European Union.

## 5.2 Economic geography estimation

At first, we estimate wage equation by single regressor.

Table 3 Economic geography

| ln(GDP per worker) | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Observations | 55 | 55 | 55 | 55 |
| Year | 1999 | 1999 | 1999 | 1999 |
| ln(MA) | 0.748** | | | |
| $= \ln(DMA + FMA)$ | (0.107) | | | |
| ln(FMA) | | 0.709** | | |
| | | (0.170) | | |
| ln(KA) | | | 2.569** | |
| $= \ln(DKA + FKA)$ | | | (0.359) | |
| ln(FKA) | | | | 1.543** |
| | | | | (0.510) |
| $R^2$ | 0.4488 | 0.2274 | 0.3468 | 0.1050 |
| F(·) | 48.72 | 17.41 | 51.18 | 9.14 |
| Prob>F | 0.0000 | 0.0000 | 0.0000 | 0.0039 |

White heteroskedasticity robust standard errors in parentheses.
*Denotes statistical significance at the 10% level. **Denotes statistical significance at the 5% level.

The result is Table3. Column (1) regresses log GDP per worker on log predicted market access by OLS. The coefficient of MA is positive and statistically significant at 5% level. Column (3) regresses log GDP per worker on log predicted foreign market access. The coefficient of that is also positive and statistically significant at the 5% level. These results are consistent with Redding and Venables (2004).

Column (2) regresses log GDP per worker on log predicted knowledge access by OLS. The coefficient of KA is positive and statistically significant at the 5% level. And, knowledge access alone explains about 35% of the cross country wage disparity. Lastly, column (4) regresses log GDP per worker on log predicted foreign knowledge access by OLS. The coefficient of FKA is also positive and statistically significant at the 5% level. And foreign knowledge access alone explains 10% of cross country wage variation.

Because this estimation is single estimation, by plotting values, we can easily understand the relationship. The results are Figure 1-4. Each country is indicated by the World Bank's three letter code (see Data Appendix for details). For MA and KA, it is clear that the effect on

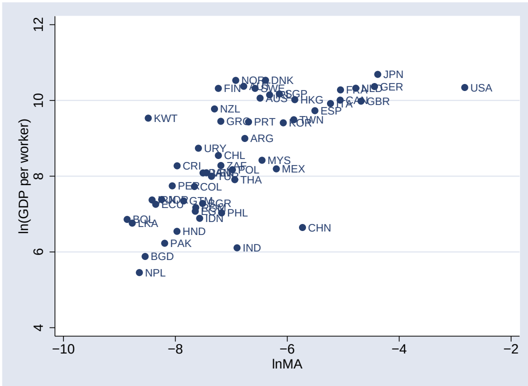



Figure 2 GDP per worker and MA.

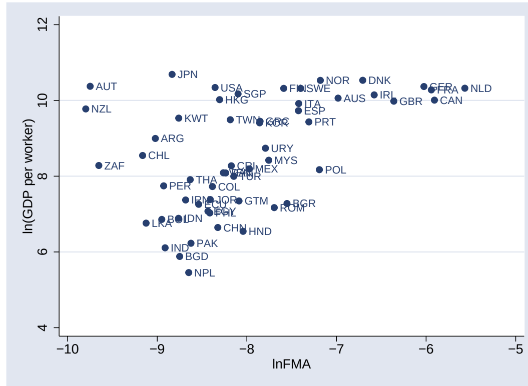



Figure 3 GDP per worker and FMA.

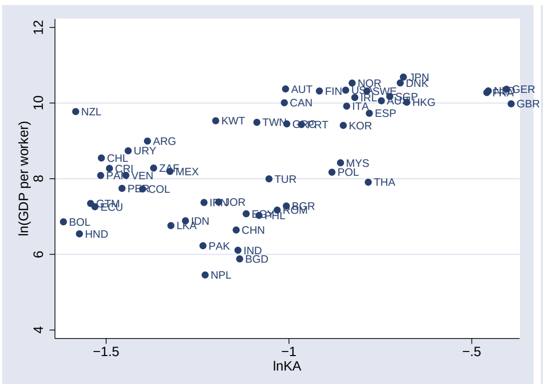



Figure 4 GDP per worker and KA.

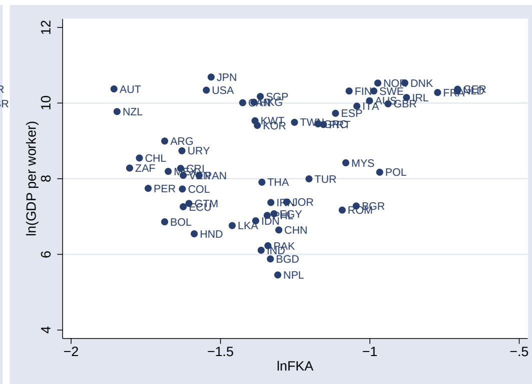



Figure 5 GDP per worker and FKA.

manufacturing wage is quite robust. But FMA and FKA, at the upper left, some outliers exist, such that Japan and United States. These countries have huge market in their own country, and they also have huge knowledge capital. Thus, we suppose these countries can increase their GDP per worker without using foreign manufacturing markets and knowledge sources.

## 5.3 The Estimation with Physical Geography and Political Factors

Next, we estimate wage equation with physical geography and political factors. Control variables are selected as same as Redding and Venables (2004). These control variables are divided into two groups. For one thing is a physical geographic group. It includes hydrocarbons per capita, arable land per capita, fraction land in geographical tropics, and prevalence of malaria. For another is a political group. It includes risk of expropriation or protection of property rights, socialist rule during 1950-1995, and the occurrence of an external war. (see Data Appendix) At first, we estimate the last estimation controlling these factors. Table4 presents the result.

Column (1) and column (3) regresses GDP per worker on log market access and log knowledge access by OLS controlled by geographical and political factors. The coefficients of MA and KA are still positive and statistically significant at the 5% level. This result shows total market access and knowledge access still affect manufacturing wage even controlled by physical geography and political factors.

Because endogeneity may exist, column (2) and column (4) regresses GDP per worker on MA and KA by instrumental variable method (IV). We instrument MA and KA by the distance from the three main markets and sources of supply for manufactures and knowledge (the United States, Japan, and the European Union), and domestic total flight passengers, and English dummy. These instruments passed the orthogonality condition by the test of Hansen's $J$ statistics on every estimation. And on the first-stage regression, these instruments are highly significant by the F-test which null hypothesis is all coefficients are zero. Thus, we conclude the using of these instruments are justified. The coefficient of lnKA is still positive and significant at the 5% level. And the coefficients of lnMA is still positive, but its significant level decreases 10% level. From these result, we can conclude that total knowledge access affects international wage variation robustly.

Next, we divide market access and knowledge access between its foreign effect and domestic effect. And we regresses GDP per worker by these regressors simultaneously. The results are column (5) and column (7). The coefficient of foreign market access and that of foreign knowledge access are positive and statistically significant at the 5% level. Column (6) and column (8) regresses by IV. We instrument DMA, FMA, DKA, and FKA by the same variables as column (2) and (4) . The use of these instruments is justified by the same tests as column (2) and (4). Same as column (5) and column (7), coefficients of FMA and FKA is positive and statistically significant at the 5% level. These results shows that foreign knowledge access plays the important role on increaseing manufacturing wage and it is considerably robust.

Note that, regression of column (6)-(8) is misspecifing from our theory. We cannot divide $\ln(KA) = \ln(DKA + FKA)$ between $\ln(DMA)$ and $\ln(FMA)$. In spite of the misspecification, the coefficients of these variables are highly statistically significant. Thus, we can conclude that these effects are highly robustness.

Table 4  Economic geography, physical geography, and political factors

| ln(GDP per worker) | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Observations | 51 | 49 | 51 | 49 | 51 | 49 | 51 | 49 |
| Year | 1999 | 1999 | 1999 | 1999 | 1999 | 1999 | 1999 | 1999 |
| ln(MA) | 0.380** | 0.186* | | | | | | |
| = ln($DMA + FMA$) | (0.124) | (0.101) | | | | | | |
| ln(DMA) | | | | | 0.225** | 0.198** | | |
| | | | | | (0.065) | (0.089) | | |
| ln(FMA) | | | | | 0.281** | 0.316** | | |
| | | | | | (0.109) | (0.132) | | |
| ln(KA) | | | 1.715** | 1.024** | | | | |
| = ln($DKA + FKA$) | | | (0.363) | (0.418) | | | | |
| ln(DKA) | | | | | | | 0.376** | 0.409** |
| | | | | | | | (0.158) | (0.204) |
| ln(FKA) | | | | | | | 0.959** | 0.644** |
| | | | | | | | (0.287) | (0.293) |
| ln(hydrocarbons | -0.009 | -0.002 | -0.003 | 0.002 | -0.016 | -0.003 | -0.012 | -0.001 |
| per capita) | (0.029) | (0.029) | (0.025) | (0.026) | (0.027) | (0.028) | (0.027) | (0.025) |
| ln(arrable land | -0.017 | -0.015 | −0.027** | −0.021** | −0.020* | −0.016* | −0.025** | −0.022** |
| per capita) | (0.010) | (0.001) | (0.010) | (0.010) | (0.010) | (0.010) | (0.011) | (0.011) |
| Fraction land in | -0.416 | -0.479 | -0.309 | -0.418 | −0.508* | -0.455 | -0.354 | -0.412 |
| geographical tropics | (0.343) | (0.375) | (0.320) | (0.341) | (0.275) | (0.356) | (0.337) | (0.311) |
| Prevalence of | −1.962** | −2.262** | −1.967** | −2.263** | −1.768** | −2.190** | −1.880** | −2.188** |
| malaria | (0.521) | (0.624) | (0.531) | (0.598) | (0.426) | (0.596) | (0.571) | (0.562) |
| Risk of | −0.257** | −0.316** | −0.262** | −0.305** | −0.225** | −0.284** | −0.282** | −0.283** |
| expropriation | (0.072) | (0.080) | (0.073) | (0.077) | (0.070) | (0.084) | (0.080) | (0.079) |
| Socialist rule | -0.658 | -0.439 | -0.517 | -0.413 | -0.779 | -0.568 | -0.404 | -0.372 |
| 1950-85 | (0.676) | (0.563) | (0.529) | (0.497) | (0.668) | (0.560) | (0.594) | (0.531) |
| External war | -0.078 | -0.143 | -0.19 | -0.238 | -0.065 | -0.084 | -0.221 | -0.224 |
| 1960-85 | (0.231) | (0.202) | (0.197) | (0.176) | (0.196) | (0.178) | (0.260) | (0.196) |
| Hansen($p$-value) | | 0.335 | | 0.764 | | 0.450 | | 0.863 |
| Estimation | OLS | IV | OLS | IV | OLS | IV | OLS | IV |
| $R^2$ | 0.8027 | - | 0.8214 | - | 0.8316 | - | 0.8024 | - |
| F(·) | 38.58 | 42.42 | 58.84 | 65.04 | 41.66 | 43.48 | 32.72 | 40.54 |
| Prob>F | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

White heteroskedasticity robust standard errors in parentheses.
*Denotes statistical significance at the 10% level. **Denotes statistical significance at the 5% level.

Among the control variables, the coefficient of prevalence of malaria and risk of expropriation or protection of property rights are negatively signed and statistically significant at the 5% level in every column.

In addition, the result about market access is also consistent with Redding and Venables (2004).

## 5.4 The relationship between knowledge access and market access

Lastly, we analyze the relationship between market access and knowledge access. We estimate market access and knowledge access simultaneously. Table5 presents the result.

Column (1) regresses GDP per worker on MA and KA by OLS. The coefficient of lnMA is positive and statistically significant at the 10% level, and the coefficient of KA is also positive and statistically significant at the 5% level. Because of the possibility of endogeneity, column (2) regresses by IV. The instrument variables are same as the last section (the distance from the United States, Japan, and the European Union, and domestic total flight passengers, and English dummy), and the use of these instruments are also justified by the same tests as the last section (Hansen's statistics and first stage regression). The coefficient of KA is still positive and statistically significant at the 5% level, but the coefficient of MA is not statistically significant at least the 10% level. From these results, we anticipate total knowledge access affects manufacturing wage at least same degree as total market access.

Column (3) and (4) regresses GDP per worker on domestic and foreign accessibility effect by OLS and IV. The coefficients of FMA and FKA are not statistically significant at least at 10% level. Moreover, on column (4), the coefficient of FMA is negatively signed. Although the variable of each of is significant in the regression of the last section, when it is estimated simultaneously, no variable is significant. Since it was suspected that correlation between regressors have started the problem in such a situation, we compute the correlation between regressors and we confirm that the regressors are highly correlated. Concretely, the correlation coefficient between FMA and FKA is 0.899, and between DMA and DKA is 0.890. From this result, we conclude that serious multicolinearity problem occurs in these regressions.

Without removing this multicolinearity, we cannot analyze the relationship between knowledge access and market access in detail. About that, it considers as a future subject. But from column (1) and column (2), knowledge access seems to be important on the international wage variation as well as market access.

Table 5 The relationship between knowledge access and market access

| ln(GDP per worker) | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Observations | 51 | 49 | 51 | 49 |
| Year | 1999 | 1999 | 1999 | 1999 |
| lnMA | 0.208* | 0.081 | | |
| $= \ln(DMA + FMA)$ | (0.119) | (0.063) | | |
| lnDMA | | | 0.215 | 0.126 |
| | | | (0.097) | (0.161) |
| lnFMA | | | 0.121 | -0.075 |
| | | | (0.174) | (0.385) |
| lnKA | 1.274** | 0.929** | | |
| $= \ln(DKA + FKA)$ | (0.427) | (0.421) | | |
| lnDKA | | | 0.063 | 0.138 |
| | | | (0.192) | (0.557) |
| lnFKA | | | 0.719 | 0.896 |
| | | | (0.525) | (1.070) |
| ln(hydrocarbons | -0.006 | 0.001 | -0.017 | -0.006 |
| per capita) | (0.025) | (0.025) | (0.026) | (0.024) |
| ln(arrable land | −0.025* | −0.021** | -0.023 | −0.021** |
| per capita) | (0.010) | (0.010) | (0.011) | (0.010) |
| Fraction land in | -0.291 | -0.394 | -0.410 | -0.410 |
| geographical tropics | (0.297) | (0.327) | (0.264) | (0.284) |
| Prevalence of | −1.896** | −2.225** | −1.798** | −2.215** |
| malaria | (0.491) | (0.577) | (0.438) | (0.596) |
| Risk of | −0.224** | −0.285** | −0.222** | −0.277** |
| expropriation | (0.068) | (0.077) | (0.073) | (0.106) |
| Socialist rule | -0.697 | -0.498 | -0.811 | -0.580 |
| 1950-85 | (0.585) | (0.495) | (0.677) | (0.678) |
| External war | -0.070 | -0.168 | -0.071 | -0.160 |
| 1960-85 | (0.184) | (0.164) | (0.187) | (0.159) |
| Hansen (*p*-value) | | 0.892 | | 0.561 |
| Estimation | OLS | IV | OLS | IV |
| $R^2$ | 0.8349 | - | 0.8390 | - |
| F(·) | 65.14 | 88.25 | 46.54 | 53.99 |
| Prob>F | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

White heteroskedasticity robust standard errors in parentheses.
*Denotes statistical significance at the 10% level. **Denotes statistical significance at the 5% level.

## 6 Concluding Remarks

By using NEG framework, we investigate the role of knowledge access and market access on the international wage disparity. We cannot analysis circumstantially the relationship between knowledge access and market access in detail because of multicolinearity, but we can conclude that knowledge accessibility plays the important role on increasing manufacturing wages as well as market accessibility. For example, foreign knowledge access alone explain 11% of the cross-country income variation. And we confirmed that the effect of foreign knowledge access is robust to controlling primary geographical factors, political factors, and institutional factors.

Interchange of people as the transportation of knowledge is very important as well as that of manufacturing goods for increasing manufacturing wage is one conclusion drawn from our research. Thus, in order to increase per capita income, developing countries should reduce travel cost from foreign countries as well as trade cost. However, from our research, though reducing tariff and improving knowledge and market openness, the wall of distance still highly remains for remote regions. This is the serious conclusion for remote countries. But it must be noted that our analysis is set the each country's expenditure and location of production are exogenous, that means if big manufacturing market or knowledge source emerged near the remote country, the country's manufacturing wage might be improved.

## Data Appendix

**Bilateral Trade**: data on bilateral trade flows are from the World Bank COMTRADE database.

**Bilateral Flight Passengers**: data on bilateral flight passengers is from ICAO's On-Flight Origin and Destination.

**GDP per Worker**: data on current price (US dollars) GDP and the number of workers are from World Development Indicators 2003.

**Bilateral Distance**: data on bilateral distance is from Western Hemispheric Research Resources[*6]

Domestic flight passengers: data on domestic flight passengers is from http://www.cid.harvard.edu/ciddata.ciddata.html.

English Dummy: English dummy is whether the residents of the country can understand English communication or not. This is available Data book of the World 2004.

Risk of Expropriation: extent of protection of property rights, measured on a scale from 1 to 5, where a higher score indicates weaker protection of property rights. The data can be downloaded from http://www.cid.harvard.edu/ciddata.ciddata.html.

Physical Geography and Institutional, Social, and Political Characteristics: data on hydrocarbons (deposits of petroleum and natural gas) per capita, fraction of land area in the geographical tropics, prevalence of malaria, socialist rule, and the occurrence of an external war are from Gallup et al. (1999). The data can be downloaded from http://www.cid.harvard.edu/ciddata.ciddata.html.

Country Codes: Argentina (ARG) Australia (AUS) Austria (AUT) Bangladesh (BGD) Bulgaria (BGR) Bolivia (BOL) Canada (CAN) Chile (CHL) China (CHN) Colombia (COL) Costa Rica (CRI) Denmark (DNK) Ecuador (ECU) Egypt (EGY) Spain (ESP) Finland (FIN) France (FRA) United Kingdom (GBR) Germany (GER) Greece (GRC) Guatemala (GTM) Hong Kong (HKG) Honduras (HND) Indonesia (IDN) India (IND) Ireland (IRL) Iran (IRN) Italy (ITA) Jordan (JOR) Japan (JPN) Korea, Republic of (KOR) Kuwait (KWT) Sri Lanka (LKA) Mexico (MEX) Malaysia (MYS) Netherlands (NLD) Norway (NOR) Nepal (NPL) New Zealand (NZL) Pakistan (PAK) Panama (PAN) Peru (PER) Philippines (PHL) Poland (POL) Portugal (PRT) Romania (ROM) Singapore (SGP) Sweden (SWE) Thailand (THA) Turkey (TUR) Taiwan (TWN) Uruguay (URY) United States (USA) Venezuela (VEN) South Africa (ZAF)

Means and Standard Deviation of Key Variables: Means and standard deviation of key variables is on Table 6-10.

[*6] http://www.macalester.edu/research/economics/PAGE/HAVEMAN/Trade.Resources/Data/Gravity/dist.txt

Table 6 Export

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Export | 2871 | 1.36E+09 | 7.25E+09 | 1 | 2.05E+11 |
| Travel | 683 | 163657.6 | 244601.7 | 673 | 1933941 |

Table 7 GDP per worker

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| GDP per worker | 55 | 25149.84 | 25005.39 | 504 | 81600 |

Table 8 MA series

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| MA | 55 | 0.003027 | 0.008202 | 0.000142 | 0.059079 |
| DMA | 55 | 0.00247 | 0.008106 | 1.04E-06 | 0.058843 |
| FMA | 55 | 0.000557 | 0.000752 | 5.56E-05 | 0.003819 |

Table 9 SA series

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| SA | 55 | 0.001885 | 0.003702 | 7.25E-05 | 0.022919 |
| DSA | 55 | 0.001558 | 0.003657 | 5.50E-08 | 0.022801 |
| FSA | 55 | 0.000327 | 0.000387 | 4.01E-05 | 0.002255 |

Table 10 KA series

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| KA | 55 | 0.360137 | 0.122106 | 0.198506 | 0.67549 |
| DKA | 55 | 0.083558 | 0.073482 | 0.003012 | 0.286795 |
| FKA | 55 | 0.27658 | 0.085869 | 0.15624 | 0.493524 |

# References


[1] Baldwin, R.E., R. Forsild, Ph. Martin, G.I.P. Ottaviano and F. Robert-Nicoud (2003) *Economic geography and public policy*, Princeton University Press, Princeton.

[2] Baldwin, R.E., Ph. Martin, (2004) Agglomeration and Regional Growth, J.V. Hender-

son, J.-F. Thisse,ed., *Handbook of Regional and Urban Economics*, 4, 2671-2711, Elsevier, Amsterdam.

[3] Baldwin, R.E., Ph. Martin, G.I.P. Ottaviano (2001) Global Income Divergence, Trade, and Industrialization: The Geography of Growth Take-Offs, *Journal of Economic Growth* 6: 5-37

[4] Barro, R.J. and Sala-i-Martin (1995) *Economic growth*, McGraw-Hill, New York.

[5] *Data Book of The World 2004* (2004) Ninomiya Shoten, Tokyo (in Japanese)

[6] Fujita, M., Krugman, P.R., Venables, A.J. (1999) *The Spatial Economy: Cities, Regions and International Trade.*, MIT Press, Cambridge.

[7] Grossman, G., E. Helpman (1991) *Innovation and Growth in the World Economy*, MIT Press, Cambridge.

[8] Gallup, J.,L., J.D.Sachs, A.D.Mellinger. (1999) Geography and Economic Development, *International Regional Science Review* 22, 2: 179-232

[9] Head, K., T. Mayer. (2004) The Empirics of Agglomeration and Trade, J.V. Henderson, J.-F. Thisse,ed., *Handbook of Regional and Urban Economics*, 4, 2609-2670, Elsevier, Amsterdam.

[10] Krugman, P.R. (1991) *Geography and Trade*. MIT Press, Cambridge.

[11] Ottaviano, G.I.P., J.-F. Thisse (2004) Agglomeration and Economic Geography, J.V. Henderson, J.-F. Thisse,ed., *Handbook of Regional and Urban Economics*, 4, 2564-2608, Elsevier, Amsterdam.

[12] Redding, S., P.K.Schott. (2003) Distance, Skill Deepening and Development: Will Peripheral Countries Ever Get Rich?, *Journal of Development Economics* 72: 515-541

[13] Redding, S., A.J.Venables. (2004) Economic geography and international inequarlity, *Journal of International Economics* 62: 53-82.