Testing the New Economic Geography's wage equation: a case study of Japan using a spatial panel model

Chian-Yue Wang* and Robert Haining**

*Graduate Institute of Urban Planning, National Taipei University, Taiwan

151, University Road, San-shia, New Taipei City, 237-41, Taiwan

andywang@mail.ntpu.edu.tw (corresponding author)

**Department of Geography, University of Cambridge, UK

Downing Place, Cambridge CB2 3EN, UK

rph26@cam.ac.uk

Acknowledgement: many thanks to Prof. Bernard Fingleton, Prof. Ron Martin, Prof. Eiichi Yamaguchi, and Prof. Hiroki Tanaka for their advice, encouragement and help.

Testing the New Economic Geography's wage equation: a case study of Japan using a spatial panel model

This paper estimates the parameters of the wage equation of the new economic geography (NEG) using a newly developed spatial panel model. The results show that wage rate variation across different prefectures in Japan can be explained by market potential, which is a key variable in NEG theory, whilst controlling for variation in labour efficiency. Spatial heterogeneity is particularly important in the context of Japan in part because of its complex physical geography and the spatial distribution of its principal urban centres. The paper considers the challenges associated with representing the spatial relationships between prefectures describing and implementing different approaches to measuring transport costs.

Keywords: new economic geography, agglomeration, spatial panel model

1. Introduction

The New Economic Geography (NEG) has emerged since the early 1990s and been recognized by economists and some geographers as a new genre of theory which explains the formation of economic agglomerations in geographical space (Krugman, 1991a, 1991b; Fujita et al., 1999). Krugman, the initiator of the New Economic Geography, attempted to internalize the idea of pecuniary externality, deriving from Marshall's threefold classification (1920, p271) which was proposed to explain the agglomeration phenomena of spatial economic activities, into a normative economic model upon which the demand and supply mechanism operates. Under the mechanism of general equilibrium, the key features of the New Economic Geography are to explain the spatial phenomena of economic activities and to explore the stability of the equilibrium, by deriving and analyzing sets of nonlinear equations.

However in order to derive sets of equations, these theoretical models have to be constructed using rather unrealistic assumptions and omitting certain factors known to be relevant in the real world. It was just recently that empirical economists and quantitative geographers started to test the theoretical frameworks against real data and estimate some of the basic parameters (Roos, 2001; Head and Mayer, 2003; Redding and Venables, 2004; Mion, 2004; Hanson, 2005; Kiso, 2005; Brakman et al. 2006; Fingleton, 2006, 2008, 2011; Baltagi et al., 2014). One of the most successful equations tested by researchers is the wage equation. This equation which illustrates the subtle relationship between nominal wage levels and market potential is the core element of the new economic geography and is widely used in studying different levels of spatial aggregation.

Drawing on a socio-economic panel dataset at the prefecture level for Japan from 1977 to 2006 not previously analysed, one of the major aims of this research is to estimate key parameters of the wage equation and test whether market potential is a significant variable in explaining wage variation across Japan. In addition, given the topographic characteristics of Japan, we contemplate whether the mountainous landscape can be reflected in the process of estimation through the specification of the transport cost variable, an essential component of market potential which makes real geography matter. Spatial dependence is a fundamental property of data collected across spatial units particularly where unit boundaries are arbitrarily drawn. The importance of this property when fitting spatial econometric models is now well established and in this research we incorporate a model for spatial dependence (Anselin, 1988; Haining, 2003; Arbia, 2014). Finally, following Fingleton (2011), endogeneity between nominal wage and market potential will be allowed for in the estimation process.

2. The NEG model

The NEG model (Fujita et al., 1999, Ch4) is built on the Dixit-Stiglitz monopolistic competition model (Dixit and Stiglitz, 1977) with the introduction of the transport cost variable, which ensures spatial relationships are incorporated. The NEG model features increasing returns to scale production technology internal to micro-economic agents, imperfect competition market structure, and multiple equilibria within which the spatial agglomeration or dispersion of economic agents are determined endogenously. In essence, the model reduces to several simultaneous non-linear equations, of which the wage equation is, as Head and Mayer (2003) stress, one of the most essential equations deriving from NEG.

Fujita et al. (1999) consider an economy with two sectors, agriculture and manufacturing. The agricultural sector is perfectly competitive and produces a single, homogeneous good, whereas the manufacturing sector provides a large variety of differentiated goods. The perfectly competitive agricultural sector is the counterpart to the action taking place in the increasing returns, imperfectly competitive manufacturing sector. The wage equation set out in Fujita et al. (1999, Ch4) is given as follows:

$$w_r^M = \left[\sum_{s=1}^{R} Y_s \left(T_{rs}^M \right)^{1-\sigma} G_s^{\sigma-1} \right]^{\frac{1}{\sigma}} = P_r^{\frac{1}{\sigma}} \qquad (1)$$

where w_r^M is the nominal wage rate in the manufacturing sector (M) of region r, Y_s is total income in region s, T_{rs}^M is transport cost in the manufacturing sector between regions r and s, G_s is the price index in region s, σ is the elasticity of substitution between any two varieties of manufacturing goods, and P_r is the market potential of region r.

The wage equation gives the manufacturing wage where manufacturing firms in each region break even, given the income levels and price indices in all regions and the costs of shipping into these regions. For the purpose of simplification, assume that the price indexes in all regions are similar. Then the wage equation implies that the nominal wage rate in region r has a propensity to be higher if incomes (demands) in region r and the neighboring regions (with low transport costs) are high. This is because firms are capable of paying higher wages if they have good access to a large market. This kind of economic relationship is generally recognized as a backward linkage (Hirschman, 1958).

3. Methodology

3.1 The KKP spatial panel model

Early papers estimating the wage equation were limited to cross-sectional analyses (see for example, Hanson, 2005; Kiso, 2005). There is nothing inappropriate in using a single cross-sectional regression if the model is correctly specified. However, such an approach has its limitations, especially in modeling spatial-unit specific heterogeneity. Recent developments in NEG empirics however show that this approach has been replaced by spatial panel data modelling (see for example, Fingleton 2011; Baltagi et al. 2014). Two of the most significant advantages of spatial panel models relative to cross sectional models are their ability to control for heterogeneity and spatial dependence, arising from the artificial delineation of spatial unit boundaries and spatial interaction effects (Hsiao, 2003; Baltagi 2008; Elhorst, 2014).

A new panel data model (Kapoor, Kelejian, and Prucha, 2007) which allows for cross-sectional spatial dependence and time dependencies in the disturbance term is used in this research (hereafter referred to as the KKP model). The fundamental specifications of the KKP model are as follows:

In each time period t = 1,…,T the data are generated according to the following model:

$$Y_N(t) = X_N(t)\beta + u_N(t) \tag{2}$$

where $Y_N(t)$ denotes the $N \times 1$ vector of observation on the dependent variable across the N spatial units in period t, $X_N(t)$ denotes the $N \times K$ matrix of observations on exogenous regressors in period t (which may contain the constant term), β is the corresponding $K \times 1$ vector of regression parameters, and $u_N(t)$ denotes the $N \times 1$ vector of disturbance terms. Kapoor et al. specify their model conditional on the realized values of the regressors and so view $X_N(t)$, t = 1,…,T as matrices of constants. A widely used approach to modeling spatial dependence was suggested by Cliff and Ord (1973, 1981). Kapoor et al (2007) follow their approach to modeling the disturbance process in each period by specifying the following first order spatial autoregressive process:

$$u_N(t) = \rho W_N u_N(t) + \varepsilon_N(t) \tag{3}$$

where W_N is an $N \times N$ spatial weights matrix of known constants which does not involve t, ρ is a scalar autoregressive parameter, and $\varepsilon_N(t)$ is an $N \times 1$ vector of innovations in period t. (The name "innovations" is given by Kapoor et al, 2007) Stacking the observations in (2) and (3) we get

$$Y_N = X_N\beta + u_N \tag{4}$$

$$u_N = \rho(I_T \otimes W_N)u_N + \varepsilon_N \tag{5}$$

where $Y_N = [Y_N^{'}(1),...,Y_N^{'}(T)]^{'}$, $X_N = [X_N^{'}(1),...,X_N^{'}(T)]^{'}$, $u_N = [u_N^{'}(1),...,u_N^{'}(T)]^{'}$, and $\varepsilon_N = [\varepsilon_N^{'}(1),...,\varepsilon_N^{'}(T)]^{'}$ where $^{'}$ denotes matrix or vector transpose.

To allow for the innovations to be correlated over time, the following error component structure for the innovation vector ε_N is assumed:

$$\varepsilon_N = (e_T \otimes I_N)\mu_N + v_N \tag{6}$$

where, e_T is a T by one vector of ones, μ_N represents the vector of unit specific error components, and $v_N = [v_N^{'}(1),...,v_N^{'}(T)]'$ contains the error components that vary over both the cross-sectional units and time periods. In scalar notation, Equation (6) can be expressed as:

$$\varepsilon_{it,N} = \mu_{i,N} + v_{it,N} \tag{7}$$

The following assumptions are made.

Assumption 1[1]: (a) $v_{it,N} \sim iid(0,\sigma_v^2)$, (b) $\mu_{i,N} \sim iid(0,\sigma_\mu^2)$, and (c) the processes $\{v_{it,N}\}$ and $\{\mu_{i,N}\}$ are independent. **Assumption2**: (a) All the diagonal elements of W_N are zero, (b) $|\rho| < 1$, and (c) The matrix $I_N - \rho W_N$ is non-singular, hence invertible.

In light of (6) or (7) it follows that Assumption 1 implies $E(\varepsilon_{it,N}) = 0$ and

$$E(\varepsilon_{it,N}\varepsilon_{js,N}) = \begin{bmatrix} \sigma_\mu^2 + \sigma_v^2 & if & i = j; t = s \\ \sigma_\mu^2 & if & i = j; t \neq s \\ 0 & otherwise & \end{bmatrix}$$

[1] the original assumption1 in full is in Kapoor et al, 2007, pp100

The innovations $\varepsilon_{it,N}$ are autocorrelated over time, but are not spatially autocorrelated across units. In matrix notation we have $E(\varepsilon_N)=0$ and the variance-covariance matrix of ε_N is

$$\begin{aligned}\Omega_{\varepsilon,N} &= E(\varepsilon_N \varepsilon'_N) = \sigma_\mu^2 (J_T \otimes I_N) + \sigma_v^2 I_{NT} \\ &= \sigma_v^2 Q_{0,N} + \sigma_1^2 Q_{1,N}\end{aligned} \tag{8}$$

Where $\sigma_1^2 = \sigma_v^2 + T\sigma_\mu^2$ and $J_T = e_T e'_T$ is a $T \times T$ matrix of unit elements.

$$Q_{0,N} = (I_T - \frac{J_T}{T}) \otimes I_N$$

$$Q_{1,N} = \frac{J_T}{T} \otimes I_N$$

The matrices $Q_{0,N}$ and $Q_{1,N}$ are symmetric, idempotent, and orthogonal to each other

Kapoor et al. (2007, p102-110) defines three sets of generalised moment estimators for ρ, σ_v^2 and σ_μ^2, or equivalently for ρ, σ_v^2 and σ_1^2. Once the parameters in the variance-covariance matrix of ε_N are estimated, then a spatially feasible generalised least squares procedure can be applied to estimate the regression parameters β.

3.2 Estimation of the wage equation

The wage equation links the nominal wage of a region to its market potential, which was defined inside the square bracket of (1). Head and Mayer (2006) augment the simple wage equation of Fujita et al (1999) by adding a labour efficiency variable to the micro-assumptions of NEG. The extended wage equation is written:

$$w_{rt}^M = P_{rt}^{\frac{1}{\sigma}} A_{rt} \tag{9}$$

A_{rt} : labour efficiency level in region r at time t

The time subscript is introduced into (9) in order to capture changes in regional shares of the country's supply of workers in the manufacturing sector and in the competitive economic sector overall which will affect Y_s , G_s and therefore P_r over time. According to NEG theory, manufacturing workers will migrate from low to high real wage areas, so these shares will vary over time. In the long run, the economic landscape will end up in several possible equilibria.

By taking the natural log of both sides of equation (9), the empirical wage equation with a disturbance term can now be written:

$$\ln w_{rt}^M = \frac{1}{\sigma} \ln P_{rt} + b_1 \ln A_{rt} + b_0 + u_{rt} \tag{10}$$

u_{rt} is the disturbance term and b_0 is the constant term. The key parameter of equation (10) is σ (here denoting the elasticity of substitution). However, σ also appears in the power terms of T_{rs}^M and G_s in the calculation of market potential. The estimating procedure will require an iterative approach.

Firstly, the researcher will assume a value of σ for constructing the value of market potential of each region over different time periods. The model is then run to obtain an estimated value of σ. If the estimated value of σ is very different from the starting value, re-define a new value of σ which is closer to the estimate at the previous step. After several iterations, the estimated value of σ will converge.

In terms of how to implement the KKP model in the estimation of the wage equation, one essential point needs to be addressed further: the endogeneity issue involved in the estimation process. The KKP model is based on an assumption that regressors are exogenous. However, when we re-examine the wage equation (1), we notice that the nominal wage rate for the manufacturing sector is closely related to the total income in each region, which is one of the major components used to construct market potential. This implies the regressor, market potential, is endogenous.

The KKP model allowing for endogeneity between dependent and independent variables is generalised by Fingleton (2008, p548-549). The estimation process involves the use of an instrumental variable (IV) approach, and is given the acronym FGS2SLS (feasible generalized spatial two stage least squares).

3.3 The fixed-effects panel model with spatially autoregressive disturbances

In the estimation of the wage equation, we also consider a fixed effects panel model with spatially autoregressive disturbances. Unlike the random effects KKP panel model which models the unobserved individual effects as a random variable, the fixed effects model actually estimates each of the unobserved individual effects (dummies). The model is run by using the Matlab software with routines provided by Elhorst (2010). The estimation method is maximum likelihood. The model specification is as follows:

$$Y_N = X_N\beta + (e_T \otimes I_N)\mu_N + u_N \tag{11}$$

$$u_N = \rho(I_T \otimes W_N)u_N + \varepsilon_N$$

where $Y_N = [Y_N^{'}(1),...,Y_N^{'}(T)]^{'}$, $X_N = [X_N^{'}(1),...,X_N^{'}(T)]^{'}$, $u_N = [u_N^{'}(1),...,u_N^{'}(T)]^{'}$, and $\varepsilon_N = [\varepsilon_N^{'}(1),...,\varepsilon_N^{'}(T)]^{'}$, μ_N is a $N \times 1$ vector representing the **unobserved fixed effects**, e_T is a $T \times 1$ vector of ones, W_N is an $N \times N$ spatial weight matrix of assumed constants which does not involve t, ρ is a scalar autoregressive parameter, $u_N(t)$ denotes the $N \times 1$ vector of **disturbance terms** in period t, $\varepsilon_N(t)$ is an $N \times 1$ vector of **innovations**[2] in period t.

The estimated results of the fixed effects panel model will serve as counterparts to those of the KKP panel model. It is worth mentioning that the fixed effects panel model cannot take account of the endogenous relationship between the wage rate and the market potential. Hence, the regressors and disturbances are not orthogonal. In this case, the estimated parameters may be biased (Greene, 2011).

4. Data and measuring market potential

4.1 Wage data

In the wage equation of Fujita et al., (1999), wage means the nominal wage rate in the manufacturing sector. The nominal wage rate in the manufacturing sector for Japan can be found in the Basic Survey on Wage Structure conducted by the Minister's Secretariat (including the Statistics and Information Department) of the Ministry of Health, Labour and Welfare (See Appendix). Wage

[2] The term innovation, was used in the KKP model, but these are often referred to as the error term in econometric textbooks. For the moment we call them innovations for the convenience of making the direct comparison with the specification of the KKP model.

data are surveyed annually, so it is feasible to build up a panel set covering several time periods. In this research, a thirty year panel dataset ranging from 1977 to 2006 will be used to estimate the key parameters of the NEG model.

In order to explore the wage data and in particular, the spatial pattern of the data, we conduct a mapping of the wage data for the manufacturing sector in 2000 (see figure 1). The wage in the manufacturing sector in Tokyo-to is the highest in Japan[3]. The average wage in Tokyo is 6290.99 thousand yen per year, while in the prefectures with lower wages, such as Aomori in the north of Japan, manufacturing workers only earn 3007.40 thousand yen per year. The wage pattern shows that the comparatively richer prefectures are located along the coastline of the Pacific Ocean between Kanto and Kansai. In the southern part of Japan, particularly Kyushu island, wages are again generally low, although Fukuoka is an exception with a comparatively high average wage.

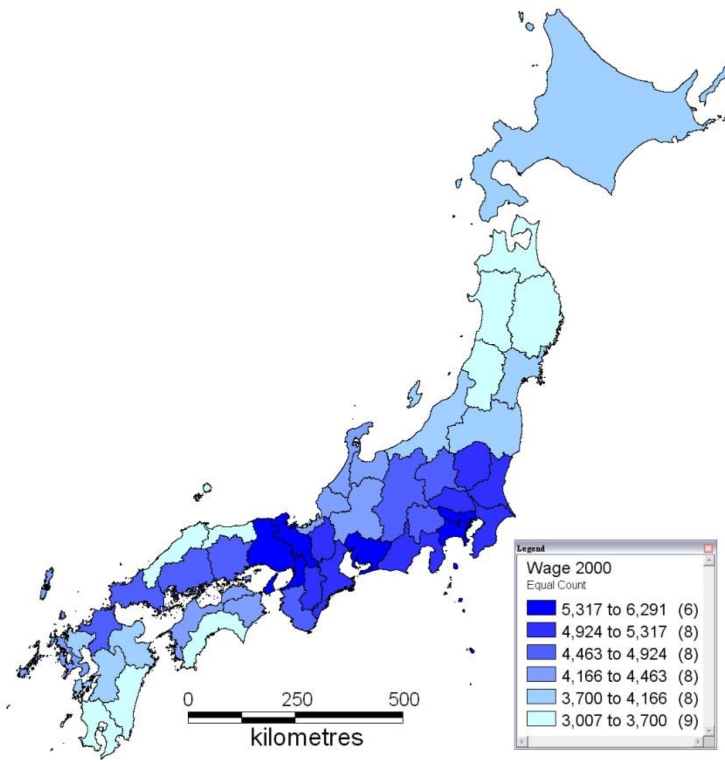


Figure 1 Wage in the manufacturing sector in 2000 (per capita).

4.2 Measuring market potential

Market potential is composed of three components, Y_s , T_{rs}^M , G_s . Y_s means total income in prefecture s. It reflects the purchasing power in each prefecture. The total taxable income of each prefecture collected by the tax office of the Ministry of Internal Affairs and Communications is used to measure Y_s . It is possible to use another data set for Y_s , namely the total residents' income of each prefecture published by the Economic and Social Research Institute of the Cabinet Office. However, total residents' income suffers from a data consistency problem. Before 1996, the data were surveyed and tabulated according to SNA68 (System of National Accounts 68). After 1996, the Japanese government adopted a new standard namely SNA93 (System of National Accounts 93). It proves to

[3] See Figure1a Japanese prefectures in Appendix

be difficult to construct a consistent panel data set of total residents' income across two accounting systems[4].

The second component, T_{rs}^{M} , means transport costs for manufactured goods between prefecture r and prefecture s. Transport costs can be measured either by traveling time or distance. According to previous empirical papers of NEG (See for example Mion, 2004; Hanson, 2005), distance data are most widely used due to their availability. In many cases, researchers assume a geometric centre for each region or state, and then calculate the crow flight distance. The approach is to use the minimum distance between two geometric centres.

However, since more than 70% of the land of Japan is covered by mountains, and the economic hubs of prefectures are mostly located along the coastline, it is particularly inappropriate to assume the majority of economic activity happens in the geometric centre of a prefecture. In addition, it is important to consider how the mountainous landscape of Japan affects the transportation route network. Using the crow flight distance as a measure of transport costs, it is possible to introduce considerable bias since real transportation networks must reconcile with the topography of Japan. On the other hand transportation networks in the real world, like road systems, are too complex to analyze. For example, from one economic centre to another, there are many alternative roads which provide different distance measures between these two points.

One feasible and pragmatic measure of transport costs is to look at the railway network of Japan. Railway transportation in Japan is one of the most important means of passenger transport, especially for mass and high-speed travel between cities as well as for commuter transport in and around metropolitan areas. Among advanced economies, the relative share of railroads in total passenger transport in Japan is high. For example, there are more than 27,000 kilometers of rail lines crisscrossing Japan. In fiscal 2006, Japan’s railways carried 22.24 billion passengers[5]. In comparison, Germany has over 40,000 kilometers of railways, but in 2008 carried only 2.2 billion passengers per year[6].

Figure 2 shows the railway network in Japan. Green lines delineate the borders of prefectures. Blue lines are the railway networks operated by the JR group. Grey lines are the rails owned by non-JR private companies. Red lines (only a very few are shown on the map) are operated by local governments. As can be seen on the map, the railways cover every prefecture. Islands such as Hokkaido, Kyushu, and Shikoku, are linked with Honshu via railways. For instance, Hokkaido and Honshu are connected through the Seikan Tunnel, a 53.85km (33.46 mi) railway tunnel under the sea bed.

Transport costs between different prefectures will be measured by the traveling distance across the railway network between two economic centres. The economic centre of each prefecture is assumed to be the major railway station of each prefecture. Usually a city develops around its central railway station and radiates towards suburban areas. This phenomenon can be seen in many cases of urban development in Japan. The distance data for the railway network are available by using route planner in the search engine of Yahoo Japan. If there exists more than one route which links two stations, the criterion of route selection is based on minimum distance.

---

[4] This issue was discussed with Prof Hiroki Tanaka of Doshisha University, Japan, who was once a researcher for the Economic and Social Research Institute of the Cabinet Office. Prof. Tanaka (personal communication) pointed out that total taxable income is unadjusted nominal data, and is not influenced by the change to the System of National Accounts, which makes it more suitable for 30 year panel research that spans the year 1996.

[5] “Annual Report of Rail Transport Statistics, 2008” by Ministry of Land, Infrastructure and Transport, Japan.

[6] “Rail transport in Germany” Wikipedia article 2008

Figure 2 Railway network in Japan (1999)
(Source: the author using the database of JMC99[7] of Mapinfo)

Admittedly, economic flows from one prefecture to another may happen through all sorts of channels, but railway transportation certainly plays an essential role in accounting for the flows between the major economic centres of different prefectures.

In addition to inter-prefecture transport costs, consideration needs to be given as to how within-prefecture transport costs are to be incorporated into determining market potential. According to the original NEG texts (Fujita et al,1999), within-prefecture transport costs are assumed to be unity (Approach1). This implies that there are no internal transport costs involved in the calculation of market potential within each prefecture. With this assumption, the total income of each prefecture, which reflects market size, would play a more important role in the calculation of its market potential, since it is not deflated by any measure of within-prefecture transport costs (eg. if $\sigma = 2$, then $1-\sigma = -1$). Another approach was first suggested by Head and Mayer (2003). They propose an assumption that consumers are supposedly located evenly on a disk with area K equal to the region in question, while manufacturing industries concentrate in the centre of the disk. In this setting, an average distance between the centre and consumers is calculated as $\frac{2}{3}\left(\frac{K}{\pi}\right)^{0.5}$ (Approach 2).

In order to have a better understanding of the implications of these two approaches to measuring internal transport costs we simulate market potential under both. Market potential is calculated using the term in square brackets in (1). Calculations are simplified by not considering the effects of the price index of each prefecture. The price index is set equal to unity everywhere in Japan. This makes the definition of market potential closer to that suggested by Harris (1954) and enables us to see more clearly how different assumptions about internal transport costs influence the spatial pattern of market potential. The elasticity of substitution (σ) is assumed to be 2. The total income data used for the simulation are total residents' income in 2000.

From Figure 3, it can be seen that the income of Tokyo-to is the highest in Japan, and is even double the income of its rich neighbours such as Saitama-ken, Chiba-ken and Kanagawa-ken. It is also evident that several prefectures show comparatively higher total income. These prefectures include Tokyo-to, Kanagawa-ken, Aichi-ken, and Osaka-fu. In Kyushu Island, Fukuoka-ken has the

[7] JMC99 by Mizsui Zosen Systems Research Inc. Japan. www.msr.co.jp/mapinfo/

highest income compared with the other prefectures on the island. The total income of Hokkaido in the north end of Japan is comparatively higher than its neighbouring prefectures.

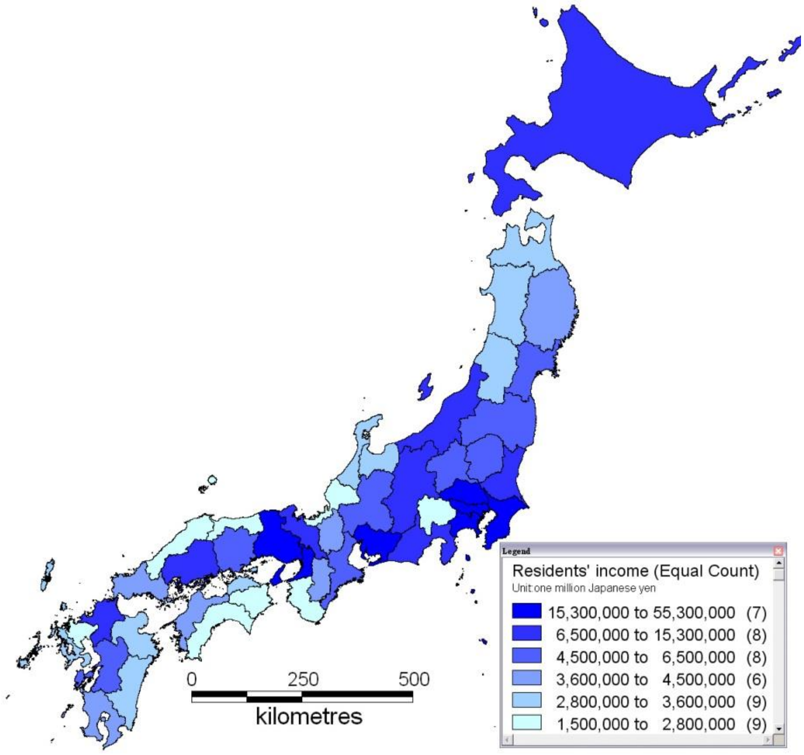


Figure 3 Total residents' income in 2000

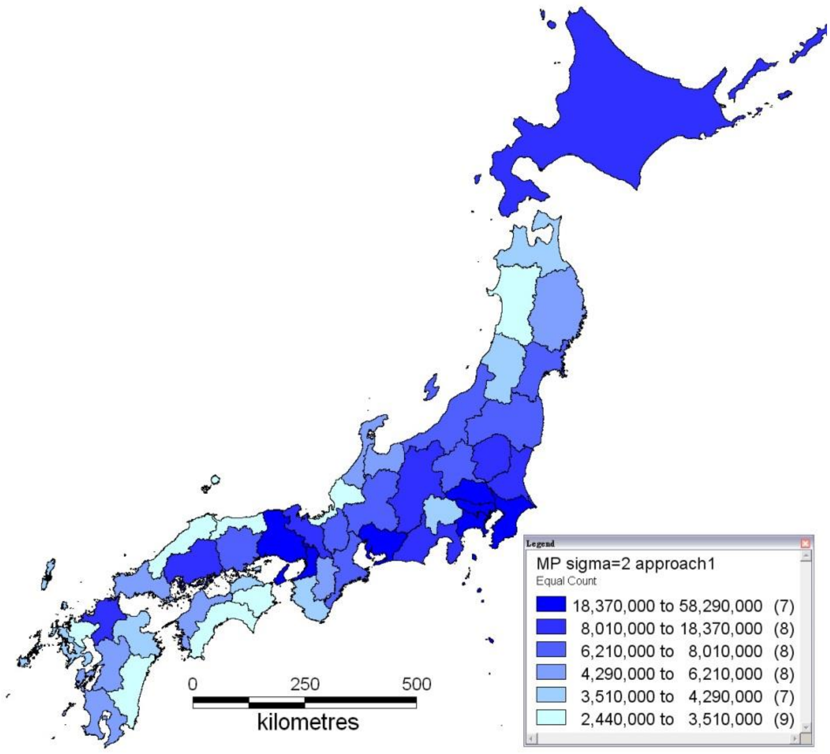


Figure 4 Thematic map of market potential in 2000 (Approach 1 $\sigma = 2$ $T_{ii} = 1$)

Figure 4 is a thematic map of market potential with internal transport costs set to unity. Tokyo-to has the highest value (58290000). This value is almost double the value of the second highest, Kanagawa-ken (31819000). Osaka-fu of Kansai area has the third highest market potential (29782000). It is worth noting that prefectures around Tokyo-to generally have higher values of market potential. These prefectures are Saitama-ken, Chiba-Ken, and Kanagawa-ken. The only exception is Yamanashi-ken. This is partly because the economic development of Yamanashi-ken is strongly limited by its mountainous topography. For example, Mt. Fuji, the highest mountain of Japan (elevation 3776 m), is located in this prefecture. The total income of Yamanashi-ken (2530883 million yen) is less than one tenth that of Kanagawa-ken (28530359 million yen). Also, Kofu station (the major railway station in Yamanashi-ken) is further from Tokyo station compared to other major stations surrounding Tokyo-to. This implies the purchasing power in Tokyo-to contributes less to the market potential of Yamanashi-ken since it is deflated by a higher transport cost.

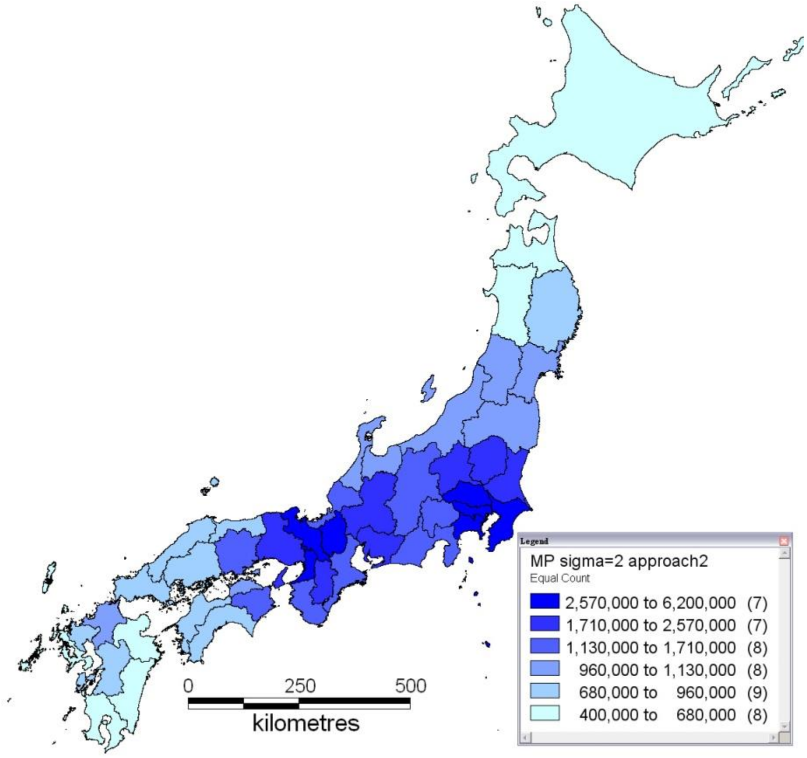
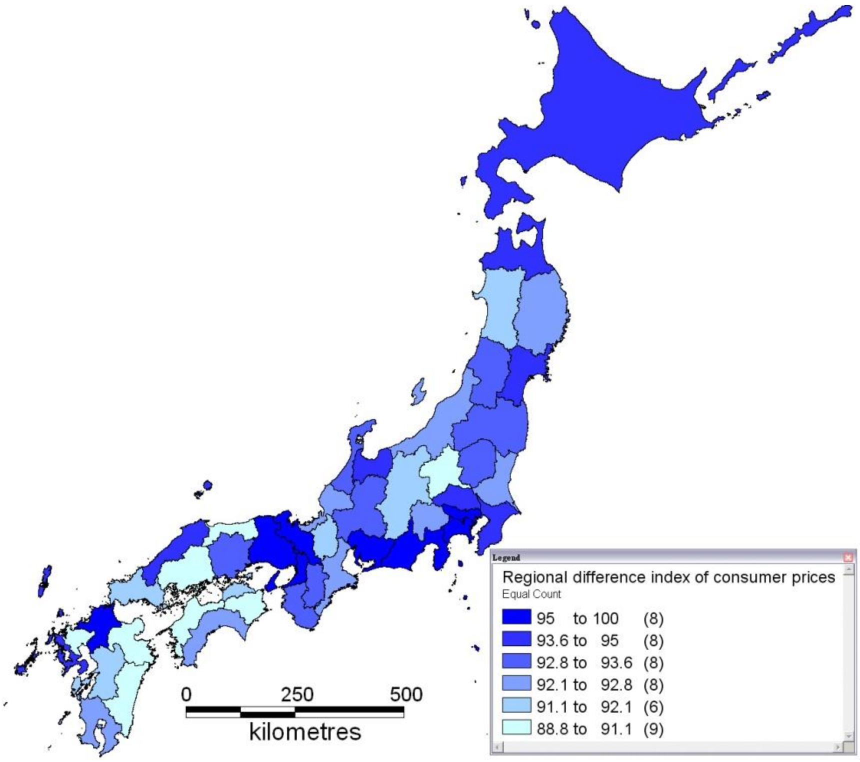


Figure 5 Thematic map of market potential (Approach 2 $\sigma = 2$ $T_{ii} = \frac{2}{3}\left(\frac{K}{\pi}\right)^{0.5}$)

Figure 5 is based on approach 2. Compared with Figure 4 which was created by applying the same parameters, but treating the internal transport costs as unity, certain parts of Figure 5 are not consistent with our general knowledge of the economic geography of Japan. For example, the market potential of Hokkaido is extremely low. Given the fact that Hokkaido has quite a significant amount of total residents' income, it should have a certain level of market potential in its own right. Since the area of Hokkaido is much larger than most of the other prefectures of Japan, its internal transport costs (105.33km) are much higher. The total residents' income of Hokkaido is deflated by a significantly larger denominator, which results in a smaller value for market potential. In addition, we find that the market potential of Shiga (2604245) is slightly higher than Kyoto (2574861). This again is contrary to our intuition, since the total residents' income of Kyoto (7766092 million yen) is much higher than that of Shiga (4481507 million yen). This is due to the fact that the internal transport costs of Kyoto (25.55km) and Shiga (23.84km) are higher than the external transport cost between these two prefectures (10km). The geographical reality is that the economic centre, Otsu city, of Shiga prefecture is located very close to the border between Shiga and Kyoto. That is why the railway transportation distance between Shiga and Kyoto is short. From this example, we conclude that if the internal transport costs are greater than the between-prefecture ones, a measurement bias will certainly appear in calculating market potential. This measurement issue of market potential is pointed out by Head and Mayer (2003) as well.

The correlation coefficient between the two sets of market potential values is 0.87 which suggests that the difference between these two approaches to measuring internal transport costs is not great. However, from Figure 5, we concluded that in certain cases, the second approach introduces more prefecture-specific measurement error than the first approach Hence, in the following empirical work, especially estimating the wage equation, we adopt the first approach to measuring market potential[8]. In this scenario, the internal transport cost of a prefecture is treated as unity.

In terms of the price index (G_s) of the wage equation, we use the regional difference index of consumer prices[9] (RDI) as a measure for each prefecture. It is worth mentioning that the price index data at the prefectural/regional level was not available in most previous research. The use of local prices releases us from dependence on Hanson's (2005) auxiliary equilibrium conditions, conditions that rest on somewhat unrealistic assumptions. However, we are aware of that the use of RDI at prefecture level necessarily introduces the prices of non-tradeables, which produces some disjuncture with the assumptions of the NEG model (Fujita et al., 1999). Nevertheless, the prices of many manufacturing items are reflected in RDI, so we think RDI is the second best proxy to the price index (G_s) . the Figure 6 is drawn on the basis of the regional difference index of consumer prices in 2000. The index of Tokyo-to is regarded as the base here, and indicates that the cost of living in Tokyo is the highest in Japan. The indices of other prefectures are less than 100. Generally, consumer prices in the southern part of Japan, including Shikoku island and Kyushu island, are lower than elsewhere in the country.



[8] where there is no clear empirical evidence to support one approach to intra-area distance calculation over another then both might be implemented to assess the sensitivity of results.

[9] A short description of the index is abstracted in Appendix.

Figure 6 Regional difference index of consumer prices in 2000

4.3 Data for labour efficiency

As noted above, it is possible to augment the original NEG model by incorporating a labour efficiency variable. In this research, two variables which reflect labour efficiency across the spatial units of Japan will be considered in the empirical models separately. The first labour efficiency variable is constructed based on the concept of the percentage of the prefectural population with a final degree of Bachelor or above (including MA and PhD). The variable is calculated by dividing the number of people in the prefectural population who got their final degree as Bachelor or above (including MA and PhD) by the total size of the prefectural population. The data are published in the report of the population census of Japan (Kokusei Chosa Hokoku). This variable is named EDU. The second labour efficiency variable is the local quotient of professional and skilled workers. This variable is named LQ. We use the number of professional and skilled workers[10] and employee data as inputs to calculate the local quotient of each prefecture. Our assumption is that a region with a higher local quotient of professional and skilled workers reflects higher labour efficiency in that region.

Table 1 Descriptive statistics of the panel data

	Wage1977-2006	LQ808590950005	EDU809000	Total taxable income1977-2006
std	1042.5872	0.1116	0.0316	3999346600
mean	3668.4529	0.955	0.0709	3148590370
max	6783.9268	1.318	0.1929	26495845682
min	1448.0772	0.741	0.0271	252353485
Time periods	30	6	3	30
Sample size	1410	282	141	1410

5 The results of the empirical model fitting

5.1 Empirical results for the wage equation of the NEG

Based on the KKP spatial panel framework, we estimate the wage equation of the NEG. In the first estimation of the wage equation (see equation 10), we assume that the elasticity of substitution (σ) is equal to two. According to the theoretical assumption that σ is at least greater than one, assuming the value of σ as two is a reasonable starting point. This value is used in the construction of the market potential variable. Market potential is composed of three elements: total income in each prefecture, a price index in each prefecture, and transport costs between prefectures. At this stage, the price index of each prefecture is not considered in the construction of the market potential. This simplification means we treat the price index as unity throughout Japan. The formula for market potential is now similar to Harris's (1954) definition of market access.

$$MP = \left[\sum_{s=1}^{R} Y_s \left(T_{rs}^{M} \right)^{1-\sigma} \right] \tag{12}$$

For example, if σ is equal to 2, equation (12) is the same as Harris' market access.

[10] "F2601 専門的技術的職業従事者數" According to the definition of professional and skilled workers, which was given by the Statistics Bureau of Japan, it includes natural science researchers, social science researchers, medical doctors, chartered accountants, professional electronic engineers, etc.

The regressor matrix (X_N) contains two variables. The first variable is a constant term and the second is the log value of market potential (ln MP) for the 47 Japanese prefectures from 1977 to 2006. Y_N is a vector which contains the log value of wages in the manufacturing sector of the 47 prefectures from 1977 to 2006. The spatial weights matrix (W_N) is a standardized contiguity matrix. The endogeneity issue surrounding wage and market potential in the wage equation has been discussed in section 3.2. In practice, we need to employ the instrumental variable (IV) technique to circumvent the inherent endogeneity problem. It has been widely suggested in the econometrics literature (see for example Davidson and Mackinnon, 1993, Greene, 2011) that under the assumption of weak exogeneity, the lagged value of the explanatory variable is a feasible candidate as an instrument. In this case, we use the log value of market potential from 1976 to 2005, with σ assumed to be 2 and a constant vector of ones as instruments.

The results for this analysis are shown in the second column of Table 2. The estimate of the regression parameter for ln MP equals the inverse value of the elasticity of substitution (σ). Hence, in the first estimation, we can calculate $\hat{\sigma}$ by taking the inverse value of the estimate. $\hat{\sigma}$ is equal to 4.16 (95% confidence interval: $\Pr(4.14 \leq \sigma \leq 4.18) = 0.95$) and is significant with t=10.95. R^{2*} is a measure of goodness of fit, which is calculated on the basis of the squared correlation between observed and fitted values. The high R^{2*} (0.7548) suggests that the spatial panel model fits the data well. However, the approximate 95% confidence interval for σ is 4.14 to 4.18, which excludes the assumed value of $\sigma = 2$ which was used to construct market potential. This indicates that it is probably more appropriate to assume elasticity is close to 4 for the purpose of re-constructing market potential for the following estimation. After a few iterations, the estimated elasticity of substitution ($\hat{\sigma}$) converges to 4.88 and is significant with t=9.67. In this case, the estimated elasticity is equal to the assumed one. The calculated confidence interval is $\Pr(4.86 \leq \sigma \leq 4.91) = 0.95$. The results are tabulated in the third column of Table 2.

Table 2 FGS2SLS[11] estimates for the NEG model

	Assumed elasticity = 2	Assumed elasticity = 4.88	Assumed elasticity = 5	Assumed Elasticity=5.55
	Estimate(t ratio)	Estimate(t ratio)	Estimate (t ratio)	Estimate (t ratio)
ln MP	0.2402 (10.95)	0.2048 (9.67)	0.1828 (9.34)	0.1799 (9.29)
Constant	2.9842 (6.3)	3.7898 (8.36)	0.9548 (1.24)	0.6251 (0.77)
ρ	0.3655	0.3941	0.4037	0.4056
σ_v^2	0.0124	0.0130	0.0146	0.0148
σ_1^2	0.4341	0.4742	0.4482	0.4443

[11] FGS2SLS (feasible generalized spatial two stage least squares) was discussed in section 3.2. Hereafter, we use FGS2SLS to represent the feasible generalized spatial two stage least squares estimator.

R^{2*}	0.7548	0.7470	0.7458	0.8638
Time periods	30	30	30	30
Sample size	1410	1410	1410	1410

In the following estimations, we adopt the original formula for market potential from Fujita et al, (1999, p.55). This formula was described in section 2. The only difference between this one and the one used in previous estimations lies in including the price index of each prefecture as one of the elements in the construction of market potential. Allowing for variation of prices reflects competition effects, which will be stronger in larger regions (where more varieties of manufactured goods are produced). Again, for the independent and dependent variables, it is still a 30 year panel data set with data ranging from 1977 to 2006 across the 47 Japanese prefectures. For the instrumental variables, we use the log value of market potential from 1976 to 2005 across the 47 prefectures and a constant vector of ones. In the first estimation, the market potential is constructed by assuming the elasticity of substitution to be 5. Following the logic of the previous estimations, we found the estimated elasticity of substitution ($\hat{\sigma}$) converges to 5.55 with the estimated 95% confidence interval, $\Pr(5.53 \le \sigma \le 5.59) = 0.95$ which includes the assumed elasticity value of 5.55 and is significant with t=9.29 (See the fourth and fifth columns of Table2). A high goodness of fit measure (0.86) shows that the wage equation through the KKP panel framework fits well to the Japanese spatial economic data.

We now add a labour efficiency variable to the model as a control variable. Since the EDU data are only available every ten years, we have to reduce our panel to three time points (1980, 1990, and 2000). The sample size is now reduced to 141. The dependent variable is the log of the wage in the manufacturing sector in 1980, 1990, and 2000. There are three independent variables including the log of market potential in 1980, 1990, and 2000, the log EDU in 1980, 1990, and 2000, plus a constant vector of ones. For the purpose of implementing instrumental variables estimation we use the following variables: the log of market potential in 1979, 1989, and 1999, the log EDU in 1980, 1990, and 2000, plus a constant vector of ones. The results are tabulated in the second column of Table 3. The major finding is that ln EDU is significant. The estimated coefficient for ln EDU can be interpreted as implying that a 1% increase in the proportion of the population awarded a Bachelor's degree or higher will contribute a 0.58% increase to the wage in that prefecture. However, the estimated coefficient for ln MP is not significant. We also found the correlation coefficient between ln MP and ln EDU to be 0.74. The high degree of correlation might be explained by reference to the literature on spatial sorting of skills, which argues that more able workers tend to move to larger cities (Glaeser and Mar´e, 2001; Puga, 2010). From the modelling point of view, the high correlation between these two vectors means that model fitting will be affected by multicollinearity between these two independent variables. The main unwanted consequence of multicollinearity is that the variances of the least squares estimates of the parameters of the collinear variables are large (Kennedy, 2003). We suspect that the lack of significance of the ln MP covariate is due to the large variance of the estimate created by this collinearity issue.

In addition, we also conducted Sargan's test of independence between the instruments and the IV residuals. The null hypothesis is that all instruments are exogenous. Sargan's statistic is calculated in two steps. Firstly, regress IV residuals on all exogenous variables (the instrumental matrix). Secondly, obtain the R^2 value from the first step regression model. The test statistic is $S = nR^2$, which is distributed as χ^2_{m-r}, where m-r is the number of instruments minus the number of endogenous variables. n denotes the size of sample. The calculated Sargan's statistic (4.8362e-022) is too small to reject the null hypothesis, hence, these instruments are appropriate.

Table 3 FGS2SLS estimates for the NEG model with a labour efficiency variable

	Assumed elasticity = 5.55	Assumed elasticity = 5.55
	Estimate (t ratio)	Estimate (t ratio)
ln MP	0.0126 (0.58) 95% C.I. = [0.0089,0.0163]	0.1472 (7.10) 95% C.I. = [0.1449,0.1495]
ln EDU ln LQ	0.5754 (10.83) 95% C.I. = [0.5667,0.5841]	0.2615 (1.36) 95% C.I. = [0.2391,0.2839]
Constant	9.1866 (9.05)	2.0304 (2.33)
ρ	0.5022	0.5092
σ_v^2	0.0029	0.0104
σ_1^2	0.0253	0.0726
R^{2*}	0.9473	0.8030
Sargan's statistic	4.8362e-022	2.7451e-023
Time periods	3	6
Sample size	141	282

Since the correlation between ln EDU and ln MP is so strong, we think there might be other variables which can also be used to reflect labour efficiency across different spatial units. A feasible variable is the local quotient (LQ) for professional and skilled workers. The data for LQ are available every five years. The data used for the estimation are as follows:

Dependent variable: log wage of manufacturing sector in 1980,85,90,95,00,05.

Independent variables: [log MP in 1980,85,90,95,00,05; log LQ in 1980,85,90,95,00,05; constant]

Instrumental variables: [log MP in 1979,84,89,94,99,04; log LQ in 1980,85,90,95,00,05; constant]

The results are shown in the third column of Table 3. Compared with the results shown in the fifth column of Table 1, the estimate of the coefficient of ln MP is slightly smaller. This suggests that the newly introduced control variable, ln LQ, takes some explanatory power from ln MP. The coefficient of ln MP is now significant with a large t ratio (7.1). The estimate for ln LQ is also significant (although not as significant as that of ln MP) with the t ratio equal to 1.36 (P-value[12]=0.087). The calculated Sargan's statistic is very small, which suggests the instrumental variables are appropriate. The goodness of fit ($R^{2*} = 0.803$) indicates that the model fits the data well. So far the results from this estimation are satisfactory with the signs of the estimates consistent with the theoretical expectations as well as being significant. This means variation in the manufacturing wage in Japan is well explained by the market potential of NEG providing the heterogeneity in labour efficiency across different prefectures is controlled.

In the next estimation, we run a fixed effects panel model with spatially autoregressive disturbances. The fixed effects are calculated as the deviations from the mean intercept. By default the intercept is automatically added by the routine. In addition to the automatically added intercept, the other input variables are defined as follows:

Dependent variable: log wage of manufacturing sector in 1980,85,90,95,00,05.
Independent variables:[log MP in 1980,85,90,95,00,05; log LQ in 1980,85,90,95,00,05]

In the previous estimations (KKP random effects model), we know that the elasticity of substitution converges to 5.55. Hence, we take this value as a start point for the fixed effects model. The results are shown in the second column of Table 4.

Table 4 ML estimates for the NEG model with LQ

	Assumed elasticity = 5.55	Assumed elasticity = 3	Assumed elasticity = 1.55
	Estimate (t ratio)	Estimate (t ratio)	Estimate (t ratio)
ln MP	0.3181 (9.18) 95% C.I. = [0.3141,0.3221]	0.5364 (18.99) 95% C.I. = [0.5349,0.5379]	0.6457 (37.21) 95% C.I. = [0.6448,0.6466]
lnLQ	0.3605 (3.05)	0.4425 (3.66)	0.6384 (5.68)
intercept	-5.1377 (-13.77)	-8.1175 (-8.12)	-8.0543 (-39.74)
ρ	0.8909 (47.32)	0.7620 (22.29)	0.5710 (11.01)
R^2	0.7883	0.9350	0.9721
Time periods	6	6	6

[12] The t ratio to P value conversion is calculated using the following website. http://www.danielsoper.com/statcalc3/calc.aspx?id=8

Sample size	282	282	282
Log likelihood	460.78	477.23	511.42

The estimated coefficients for both ln MP and ln LQ are significant and signed in line with theoretical expectations. The estimated elasticity of substitution (1/0.3181=3.14) is smaller than the assumed value. The estimated spatial autoregressive parameter (ρ =0.8909; t=47.32) is significant indicating a significant spatial autoregressive process in the model's disturbances. The R^2 value (0.7883) shows that the model fits the data well. However, according to the calculated t-ratios, the fixed effects (not fully listed here), in general, are not significant at the $\alpha = 0.05$ significance level.

In the next estimation, the elasticity is assumed to be 3. The empirical results are better than the previous ones (assumed elasticity=5.55) with a higher R-squared value (0.9350) and Log-likelihood value (477.23). However, the estimated elasticity (1/0.5364=1.86) is still too distant to the assumed value. After several iterative processes, we found the elasticity of substitution converges to 1.55 (1/0.6457=1.55) with a very large t-ratio (37.21). The goodness of fit is 0.9721 and the Log-likelihood value equals 511.42. The spatial autoregressive parameter ρ and the coefficient of ln LQ are both significant with large t-ratios. The only concern with this fixed effects model is that the majority of the fixed effects are not significant (See Table 1a Fixed effects and t-ratios in Appendix) given a significance level $\alpha = 0.05$. This implies that fixed effects play little role in explaining the wage variation across Japanese prefectures.

When reviewing the estimated results from both the KKP model and the fixed effects panel model with a spatial autoregressive error process, we conclude that the KKP model is the more appropriate choice since it models the endogenous relationship between wages and market potential. In addition, the estimated coefficients for both ln MP and ln LQ are significant and signed in line with theoretical expectations. The goodness of fit is high (0.8030). On the other hand, the fixed effects panel model also fits the data well, and market potential proves to be a significant variable in explaining the wage variation across Japanese prefectures, but the convergent estimate of the elasticity of substitution (1.55) departs from that of the KKP model (5.55). This is not beyond our expectation since maximum likelihood estimation does not account for endogeneity so that estimates are likely to be biased.

Finally, the structures of these two spatial panel models are quite different. The KKP model structure involves three levels or hierarchies. The first level is a linear regression. The disturbance terms of the linear regression contain the spatial autoregressive process. This is the second level of the model. From the innovations of the spatial autoregressive process, the individual random effects and the error terms which vary over time and space are specified. This is the third level of the model. However, the fixed effects model has only two levels. The first level is a linear regression with individual fixed effects. The spatial autoregressive process is specified in the second level from the disturbance terms of the regression model at the first level. Hence, due to these structural differences, we are reluctant to jump into the debate on random versus fixed effects preferring the results of these two models to speak for themselves.

6. Conclusion

This research has re-examined the wage equation using a new approach, namely a spatial panel model, a new panel dataset of Japanese prefectures across thirty years not previously analysed, and a novel measure of transport costs between prefectures which can reflect the topographic reality of Japan as it impacts on the movement of goods and people. Through inclusion of such a measure of transport costs, this research has attempted to respond to the criticism of NEG levelled by "proper economic geographers" (Martin and Sunley, 2010). These economic geographers argue that NEG "neglects real places" (Martin, 1999 p.77). Whilst their criticism of NEG extends well beyond how transport costs are handled, the work here suggests how it may be possible to enrich research within the NEG tradition drawing on the insights of "proper economic geographers" and in so doing achieve a measure of synthesis between these two areas of research.

The estimates of the elasticity of substitution (σ) range from 4.88 to 5.55 which are quite close to the estimated results (ranging from 4.9 to 7.6) of Hanson (2005) based on the cross sectional county-level data (3075 counties) of the United States of America. Both estimates are statistically significant which indicate that the NEG's concept of market potential provides a plausible explanation of wage variation at the prefecture or county level. Whether the success of the NEG's wage equation can be used to explain local level variation in the case of data recorded for smaller spatial units and at what level the true spatial process of wage variation can be properly depicted, calls for further empirical analysis.

References

Anselin, L. (1988). Spatial Econometrics: Methods and Models. The Netherlands: Kluwer Academic Publishers.

Arbia, G. (2014). A Primer for Spatial Econometrics with applications in R. London: Palgrave Macmillan.

Baltagi, B.H. (2008). Econometric Analysis of Panel Data, Fourth edition. John Wiley

Baltagi, B.H., Fingleton, B., and Pirotte, A. (2014) 'Estimating and Forecasting with a Dynamic Spatial Panel Data Model' Oxford Bulletin of Economics and Statistics 76(1): 112-138

Brakman, S.H., Garretsen, H., Schramm, M. (2006). 'Putting new economic geography to the test.' Regional Science and Urban Economics 36: 613–635.

Chamberlin, E. H. (1933). The Theory of Monopolistic Competition, First Edition. Harvard University Press, Cambridge, MA.

Cliff, A., and Ord, J. (1973). Spatial autocorrelation, Pion Limited, London, UK.

Cliff, A., and Ord, J. (1981). Spatial Processes, Models and Applications. Pion, London.

Davidson, R. and MacKinnon, J. G. (1993). Estimation and inference in econometrics. Oxford: Oxford University Press

Dixit, A. K., and J. E. Stiglitz. (1977). 'Monopolistic competition and optimum product diversity.' American Economic Review 67 (3): 297-308.

Elhorst, J. P. (2010). Spatial Panel Data Models. In M. M. Fischer and A. Getis (eds) Handbook of Applied Spatial Analysis. Berlin: Springer-Verlag, 377-405. Matlab routines are downloadable from the following webpage. http://www.rug.nl/staff/j.p.elhorst/

Elhorst, J. P. (2014). Spatial Econometrics From Cross-Sectional Data to Spatial Panels. Springer

Fingleton, B. (2006). 'The new economic geography versus urban economics: an evaluation using local wage rates in Great Britain.' Oxford Economic papers, 58: 501-530

Fingleton, B. (2008). 'Competing models of global dynamics: Evidence from panel models with spatially correlated error components.' Economic Modelling 25: 542–558.

Fingleton, B. (2011). ‘The empirical performance of the NEG with reference to small areas.’ Journal of Economic Geography 11: 267-279

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

Glaeser, E. L. and Mar´e D. C. (2001). 'Cities and Skills' Journal of Labor Economics, 19(2): 316–342.

Greene, W. H. (2011). Econometric Analysis, Seventh Edition, Pearson

Haining, R. (2003). Spatial Data Analysis: Theory and Practice. Cambridge: Cambridge University Press.

Hanson, G. (2005). ‘Market potential, increasing returns and geographic concentration.’ Journal of International Economics 67: 1–24.

Harris, C. (1954). ‘The market as a factor in the localization of industry in the United States.’ Annals of the Association of American Geographers 44, 315–348.

Head, K., and Mayer, T. (2003). ‘The empirics of agglomeration and trade.’ Centre for Economic Policy Research (CEPR) Discussion Paper No. 3985.

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

Hirschman, A. (1958). The strategy of economic development. Yale University Press, New Haven, CT.

Hsiao, C. (2003). Analysis of Panel Data, Second Edition Cambridge: Cambridge University Press

Kapoor, M., Kelejian, H. H., Prucha, I. R. (2007). ‘Panel Data Models with Spatially Correlated Error Components.’ Journal of Econometrics 140, 97–130.

Kiso, T. (2005). ‘Does New Economic Geography Explain the Spatial Distribution of Wages in Japan?’ Mimeo, University of Tokyo.

Krugman, P. R. (1991a). ‘Increasing returns and economic geography.’ Journal of Political Economy 99: 483-499.

Krugman, P. R. (1991b). Geography and Trade. Cambridge: MIT Press.

Marshall, A. (1920). Principles of Economics. London: Macmillan (8th ed.).

Martin, R.L. (1999). ‘The New ‘Geographical Turn’ in Economics: Some Critical Reflections.’ Cambridge Journal of Economics 23: 65-91.

Martin, R.L. (2008), The ‘New Economic Geography’: Credible Models of the Economic Landscape? Chapter in A Compendium of Economic Geography Edited by: Roger Lee, Andrew Leyshon, Linda McDowell and Peter Sunley, Sage Publications, London and New York.

Martin, R.L. and Sunley, P. J. (2011). ’The New Economic Geography and Policy Relevance.’ Journal of Economic Geography, 11: 357-370.

Mion, G. (2004). ‘Spatial externalities and empirical analysis: The case of Italy.’ Journal of Urban Economics 56: 97–118.

Puga, Diego. (2010). 'The Magnitude and Causes of Agglomeration Economies' Journal of Regional Science 50(1): 203-219

Redding, S., Venables, A.J. (2004). 'Economic geography and international inequality.' Journal of International Economics 62: 53–82.

Roos, M. (2001) 'Wages and market potential in Germany.' Jahrbuch für Regionalwissenschaft (Review of Regional Research) 21: 171–95.

Appendix

"The survey aims at obtaining a clear picture of the wage structure of employees in major industries i.e., wage distribution by type of employment, type of work, occupation, sex, age, school career, length of service and occupational career, etc."[13] Since wage data are collected following a detailed classification, the nominal wage rate for manufacturing industry per capita per year is calculated as follows:

The nominal wage rate of manufacturing industry per capita per year[14]={[regular monthly cash income per male worker of manufacturing industry*12+Bonus(male worker of manufacturing industry)]*the number of male workers in manufacturing industry+ [regular monthly cash income per female worker of manufacturing industry*12+Bonus(female worker of manufacturing industry)]*the number of female workers in manufacturing industry}/[the number of male workers in manufacturing industry+ the number of female workers in manufacturing industry].

The price index (G_s)

The Regional Difference Index of Consumer Prices (RDI) is an index that indicates the regional differences of the price level based on the average prices of Japan of goods and services purchased by households nationwide. The RDI is calculated from the result of the Retail Price Survey (RPS) (the Trend Survey and the Structural Survey). The items to perform the calculation of the RDI (hereinafter "RDI items") are the items used in the calculation of the CPI, except for the "imputed rent" and the "items surveyed only in Okinawa-ken". (The Calculation method of the Regional Difference Index of Consumer Prices is downloadable from the website of the Statistics Bureau, Ministry of Internal Affairs and Communications, Japan)[15]

---

[13] http://www.mhlw.go.jp/english/database/db-slms/dl/slms-04.pdf

[14]The calculation method was further vindicated by Prof Eiichi Yamaguchi of Doshisha University
The original Japanese: {[きまって支給する現金給与額(男性労働者)*12+年間賞与その他特別給与額(男性労働者)]* 男性労働者数+[きまって支給する現金給与額(女性労働者)*12+年間賞与その他特別給与額(女性労働者)]* 女性労働者数}/[男性労働者数+女性労働者]

[15]http://www.stat.go.jp/english/data/kouri/kouzou/pdf/estimation_e.pdf

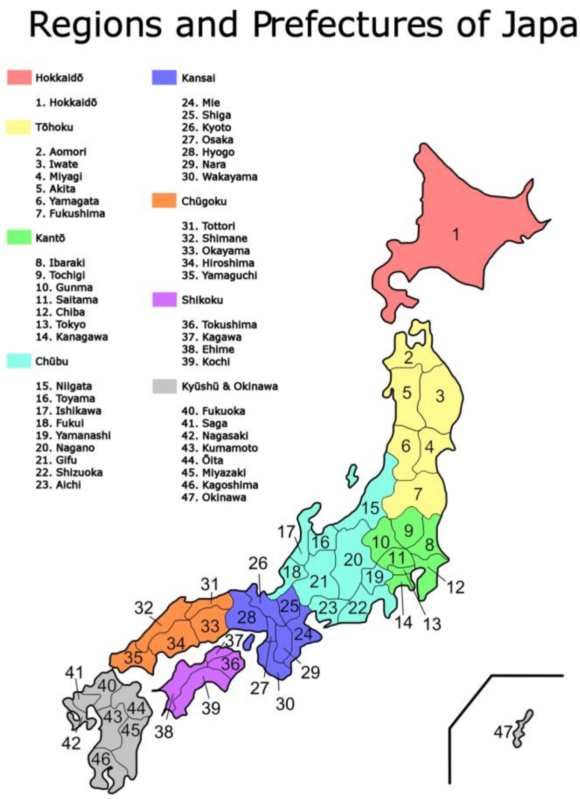


Figure 1a Japanese prefectures (source: Wikipedia)

Table1a   Fixed effects and t-ratios (Assumed elasticity = 1.55)

1.Hokkaido	-0.1668 (-0.82)	17.Ishikawa	0.0508 (0.25)	33.Okayama	0.0670 (0.33)
2.Aomori	0.0288 (0.14)	18.Fukui	0.0658 (0.32)	34.Hiroshima	0.0758 (0.38)
3.Iwate	0.0546 (0.27)	19.Yamanashi	0.1021 (0.50)	35.Yamaguchi	0.3041 (1.51)
4.Miyagi	-0.0347 (-0.17)	20.Nagano	0.0429 (0.21)	36.Tokushima	0.0484 (0.24)
5.Akita	0.0555 (0.27)	21.Gifu	-0.0511 (-0.25)	37.Kagawa	0.0936 (0.47)
6.Yamagata	0.0224 (0.11)	22.Shizuoka	-0.0135 (-0.06)	38.Ehime	0.1447 (0.72)
7.Fukushima	0.0315 (0.15)	23.Aichi	-0.1793 (-0.85)	39.Kochi	-0.0195 (-0.10)
8.Ibaraki	0.0549 (0.27)	24.Mie	0.1143 (0.55)	40.Fukuoka	-0.0885 (-0.44)
9.Tochigi	0.0977 (0.47)	25.Shiga	0.0385 (0.19)	41.Saga	0.1693 (0.85)
10.Gumma	0.0678 (0.33)	26.Kyoto	-0.0874 (-0.43)	42.Nagasaki	0.1947 (0.98)
11.Saitama	-0.3994 (-1.91)	27.Osaka	-0.2883 (-1.38)	43.Kumamoto	0.0947 (0.48)
12.Chiba	-0.3116 (-1.50)	28.Hyogo	-0.2035 (-0.99)	44.Oita	0.2131 (1.07)
13.Tokyo	-0.6741 (-3.23)	29.Nara	-0.0859 (-0.43)	45.Miyazaki	0.1752 (0.88)
14.Kanagawa	-0.5282 (-2.56)	30.Wakayama	0.1811 (0.89)	46.Kagoshima	0.1089 (0.55)
15.Niigata	-0.0429 (-0.21)	31.Tottori	-0.0063 (-0.03)	47.Okinawa	0.3161 (1.63)
16.Toyama	0.0814 (0.40)	32.Shimane	0.0852 (0.42)		