Can NEG explain the spatial distribution of wages of Chile?

Dusan Paredes

REAL 11-T-04 May, 2011

Can NEG explain the spatial distribution of wages of Chile?

Dusan Paredes[a]
[a] *Department of Agricultural and Consumer Economics,*
University of Illinois at Urbana-Champaign, United States of America

April 27, 2011

SUMMARY

The New Economic Geography (NEG) has been tested to explain the spatial concentration of wages in developed countries, but it has not been evaluated for developing countries where the excessive spatial concentration seems to be related with negative consequences on the economic development. This paper covers this gap in the literature estimating by first time a NEG model for a developing country such as Chile, pursuing two research questions : 1) Can the NEG explain the spatial distribution of wages in a developing country as Chile?, and 2) How can the NEG be used to infer information about the future level of spatial concentration of wages in Chile? The results suggest that the case of Chile is poorly explained by the NEG and even higher level of spatial concentration should be expected in the future. These results indicate that the empirical application of NEG is not trivial for developing countries, and some considerations such as inclusion of the first nature or analysis at micro data level must be incorporated by future researches.

Keywords: New Economic Geography, spatial distribution of wages, market-access effect, housing-price effect.

1. INTRODUCTION

Developing countries are characterized by alarming levels of spatial concentration (Puga, 1998, Rosenthal and Strange, 2004). Chile illustrates this case; several descriptive studies show high spatial concentration of wages around the Metropolitan Region (MR), which is reinforced by tremendous regional disparities in several economic variables, such as GDP and population[1] (Aroca, 2009). This situation could be a complex problem considering that the literature identifies negative consequences of excessive spatial concentration, such as low productivity, social dissatisfaction, negative externalities and high levels of regional inflation, among others (Henderson, 2003, Armstrong and Taylor, 2000). In spite of this, there is no literature dedicated to formally evaluating whether

[1] A deeper analysis of the Chilean case is detailed in the next section.

Chile is characterized by excessive spatial concentration of wages. This paper provides a first-time estimation of an economic theory to evaluate whether Chile presents this situation. The theoretical background to evaluate the case of Chile is provided by the NEG, an economic theory designed to explain the spatial concentration of wages. This evaluation is focused on two critical questions: 1) whether the NEG explains the spatial distribution of wages in a developing country like Chile, and 2) how the NEG can be used to infer information about future levels of spatial concentration of wages in Chile. Through these questions, this paper provides an evaluation necessary for the design of future regional policies oriented to reducing the potential negative consequences of excessive spatial concentration.

To make the existence of a spatial distribution of wages evident, we must first recognize the potential sources of this concentration. Combes et al. (2008a) suggested three sources to explain spatial wage disparities: the skill composition of the workforce, non-human endowments (first nature) and local interaction (second nature[2]). While the NEG explains the disparities produced by second nature, this theory does not make reference to the other two sources. Of the sources left out, first nature is a key source to explain the spatial distribution of wages in developing countries. Such countries are mainly exporters of natural resources (a type of non-human endowment), which presents a kind of paradox: on the one hand, developing countries show alarming levels of spatial concentration, the explanation of which is the main objective of the NEG. On the other hand, developing countries presents wage differentials linked to the existence of natural resources, a source left out by the NEG. This paradox raises the question of whether the NEG can replicate the spatial distribution of wages in countries where several sources of wage disparities are mixed. This paper answers this question by estimating the NEG for the case of Chile, a country with one of the highest levels of income inequality in the world, where a fundamental role is played by first nature.

Given the lack of empirical discussion about how the NEG explains the spatial disparities of wages in developing countries, there is less information on how these spatial disparities evolve over time. This paper proposes a simple method to predict future levels

[2]An early conceptualization of second nature was discussed by Marshall (1890), while the modern term "economics of agglomeration" was discussed by (Fujita and Thisse, 2002)

of spatial concentration using the well known "bell-shaped curve of spatial development" (Puga, 1999). Bell-shaped spatial development is characterized by three stages: dispersion, concentration and dispersion again, as Figure (2) describes. In the early stage of development in a particular country, high transport costs encourage dispersed centers with similar spatial concentration. The reduction in transport costs encourages concentration and increases wage disparities, which could represent the current realities of developing countries. At some point, negative externalities outweigh the benefits of concentration, leading to dispersion again. This last phenomenon could represent the situatiion of countries that have already considerably reduced transportion costs, such as developed countries.

<Figure 2 about here>

The literature identifies a statistical regularity where developing countries and developed countries can be located on the left and right side, respectively, of a bell-shaped curve representing spatial development (Combes et al., 2008b). I used this intuition to develop a simple method to determine on what side Chile is located. To carry out this objective, I applied two models, one developed by Helpman (1998) and the other by Sudekum (2006) (hereafter known respectively as Helpman's and Sudekum's models) to evaluate whether the spatial distribution of wages can be explained by the factors described by the NEG, namely the tension of the market access-effect (agglomeration force) and the housing-price effect (dispersion force). However, the two models can be assumed as two sides of bell-shaped spatial development because they represent different concentration processes, as described by Figure (2). Sudekum's model starts with total dispersion, and a high level of spatial agglomeration appears as an efficient equilibrium. Helpman's model starts with a high level of concentration and a dispersion phenomenon then takes place, replicating the case of developed countries such as the United States. I show that Helpman's model is a special case of Sudekum's model, and that the two can be seen as two sides of the bell-shaped curve. As far as I know, this paper is the first attempt to provide this kind of information, which is crucial for developing countries.

Additionally, this paper presents several contributions to the standard empirical literature of the NEG. I contribute to reducing the empirical gap on developing countries by assessing an NEG model to analyze the spatial distribution of wages in Chile for 334 spatial units, called communes, representing the first estimation of the NEG in this country and at this spatial level. While Helpman's model has already been estimated (Brakman et al., 2004, Kiso, 2005, Mion, 2004, Hanson, 2005), this paper is the first to estimate Sudekum's model. This estimation provides two nested models to evaluate whether there will be higher levels of concentration in Chile in the future. This information at the low spatial scale is relevant to the development of regional policy to reduce potential disparities, especially in cases where there is a lack of information. While some data sets are availalbe for developed countries (from a temporal perspective), this is not the case for developing countries. This paper proposes a novel way to carry out this evaluation, but care must be given in considering the conclusions given the lack of an explicit temporal dimension.

2. SPATIAL DISTRIBUTION OF WAGES IN CHILE

Chile has one of the highest levels of income inequality in the world (Ferreira and Litchfield, 1999, Contreras, 2003), and this inequality has a spatial dimension. Aroca and Bosch (2000) used spatial econometrics to reject the hypothesis of convergence in Gross Regional Product (GRP) per capita between 1990-98. This result was supported by Echeverria and Gopinath (2007) who noted that the Metropolitan Region of Santiago (MR), with only 2% of the total territory, accounts for 50% of production and 40% of the population[3]. According to Aroca (2009), by 1992 only two regions had a higher GRP than that of the MR, but the MR had the highest wage levels. By 2003, all the regions showed positive growth of GRP relative to the MR, but none of them had higher wages than those of the MR, indicating an "absorption" of wages by the MR.

This evidence indicates the existence of spatial inequality in per capita GRP and wages. Unlike other articles, this paper focuses on wages instead of GRP. From a statistical per-

[3]Before October, 2007, Chile was divided into 13 regions, including the MR. In October, 2001, two new regions were added, making 15 regions, including the MR. The regions are divided into 54 provinces and provinces are further divided into 346 communes. The communes can be considered equivalent to counties in the United States.

spective, wage data are more readily available than GRP data for more spatial units. While Morande et al. (1997), Aroca and Bosch (2000) and Duncan and Fuentes (2006) estimated their models with GRP data available for 13 Chilean regions (regional level), I increased the size of the sample by using wage data for 334 communes (communal level). Moreover, the NEG assumes that wage differentials are generated by economies of agglomeration, or second nature, that are mostly localized at a low spatial scale (Rosenthal and Strange, 2004). Thus, the communal, rather than the regional level, properly represents these theoretical requirements. Additionally, Aroca (2009) identifies a spatial divergence in wages, with the wage of the MR increasing more rapidly than those in any other region. From this perspective, regional divergence in wages is a potential concern for policy makers, therefore empirical studies are needed to evaluate future policies to reduce spatial inequality.

The arguments described previously are supported by the spatial distribution of wages for 2006[4] (See Figure 1). Wages around the MR are between 95% and 132% of the mean national wage, with two small clusters of high wages located in the extreme north and south. The first is a mining cluster based on cooper mining, while the second in the south is termed the salmon cluster. Aroca and Bosch (2000) provided statistical evidence for the mining cluster, but they described the salmon cluster as being in an initial stage. Some years later, the salmon cluster is established, making the spatial distribution of wages in Chile more uneven. However, the two clusters are not strong enough to reduce the core-periphery structure generated by the wages around the MR (see the zoom map of the MR in Figure 1).

<Figure 1 about here>

3. CAUSAL RELATIONSHIP OF THE NEG FOR TWO REGIONS

The NEG is a highly nonlinear general equilibrium model in which the causal mechanism is not made clear solely by observing the equations. To avoid this problem, I follow

[4]This spatial distribution was constructed using the Chilean Household Survey CASEN 2006. A complete description of the data is in Section 5.

Baldwin et al. (2003) who provided a simple example to illustrate the standard forces of the NEG: the market-access effect, the cost of living effect and the market-crowding effect, but I complement this with a new force: the housing-price effect. Let two symmetric regions be called the northern (N) and southern (S) regions, assuming a small migration of workers from S to N, which increases the relative market size of N[5]. Because of transport costs and increasing returns, firms prefer to locate near large markets following the initial migration of workers from S to N[6]. Migration of workers and firms represents the first agglomeration force, called the market-access effect or demand-linked circular causality, which increases the relative wage in N. The word circular is added because after the "first round" of migration, new jobs and higher wages are generated in N, attracting new workers again[7].

Growth in the number of firms in N increases the variety of industrial goods produced locally. All other things being equal, this increase reduces the industrial price index and consequently reduces the cost of living (COL)[8]. Reduction in the COL, ceteris paribus, increases relative real wages in N, attracting workers to the agglomeration. This process is called the cost-of-living effect or cost-linked circular causality. Again, the word circular is added because the migration of workers increases the market, attracting new firms and further reducing the COL. In summary, both agglomeration forces increase wages in N compared to S. However, the NEG also establishes forces working in the opposite direction.

The migration of firms and workers increases the level of competition in N, thus reducing profits. To compensate for the loss, firms must pay lower wages, reducing the relative wage in N. This dispersion force is called the market-crowding effect or the market-crowding dispersion force, which does not involve a circular process. The second dispersion force is called the housing-price effect, which was formalized by Helpman (1998)

[5]If the initial "random" migration seems to be imposed, then assume two asymmetrical regions. The causal mechanism works in a similar way.

[6]Increasing returns imply that a firm can reduce the average cost by producing large quantities. Thus, firms prefer to concentrate production in just one plant. In this scenario, firms are indifferent to the specific location (N or S). However, transportion costs imply that firms must pay to export industrial goods. Thus, firms prefer to be located in large markets to reduce transport costs.

[7]This circular process refers to the cumulative causation addressed by Myrdal (1957)

[8]The market is assumed to be in imperfect competition (Dixit and Stiglitz, 1977) and the competition is different from that of markets with perfect competition. Firms compete with varieties, therefore an increase in the number of firms implies a reduction in the industrial price index.

and Sudekum (2006). After the initial migration, N has a higher demand for housing, pushing up housing prices and reducing the incentives to stay agglomerated in N because real wages are lower than in S[9]. Whether one force is stronger than the other depends on transport costs.

<Figure 2 about here>

A summary is presented in figure 2, where the magnitude of the forces are a function of the freeness of trade ϕ. Transport costs are inverse to ϕ such that high transport costs imply a low level of ϕ (equal to 0) and low transport costs imply a high level of ϕ (equal to 1). An increase in ϕ reduces the magnitude of both forces, but dispersion forces decrease more rapidly than agglomeration forces. To analyze this difference, a mathematical analysis is needeed, but some intuitions can be derived from the previous example where agglomeration forces are characterized by a reinforcing process, while the dispersion forces are not. This implies that dispersion forces fall more rapidly than agglomeration forces with an increase in ϕ.

In figure 2 the threshold ϕ_B is termed the "break-even point". To the left of ϕ_B (lower levels of ϕ), dispersion forces are stronger than agglomeration forces, shaping similar wages among spatial units. To the right of ϕ_B (higher levels of freeness of trade), agglomeration forces are stronger than dispersion forces, generating a spatial distribution of wages dominated by a core-periphery structure. This paper proposes that ϕ is high among communes within Chile, using three arguments to support this claim. First, the literature supports the existence of a core-periphery structure in Chile (Echeverria and Gopinath, 2007, Aroca, 2009). Second, imports and exports among communes do not have costs, increasing the level of ϕ. This implies that ϕ within the country is pushed toward 1. Finally, migration of firms and workers is free of transaction costs, decreasing frictions for trade. I argue that the spatial distribution of wages is observed because agglomeration forces are stronger than dispersion forces. The empirical literature has focused on agglomeration and disper-

[9]A third dispersion force discussed by Fujita et al. (1999) is called the immobile-sector dispersion force. However, Baldwin et al. (2003) do not include it as a dispersion force. There are workers in the southern region who cannot migrate to the north. In the setup of the model, these are agricultural workers. However, in real life they can be considered as workers who cannot move because of economic, cultural or other reasons. They demand industrial goods, therefore they present an incentive for firms to stay in the southern region.

sion forces that can be measured, namely the market-access and housing-market effects. Details about the forces considered explicitly in the estimation process are discussed in the next section.

<Figure 3 about here>

This image represents the well known bell-shaped curve of spatial development that describes the situation for developing and developed countries Puga (1999). Starting from the left, low levels of ϕ (high transport costs) imply that dispersion is not an efficient equilibrium, spatial concentration being the optimal result. High transport costs are present in developing countries, where high levels of concentration are an accepted stylized fact Puga (2002). High levels of ϕ (low transport costs) imply that concentration is no longer an efficient equilibrium, dispersion being the optimal result. This result is common for developed countries with multi-centered urban systems. I propose a simple method to determine where Chile is located in this picture. While Sudekum's model represents the first part of this picture, Helpman's model implies the reality for developed countries; therefore the simultaneous estimation of both models can give us information to evaluate the particular position of Chile on the bell-shaped curve of spatial development. To carry out this goal, I estimate Sudekum's model to represent the reality for developing countries. This simple test can give a new perspective to classify developing and developed countries from a spatial perspective.

4. MODEL

The literature has considered the roles of the market-access and housing-price effects through estimation of Helpman's model (Brakman et al., 2004, Mion, 2004, Kiso, 2005, Hanson, 2005). However, the literature has not discussed the role of the housing-price effect under alternative extensions. I address this gap by estimating Sudekum's model, as well as Helpman's model. Using both models, this paper establishes if the market-access and housing-price effects are relevant forces under different functional forms.

Similarly, Helpman's and Sudekum's models consider the role of the market-access effect, but they differ in the way the housing-price effect is incorporated[10]. Each time the models differ in a variable, a superscript h (Helpman) or s (Sudekum) is added. The consumers located in the spatial unit j have identical preferences defined by a well behaved Cobb-Douglas utility function:

$$U_j^h = M_j^{\mu} H_j^{1-\mu}, \qquad U_j^s = M_j^{\mu} H_j^{\gamma} A_j^{1-\mu-\gamma}, \tag{1}$$

where M_j represents the consumption of manufactured goods, H_j are units of nontradable domestically-produced goods and A_j is the standard agricultural good. The parameter μ represents the share spent on manufactured goods and γ represents an equivalent measure for domestically-produced goods. The agricultural sector is the numerarie good ($p^A = 1$), which is characterized by a competitive market with immobile workers, while the manufacturing sector follows a monopolistic competition model[11] (Dixit and Stiglitz, 1977). M_j is a composite good represented by a symmetrical CES function:

$$M_j = \left(\sum_{i=1}^{R} \int_0^{n_i} m_{ij}\,(z)^{\rho}\, dz \right)^{\frac{1}{\rho}}, \tag{2}$$

where ρ is the parameter of substitution among z varieties of the industrial good and m_{ij} is the manufactured good consumed in (j), but produced in (i)[12]. The imports in (j) can be obtained from R spatial units, where each region has a number of varieties equal to n_i. The Marshallian demand of the consumer located in (j) for a good produced in (i) is specified by:

$$m_{ij} = p_{ij}^{-\sigma} \mu Y_j G_j^{\sigma-1}, \qquad G_j = \left(\sum_{i=1}^{R} n_i p_{ij}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{3}$$

where p_{ij} is the price of the manufactured good consumed in location (j), but produced

[10] General derivations are specified in this section and more details can be obtained in the Appendix.

[11] In the original model, Krugman (1991) used the terms agricultural and industrial workers to refer to mobile and immobile workers. He established that a portion of workers in the southern region are not able to move because of linkage to assets generated by agricultural activity. However, in real life this group can be assumed as workers who cannot migrate to agglomerations. From a strict perspective, it is dispersion force, but not all textbooks include it.

[12] The core of the monopolistic model is represented by the effect of variety in the utility function. In contrast to competitive markets, the consumer prefers diversity. If $\rho < 0$, the varieties are complementary. If $\rho = 1$, variety as such does not matter. Finally, $\rho < 1$ ensures that the varieties are imperfect substitutes.

in location (i), Y_j represents the consumer purchasing power (CPP) of the market located in (j), G_j is the composite price index for the industrial goods of location (j) and σ is the constant elasticity of the CES function, which must be $\sigma = \frac{1}{(1-\rho)} > 1$.

The price differential for industrial goods between the origin (i) and destination (j) is attributed to transport costs, which are modeled using the iceberg transport costs proposed by Samuelson (1952). This type of transport cost implies that only a fraction of manufactured goods reach their final destination. If location (j) sends one unit of the industrial good, just $1/T_{ij}$ are received in location (j). If T_{ij} is considered the transport costs, price p_{ij} can be defined as:

$$p_{ij} = T_{ij} p_i, \quad T_{ij} = 1 \quad \forall \quad i = j \tag{4}$$

Using the equilibrium conditions found in the appendix[13], the wage equation can be specified by:

$$w_i = \left(\sum_{j=1}^{R} Y_j T_{ij}^{1-\sigma} G_j^{\sigma-1} \right)^{\frac{1}{\sigma}} \tag{5}$$

Equation (5) shows the role played by the forces described in the Section 3. To identify the market-access effect, the equation (5) can be redefined as:

$$w_i = \left(\frac{Y_i}{G_i^{1-\sigma}} + \sum_{j \neq i}^{R} \frac{Y_j T_{ij}^{1-\sigma}}{G_j^{1-\sigma}} \right)^{\frac{1}{\sigma}} \tag{6}$$

The market-access effect is composed of two parts: direct and indirect effects. w_i is a direct and positive function of the market-access effect, represented by real CPP (Y_i). Simultaneously, w_i is an indirect and positive function of market-access on surrounding locations determined by transport costs[14]. However, there are two problems with equation (6): the housing-price effect is missed in the specification and the estimation is not possible because the real data of G_j are not available. To overcome these problems, Hanson (2005)

[13]Appendix can be downloaded from `https://sites.google.com/a/ucn.cl/dusanparedes/`

[14]This result represents the geography law in NEG: *"Everything is related to everything else, but near things are more related than distant things."*, which represents the geographical dimension incorporated by the NEG (Tobler, 1970).

argues that, while nominal wages can be spatially different, real wages equalize labor markets among spatial units through inter-regional migration. If this condition is fulfilled, real wages are defined as:

$$\varpi = \frac{w_i}{\psi_i} = \frac{w_j}{\psi_j}, \tag{7}$$

where ϖ is a constant real wage and ψ_i and ψ_j represent the COL for location (i) and (j), respectively. The COL is the aggregated prices of agricultural $\left(P^A = 1\right)$, manufactured good (G) and housing (p^H). Helpman's and Sudekum's models incorporate housing prices through different specifications of the utility function (See equation (1)). Using the definition of the price index for the Cobb-Douglas utility function, the respective COLs for the two models are defined by:

$$\psi_i^h = G_i^\mu \left(p_i^H\right)^{(1-\mu)}, \qquad \psi_i^s = G_i^\mu \left(p_i^H\right)^\gamma \tag{8}$$

Inserting (8) into equation (7), the COL is specified as a function of w and p^H, and a constant ϖ. Inserting this price index G_i into equation (5), applying logarithms and adding a well-behaved error term ϵ_i; the estimable wage equations are specified as follows:

$$\ln\left(w_i^h\right) = k^{'} + \frac{1}{\sigma}\ln\left(\frac{Y_i w_i^{\frac{\sigma-1}{\mu}}}{\left(p_i^H\right)^{\frac{(1-\mu)(1-\sigma)}{\mu}}} + \sum_{j\neq 1}^{R} \frac{Y_j w_j^{\frac{\sigma-1}{\mu}} T_{ij}^{1-\sigma}}{\left(p_j^H\right)^{\frac{(1-\mu)(1-\sigma)}{\mu}}}\right) + \epsilon_i \tag{9}$$

$$\ln\left(w_i^s\right) = k^{'} + \frac{1}{\sigma}\ln\left(\frac{Y_i w_i^{\frac{\sigma-1}{\mu}}}{\left(p_i^H\right)^{\frac{\gamma(1-\sigma)}{\mu}}} + \sum_{j\neq 1}^{R} \frac{Y_j w_j^{\frac{\sigma-1}{\mu}} T_{ij}^{1-\sigma}}{\left(p_j^H\right)^{\frac{\gamma(1-\sigma)}{\mu}}}\right) + \epsilon_i, \tag{10}$$

where $k^{'}$ represents a constant. Now equations (9) and (10) represent the direct and indirect effects of the market-access and housing-price effects. The market-access effect has already been explained, so I will focus on the housing-price effect. w_i is a direct and negative function of housing prices p_i^H. Simultaneously, w_i is a negative and indirect function of the housing-price effect of other locations p_j^H.

The NEG represents a spatial distribution of wages if the estimated parameters imply that agglomeration forces outweigh dispersion forces (Redding, 2010). If this is the case, the market-access effect is stronger than the housing-price effect, implying a spatial distribution of wages. These conditions are specified in the table I. σ represents the nature of imperfect substitution among industrial goods, making the value of this parameter greater than 1. μ and γ are expenditures shares, therefore they must lie between 0 and 1. The parameter τ is crucial to evaluate the role of transport costs (See section 5). If $\tau > 0$, the distance is a correct variable to proxy transport costs. Moreover, $\tau > 0$ warrants that the indirect effect described in the equations (9) and (10) exists.

<Table I about here>

Additionally, two relationships among parameters should be considered. The first, $\sigma / (\sigma - 1) > 1$, reflects the market power given by imperfect competition, which is crucial to support the market-access effect[15]. The second relationship is derived from the the nonlinear equation system and is called the "no-black-hole" condition. This condition ensures the existence of the break-even point ϕ_B in a possible range of transport costs (See figure 2). If the "no-black-hole" condition is not fulfilled, then *"the forces working toward agglomeration (regional divergence) always prevail and the economy tends to collapse to a point"* (Fujita et al., 1999). Following Brakman et al. (2009) the no-black-hole condition is specified by $\sigma (\mu - 1) < 1$ and $\sigma (\mu - 1) > 1$ for Helpman's and Sudekum's models, respectively[16].

5. DATA, PROXIES AND ECONOMETRIC ISSUES

The estimation of (9) and (10) requires information about CPP (Y), transport cost (T), wages (w) and housing prices (p^H). The variables Y, w and p^H were obtained from the Chilean Household Survey CASEN 2006, which was conducted by the Ministry of Planning (MIDEPLAN). CASEN 2006 is a survey with representation at the national, regional, communal, urban and rural levels.

[15] See the Appendix, the condition of marginal benefit equal to marginal cost is $p_i \left(\frac{\sigma-1}{\sigma}\right) = \beta w_i$. Imperfect competition implies that $\sigma / (\sigma - 1) > 1$, otherwise there is no market power and the competitive market structure applies.

[16] For Helpman (1998) the sign is reversed because the interpretation of a reduction in transport cost is inverse: A reduction in transport cost increases dispersion. For details about this change, see Brakman et al. (2009).

I estimated equations (9) and (10) with 334 communes, which represent the lowest level of spatial unit available in the CASEN 2006. This choice implies that Y, w and p must be aggregated from individual observations at the communal level[17]. With respect to Y, Brakman et al. (2004), Brakman et al. (2009) and Hanson (2005) use GRP, but this information is not available at the communal level in Chile. However, CASEN 2006 contains information about disposable income for each observation in a commune. Following the strategy of Mion (2004), I proxy Y as the weighted sum of disposable income for a sample of observations in a commune[18]. I argue that this measure better represents the CPP than GRP because it is the real income that people spend on goods, instead of GRP, which includes money spent outside the commune. Table II shows the descriptive statistics for CPP 2006. The correlation coefficient between wages and CPP is significant and positive (0.576), showing the relationship between the market-access effect and wages. The Moran's test (0.158) shows that CPP is positively correlated across space. This implies that rich communes are surrounded by rich communes and poor communes surrounded by poor communes, supporting the hypothesis that economic growth is spatially clustered.

<Table II about here>

Monetary information of T_{ij} is not available. Following Hanson (2005) and Mion (2004), an exponential decay and power distance function can be defined as proxies:

$$T_{ij} = \begin{cases} \exp(\tau d_{ij}) & \text{Specification 1 (Hanson, 2005)} \\ d_{ij}^{\tau} & \text{Specification 2 (Mion, 2004)} \end{cases}$$

where d_{ij} represents the geodesic distance (in kilometers) between the location (i) and (j), obtained from ArcGIS[19]. The parameter τ represents the gradient of reduction in transport

[17] While Chile has 346 communes, only 335 are statistically represented in CASEN 2006. The ArcGIS map presented a problem with the commune called "La Higuera", therefore this paper considers 334 communes for analysis.

[18] Each observation of CASEN 2006 has a weight to expand the sample to the population. Using these weights, this analysis is expanded to the population level.

[19] Transport cost is better represented by average car travel time between locations, however this information is not available. Geodesic distance can be assumed as the flight-distance between two spatial units.

cost according to distance; therefore $\tau > 0$. Geodesic distance means that $d_{ij} = 0$ when $i = j$, but the assumption of zero transport cost within a spatial unit is not realistic. I follow the recommendation of Head and Mayer (2000) by specifying $d_{ii} = 0.667\sqrt{\frac{area}{\pi}}$. This implies that distance between a spatial unit and itself depends on its size.

The choice of the proxy w is crucial to represent the causal-mechanism described in section 3. Following the insights specified by Helpman (1998) and Sudekum (2006), w represents the level of attractiveness of a location to encourage the migration of workers and firms. The literature recognizes that workers with low education tend not to migrate due to the high cost involved, therefore they do not look at the complete distribution of wages to take the migration decision[20]. To aggregate wages at the communal level, I calculated the average wage of the economically active population (EAP) considering the wages between 25% and 99% of the wage distribution of each commune. The upper boundary was choosen to reduce the impact of outliers with extremely high wages.

Figure 1 presents the core-periphery structure of wages[21]. The mining and salmon clusters are clearly identified in the extremes of the country, but the map does not show the concentration of high wages around the MP so a zoom map is needed. In the zoom map, the MR reveals a core-periphery pattern that is supported by the Moran's test result of 0.189. This value indicates that the spatial relationship of w is positive, with high wages surrounded by high wages. Finally the spatial autocorrelation of wages is greater than any other variables, indicating that wages are strongly clustered, supporting the NEG analysis.

Following Brakman et al. (2004) and Kiso (2005), I represent p^H as the average housing price for each commune. Table II indicates that p^H correlates strongly with higher wages (0.595), suggesting that communes with higher wages also have higher housing prices. Simultaneously, p^H has a high spatial autocorrelation (0.156), suggesting that communes with high p^H are surrounded with communes with high p^H. This supports the evidence that the market-access effect is spatially clustered. In summary, descriptive statistics appear to support the theoretical relationship among wages, income and housing prices.

[20]This consideration is equivalent to the third dispersion force proposed by Krugman (1991), which is not discussed by Baldwin et al. (2003)

[21]Shape file of the map of Chile can be downloaded free of charge from http://www.diva-gis.org/. The map shows the 334 communes with statistical representativity, leaving empty spaces for the communes without statistical representativity in CASEN 2006.

Finally, the estimation of equations (9) and (10) is affected by endogeneity. The first source of endogeneity is measurement-error in housing prices and CPP. The careful choice of proxy variables for p^H and Y should reduce this bias. The second source of endogeneity is related to the variable w, which is present on both sides of the wage-equations, implying that $E[w_j|\epsilon_i]\neq 0$ for $j = 1 \ldots R$ each time that wages are evaluated in the same year. However, the wage of a location j in 2006 (w_j^{06}) can be broken down into a variation from 2006 (Δ^{06}) and a variation from 2003 (Δ^{03}). Technically, the main source of endogeneity is derived from (Δ^{06}), but (Δ^{03}) must be uncorrelated with the (w_j^{06}). Thus, using the the predicted values $\hat{w}_j^{06}$ derived from the regression between w^{2006} and w^{2003} should reduce endogeneity bias.

<Table III about here>

A limitation of this methodology is the availability of wage data at a communal level for 2003. Using the CASEN 2003, wages were estimated using the same procedure described for CASEN 2006. However, CASEN 2003 has a statistical representation of 302 communes; so imputation is needed for the 32 remaining communes. In the communes without information for 2003, I replaced the fitted values with the w of 2006. Another way to deal with endogeneity problems is to use GMM for the estimation, but the functional forms of equations (9) and (10) make this process extremely complicated. I estimated equations (9) and (10) using GMM, but the logarithm function form implies a flat objective function and convergence was not achieved. However, the procedure that I present with the fitted value of wages must reduce the inherent problems of endogeneity attributed to NEG models. In summary, to estimate equations (9) and (10), the wages on the right-hand side were replaced by the predicted value derived from the regression detailed in table III. To control for endogeneity, a non-linear estimation was carried out using Stata software[22].

6. RESULTS

[22]The codes and data are available upon request.

The results of the estimation of equations (9) and (10) are shown in table IV. Columns 1 and 2 contain the structural parameters for Helpman's model under both specifications of transport costs defined in equation (11). Comparing the structural parameters to table I, all of them are significant and have the expected sign . In Helpman's model, the value $(1 - \mu)$ represents the expenditure of household goods. According the the National Expenditure Survey 1996, the expenditure on household goods is around 14% of the total income, which is close to the value estimated (15.9%). A small change in μ is observed for the distance functional form, but still highly consistent with theoretical requirements (13.9%).

A key parameter is the elasticity of substitution σ, the estimation of which is significantly greater than one for Helpman's model. This implies a profit margin of 2.3% attributed to imperfect competition. In spite of the correct sign of σ, the profits derived from imperfect competition are low compared to estimations of the US (14%) and Germany (38%) (See Hanson (2005) and Brakman et al. (2004) respectively). The profit margin does not change for different specifications of transport costs, supporting the consistency of its estimation. As discussed in section 4, the no-black-hole condition is another relevant condition, which is represented by $\sigma(1 - \mu) < 1$. For Helpman's model the no-black-hole condition is rejected for both transport costs, implying that agglomeration is a stable equilibrium, independent of the level of transport costs. Given that the no-black-hole condition is rejected, any comment about τ is meaningless. In summary, Helpman's model shows parameters that support the spatial distribution of wages in Chile, but reflects a reduced profit margin derived from imperfect competition.

The conclusions are slightly different for Sudekum's model (column 3 and 4). The parameter γ, namely the share of housing, is estimated at around 16% for both transport costs. The parameter μ does not change considerably when using an exponential specification, but it increases to 0.88 for the distance function. The parameter σ is very similar to that in Helpman's model, implying a similarly low margin profit of 2.3%. The no-black-hole condition is fulfilled for both specifications of transport cost, indicating that an increase in the level of freeness among communes implies a higher heterogeneity of wages. Columns 3 and 4 show that τ is significant and positive, suggesting that the market-access effect has a positive impact on wages, but is greater when locations are closer. In summary, both models suggest a core-periphery for Chile with a market-access effect that is stronger than the housing price-effect. However, the level of freenees is only relevant for Sudekum's model.

As discussed in the previous sections, the case of developing countries is a mix between the skill composition of workforce and first and second nature. To control for skill and first nature, several variables are added, such as the number of workers participating in the mining sector, average education and proportion of highly skilled workers, such as Brakman et al. (2004). The incorporation of these controls variables are shown in the columns 5, 6, 7 and 8 of the table IV. The set of control variables are described in table II, but just the significant ones are reported in table IV. They have the expected signs. For example, the proportion of male workers positively affects the local wage, supporting the gender gap in wages. The relationship between the proportion of highly skilled workers and wages is positive and significant. Finally, the proportion of agricultural workers has a negative impact on local wages, representing the negative effect of low skills. The incorporation of these variables implies some changes in the parameters, especially in σ which lose its significance every time that the power distance function is used. Additionally, the σ coefficient is unstable for different initial values under the distance function, reducing the stability of this functional form. The p-value goes up for the power distance function specification in both models, reducing its level of fitness. Using this result, I choose to keep the analysis with the more stable specification derived from the exponential distance function.

<Table IV about here>

In spite of controlling for several sources of agglomeration, the concern about endogeneity still remains. Using the methodology discussed in section 5, the equations (9) and (10) are estimated using the fitted variable $\hat{w}_j^{2006}$ in the right-hand side of the equation. As Table IV suggests, endogeneity generates an under-estimation of σ. The parameter σ increases significantly for both models, indicating a higher imperfect substitution of manufactured goods. μ and γ are within the theoretical requirements. Columns 9 and 10 of Table IV show the importance of the neighborhood to determine the wage level through the positive value of τ. Locations with high wages are surrounded by locations that also have high wages, supporting the earlier results suggested by the Moran's test. This parameter is very stable, independently of the model estimated. Control variables have the right sign. The profit rule is fixed around 2.3% for both models, confirming the presence of imperfect competition. However, only Sudekum's model supports the no-black-hole condition, suggesting that the core-periphery structure can increase even for a continuous reduction of transport costs. Helpman's model does not fulfill the no-black-hole condition, indicating that total agglomeration is a stable equilibrium independently of dispersion forces. Although comparison of the models is not the main goal of this paper, the Wald test suggests that the parameter γ is significantly different from $(1-\mu)$, arguing that both models explain the core-periphery structure. However, only Sudekum's model fulfills the no-black-hole condition and this model should represent more properly the spatial distribution of wages in Chile. This result implies that Chile seems to be located on the right side of the bell-shaped curve of spatial development, fitting with the case of developing countries and supporting the empirical literature (Echeverria and Gopinath, 2007, Aroca, 2009).

7. CONCLUSIONS

The NEG uses the combination of increasing returns, transport costs and imperfect competition to explain the spatial distribution of wages. In this paper I suggest that Chile fits with the predictions derived from the NEG, but with weak evidence of imperfect competition. The spatial distribution of wages is mainly explained by the market-access effect being stronger than the housing-price effect. This implies that the MR, even with congestion costs derived from agglomeration generates higher agglomeration economies with respect to the other regions in Chile. This result is supported by the estimation of the models proposed by Helpman (1998) and Sudekum (2006). Simultaneously, the empirical evidence suggests that Sudekum (2006) fulfills the no-black-hole condition, suggesting the case of Chile as a developing country. From the policy perspective, this suggests there will be higher levels of wage disparities in the future because agglomeration is a stable equilibrium. These results are in line with previous research of Aroca and Bosch (2000), Echeverria and Gopinath (2007) and Aroca (2009).

With respect to the structural estimation of Helpman's and Sudekum's models, both are theoretically consistent for almost the complete set of theoretical requirements. The coefficients σ, μ, γ and τ lie in the theoretical ranges specified by the NEG model. This consistency is robust for two functional forms of transportation costs, namely the exponential distance function and the power distance function. The shares μ, γ are precisely estimated according to the evidence provided by the National Expenditure Survey 1996. However, the control of the skill composition of the workforce and first nature do not appear to modify the estimations for both models. This could be explained by the high levels of nonlinearity of equations (9) and (10), where the ad hoc incorporation of a linear set of variables has a marginal impact on an extremely flat function. These control variables only changed the significance of the parameter σ, specifically for the power distance function form of transport costs. This result is in line with the literature, where the exponential decay functional form has been widely used in empirical estimations. Considering the exponential decay function, both models show precise estimations for the structural parameters.

A final adjustment was the control of endogeneity. Using the fitted values $\hat{w}_j^{2006}$, the structural parameters are consistent for both models. However, the no-black-hole condition is rejected for Helpman's model. This implies that concentration of wages around the MR is a fixed result, independent of the level of freeness trade. However, Sudekum's model satisfies the no-black-hole condition, giving evidence about the role played by the level of freeness trade. This opens a new empirical alternative to evaluate the market-access and housing-price effects. Sudekum's model was estimated for the first time and its performance was suitable for the case of Chile, providing evidence about the spatial distribution of wages at the communal level. This implies that Chile should be located on the left side of the bell shape. With this evidence, higher levels of agglomeration can be expected, consistent with a developing country.

In spite of the theoretical requirements, intuition must play a role in interpreting the results derived from the econometric exercise. The core-periphery structure rests on the crucial assumption of imperfect competition. For the complete set of models, the power market derived from imperfect competition implies only a 2.3% margin on the competitive equilibrium price. This is extremely low with respect to the literature and the evidence must be considered carefully. This result can be related directly to the combination of characteristics present in developing countries. Ad hoc incorporation at the aggregated level (communal) does not seem to properly capture the additional variation of wages. New directions are needed for the special case of developing countries. One is the estimation of the models at microdata, where the controls of skill composition and first nature can be incorporated in the estimation of the causal mechanism. Future research in this area will provide information on where Chile is located on the bell-shaped curve of spatial development, as well as the role played by second nature in the spatial distribution of wages.

REFERENCES

Armstrong H, Taylor J. 2000. *Regional Economics and Policy.* Oxford: Blackwell Publishers.

Aroca P. 2009. Las desigualdades regionales en chile. *Foreign Affairs* **9**: 53–62.

Aroca P, Bosch M. 2000. Crecimiento, convergencia y espacio en las regiones chilenas: 1960-1998. *Estudios de Economia* **27**: 199–224.

Baldwin R, Forslid R, Martin P, Ottaviano G, Robert-Nicoud F. 2003. *Economic Geography and Public Policy.* New Jersey: Princeton University Press.

Brakman S, Garretsen H, Schramm M. 2004. The spatial distribution of wages: estimating the Helpman-Hanson model for Germany. *Journal of regional Science* **44**: 437–466.

Brakman S, Garretsen H, van Marrewijk C. 2009. *The New Introduction to Geographical Economics.* Cambridge: Cambridge University Press.

Combes P, Duranton G, Gobillon L. 2008a. Spatial wage disparities: Sorting matters! *Journal of Urban Economics* **63**: 732–742.

Combes P, Lafourcade M, Thisse J, Toutain J. 2008b. The rise and fall of spatial inequalities in france: a long-run perspective.
URL http://www.cepr.org/pubs/dps/DP7017.asp

Contreras D. 2003. Poverty and inequality in a rapid growth economy: Chile 1990-96. *Journal of Development Studies* **39**: 181–200.

Dixit A, Stiglitz J. 1977. Monopolistic competition and optimum product diversity. *American Economic Review* : 297–308.

Duncan R, Fuentes J. 2006. Convergencia regional en chile: nuevos test, viejos resultados. *Cuadernos de Economia* **43**: 81–112.

Echeverria R, Gopinath M. 2007. *Geographic concentration of Economic Actvity: The case of Chile*.
URL http://aede.osu.edu/Programs/Anderson/trade/65EcheverriaGopinath.pdf

Ferreira F, Litchfield J. 1999. Calm after the storms: income distribution and welfare in chile 1987-1994. *World Bank Economic Review* **13**: 509–538.

Fujita M, Krugman P, Venables A. 1999. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge: MIT Press.

Fujita M, Thisse J. 2002. *Economics of Agglomeration*. Cambridge: Cambridge University Press.

Hanson G. 2005. Market potential, increasing returns and geographic. *Journal of International Economics* **67**: 1–24.

Head K, Mayer T. 2000. Non-Europe: the magnitude and causes of market fragmentation in Europe. *Weltwirtschaftliches Archiv* **136**: 285–314.

Helpman E. 1998. The size of regions. In Pines D, Zadka E, Zilcha I (eds.) *Topics in Public Economics*. Cambridge: Cambridge University Press, 33–54.

Henderson V. 2003. The urbanization process and economic growth: the so-what question. *Journal of Economic Growth* **8**: 47–71.

Kiso T. 2005. Does new economic geography explain the spatial distribution of wages in japan. *mimeo* .

Krugman P. 1991. Increasing returns and economic geography. *Journal of Political Economy* : 483–499.

Marshall A. 1890. *Principles of Economics*. New York: Prometheous Books.

Mion G. 2004. Spatial externalities and empirical analysis: the case of Italy. *Annals of the Association of American Geographers* **56**: 97–118.

Morande F, Soto R, Vergara R. 1997. Achilles, the tortoise and regional growth in chile. In Morande F, Vergara R (eds.) *Estudios Empíricos sobre el Crecimiento en Chile*. 141–170.

Myrdal G. 1957. *Economic Theory and Underveloped Regions.* Methuen: University Paperbacks.

Puga D. 1998. Urbanisation patterns: European versus less developed countries. *Journal of Regional Science* **38**: 231–252.

Puga D. 1999. The rise and fall of regional inequalities. *European Economic Review* **43**: 303–334.

Puga D. 2002. European regional policies in the light of recent location theories. *Journal of Economic Geography* **2**: 372–406.

Redding S. 2010. The empirics of new economic geography. *Journal of Regional Science* **50**: 297–311.

Rosenthal S, Strange W. 2004. Evidence on the nature and sources of agglomeration economies. In Henderson J, Thisse J (eds.) *Handbook of Regional and Urban Economics.* 2120–2171.

Samuelson P. 1952. The transfer problem and transport costs: The terms of trade when impediments are absents. *Economic Journal* : 278–304.

Sudekum J. 2006. Agglomeration and regional costs of living. *Journal of Regional Science* **46**: 529–543.

Tobler W. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* : 234–240.

7.1 Figures

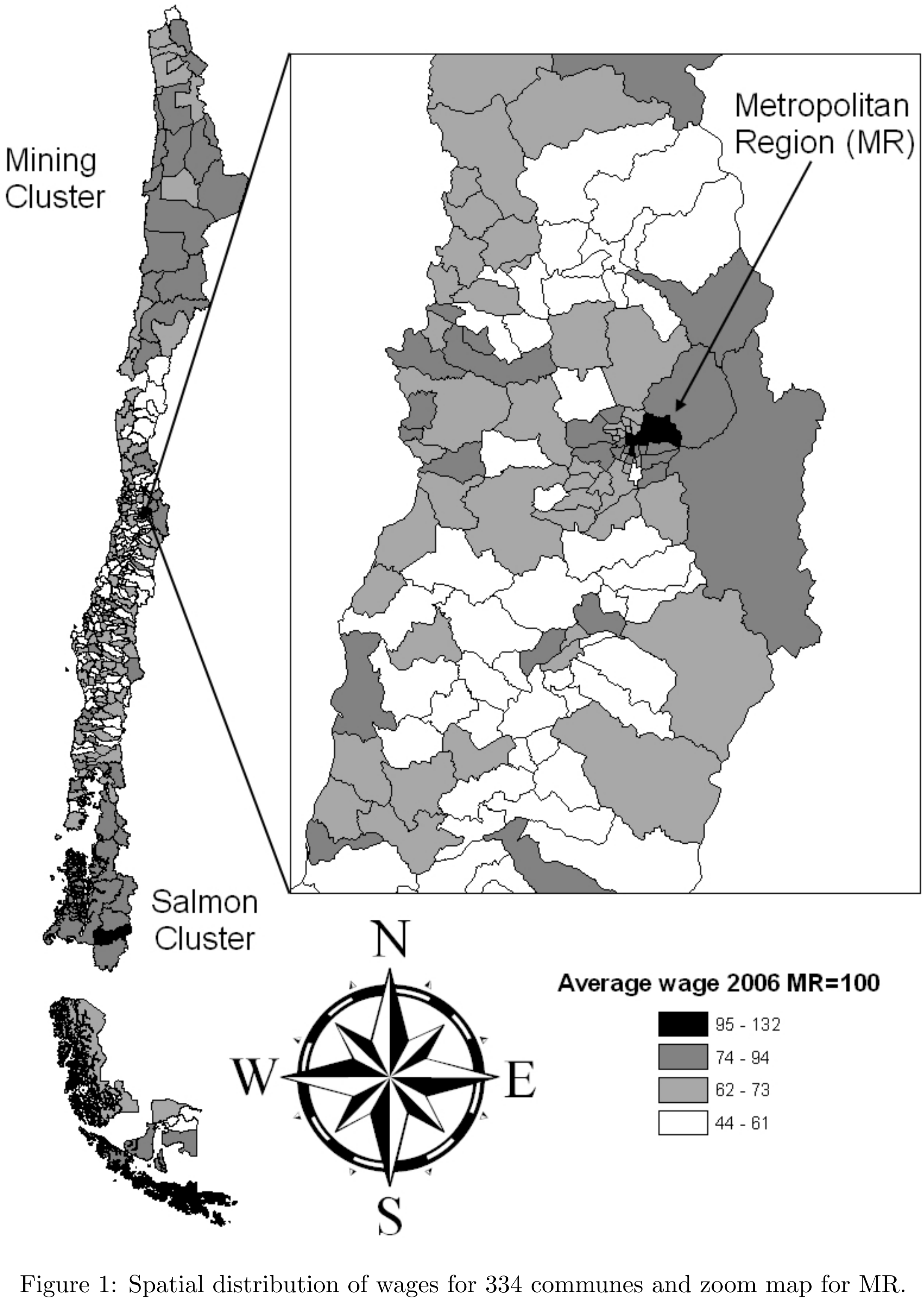

Figure 1: Spatial distribution of wages for 334 communes and zoom map for MR.

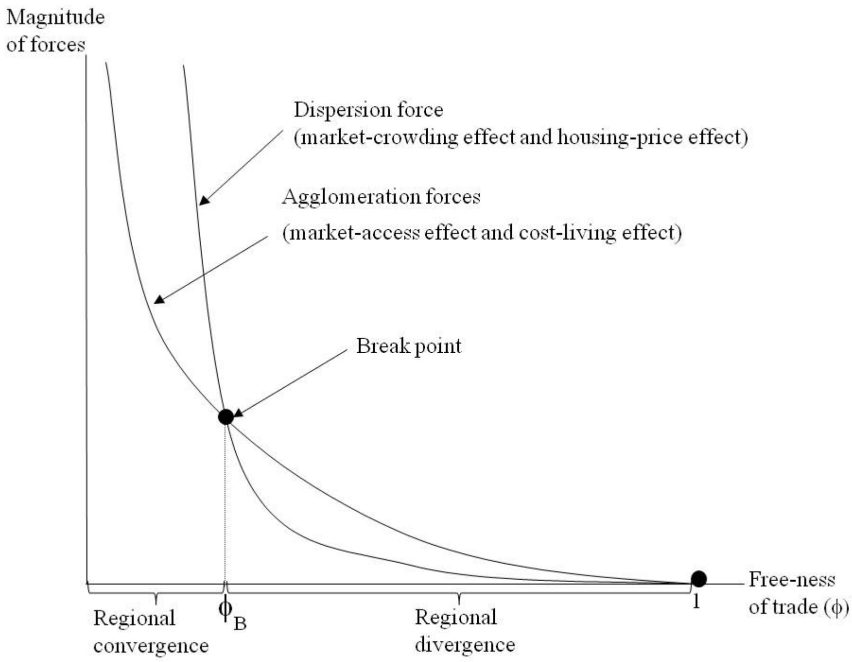

Figure 2: Agglomeration and dispersion forces erode with trade freeness (Baldwin et al., 2003).

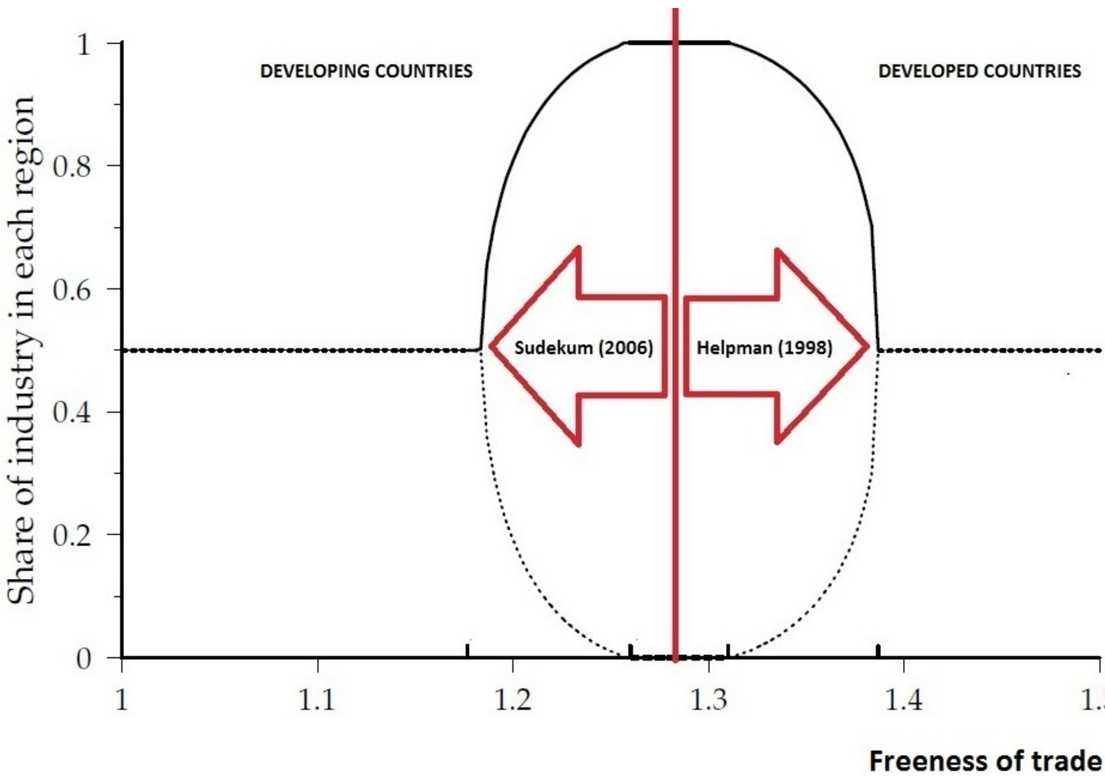

Figure 3: Bell-shape curve of spatial development.

7.2 Tables

Table I: Theoretical constraint of parameters

	Helpman	Sudekum
$\sigma > 1$	Substitution elasticity	Substitution elasticity
$0 < \mu < 1$	Share of industrial goods	Share of industrial goods
$\tau > 0$	Unit transportation cost	Unit transportation cost
$0 < \gamma < 1$	-	Share of housing
$\sigma / (\sigma - 1) > 1$	Imperfect competition	Imperfect competition
$\sigma (\mu - 1) < 1$	No-black-hole condition	-
$\sigma (\mu - 1) > 1$	-	No-black-hole condition

Table II: Summary statistics

Variable	Mean	Std. Dev.	N
Wage	243118	43933	334
Consumer Purchasing Power (CPP)	5839	13589	334
IHPI	71734	45865	334
Sex	0.494	0.019	334
Education	8.819	1.401	334
Share highly skilled workers	0.122	0.092	334
Share fully employed workers	0.893	0.048	334
Share agricultural workers	0.315	0.214	334
Share mining workers	0.022	0.053	334
Share industrial workers	0.108	0.06	334
Share construction workers	0.084	0.044	334
Share service workers	0.036	0.037	334
Wage 2003	233077	41379	334
Correlation coefficients			
	Wage	CPP	IHPI
Wage	1.000		
CPP	0.576	1.000	
IHPI	0.595	0.528	1.000
Moran's test			
	Wage	CPP	IHPI
Wage	0.142*		
CPP		0.158*	
IHPI			0.156*

*Significant at 1% level.

1) All variables are estimated for 2006 unless indicated.

2) Wage and IHPI are expressed in Chilean pesos for 2006.

3) CPP is expressed in millions of Chilean pesos for 2006.

4) Sex is 0 for woman and 1 for man.

5) Education is the average of years for all workers at the commune level.

6) Moran's test was computed using the inverse distance matrix.

Table III: Fitted wage

	(1) Wage 2006
Constant	49432.12
	(0.000)
Wage 2003	0.83
	(0.000)
N	334
R^2	0.61

1) p-values in parentheses

2) p-values are based on White's heteroscedasticity-adjusted standard errors.

3) Estimation with OLS.

Table IV: Structural estimation

	Helpman (1)		Sudekum (2)		Helpman (3)		Sudekum (4)		Final model (5)	
	Exponential	Distance	Exponential	Distance	Exponential	Distance	Exponential	Distance	Sudekum	Helpman
σ	44.96	46.01	44.96	47.38	40.44	41.80	33.57	40.54	42.14	41.06
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.045)	(0.000)	(0.854)	(0.001)	(0.000)
μ	0.841	0.861	0.841	0.887	0.754	0.780	0.623	0.756	0.780	0.759
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
τ	0.103	0.029	0.104	0.061	0.000206	0.0135	0.00267	0.0171	0.003	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.014)	(0.000)	(0.000)
γ			0.159	0.158			0.429	0.310		0.308
			(0.000)	(0.000)			(0.000)	(0.328)		(0.000)
$\sigma/(\sigma-1)$	1.023	1.022	1.023	1.022	1.025	1.025	1.031	1.025	1.024	1.025
$\sigma(1-\mu)$	7.149	6.395	7.149	5.354	9.95	9.20	12.66	9.89	9.271	9.900
Wald test			0.25	1.33			17.69	0.32		18.90
$(1-\mu)=\gamma$			(0.620)	(0.000)			(0.000)	(0.572)		(0.000)
Sex					0.107	0.471	0.346	0.635	0.370	1.015
					(0.663)	(0.100)	(0.317)	(0.036)	(0.274)	(0.005)
High skill					1.655	0.895	1.088	1.436	1.333	1.316
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Full time					-0.325	-0.300	-0.104	-0.0465		
					(0.003)	(0.006)	(0.427)	(0.687)		
Agriculture					-0.376	-0.390	-0.298	-0.537	-0.394	-0.425
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\sqrt{MSE}$	0.444	0.911	0.446	0.176	0.457	0.790	0.347	0.121	0.448	0.221
N	334	334	334	334	334	334	334	334	334	334

1) p-values in parentheses.
2) p-values are based upon White's heteroscedasticity-adjusted standard errors.
3) Estimation with NLS.
4) The coefficients for the rest of the control variables are available upon request to the author.