

# Regional Unemployment Rates in an Agglomeration Economy: A Theoretical and Empirical Analysis\*

Keisuke Kondo<sup>†</sup>

First Version: February 27, 2013

This Version: January 29, 2014

## Abstract

This paper theoretically and empirically analyzes the relationship between regional unemployment rates and agglomeration by introducing the standard search and matching framework into a new economic geography model. Furthermore, we incorporate agglomeration externalities into a search and matching framework. After our theoretical analysis, we empirically examine relationships between regional unemployment rates and agglomeration and between matching efficiency and agglomeration by using Mexican data. An important prediction of our theory is that regional unemployment rates can be positively or negatively correlated with agglomeration under negative agglomeration externalities on matching efficiency. We empirically find that denser areas have comparatively low unemployment rates under negative agglomeration externalities on matching efficiency. Considering our theoretical predictions, we conclude that in Mexico, the agglomeration effect lowering the unemployment rates is much stronger than that increasing the rates.

*JEL classifications:* F12, J61, J64, R12, R23.

*Keywords:* Agglomeration, Regional Unemployment Rates, Search and Matching, Trade, Migration

---

\*I would like to specially thank Kensuke Teshima and Yasuhiro Sato for their very helpful comments and suggestions. I also thank Jorge Alonso, Hiroshi Goto, Nobuaki Hamaguchi, Naoto Jinji, Hisaki Kono, Toshihiro Okubo, Mun Se-il, Koji Shintaku, and all the participants in the CIE Brown Bag Seminar at the Instituto Tecnológico Autónomo de México (ITAM), the third spring meeting of the Japan Society of International Economics at Fukuoka University, and the Brown Bag Lunch Seminar at Kyoto University for their useful comments and suggestions. Any remaining errors are naturally my own. I am grateful to those at the ITAM for all the support that I received during my stay there. This research was carried out under a scholarship granted by the Government of Mexico through the Ministry of Foreign Affairs of Mexico. The Web supplement file for this work is available on the following Web site:  
(URL: <http://www2.kobe-u.ac.jp/~080e702e/>).

<sup>†</sup>Graduate School of Economics, Kobe University. 2-1 Rokkodaicho, Nada-ku, Kobe-shi, Hyogo-ken, 657-8501, Japan. (e-mail: [keisuke.kondo@stu.kobe-u.ac.jp](mailto:keisuke.kondo@stu.kobe-u.ac.jp)).

# 1 Introduction

Since the publication of Krugman (1991), new economic geography (NEG) studies have examined the agglomeration mechanism of economic activities, with particular attention to the increasing returns to scale, monopolistic competition, transport costs, and mobile labor across regions (e.g., Fujita et al., 1999). With regarding to regional labor markets and agglomeration, Marshall (1890) observes that the concentration of economic activities facilitates the job search and matching between employers and job seekers in terms of industry-specific skills. Similarly, Rosenthal and Strange (2001), investigating the determinants of agglomeration, find that labor market pooling fosters agglomeration. Despite such observations and studies, only limited attention has been paid to job search and regional unemployment issues in the NEG literature.<sup>1</sup> Thus, we do not fully understand the underlying mechanism acting between regional unemployment rates and the agglomeration of economic activities.

In the recent NEG literature, attempts have been made to tackle job search and unemployment issues.<sup>2</sup> For example, Epifani and Gancia (2005) and Francis (2009) developed a dynamic NEG model by introducing a search and matching mechanism.<sup>3</sup> Their models predict a lower unemployment rate in agglomerated regions in the long-run. On the other hand, motivated by that unemployment rates in high-density regions seem to be higher than in low-density regions from developed countries data, vom Berge (forthcoming) extended Krugman's (1991) model by introducing a search and matching framework.<sup>4</sup> His model shows

---

<sup>1</sup>Note that the NEG literature has also contributed to uncovering wage inequality from the perspective of geographical networks. For example, many empirical papers have shown that market potential leads to higher regional nominal wages (e.g., Redding and Venables, 2004; Hanson, 2005; Hering and Poncet, 2010). However, these studies are based on theoretical models under perfectly competitive labor markets.

<sup>2</sup>Some theoretical mechanisms that generate unemployment need to be introduced (e.g., efficiency wage or search and matching frameworks). This paper employs the search and matching model proposed by Pissarides (2000). Rogerson et al. (2005) offer a review of this literature. In the literature of international trade, Helpman and Itskhoki (2010) developed an international trade model to analyze the effect of labor market rigidity on trade flow. Unlike those studies, we focus on the trade model dealing with migration between regions.

<sup>3</sup>Unlike the search and matching framework, Zierahn (forthcoming) introduces the efficiency wage and congestion costs due to agglomeration into Krugman's (1991) model.

<sup>4</sup>Unlike vom Berge (forthcoming), we find both positive and negative relationships between unemployment and agglomeration, expressed as population size or population density in empirical studies. See for example Simon (1988), Izraeli and Murphy (2003), and Chiang (2009).

that the unemployment rates in agglomerated regions are comparatively high.<sup>5</sup> However, as mentioned by Zierahn (forthcoming), when NEG models show full agglomeration under the spatial equilibrium, it indicates that unemployed workers do not live in the periphery region.<sup>6</sup> That is, under full agglomeration, the unemployment rate in the periphery region virtually becomes zero (or cannot be defined), whereas it is always positive in the agglomerated region. As such, the results obtained from *full* agglomeration models cannot exactly capture situations in the periphery regions. Therefore, we investigate the relationship between regional unemployment rates and agglomeration by using an NEG model with *partial* agglomeration.

Following the framework proposed by vom Berge (forthcoming), we develop a multi-region model of Helpman (1998) by incorporating a search and matching mechanism.<sup>7</sup> Unlike Krugman (1991), Helpman (1998) lays more emphasis on the dispersion force arising from non-tradable local services. For example, the concentration of economic activities raises the prices of land and housing owing to the increased demand for them. Consequently, this type of dispersion force leads to partial agglomeration. Thus, focusing on Helpman's (1998) model, we offer fresh insight into the regional distribution of unemployment rates in an agglomeration economy. Furthermore, to analyze how transport costs affect the relationship between regional unemployment rates and agglomeration, we carry out a numerical analysis of the theoretical model.<sup>8</sup>

A contribution of this paper is to incorporate agglomeration externalities into a search and matching framework. As observed in Marshall (1890), denser areas seem to promote job matching between job seekers and firms. However, this is not necessarily true in the

---

<sup>5</sup>vom Berge (forthcoming) introduces regions into the model developed by Zieseimer (2005), who extended Pissarides (2000, Chap. 3) model by introducing monopolistic competition.

<sup>6</sup>Agricultural workers still live there in the case of Krugman-type models.

<sup>7</sup>An extension of Helpman (1998) can be found in Pflüger and Tabuchi (2010). They assume that a firm uses land as a production input.

<sup>8</sup>Although NEG models provide insightful policy implications, their theoretical and numerical analyses are usually limited to two-region cases to avoid mathematical difficulties, which are also known as *three-ness* (Combes et al., 2008b, Chap. 4). Although we build a multi-region model for the theoretical part of our study, our numerical analysis is restricted to a case of two symmetric regions.

current economy. Recent empirical studies provide two contradictory evidences. Hynninen and Lahtonen (2007) show a positive relationship between matching efficiency and population density, whereas Kano and Ohta (2005) show a negative one. Therefore, our theoretical model assumes positive or negative agglomeration externalities on matching efficiency. Consequently, our model is able to describe a wide variety of relationships between unemployment rates and agglomeration.

Our study also contributes to the literatures of development economics and wage curve. Beginning with Harris and Todaro (1970), the literature of development economics has studied urban unemployment and migration. Given the exogenously high wage in urban area, Harris and Todaro (1970) showed that urban unemployment rate increases on account of excessive workers immigrating into a city in response to higher expected wage. Therefore, a positive relationship between wages and unemployment rates can be expected.<sup>9</sup> In contrast, the literature of the wage curve, beginning with Blanchflower and Oswald (1994), has studied the negative relationship between regional wages and unemployment rates.<sup>10</sup> Our theoretical model therefore attempts to uncover this contradictory observation.

We specifically describe three relationships between nominal wages, unemployment rates, and agglomeration across regions, as clearly illustrated in Figure 1. Note that a consensus already exists on the positive relationship between wages and agglomeration (e.g., Combes et al., 2008a; Mion and Naticchioni, 2009; Combes et al., 2010; de la Roca and Puga, 2012), and this always holds in our model. In addition, previous studies show that agglomeration has a decreasing effect on unemployment rates in the production side.<sup>11</sup> Further, if agglomeration is assumed to have positive/negative externalities on matching efficiency, it would also lead to negative/positive effects on the unemployment rate. Consequently, the advantage of our model is that we explain both the positive and negative relationships between nominal

---

<sup>9</sup>Contrary to the prediction of Harris and Todaro (1970), Suedekum (2005) showed a lower unemployment rate and higher wage in agglomerated region by endogenously expressing higher urban wage within the NEG framework.

<sup>10</sup>See Card (1995) for a literature review of the wage curve.

<sup>11</sup>See also Suedekum (2005) and Zierahn (forthcoming).

wages and unemployment rates, while also endogenously explaining the higher wages in agglomerated regions. Therefore, we believe that this paper makes a valuable contribution to the Harris–Todaro model and the wage curve literature.

[Figure 1 about here]

This paper also includes an empirical analysis of the relationship between regional unemployment rates and agglomeration. We use Mexican municipal data and control for spatial dependence within the municipal data by using spatial econometric methods. We also estimate the matching function to examine the relationship between matching efficiency and agglomeration. Finally, we draw a conclusion about the relationship between unemployment rates and agglomeration by taking into account both the estimation results.

As mentioned in Krugman and Livas-Elizondo (1996), Mexico has experienced a dynamic change in economic activities since the trade liberalization in the 1980s and 1990s. In the meantime, this movement brought about drastic changes in the country’s domestic distributional employment pattern. According to Hanson (1998), the Mexico–US border states attracted more manufacturing workers. For example, Hanson (1998) shows that the share of regional employment in the Mexico-US border states was 21.0% in 1980 but 29.8% in 1993. On the other hand, the manufacturing workers tend to leave the Mexico City metropolitan area (their share came down from 46.4% in 1980 to 28.7% in 1993). However, little attention has been paid to the relationship between regional unemployment rates and agglomeration in the Mexican literature; therefore, we try to examine whether the agglomerated regions have higher or lower unemployment rates.

The remainder of this paper is organized as follows. Section 2 builds a multi-region model of Helpman (1998) consisting of a standard search and matching framework. Section 3 numerically analyzes a case of two symmetric regions. Section 4 details the empirical strategy used for this study. Section 5 explains the data used. Section 6 discusses the estimation results. Finally, Section 7 concludes.

## 2 The Model

Following vom Berge (forthcoming), we extend the multi-region model of Helpman (1998) by introducing a search and matching framework. We consider an economy with  $R$  regions having both manufacturing and land/housing sectors. The manufacturing sector is monopolistically competitive, and each firm produces one variety of a differentiated good under increasing returns to scale. Labor is a unique production input. On the other hand, the land/housing sector is perfectly competitive; land endowment in each region is fixed, so that the supply of land/housing services is also given; consumers have their own land equally. There are two types of workers, the employed and the unemployed. We assume that both types of the worker are mobile across regions in the long-run, and that there are no migration costs. We introduce job search and matching frictions into the regional labor markets. Unemployed workers search for jobs in their own living regions, and spatial job search is not allowed. For the present purpose, we focus on steady state analysis.

### 2.1 Matching Function

We first assume that there are search and matching frictions in the regional labor markets. The number of matches existing between the job seekers and vacancies is determined by the following matching function:

$$m_i L_i = A_i m(u_i L_i, v_i L_i), \quad i = 1, 2, \dots, R \quad (1)$$

where  $m_i$  is the matching rate,  $u_i$  is the unemployment rate,  $v_i$  is the vacancy rate in terms of labor,  $A_i$  is the matching efficiency, and  $L_i$  is the labor force, with the subscript  $i$  indicating region  $i$ . Note that job matches are made only within region  $i$ . We further assume that the matching function is increasing in both variables, homogeneous of degree one, concave, and twice continuously differentiable, and that  $m(u_i L_i, 0) = m(0, v_i L_i) = 0$ .<sup>12</sup> As mentioned earlier, we assume that agglomeration of economic activity has externalities on the matching

---

<sup>12</sup>See Petrongolo and Pissarides (2001) for details of the matching function, including empirical findings.

efficiency  $A_i$ ; our specification is as follows:

$$A_i = A(L_i/\bar{S}_i)^\xi, \quad (2)$$

where  $A$  is constant,  $\bar{S}_i$  represents land endowment (or fixed supply of land/housing services), and  $\xi$  is the elasticity of agglomeration to matching efficiency. Thus,  $L_i/\bar{S}_i$  can be interpreted as a kind of population density in region  $i$ .

Given the matching function (1), the rates at which vacancies are filled and unemployed workers leave unemployment can be expressed respectively as

$$q_i(\theta_i) \equiv \frac{A_i m(u_i L_i, v_i L_i)}{v_i L_i} \quad \text{and} \quad \theta_i q_i(\theta_i) \equiv \frac{A_i m(u_i L_i, v_i L_i)}{u_i L_i},$$

where  $\theta_i \equiv v_i/u_i$  denotes the labor market tightness. From the above assumptions, we can easily verify that both  $q_i(\theta_i) > 0$  and  $q'_i(\theta_i) < 0$  hold for a given value of  $A_i$ .

## 2.2 Consumer and Worker

For simplicity, we assume a static consumer problem; consumers do not save any part of their income but spend all of it in each period.<sup>13</sup>

Further, each consumer has identical Cobb–Douglas preferences for two goods; that is,

$$U_i = \frac{1}{\mu^\mu (1-\mu)^{1-\mu}} M_i^\mu H_i^{1-\mu}, \quad (3)$$

where  $0 < \mu < 1$  is the expenditure share for manufactured goods,  $M_i$  is the composite consumption of manufactured goods in region  $i$ , and  $H_i$  is the consumption of land/housing service in region  $i$ .<sup>14</sup> The composite consumption of manufactured goods is given by the constant elasticity of substitution (CES) function

$$M_i = \left( \sum_{j=1}^R \int_0^{n_j} m_{ji}(\nu)^{(\sigma-1)/\sigma} d\nu \right)^{\sigma/(\sigma-1)},$$

<sup>13</sup>This simplification, however, does not change the essential results of our model.

<sup>14</sup>We modify the methodology of Pflüger and Tabuchi (2010) to describe a land/housing market.

where  $m_{ji}(\nu)$  is region  $i$ 's consumption of variety  $\nu$  produced in region  $j$ ,  $n_j$  the number of varieties produced in region  $j$ , and  $\sigma > 1$  the elasticity of substitution between any two varieties. The budget constraint of region  $i$  is given by  $G_i M_i + p_i^H H_i = Y_i$ , where  $G_i$  is the price index for manufactured goods,  $p_i^H$  is the price of land/housing services, and  $Y_i$  is the regional income.

From utility maximization, we obtain the following demand functions:

$$H_i = \frac{(1 - \mu)Y_i}{p_i^H}, \quad M_i = \frac{\mu Y_i}{G_i}, \quad \text{and} \quad m_{ji}(\nu) = \mu p_{ji}(\nu)^{-\sigma} G_i^{\sigma-1} Y_i, \quad (4)$$

where  $p_{ji}(\nu)$  is region  $i$ 's consumer price for variety  $\nu$  imported from region  $j$ ; the price index in region  $i$  takes the following form:

$$G_i = \left( \sum_{j=1}^R \int_0^{n_j} p_{ji}(\nu)^{1-\sigma} d\nu \right)^{1/(1-\sigma)}. \quad (5)$$

By substituting demand functions (4) into utility function (3), we obtain the indirect utility  $\mathbb{V}_i$  of an individual living in region  $i$ :

$$\mathbb{V}_i = \frac{I_i}{G_i^\mu (p_i^H)^{1-\mu}}, \quad (6)$$

where  $I_i$  is the income of the individual living in region  $i$ . Indirect utility can be interpreted as the real income, that is, the individual's income  $I_i$  deflated by the cost-of-living index  $G_i^\mu (p_i^H)^{1-\mu}$ .

As mentioned earlier, there are two types of workers in the economy, the employed and the unemployed. Let  $\mathbb{V}_i^e$  and  $\mathbb{V}_i^u$  denote the indirect utilities of the employed and the unemployed, respectively. We assume that while the employed earns  $w_i$ , the unemployed receives unemployment benefit  $z$  from the government. The unemployment benefit is exogenously given. The government imposes a tax  $\tau$  on all the workers in order to finance the unemployment benefits. Further, we assume that the rate of interest  $r$  is common across all regions. Thus, the steady state Bellman equations for the employed and the unemployed



are, respectively, given as follows:

$$\begin{aligned} rE_i &= \mathbb{V}_i^e + \delta(U_i - E_i), \\ rU_i &= \mathbb{V}_i^u + \theta_i q_i(\theta_i)(E_i - U_i), \end{aligned} \tag{7}$$

where  $E_i$  and  $U_i$  are the present discounted values (PDV) of the expected real income stream for the employed and the unemployed, respectively, and  $\delta$  is the job destruction rate. In the long-run, individuals decide to migrate depending on the expected PDV from continuing to live in the region.

### 2.3 Producer Behavior

We assume that the prices of all the varieties produced within a region are identical in view of the same production technology used and therefore denote the price of all the varieties produced in region  $i$  as  $p_i$ . We assume that a manufactured good is traded between regions  $i$  and  $j$  with iceberg transport cost  $T_{ij}$ . Thus, if one unit of any variety of manufactured goods is shipped from region  $i$  to region  $j$ , only  $1/T_{ij}$  of the unit arrives. A variety of manufactured goods produced in region  $i$  is sold at price  $p_i$  in that region. If this variety is shipped from region  $i$  to region  $j$ , the delivered price is

$$p_{ij} = p_i T_{ij}, \quad T_{ij} = T_{ji} \geq 1, \quad T_{ii} = 1, \quad i, j = 1, 2, \dots, R.$$

The total amount of goods that a firm produces to satisfy the consumption demand of all the regions therefore becomes

$$x_i = \sum_{j=1}^R m_{ij} T_{ij}. \tag{8}$$

Next, all the firms require not only fixed and marginal labor input for producing the varieties but also recruiters for hiring their workers.<sup>15</sup> Thus, the total labor input in region

---

<sup>15</sup>This formulation is developed by vom Berge (forthcoming), following Pissarides (2000, Chap. 3) and Ziesemer (2005).

$i$  is

$$\ell_i = F + cx_i + \gamma N_i \quad (9)$$

where  $F$  and  $c$  are respectively the fixed and marginal labor requirements for production,  $\gamma$  is the marginal labor requirement for recruiting per vacancy, and  $N_i$  is the number of vacancies that a firm needs to post. The first two terms correspond to the standard Dixit–Stiglitz assumption of increasing returns to scale. The third term indicates that a firm needs to hire recruiters to keep their workers from decreasing because the workers quit their jobs at a job destruction rate of  $\delta$ . The same wage  $w_i$  is paid to both workers and recruiters. The total cost is therefore  $w_i \ell_i$ .

A vacant job is filled with a probability of  $q_i(\theta_i)$ , and an occupied job is destroyed with a probability of  $\delta$ . Thus, the dynamics of total labor input is given by

$$\dot{\ell}_i = q_i(\theta_i)N_i - \delta \ell_i. \quad (10)$$

A firm maximizes the PDV of its expected profit with respect to the produced quantity  $x_i$  and number of vacancies  $N_i$  as follows:<sup>16</sup>

$$\begin{aligned} \max_{x_i, N_i} \int_0^{\infty} e^{-rt} [p_i(x_i)x_i - w_i(F + cx_i + \gamma N_i)] dt \\ \text{s.t. } \dot{x}_i = \frac{1}{c} [(q_i(\theta_i) - \gamma\delta)N_i - \delta(F + cx_i)] \\ \lim_{t \rightarrow \infty} [\lambda(t)e^{-rt}x_i(t)] = 0 \end{aligned} \quad (11)$$

where  $p_i(x_i)$  is the mill price in region  $i$ , and  $\lambda(t)$  the Lagrange multiplier. Solving the current value Hamiltonian, we obtain the optimal mill price with a constant markup on marginal costs as follows:

$$p_i = \frac{\sigma}{\sigma - 1} cw_i \left( 1 + \frac{r\gamma}{q_i(\theta_i)} \right) \left( 1 - \frac{\gamma\delta}{q_i(\theta_i)} \right)^{-1}. \quad (12)$$

Note that this price is higher than that of the standard Dixit–Stiglitz monopolistic compe-

---

<sup>16</sup>Using (9), (10), and the envelop theorem, we obtain the dynamic equation on production  $\dot{x}$  in (11).

tition model because the multiplication of the second and third terms is greater than one. Intuitively, the marginal cost consists of three parts. The first two terms give the workers' wage for producing the additional quantity  $x_i$  and the expected cost of hiring a worker, and the third term captures the cost of hiring the workers engaged in production and recruitment.<sup>17</sup> If the job search cost is zero ( $\gamma = 0$ ), this price takes the same form obtained for the standard Dixit–Stiglitz model.

Let  $V_i$  and  $J_i$  be the PDVs of the expected profit from the vacant and occupied jobs respectively. Then, the steady state Bellman equation for a vacancy is given by

$$rV_i = -\gamma\tilde{w}_i + q_i(\theta_i)(J_i - V_i), \quad (13)$$

where  $\tilde{w}_i \equiv w_i/p_i$  is the real wage defined in terms of firm.

All the profit opportunities from creating new jobs are exploited in equilibrium, and the value of the vacant jobs becomes zero ( $V_i = 0$ ). Hence, the equilibrium condition yields

$$J_i = \frac{\gamma\tilde{w}_i}{q_i(\theta_i)}. \quad (14)$$

From this equation, since  $1/q_i(\theta_i)$  is the expected duration of a vacant job, the expected profit from a new job is equal to the expected cost of hiring a worker in equilibrium.

## 2.4 Wage Bargaining

In a wage bargaining process, we endogenize the labor market tightness  $\theta_i$ . Each firm in a standard search and matching model is assumed to have only one job. Although a firm in our model employs multiple workers, we consider the bargaining process in a similar manner.<sup>18</sup>

---

<sup>17</sup>To understand the third term, we manipulate (17) to obtain

$$\frac{\delta(F + cx_i)}{q_i(\theta)N_i} = 1 - \frac{\gamma\delta}{q_i(\theta_i)} < 1.$$

The left-hand side shows how the quitting workers engaged in production are filled up from among the newly hired workers, implying that a part of the newly hired workers are engaged in recruitment.

<sup>18</sup>Stole and Zwiebel (1996a,b) consider an extended version of Nash bargaining for multiple workers, in which the firm and a worker divide the marginal surplus obtained from the firm producing goods by hiring additional worker and the worker leaving the unemployed status. This assumption reflects the case in which

Following Pissarides (2000, Chap. 3), we assume that the wages of workers are fixed in Nash bargains, in which the firm gets involved with each worker separately, considering the wages of all the other workers as given. This assumption results in a one-to-one relationship between a worker and a job. The total surplus arising from a job match (i.e., the net benefit of the worker and the firm from the unemployed worker starting to work and the firm producing additional goods) is shared through Nash bargaining between the worker and the firm:

$$\tilde{w}_i = \arg \max (E_i - U_i)^\beta (J_i - V_i)^{1-\beta},$$

where  $0 \leq \beta \leq 1$  is the bargaining power of the workers. From the first-order condition, the result of the bargaining is given by

$$(1 - \beta)(E_i - U_i)J'_i = \beta(J_i - V_i)E'_i.$$

By substituting (7) and (14) and imposing the equilibrium condition  $V_i = 0$ , we obtain the following equation

$$\tilde{w}_i = rU_i + \beta \left( \frac{\sigma - 1}{c\sigma} - rU_i \right).$$

With some manipulations, we obtain the following relationship between the nominal wage and labor market tightness:

$$g_i(w_i, \theta_i) \equiv (1 - \beta) \left( 1 - \frac{z_i}{w_i} \right) - \beta \frac{\gamma[r + \delta + \theta_i q_i(\theta_i)]}{q_i(\theta_i) - \gamma\delta} = 0. \quad (15)$$

This corresponds to the wage-setting curve in Pissarides (2000), but shows a nonlinear function with regard to labor market tightness and wages in our case. From the implicit function theorem, we obtain

$$\frac{d\theta_i}{dw_i} = -\frac{\partial g_i / \partial w_i}{\partial g_i / \partial \theta_i} > 0,$$

where a homogeneous degree one is assumed in the matching function.<sup>19</sup> Since the unem-

---

additional employment additionally affects the wages of the remaining workers. In this paper, we use a simpler methodology employed by Pissarides (2000, Chap. 3).

<sup>19</sup>Under the assumption of a homogeneous degree one in matching function, we confirm that  $q_i(\theta_i) + \theta_i q'_i(\theta_i) > 0$  holds.

ployment rate  $u_i$  and labor market tightness  $\theta_i$  are negatively correlated, this result indicates a negative relationship between wage and unemployment rate.<sup>20</sup>

## 2.5 Short-Run Equilibrium

We now consider a short-run equilibrium, characterized by a general equilibrium in each region without migration.<sup>21</sup> By substituting the price in (12) into the current profit in (11) and imposing a zero-profit condition, the equilibrium output is given by

$$x_i = \frac{F(\sigma - 1)}{c} \left( 1 + \frac{\sigma r \gamma}{q_i(\theta_i)} \right)^{-1}. \quad (16)$$

Note that the equilibrium output is lower than the output of a standard Dixit–Stiglitz monopolistic competition model.

Since  $\dot{\ell}_i = 0$  in the steady state, by substituting (9), the number of vacancies in the steady state becomes

$$N_i = \frac{\delta(F + cx_i)}{q_i(\theta_i) - \gamma\delta}, \quad (17)$$

where we assume  $q_i(\theta_i) > \gamma\delta$  so that the number of vacancies takes a positive value. Substituting the equilibrium output (16) and the number of vacancies (17) into the total labor input (9), we obtain the equilibrium total labor input in region  $i$  as follows:

$$\ell_i = F\sigma \left( 1 + \frac{r\gamma}{q_i(\theta_i)} \right) \left( 1 + \frac{\sigma r \gamma}{q_i(\theta_i)} \right)^{-1} \left( 1 - \frac{\delta\gamma}{q_i(\theta_i)} \right)^{-1}. \quad (18)$$

Further, from the labor market clearing condition  $n_i\ell_i = (1 - u_i)L_i$ , the number of firms is given by

$$n_i = \frac{(1 - u_i)L_i}{F\sigma} \left( 1 + \frac{r\gamma}{q_i(\theta_i)} \right)^{-1} \left( 1 + \frac{\sigma r \gamma}{q_i(\theta_i)} \right) \left( 1 - \frac{\delta\gamma}{q_i(\theta_i)} \right). \quad (19)$$

---

<sup>20</sup>This result implies the existence of wage curve (Blanchflower and Oswald, 1994). In case the regional labor markets are homogeneous with regard to job destruction rates and job matches, a negative correlation could arise between the regional unemployment rates and nominal wages. This result is quite similar to Sato (2000), who shows that even when workers are mobile, the wage curve can be observed by using a theoretical search framework assuming different productivities across the regions and a monocentric city structure.

<sup>21</sup>For ease of expression and interpretation, a numéraire good is not particularly set up. This is not to lose generality of our model analysis and draw model implications for numerical analysis.

From (8), the total sales of the variety produced in region  $i$  amount to

$$x_i = \mu \sum_{j=1}^R p_i^{-\sigma} G_j^{\sigma-1} Y_j T_{ij}^{1-\sigma}. \quad (20)$$

Choosing the convenient units of measurement for marginal labor requirement  $c = (\sigma - 1)/\sigma$  and fixed labor requirement  $F = \mu/\sigma$ , we simplify the model outcomes. Thus, from (12), (16), and (20), we obtain the NEG wage equation:

$$w_i = \Gamma(\theta_i) \left[ \mu \sum_{j=1}^R Y_j G_j^{\sigma-1} T_{ij}^{1-\sigma} \right]^{1/\sigma}, \quad (21)$$

where

$$\Gamma(\theta_i) = \left( 1 + \frac{\sigma r \gamma}{q_i(\theta_i)} \right)^{1/\sigma} \left( 1 + \frac{r \gamma}{q_i(\theta_i)} \right)^{-1} \left( 1 - \frac{\delta \gamma}{q_i(\theta_i)} \right). \quad (22)$$

The sum in brackets gives the RMP  $\equiv \mu \sum_{j=1}^R Y_j G_j^{\sigma-1} T_{ij}^{1-\sigma}$ , expressing the sum of the regional income discounted by the price index, and weighted by the transport cost. Even if we assume the frictions in the regional labor markets, the standard implication for NEG holds; that is, the goodness of accessibility to other markets increases the nominal wages.

Following the assumption of an identical price for all the varieties produced within a region, the price index takes the following form:

$$G_i = \left[ \sum_{j=1}^R n_j (p_j T_{ji})^{1-\sigma} \right]^{1/(1-\sigma)}. \quad (23)$$

By substituting (12) and (19) into (23) and with normalization, we obtain

$$G_i = \left[ \sum_{j=1}^R (1 - u_j) L_j \Gamma(\theta_j)^\sigma w_j^{1-\sigma} T_{ji}^{1-\sigma} \right]^{1/(1-\sigma)}. \quad (24)$$

As mentioned earlier, wage equation, RMP, and price index are essentially identical with vom Berge (forthcoming).

The price of land/housing services  $p_i^H$  is determined at equilibrium, where land endowment (or fixed supply of land/housing services)  $\bar{S}_i$  and the regional demand for land/housing

services  $H_i$  are equal. Thus, the price of land/housing services in region  $i$  is as follows:

$$p_i^H = \frac{(1 - \mu)Y_i}{\bar{S}_i}. \quad (25)$$

The regional income,  $Y_i$ , includes the income of every employed and unemployed worker living in region  $i$ . The respective disposable income of the employed and unemployed workers are given by  $I_i^e = w_i + h - \tau$  and  $I_i^u = z + h - \tau$ , where  $h$  is the land rent and  $\tau$  is the tax rate. Since all the individuals have their own land equally, the land rent is equally redistributed.<sup>22</sup> Thus, the land rent is given by

$$h = \frac{1 - \mu}{\mu} \frac{\sum_{j=1}^R [w_j(1 - u_j) + zu_j - \tau]L_j}{\sum_{j=1}^R L_j}. \quad (26)$$

Therefore, the regional income  $Y_i$  becomes

$$Y_i = [w_i(1 - u_i) + zu_i - \tau]L_i + \frac{1 - \mu}{\mu} \frac{L_i}{\sum_{j=1}^R L_j} \left[ \sum_{j=1}^R (w_j(1 - u_j) + zu_j - \tau)L_j \right]. \quad (27)$$

Further, the individual real income takes the following forms:

$$\mathbb{V}_i^e = \frac{w_i + h - \tau}{G_i^\mu (p_i^H)^{1-\mu}}, \quad \text{and} \quad \mathbb{V}_i^u = \frac{z + h - \tau}{G_i^\mu (p_i^H)^{1-\mu}}. \quad (28)$$

Next, we consider labor market tightness and unemployment rates. Given  $w_i$ , labor market tightness is determined in (15). Since the inflow and outflow of unemployment are equalized in steady state equilibrium, we obtain  $\delta(1 - u_i)L_i = \theta_i q_i(\theta_i)u_i L_i$ . Solving this with respect to  $u_i$ , we obtain the so-called Beveridge curve:

$$u_i = \frac{\delta}{\delta + \theta_i q_i(\theta_i)}. \quad (29)$$

The tax rate  $\tau$  is determined to balance the budget for tax revenue and expenditure for unemployment benefits as follows:

$$\tau \sum_{j=1}^R L_j = z \sum_{j=1}^R u_j L_j. \quad (30)$$

---

<sup>22</sup>See Appendix A for details of the derivation.

Finally, the matching function is assumed to take the Cobb–Douglas form with constant returns to scale

$$A_i m(u_i L_i, v_i L_i) = A_i (u_i L_i)^\alpha (v_i L_i)^{1-\alpha}, \quad (31)$$

where  $\alpha$  is the matching elasticity. This specification of the matching function is also used in the empirical analysis.

### 3 Long-Run Equilibrium: A Two-Region Case

In this section, we numerically analyze the properties of our model.<sup>23</sup> We limit our numerical analysis to a two-region case ( $R = 2$ ) owing to mathematical difficulties.

#### 3.1 Spatial Equilibrium

We assume that workers are mobile across regions in response to the expected PDV differentials in the long-run. For convenience of notation, we denote the shares of labor force in regions 1 and 2 as  $s_1 = L_1/(L_1 + L_2)$  and  $s_2 = 1 - s_1$ , respectively. The regional differentials in the expected PDVs are then expressed as follows:

$$\Delta\omega(s_1) \equiv \omega_1(s_1) - \omega_2(s_1), \quad (32)$$

where the expected PDV from living in region  $i$  is expressed as  $\omega_i(s_1) = (1 - u_i(s_i))E(s_1) + u_i(s_i)U(s_1)$ , with the PDVs of the employed and the unemployed worker living in region  $i$  given respectively as

$$E_i(s_1) = \frac{(r + \theta_i q_i(\theta_i))\mathbb{V}_i^e + \delta \mathbb{V}_i^u}{r(r + \delta + \theta_i q_i(\theta_i))} \quad \text{and} \quad U_i(s_1) = \frac{\theta_i q_i(\theta_i)\mathbb{V}_i^e + (r + \delta)\mathbb{V}_i^u}{r(r + \delta + \theta_i q_i(\theta_i))}. \quad (33)$$

Note that the wage  $w_i$ , price index  $G_i$ , price of land/housing services  $p_i^H$ , land rent  $h$ , labor market tightness  $\theta_i$ , unemployment rate  $u_i$ , and tax  $\tau$  are functions of  $s_i$ . A spatial equilibrium arises at  $s_1^* \in (0, 1)$  when  $\Delta\omega(s_1) = 0$ , at  $s_1 = 0$  when  $\Delta\omega(0) \leq 0$ , or at  $s_1 = 1$

---

<sup>23</sup>Numerical analysis is conducted using the Ox Console 7.01 (Doornik and Ooms, 2006).



when  $\Delta\omega(1) \geq 0$ . Any adjustment process over time  $t$  is governed by the following differential equation:

$$\frac{ds_1}{dt} \equiv \dot{s}_1 = \Delta\omega(s_1)s_1(1 - s_1), \quad (34)$$

where the equilibrium is stable when the slope of  $\dot{s}_1$  is negative. The parameter setting for the numerical analysis is shown in Table 1.

[Table 1 about here]

### 3.2 Regional Labor Markets When Agglomeration Has No Externalities on Matching Efficiency

We first consider the benchmark case in which agglomeration has no externalities on the matching efficiency. Panel (a) of Figure 2 illustrates the regional differentials in PDVs of the employed and the unemployed for three cases of transport costs ( $T = 1.5, 1.6, 1.7$ ). When  $T = 1.7$ , we have three equilibria, two stable at  $s_1 = 0.04, 0.96$  and one unstable at  $s_1 = 0.50$ . When  $T = 1.6$ , we have two stable equilibria at  $s_1 = 0.11, 0.89$  and one unstable equilibrium at  $s_1 = 0.50$ . However, the stable equilibria shift inward. When  $T = 1.5$ , we have a unique and stable equilibrium at  $s_1 = 0.5$ .

Panel (b) of Figure 2 describes the unemployment differentials between regions 1 and 2 under the short-run equilibrium. When  $s_1 > 0.5$ , the unemployment rate in region 1 is always lower than that in region 2, where the relationship is robust under different values of transport costs. This result derives from the fact that the nominal wage in a denser region is always higher, resulting in a lower unemployment rate. In contrast, vom Berge (forthcoming) shows opposite results. This is because the nominal wage in a denser region is lower in the Krugman (1991) model.

Panel (c) of Figure 2 summarizes the spatial equilibria with respect to transport costs. The solid and dashed lines indicate stable and unstable equilibria respectively. A partial

agglomeration arises when the transport costs are high.<sup>24</sup> In our model, the break and sustain points coincide with each other. These points are at  $T = 1.53$  in Panel (c) of Figure 2. Contrary to our results, vom Berge (forthcoming) shows a full agglomeration when the transport costs are low.

[Figure 2 about here]

Following our numerical results, we discuss mainly the regional labor market outcomes in spatial equilibrium.<sup>25</sup> We assume that region 1 has at least half of the labor force ( $0.5 \leq s_1 < 1$ ). Figure 3 illustrates how the regional shares of the employed workers, unemployment rates, and labor market tightness vary depending on transport costs.

Panel (a) of Figure 3 shows that when transport costs are high, region 1 has a larger share of the employed than region 2. In such a case, we call region 1 an employment cluster, a core region, or an agglomerated region. Panel (b) of Figure 3 presents a lower unemployment rate in the employment cluster. From the negative relationship between unemployment rate and labor market tightness, as shown in Panel (c) of Figure 3, labor market tightness in the employment cluster takes a higher value than that in a less dense region, suggesting that the unemployed can easily find jobs, thus lowering the unemployment rate in an agglomerated region.

[Figure 3 about here]

### 3.3 Regional Labor Markets When Agglomeration Has Externalities on Matching Efficiency

We further explore three cases in which the agglomeration has externalities on the matching efficiency. Figures 4, 5, and 6 present the results of numerical analysis for the three cases,

<sup>24</sup>As shown in Pflüger and Tabuchi (2010), a full agglomeration is never a stable spatial equilibrium in a typical Helpman (1998) model. Intuitively, this is because if all the workers gather in one region, the price of land/housing services in the other region becomes zero. Consequently, workers have an incentive to move to the vacant region to enjoy higher utility; thus, a full agglomeration never arises.

<sup>25</sup>The figures for coefficient of variation of unemployment rates, labor market tightness  $\theta_i$  ( $i = 1, 2$ ), relative nominal wage  $w_1/w_2$ , relative cost-of-living index  $(G_1)^\mu(p_1^H)^{1-\mu}/(G_2)^\mu(p_2^H)^{1-\mu}$ , relative price index for manufactured goods  $G_1/G_2$ , and relative price of land/housing services  $p_1^H/p_2^H$  are available on the Web supplement file.

respectively. For ease of comparison with the benchmark case, each panel of the figures corresponds to respective panels of Figure 2 and Panel (b) of Figure 3.

First, Figure 4 presents the results of numerical analysis for the case in which the agglomeration has positive externalities on matching efficiency. We see that the spatial distribution of workers does not change qualitatively compared to the benchmark case. However, the positive agglomeration externalities on matching efficiency lower the dispersion force from congestion costs and widen the gap in unemployment rates. In both the short- and long-run, the unemployment rate in the employment cluster is relatively low.

Second, Figure 5 presents the results of numerical analysis for the case in which the agglomeration has negative externalities on matching efficiency, but the relationship is comparatively weak. In this case as well, the spatial distribution of workers does not change qualitatively compared to the benchmark case. The negative and weak agglomeration externalities on matching efficiency increase the dispersion force from congestion costs and narrow down the regional gap of unemployment rates partly. Note that the unemployment rate in the employment cluster becomes either lower or higher in the short-run depending on the degree of agglomeration ( $s_1$ ).

Third, Figure 6 presents the results of numerical analysis for the case in which the agglomeration has negative externalities on matching efficiency, but the relationship is comparatively strong. The negative and strong agglomeration externalities on matching efficiency increase the dispersion force from congestion costs and gradually widen the regional gap of unemployment rates above a certain degree of the negative relationship. The unemployment rate in the employment cluster is relatively high in the short- and long-run. Another important result is that the nominal wage in the employment cluster is always relatively high in all cases, which is consistent with the stylized facts of this literature.<sup>26</sup>

The theoretical predictions of this study are as follows. In the benchmark case in which agglomeration has no externalities on the matching efficiency, the unemployment rate in

---

<sup>26</sup>See the Web supplement file for numerical simulation results of the relative nominal wages in each case.

the employment cluster is relatively low so that agglomeration has a decreasing effect on unemployment rates in the production side.<sup>27</sup> However, agglomeration has a positive effect on regional unemployment rates in a search and matching process when agglomeration gives rise to negative externalities on the matching efficiency. When the negative agglomeration externalities on matching efficiency are comparatively weak, the unemployment rate in the employment cluster still remains partly low. When these externalities are comparatively strong, the unemployment rate in the employment cluster becomes higher.

Some predictions of our model are different from vom Berge (forthcoming), who incorporated a search and matching framework into Krugman's (1991) model. vom Berge (forthcoming) showed a positive relationship between regional unemployment rates and agglomeration through a negative relationship between nominal wages and agglomeration.<sup>28</sup> However, the latter relationship is clearly inconsistent with empirical evidence. The advantage of our model is that we describe a wide variety of relationships between regional unemployment rates and agglomeration, with the relationship between nominal wage and agglomeration positive. Consequently, our unifying framework contains aspects of both Harris and Todaro (1970) and Blanchflower and Oswald (1994). From our theoretical predictions, we empirically examine the relationship between unemployment rates and agglomeration, and between matching efficiency and agglomeration.

[Figures 4–6 about here]

---

<sup>27</sup>This result is essentially the same as Suedekum (2005) and Zierahn (forthcoming).

<sup>28</sup>This difference arises from the sector generating a dispersion force. Krugman's (1991) model deals with freely tradable agricultural goods, but the agricultural workers are not mobile. Helpman's (1998) model deals with the land/housing sector, whose services are consumed locally. Intuitively, in a Krugman-type model, a full agglomeration emerges and no manufacturing worker lives in the periphery region. Therefore, the unemployment rate in a periphery region is virtually zero. In other words, the nominal wage given by equation (21) can be defined even in regions with no manufacturer and is lower in agglomerated region; so the implicit unemployment rate also can be calculated. In contrast, in a Helpman-type model, there is a partial agglomeration, and so manufacturing workers always live in the periphery region. Therefore, higher nominal wage in the core region generates higher labor market tightness, leading to a further lower unemployment rate.

## 4 Empirical Analysis

### 4.1 Unemployment Rates and Agglomeration

First, we attempt to examine the relationship between regional unemployment rates and agglomerations. As a proxy for agglomeration, we use employment density. We use municipal data from Mexico for this analysis. More attention should be paid to spatial autocorrelation when spatially small regional units are used. In this case, the observations are closely related to each other. If the spatial dependence across observations is ignored, the estimators will be inconsistent or not efficient.<sup>29</sup> To solve this problem, we use spatial econometric methods. Thus, our regression models for unemployment rates are given by

$$\log(u_{i,t}) = \rho \sum_{j=1}^R b_{ij} \log(u_{j,t}) + \psi \log(\text{Dens}_{i,t}^s) + \mathbf{Z}_{i,t}^s \boldsymbol{\phi} + \varepsilon_{i,t}, \quad (35)$$

and

$$\log(u_{i,t}) = \psi \log(\text{Dens}_{i,t}^s) + \mathbf{Z}_{i,t}^s \boldsymbol{\phi} + e_{i,t}, \quad e_{i,t} = \lambda \sum_{j=1}^R b_{ij} e_{j,t} + \varepsilon_{i,t}, \quad (36)$$

where  $u_{i,t}$  is municipality  $i$ ' unemployment rate in year  $t$ ,  $b_{ij}$  is the  $ij$ th element of the spatial weight matrix (SWM),  $\psi$  is the key parameter of our interest,  $\text{Dens}_{i,t}^s$  is the log of spatially smoothed employment density,  $\mathbf{Z}_{i,t}^s$  is a row vector of spatially smoothed control variables,  $\boldsymbol{\phi}$  is a column vector of parameters for control variables, and  $e_{i,t}$  and  $\varepsilon_{i,t}$  are error terms. The control variables include the average years of schooling, rates of male and female labor force participation, and shares of the population aged 15–24, 25–59, and 60 and above.

Note that raw municipal data are not appropriate because the commuting that flows across municipal borders are not negligible at the municipality level and the local labor markets do not necessarily coincide with the administrative areas. Therefore, we use spatially smoothed municipal data in terms of the neighboring municipalities. See Section 5 for

---

<sup>29</sup>Regardless of endogeneity problem from employment density, OLS estimators are biased due to the omitted variable when  $\rho \neq 0$ . In addition, the covariance matrix of OLS estimators are no more efficient when  $\lambda \neq 0$ . See LeSage and Pace (2009) for detailed discussions

calculation of the spatially smoothed variables. To control for the endogeneity problem of employment density, we estimate equations (35) and (36) by using the method of instrumental variable (IV) and generalized method of moments (GMM). Our estimation methodology is based on Kelejian and Prucha (1998).

## 4.2 Matching Efficiency and Agglomeration

We furthermore examine agglomeration externalities on matching efficiency. The estimation procedure takes a two-step approach. In the first step, we estimate the regional matching efficiencies by estimating the matching function. From the logarithm of (31), the regression model to be estimated is given by

$$\log(\text{Match}_{s,t}) = \alpha_1 \log(\text{Seeker}_{s,t}) + \alpha_2 \log(\text{Vacancy}_{s,t}) + a_s + \text{year}_t + \varepsilon_{s,t}, \quad (37)$$

where  $\text{Match}_{s,t}$  is the number of matched jobs in state  $s$  at time  $t$ ,  $\text{Seeker}_{s,t}$  is the number of job seekers,  $\text{Vacancy}_{s,t}$  is the number of vacancies,  $\alpha_1$  and  $\alpha_2$  are the elasticities of matching,  $a_s = \log(A_s)$  is the state fixed effect,  $\text{year}_t$  is the year dummy, and  $\varepsilon_{s,t}$  is the error term. Note that our data set of job seeker, vacancy, and matched job is at the state level owing to the data limitations, and that subscript  $s$  is used instead of  $i$ . The state fixed effect  $a_s$  represents the regional differences in matching efficiency.<sup>30</sup> If we assume constant returns to scale in the matching function, then  $\alpha_2 = 1 - \alpha_1$ . In the estimation, we test the null hypothesis of constant returns to scale.

In the second step, the estimated matching efficiency is regressed on employment density as follows:

$$\hat{a}_s = \varphi + \xi \log(\text{Dens}_s) + \varepsilon_s, \quad (38)$$

where  $\hat{a}_s$  is the estimated matching efficiency,  $\varphi$  is the parameter for a constant term,  $\xi$  is a parameter of our interest, the elasticity of employment density to matching efficiency in equation (2),  $\text{Dens}_s$  is the employment density of state  $s$ , and  $\varepsilon_s$  is an error term. Therefore,

---

<sup>30</sup>The state fixed effects are estimated by `areg` command in Stata.

we examine the relationship between matching efficiency and agglomeration by inspecting the coefficient estimate of  $\text{Dens}_s$ .

## 5 Data

### 5.1 Unemployment Rates and Agglomeration

We use the 2000 and 2010 Mexican population censuses.<sup>31</sup> From the censuses, the National System of Municipal Information (*Sistema Nacional de Información Municipal*, SNIM) provides its summarized municipal data on area, labor force (the employed and unemployed), average years of schooling, labor force participation rate by gender, and the population aged 15–24, 25–59, and 60 and above.<sup>32</sup>

We construct our data set as follows. The unemployment rate of municipality  $i$  is calculated by the ratio of the employed to the labor force living in the municipality. Let  $z_{i,t}^s$  denote the spatially local sum data of municipality  $i$  in year  $t$ , calculated as  $z_{i,t}^s = \sum_{j=1}^R \mathbf{1}_{ij}(d) z_{j,t}$ , where  $R$  stands for the number of municipalities,  $z_{j,t}$  the raw data of municipality  $j$ , and  $\mathbf{1}_{ij}(d)$  the  $ij$ th element of the indicator matrix, in which the  $ij$ th element takes the value of 1 if the distance between municipalities  $i$  and  $j$  is less than  $d$ km and 0 otherwise.<sup>33</sup> We set  $d = 40$ km. Thus, the spatially smoothed employment density is  $\text{Dens}_{i,t}^s = \text{Emp}_{i,t}^s / \text{Area}_{i,t}^s$ , where  $\text{Emp}_{i,t}^s$  and  $\text{Area}_{i,t}^s$  are spatially local sum of employed worker and area, respectively, of municipality  $i$  in year  $t$ . Further, the other variables are also calculated using the same method.<sup>34</sup> We drop the lowermost 1% and the uppermost 99% of the distribution of unemployment rates.<sup>35</sup> We use the spatially smoothed employment density of 1990 for IV, and so

<sup>31</sup>In population censuses, labor data are available for every ten years. The 1990 population census data are also used for instrumental variables. We exclude Nicolás Ruíz in the state of Chiapas from the 2000 data owing to lack of labor data. Furthermore, we found some municipalities were originally lacking in the 2000 population census data.

<sup>32</sup>The data are available at the following Web site (URL: <http://www.snim.rami.gob.mx/>).

<sup>33</sup>SNIM also offers the latitude and longitude data of municipalities, from which the bilateral distances between any two municipalities can be calculated by using the formula of Vincenty (1975).

<sup>34</sup>The average years of schooling is calculated as the spatially local sum of years of schooling divided by the number of municipalities within a radius of  $d$ km from municipality  $i$ .

<sup>35</sup>Observations of zero are excluded because they are included in the lowermost 1 percent. The municipality

use the 1990 population census as well.<sup>36</sup> Table 2 gives the descriptive statistics of municipal data by year.

[Table 2 about here]

For our estimation, we use distance-based SWMs, which take the following form:

$$b_{ij} = \frac{d_{ij}^{-\eta}}{\sum_{j=1}^R d_{ij}^{-\eta}}$$

where  $b_{ij}$  is the  $ij$ th element of an SWM,  $d_{ij}$  is the bilateral distance between municipalities  $i$  and  $j$ ,  $R$  is the number of municipalities, and  $\eta$  is a distance decay parameter. The bilateral distance is calculated as the great-circle distance between two municipalities measured by latitude and longitude (Vincenty, 1975). The SWMs are row-standardized. In this paper, our estimation results are obtained from using distance-based SWMs ( $\eta = 5$ ).<sup>37</sup>

## 5.2 Matching Efficiency and Agglomeration

The yearly job seeker, vacancy, and matched job data are available from the Secretariat of Labor and Social Welfare (*Secretaría del Trabajo y Previsión Social*, STPS). The time span is from 2001 to 2011. The STPS offers services for the promotion of job matching in job placement offices (*Bolsa de Trabajo*). The data include the number of applications registered both for the first time and on subsequent occasions, the number of job vacancies, and the number of matched jobs out of the vacant jobs registered.<sup>38</sup> Table 3 presents the descriptive statistics of job seeker, vacancy, and matched job by year.

[Table 3 about here]

We then calculate the employment density at the state level. For this, we use the 2000 population census. In the regression analysis at the second step, the dependent variable is of *Nicolás Ruíz* located in the state of Chiapas is also excluded owing to lack of data.

<sup>36</sup>There is no information of municipal area in 1990 population census. Therefore, we complement municipal areas in 1990 with the 2000 population census. In that case, separated municipalities between 1990 and 2000 are added to original municipalities.

<sup>37</sup>Our main results do not change even if different values of  $\eta$  are used.

<sup>38</sup>A person can be hired once more depending on the type of employment (casual, temporary, or permanent).



the matching efficiency by state estimated between 2001 and 2011. To avoid endogeneity issues, we use the employment density of 2000. For robustness, we also use the employment density of 1990 as an instrumental variable. A problem with employment density at the state level is that some states have vast uninhabitable regions, leading to underestimated employment densities. To mitigate this issue, we calculate the employment density as follows. The municipal employment density is first simply calculated and sorted by size. Then, the number of the employed in municipalities and municipal areas are summed up respectively until the share of the employed by state reaches 80%. Finally, the state employment density is calculated as the employed–area ratio.

## 6 Empirical Results

### 6.1 Unemployment Rates and Agglomeration

Table 4 shows the estimation results for equations (35) and (36). Columns (1) and (4) of Table 4 presents the ordinary least squares (OLS) estimates for 2000 and 2010, respectively. In Column (1), employment density has a significantly negative impact on unemployment rates at the 5% level, but is insignificant even at the 10% level in 2010. According to the robust LM tests for spatial dependence in the dependent variable and error terms, the null hypotheses  $\rho = 0$  and  $\lambda = 0$  are rejected at least at the 5% level in 2000 and 2010, respectively, and we need to control for spatial dependence.<sup>39</sup> The estimation results for 2000 and 2010 are given in Columns (2) and (3) and Columns (5) and (6), respectively. As expected, the parameter estimates measuring spatial dependence in the dependent variable and error terms are significantly positive in both years. The coefficient estimates of employment density remain significantly negative even after controlling for spatial dependence in 2000. However, employment density is no longer significant in 2010.

---

<sup>39</sup>We follow the hypothesis-testing methodology for spatial dependence proposed by Anselin et al. (1996). See also Anselin (2006) for a brief summary.

For robustness, we control for the endogeneity of employment density. Table 5 presents the IV/GMM estimation results. In Columns (1) and (4), we control for the endogeneity of employment density but do not control for spatial dependence in the dependent variable and error terms. The Dubin–Wu–Hausman test shows that there exists an endogeneity problem in regression model. From IV/GMM estimation, we find the coefficient estimates of employment density significantly negative in 2000 and 2010. However, Robust LM Tests suggest that spatial dependence should be controlled for. As earlier, Columns (2) and (3) and Columns (4) and (5) show, respectively, the estimation results when spatial dependence in the dependent variable and error terms are controlled for. In 2000, employment density shows a significantly negative impact on unemployment rate. However, this is not the case in 2010.

[Tables 4 and 5 about here]

Our evidence on the negative relationship between unemployment rates and agglomeration is robust for 2000, but not for 2010. When based exactly on our model, the unemployment differentials decreased as the transport costs fell, and the magnitude of the estimated coefficient became smaller and the statistical significance might not be confirmed. Another important implication is that the negative relationship between unemployment and agglomeration can be observed only when agglomeration has positive or weakly negative externalities on the matching efficiency. In the next subsection, we examine the relationship between matching efficiency and agglomeration.

## 6.2 Matching Efficiency and Agglomeration

Table 6 presents the estimation results of regression models (37) and (38). Column (1) shows the estimation results of the matching function. The elasticity of job match to job seeker is significant at the 5% level and takes the value of 0.33. The elasticity of job match to vacancy is also significant at the 1% level and takes the value of 0.71. The null hypothesis of constant returns to scale for the matching function is not rejected. Our estimates are consistent with

the results of most of the empirical studies on the matching function. According to a survey of Petrongolo and Pissarides (2001), the estimate of plausible elasticity on job seeker lies in the range between 0.5 and 0.7.

Column (2) shows the estimation results of the agglomeration effect on matching efficiency. The elasticity of employment density estimated by OLS is significantly negative at the 1% level. The elasticity of employment density estimated by GMM is also significantly negative. Figure 7 clearly illustrates the negative relationship between matching efficiency and employment density.

[Table 6 and Figure 7 about here]

To sum up, the Mexican data examined show comparatively low unemployment rates as well as matching efficiency in agglomerated regions. Taking into account the theoretical prediction that when agglomeration has negative externalities on matching efficiency it will have both positive and negative effects on regional unemployment rates, the agglomeration effect lowering the unemployment rates in Mexico is much stronger than that increasing the unemployment rates.

## 7 Concluding Remarks

In this paper, we theoretically and empirically analyzed the relationship between regional unemployment rates and agglomeration. In the theoretical part of our analysis, we extended a multi-region model of Helpman (1998) by incorporating search and matching frictions in regional labor markets. In addition, we incorporate agglomeration externalities into a search and matching framework. In the empirical part of our analysis, we examined the relationship between regional unemployment rates and agglomeration (expressed in employment density) by using Mexican municipal data. We also estimated the matching function by using the data of job seekers, vacancies, and matched jobs.

An important prediction of our theory is that agglomeration can be positively or negatively related with regional unemployment rates under negative agglomeration externalities

on matching efficiency. Thus, our theoretical framework with agglomeration externalities on matching efficiency can describe a wide variety of relationships between regional unemployment rates and agglomeration, with the relationship between nominal wages and agglomeration positive, as supported by most empirical studies. Therefore, our model can lead to predictions on unemployment rates and wages of both Harris and Todaro (1970) and Blanchflower and Oswald (1994) within a unified framework.

From our empirical results obtained with Mexican data, we found that denser areas have comparatively low unemployment rates under negative agglomeration externalities on matching efficiency. Considering our theoretical predictions, we conclude that in Mexico, the agglomeration effect lowering the unemployment rates is much stronger than that increasing the rates.

## References

- [1] Anselin, Luc (2006) “Spatial Econometrics,” in Mills, Terence C. and Kerry Patterson eds. *Palgrave Handbook of Econometrics* Vol. 1 Econometric Theory, Basingstoke: Palgrave Macmillan, Chap. 26, pp. 901–969.
- [2] Anselin, Luc, Anil K. Bera, Raymond Florax, and Mann J. Yoon (1996) “Simple diagnostic tests for spatial dependence,” *Regional Science and Urban Economics* 26(1), pp. 77–104.
- [3] Blanchflower, David G. and Andrew J. Oswald (1994) *The Wage Curve*, Cambridge, MA: MIT Press.
- [4] Card, David (1995) “The wage curve: A review,” *Journal of Economic Literature* 33(2), pp. 785–799.
- [5] Chiang, Shu-hen (2009) “The effects of industrial diversification on regional unemployment in Taiwan: Is the portfolio theory applicable?” *Annals of Regional Science* 4(43), pp. 947–962.

- [6] Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon (2008a) “Spatial wage disparities: Sorting matters!,” *Journal of Urban Economics* 63(2t), pp. 723–742.
- [7] Combes, Pierre-Philippe, Thierry Mayer, and Jacques-François Thisse (2008b) *Economic Geography: The Integration of Regions and Nations*, New Jersey: Princeton University Press.
- [8] Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux (2010) “Estimating agglomeration economies with history, geology, and worker effects,” in Glaeser, Edward L. ed. *Agglomeration Economics*: University of Chicago Press, Chap. 1, pp. 15–66.
- [9] de la Roca, Jorge and Diego Puga (2012) “Learning by working in big cities,” CEPR Discussion Papers No. 9243.
- [10] Doornik, Jurgen A. and Marius Ooms (2006) *Introduction to Ox*, London: Timberlake Consultants Press. (URL: <http://www.doornik.com/>).
- [11] Epifani, Paolo and Gino A. Gancia (2005) “Trade, migration and regional unemployment,” *Regional Science and Urban Economics* 35(6), pp. 625–644.
- [12] Francis, John (2009) “Agglomeration, job flows and unemployment,” *Annals of Regional Science* 43(1), pp. 181–198.
- [13] Fujita, Masahisa, Paul Krugman, and Anthony J. Venables (1999) *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, MA: MIT Press.
- [14] Hanson, Gordon H. (1998) “North American economic integration and industry location,” *Oxford Review of Economic Policy* 14(2), pp. 30–44.
- [15] ——— (2005) “Market potential, increasing returns and geographic concentration,” *Journal of International Economics* 67(1), pp. 1–24.
- [16] Harris, John R. and Michael P. Todaro (1970) “Migration, unemployment and development: a two-sector analysis,” *American Economic Review* 60(1), pp. 126–142.
- [17] Helpman, Elhanan (1998) “The size of regions,” in Pines, David, Efraim Sadka, and Itzhak Zilcha eds. *Topics in Public Economics: Theoretical and Applied Analysis*, Cam-

- bridge: Cambridge University Press, Chap. 2, pp. 33–54.
- [18] Helpman, Elhanan and Oleg Itskhoki (2010) “Labour market rigidities, trade and unemployment,” *Review of Economic Studies* 77(3), pp. 1100–1137.
- [19] Hering, Laura and Sandra Poncet (2010) “Market access impact on individual wages: Evidence from China,” *Review of Economics and Statistics* 92(1), pp. 145–159.
- [20] Hynninen, Sanna-Mari and Jukka Lahtonen (2007) “Does population density matter in the process of matching heterogeneous job seekers and vacancies?” *Empirica* 34(5), pp. 397–410.
- [21] Izraeli, Oded and Kevin J. Murphy (2003) “The effect of industrial diversity on state unemployment rate and per capita income,” *Annals of Regional Science* 37(1), pp. 1–14.
- [22] Kano, Shigeki and Makoto Ohta (2005) “Estimating a matching function and regional matching efficiencies: Japanese panel data for 1973–1999,” *Japan and the World Economy* 17(1), pp. 25–41.
- [23] Kelejian, Harry H. and Ingmar R. Prucha (1998) “A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances,” *Journal of Real Estate Finance and Economics* 17(1), pp. 99–121.
- [24] Krugman, Paul (1991) “Increasing returns and economic geography,” *Journal of Political Economy* 99(3), pp. 483–499.
- [25] Krugman, Paul and Raul Livas-Elizondo (1996) “Trade policy and the Third World metropolis,” *Journal of Development Economics* 49(1), pp. 137–150.
- [26] LeSage, James P. and R. Kelley Pace (2009) *Introduction to Spatial Econometrics*, Boca Raton: CRC Press.
- [27] Marshall, Alfred (1890) *Principles of Economics*, London: Macmillan.
- [28] Mion, Giordano and Paolo Naticchioni (2009) “The spatial sorting and matching of skills and firms,” *Canadian Journal of Economics* 42(1), pp. 28–55.
- [29] Petrongolo, Barbara and Christopher A. Pissarides (2001) “Looking into the black

- box: A survey of the matching function,” *Journal of Economic Literature* 39(2), pp. 390–431.
- [30] Pflüger, Michael and Takatoshi Tabuchi (2010) “The size of regions with land use for production,” *Regional Science and Urban Economics* 40(6), pp. 481–489.
- [31] Pissarides, Christopher A. (2000) *Equilibrium Unemployment Theory*, Cambridge, MA: MIT Press, 2nd edition.
- [32] Redding, Stephen and Anthony J. Venables (2004) “Economic geography and international inequality,” *Journal of International Economics* 62(1), pp. 53–82.
- [33] Rogerson, Richard, Robert Shimer, and Randall Wright (2005) “Search-theoretic models of the labor market: A survey,” *Journal of Economic Literature* 43(4), pp. 959–988.
- [34] Rosenthal, Stuart S. and William C. Strange (2001) “The determinants of agglomeration,” *Journal of Urban Economics* 50(2), pp. 191–229.
- [35] Sato, Yasuhiro (2000) “Search theory and the wage curve,” *Economics Letters* 66(1), pp. 93–98.
- [36] Simon, Curtis J. (1988) “Frictional unemployment and the role of industrial diversity,” *Quarterly Journal of Economics* 103(4), pp. 715–728.
- [37] Stole, Lars A. and Jeffrey Zwiebel (1996a) “Intra-firm bargaining under non-binding contracts,” *Review of Economic Studies* 63(3), pp. 375–410.
- [38] ——— (1996b) “Organizational design and technology choice under intrafirm bargaining,” *American Economic Review* 86(1), pp. 195–222.
- [39] Suedekum, Jens (2005) “Increasing returns and spatial unemployment disparities,” *Papers in Regional Science* 84(2), pp. 159–181.
- [40] Vincenty, Thaddeus (1975) “Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations,” *Survey Review* 23(176), pp. 88–93.
- [41] vom Berge, Philipp (forthcoming) “Search unemployment and new economic geography,” *Annals of Regional Science* Online First.
- [42] Zierahn, Ulrich Theodor (forthcoming) “Agglomeration, congestion, and regional un-

employment disparities,” *Annals of Regional Science* Online First.

- [43] Zieseimer, Thomas (2005) “Monopolistic competition and search unemployment,” *Metroeconomica* 56(3), pp. 334–359.

## Appendix A Derivation of Land Rent

The aggregate income of all regions is equal to the sum of their disposable labor income and income obtained from land/housing services:

$$\sum_{j=1}^R Y_j = \sum_{j=1}^R [(w_i - \tau)(1 - u_i)L_i + (z - \tau)u_i L_i] + (1 - \mu) \sum_{j=1}^R Y_j.$$

Thus, the aggregate income from land/housing services in the economy becomes

$$(1 - \mu) \sum_{j=1}^R Y_j = \frac{1 - \mu}{\mu} \sum_{j=1}^R [(w_i - \tau)(1 - u_i)L_i + (z - \tau)u_i L_i]$$

Dividing this by the share of regional labor force, the aggregate land rent in region  $i$  becomes

$$\frac{L_i}{\sum_{j=1}^R L_j} (1 - \mu) \sum_{j=1}^R Y_j = \frac{L_i}{\sum_{j=1}^R L_j} \frac{1 - \mu}{\mu} \sum_{j=1}^R [(w_i - \tau)(1 - u_i)L_i + (z - \tau)u_i L_i].$$

Furthermore, dividing this by the workers living in region  $i$ , the land rent that individuals receive becomes

$$h = \frac{1}{\sum_{j=1}^R L_j} (1 - \mu) \sum_{j=1}^R Y_j = \frac{1}{\sum_{j=1}^R L_j} \frac{1 - \mu}{\mu} \sum_{j=1}^R [(w_i - \tau)(1 - u_i)L_i + (z - \tau)u_i L_i].$$

See Helpman (1998) for more details.



Table 1: Parameter Setting for Numerical Analysis

Parameter	Explanation
$1 \leq T \leq 2$	Transport Cost
$\sigma = 6$	Elasticity of Substitution between Varieties
$\mu = 0.86$	Expenditure Share for Manufactured Goods
$\delta = 0.03$	Job Destruction Rate ( $i = 1, 2$ )
$\gamma = 0.5$	Marginal Labor Input for Recruiter per Vacancy
$\beta = 0.5$	Bargaining Power of Worker
$\bar{S}_i = 1$	Land Endowment ( $i = 1, 2$ )
$r = 0.01$	Interest Rate
$z = 0.4$	Unemployment Benefit
$L_1 + L_2 = 1$	Total Labor Force
$A = 0.6$	Constant of Matching Efficiency
$\alpha = 0.5$	Matching Elasticity on Job Seekers
$\xi = 0$	Elasticity of Agglomeration to Matching Efficiency (Benchmark)
$\xi = 0.02$	Elasticity of Agglomeration to Matching Efficiency (Positive)
$\xi = -0.02$	Elasticity of Agglomeration to Matching Efficiency (Negative and Weak)
$\xi = -0.06$	Elasticity of Agglomeration to Matching Efficiency (Negative and Strong)

Notes: The matching function is  $A_i m(u_i L_i, v_i L_i) = A_i (u_i L_i)^\alpha (v_i L_i)^{1-\alpha}$ , where  $A_i = A(L_i/\bar{S})^\xi$ . We set  $w_i = 1$ , ( $i = 1, 2$ ) as the initial value for the derivation of short- and long-run equilibria.

Table 2: Descriptive Statistics for Unemployment Analysis

Variable	Mean	Std. Dev.	Min	Max
<i>Year 2000</i>				
Unemployment Rate (%)	0.953	0.608	0.031	3.815
Employment Density (person/km <sup>2</sup> )	67.439	180.959	0.052	1386.694
Employment Density (person/km <sup>2</sup> ) in 1990	46.685	130.523	0.068	1020.300
Years of Schooling	5.473	1.178	2.910	9.140
Male Labor Force Participation Rate (%)	68.691	6.173	30.764	84.692
Female Labor Force Participation Rate (%)	25.163	6.537	6.456	40.714
Share of Population Aged 15–24 (%)	19.137	1.337	13.225	23.186
Share of Population Aged 25–59 (%)	35.143	3.740	22.531	43.705
Share of Population Aged 60 and above (%)	8.278	2.102	2.589	16.767
<i>Year 2010</i>				
Unemployment Rate (%)	4.037	2.626	0.067	16.266
Employment Density (person/km <sup>2</sup> )	79.295	203.179	0.056	1572.537
Employment Density (person/km <sup>2</sup> ) in 1990	44.951	127.496	0.068	1020.300
Years of Schooling	6.689	1.149	4.081	10.360
Male Labor Force Participation Rate (%)	72.794	3.605	45.476	84.537
Female Labor Force Participation Rate (%)	27.180	8.085	4.212	48.474
Share of Population Aged 15–24 (%)	18.820	1.076	13.674	22.312
Share of Population Aged 25–59 (%)	39.422	3.661	27.211	46.881
Share of Population Aged 60 and above (%)	10.414	2.598	3.278	22.720

Notes: The numbers of observations in 2000 and 2010 are 2255 and 2387, respectively. The lowermost 1% and uppermost 99% of the distribution of unemployment rates are dropped. These municipal data are spatially smoothed except for unemployment rates.

Table 3: Descriptive Statistics of Matched Jobs, Job Seekers, and Vacancies

Variables	Year										
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
<i>Matching:</i>											
Mean	5419.7	6579.3	6434.8	7155.2	7212.6	7441.5	8560.1	9479.8	8549.4	8270.0	7968.5
Std. Dev.	10087.8	11791.7	10904.6	11959.3	10897.4	10893.0	11703.5	16663.9	14330.2	11337.9	11724.0
Min	607.0	748.0	790.0	1416.0	1769.0	957.0	1503.0	852.0	973.0	1232.0	920.0
Max	59102.0	68554.0	63704.0	69903.0	62221.0	61065.0	63541.0	91826.0	74760.0	49139.0	55518.0
<i>Job Seeker:</i>											
Mean	18199.6	20638.7	21856.6	23665.6	24071.5	24606.9	27304.4	33273.2	37362.8	32245.5	29782.7
Std. Dev.	24722.6	27026.4	28682.3	34722.2	30762.6	33742.2	35747.7	43188.4	47648.0	31384.3	28800.9
Min	4017.0	4200.0	4479.0	4970.0	4473.0	4981.0	5130.0	7170.0	6396.0	5715.0	3575.0
Max	142235.0	154099.0	162969.0	199075.0	173291.0	192490.0	196705.0	240831.0	254189.0	161809.0	146593.0
<i>Vacancy:</i>											
Mean	14147.8	14331.6	14174.0	15758.7	16986.2	18302.4	19405.4	24822.8	25389.1	22295.9	20064.7
Std. Dev.	31392.9	35818.6	32442.8	39032.4	37674.0	41262.8	39293.2	46842.2	41436.4	28689.2	31163.5
Min	1177.0	1415.0	1359.0	1315.0	1753.0	1462.0	1695.0	2987.0	4004.0	3727.0	3021.0
Max	178649.0	207282.0	187916.0	225711.0	217825.0	236441.0	218023.0	256896.0	219377.0	134007.0	136153.0

Notes: The number of states in each year is 32.

Table 4: Estimation Results for Regional Unemployment Rates and Agglomeration

Explanatory Variable	Dependent Variable: $\log(u_{i,t})$					
	Year: 2000			Year: 2010		
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Employment Density</b>						
Years of Schooling	-0.058** (0.023)	-0.032*** (0.011)	-0.051* (0.028)	-0.035 (0.022)	0.006 (0.015)	-0.035 (0.028)
Labor Force Participation Rate (Male)	0.201*** (0.032)	0.154*** (0.021)	0.197*** (0.034)	-0.038 (0.034)	-0.029 (0.025)	-0.023 (0.039)
Labor Force Participation Rate (Female)	-2.149*** (0.247)	-1.288*** (0.205)	-2.005*** (0.248)	-2.757*** (0.461)	-1.201*** (0.384)	-2.303*** (0.484)
Share of Population Aged 15-24	-0.185* (0.100)	-0.081 (0.076)	-0.195* (0.102)	-0.180* (0.101)	-0.128* (0.076)	-0.132 (0.104)
Share of Population Aged 25-59	0.747** (0.362)	0.383 (0.246)	0.753** (0.370)	1.033** (0.484)	0.732* (0.382)	0.920* (0.528)
Share of Population Aged 60 and above	1.347*** (0.341)	0.557** (0.273)	1.287*** (0.357)	2.570*** (0.495)	1.398*** (0.392)	2.257*** (0.523)
Constant	-0.200* (0.106)	-0.090 (0.080)	-0.156 (0.113)	-0.316*** (0.112)	-0.079 (0.086)	-0.278** (0.130)
<b>Spatially Lagged Dependent Variable (<math>\rho</math>)</b>	2.027 (1.519)	1.916 (1.251)	1.568 (1.630)	3.080 (2.648)	-0.259 (2.190)	2.271 (2.765)
<b>Spatially Lagged Error Terms (<math>\lambda</math>)</b>		0.365*** (0.047)	0.153*** (0.020)		0.509*** (0.052)	
State Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	2255	2255	2255	2387	2387	2387
Adjusted $R^2$	0.243			0.253		
Robust LM Test ( $\rho$ ), $p$ -value	0.000			0.000		
Robust LM Test ( $\lambda$ ), $p$ -value	0.008			0.031		

Notes: Standard errors are in the parenthesis. Columns (1), (2), (4), and (5) consider heteroskedastic errors. The explanatory variables are expressed in logarithm except years of schooling. Spatially smoothed municipal data are used. The instrumental variable for spatially lagged dependent variable is  $B\bar{Z}, \dots, B^5\bar{Z}$ , where  $B$  is the spatial weight matrix and  $\bar{Z}$  a matrix consisting of employment density and control variables. Robust LM Test ( $\rho$ ) indicates the testing of the null hypothesis  $\rho = 0$  against alternative hypothesis  $\rho \neq 0$ . Robust LM Test ( $\lambda$ ) indicates the testing of the null hypothesis  $\lambda = 0$  against alternative hypothesis  $\lambda \neq 0$ . \*, \*\*, and \*\*\* denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 5: Robustness Checks for Endogeneity of Employment Density

Explanatory Variable	Dependent Variable: $\log(u_{i,t})$					
	Year: 2000			Year: 2010		
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Employment Density</b>	-0.075*** (0.024)	-0.048*** (0.013)	-0.072** (0.029)	-0.047** (0.023)	0.006 (0.016)	-0.042 (0.029)
Years of Schooling	0.204*** (0.032)	0.161*** (0.020)	0.199*** (0.034)	-0.035 (0.034)	-0.027 (0.025)	-0.026 (0.039)
Labor Force Participation Rate (Male)	-2.160*** (0.245)	-1.371*** (0.201)	-2.093*** (0.247)	-2.770*** (0.458)	-1.166*** (0.384)	-2.505*** (0.485)
Labor Force Participation Rate (Female)	-0.159 (0.099)	-0.071 (0.075)	-0.163 (0.103)	-0.165* (0.100)	-0.125* (0.076)	-0.151 (0.104)
Share of Population Aged 15-24	0.771** (0.359)	0.456** (0.232)	0.751** (0.368)	1.000** (0.479)	0.751** (0.383)	0.940* (0.529)
Share of Population Aged 25-59	1.408*** (0.340)	0.670** (0.270)	1.414*** (0.356)	2.600*** (0.493)	1.359*** (0.398)	2.409*** (0.524)
Share of Population Aged 60 and above	-0.222** (0.105)	-0.119 (0.079)	-0.203* (0.113)	-0.344*** (0.113)	-0.073 (0.087)	-0.324** (0.131)
Constant	1.783 (1.516)	1.671 (1.237)	1.541 (1.624)	3.160 (2.628)	-0.376 (2.196)	2.733 (2.765)
<b>Spatially Lagged Dependent Variable (<math>\rho</math>)</b>		0.332*** (0.043)			0.516*** (0.051)	
<b>Spatially Lagged Error Terms (<math>\lambda</math>)</b>			0.153*** (0.020)			0.245*** (0.019)
State Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	2255	2255	2255	2387	2387	2387
Robust LM Test ( $\rho$ ), $p$ -value	0.001			0.000		
Robust LM Test ( $\lambda$ ), $p$ -value	0.018			0.049		
Dubin-Wu-Hausman Test, $p$ -value	0.013			0.103		

Notes: Standard errors are in parenthesis. Columns (1), (2), (4), and (5) consider heteroskedastic errors. The explanatory variables are expressed in logarithm except years of schooling. Spatially smoothed municipal data are used. The instrumental variable for employment density is the spatially smoothed 10-year lagged employment density. The instrumental variable for spatially lagged dependent variable is  $BZ, \dots, B^5Z$ , where  $B$  is the spatial weight matrix and  $Z$  a matrix consisting of the spatially smoothed employment density in 1990 and control variables. Robust LM Test ( $\rho$ ) indicates the testing of null hypothesis  $\rho = 0$  against alternative hypothesis  $\rho \neq 0$ . Robust LM Test ( $\lambda$ ) indicates the testing of the null hypothesis  $\lambda = 0$  against alternative hypothesis  $\lambda \neq 0$ . Dubin-Wu-Hausman Test indicates the hypothesis testing of endogeneity. \*, \*\*, and \*\*\* denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 6: Estimation Results for Matching Function

Explanatory Variable	(1)	(2)	(3)
	Dependent Variable		
	$\log(\text{Match}_{s,t})$	State Fixed Effects in Column (1) (OLS)	State Fixed Effects in Column (1) (GMM)
<b>Log of Employment Density</b>		−0.086*** (0.026)	−0.081*** (0.027)
<b>State Fixed Effects</b>	Yes		
Log of Job Seeker	0.332** (0.150)		
Log of Vacancy	0.714*** (0.161)		
Year Dummy	Yes		
Number of Observations	352	32	32
Adjusted $R^2$	0.886	0.200	
CRS Test, $p$ -value	0.700		
Dubin-Wu-Hausman Test, $p$ -value			0.295
$F$ -Statistic (Weak IV)			1536.740

Notes: Heteroskedasticity-consistent standard errors are in parenthesis. Column (1) gives heteroskedasticity-consistent standard errors clustered by state. All regressions contain a constant term. The instrumental variable for employment density shown in Column (3) is the employment density in 1990. CRS Test indicates the hypothesis testing of constant returns to scale for the matching function. Dubin-Wu-Hausman Test indicates hypothesis testing of endogeneity.  $F$  Statistic (Weak IV) is Cragg–Donald Wald  $F$  statistic for test of weak instruments. \*, \*\*, and \*\*\* denote statistical significance at the 1%, 5%, and 10% level, respectively.

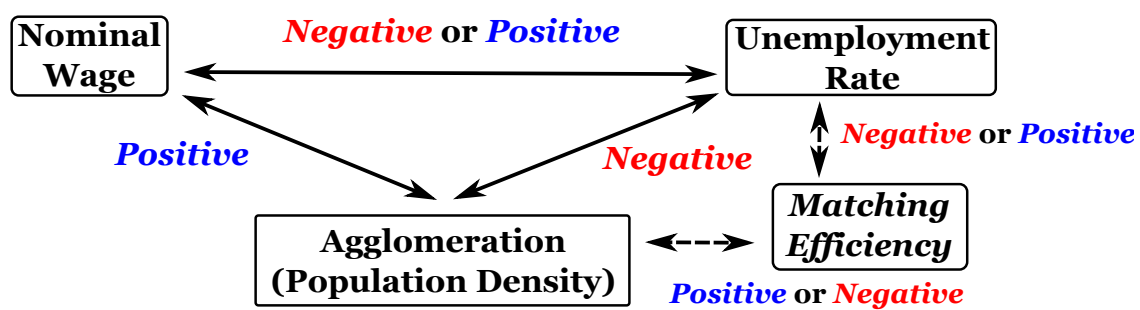
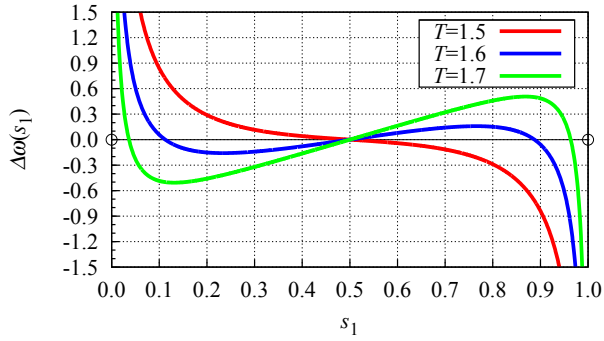
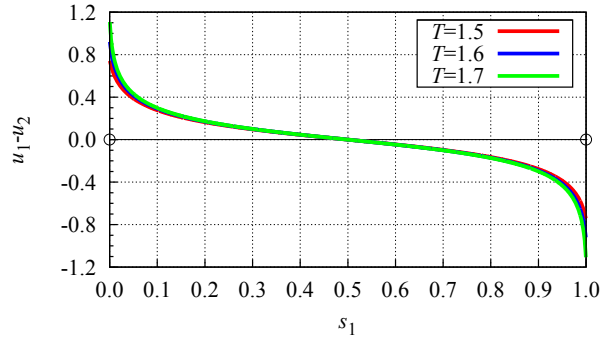


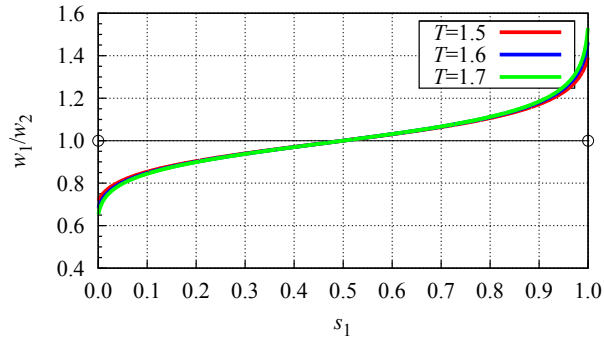
Figure 1: Relationship between Wage, Unemployment Rate, and Agglomeration



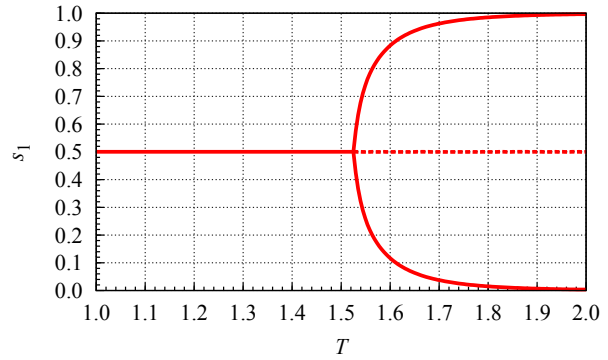
(a) Dynamics



(b) Unemployment Differentials



(c) Relative Nominal Wage

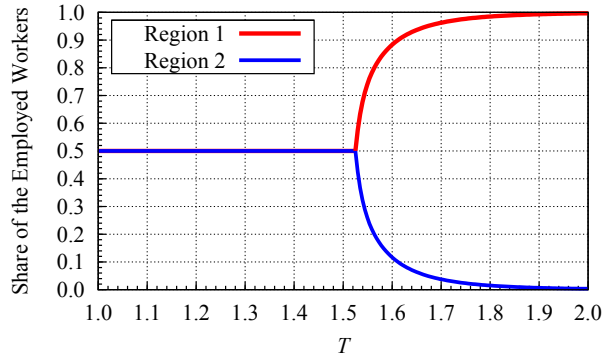


(d) Spatial Equilibrium

Figure 2: Numerical Analysis Results When Agglomeration Has No Externalities on Matching Efficiency

Notes: The solid and dashed lines in Panel (d) denote stable and unstable equilibria, respectively. The parameters used in this numerical analysis are shown in Table 1.





(a) Regional Share of the Employed Workers

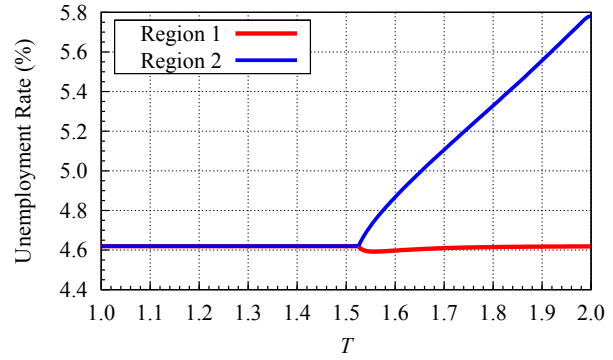
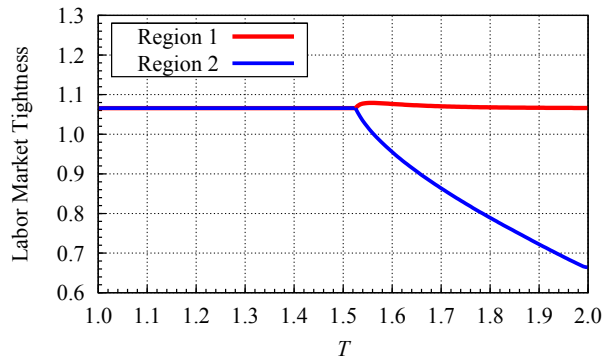
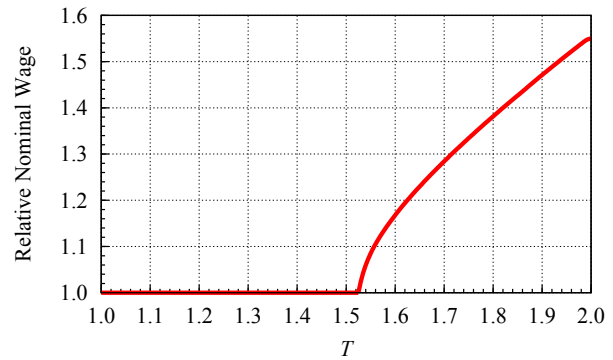
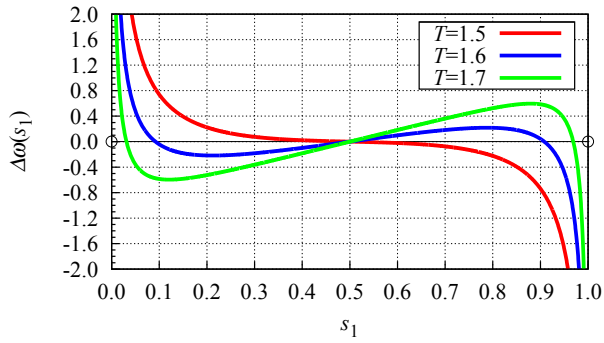
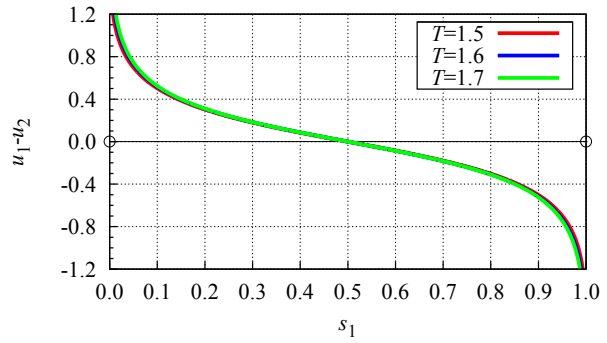
(b) Unemployment Rate,  $u_i$ (c) Labor Market Tightness,  $\theta_i$ (d) Relative Nominal Wage,  $w_1/w_2$ 

Figure 3: Numerical Simulation in Spatial Equilibrium When Agglomeration Has No Externalities on Matching Efficiency

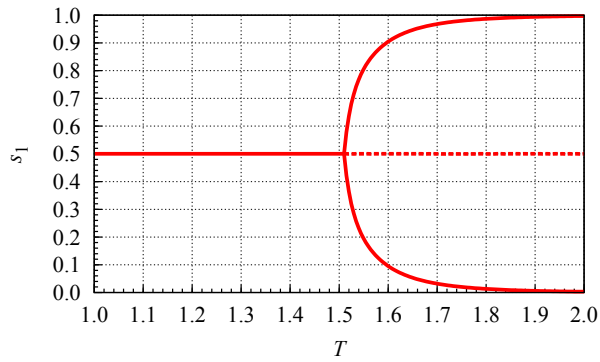
Notes: The parameters used in this numerical analysis are in shown Table 1.



(a) Dynamics



(b) Unemployment Differentials



(c) Spatial Equilibrium

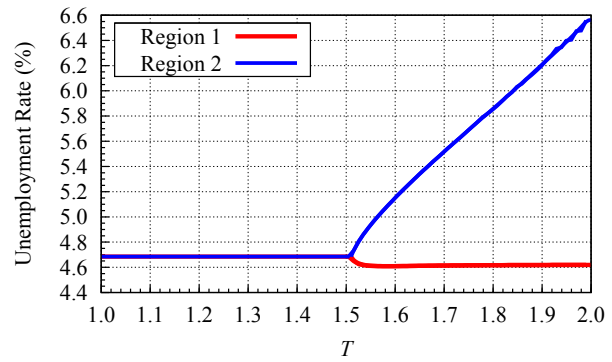
(d) Unemployment Rate,  $u_i$ 

Figure 4: Numerical Simulation Results When Agglomeration Has Positive Externalities on Matching Efficiency

Notes: The solid and dashed lines in Panel (c) denote stable and unstable equilibria, respectively. The parameters used in this numerical analysis are shown in Table 1.

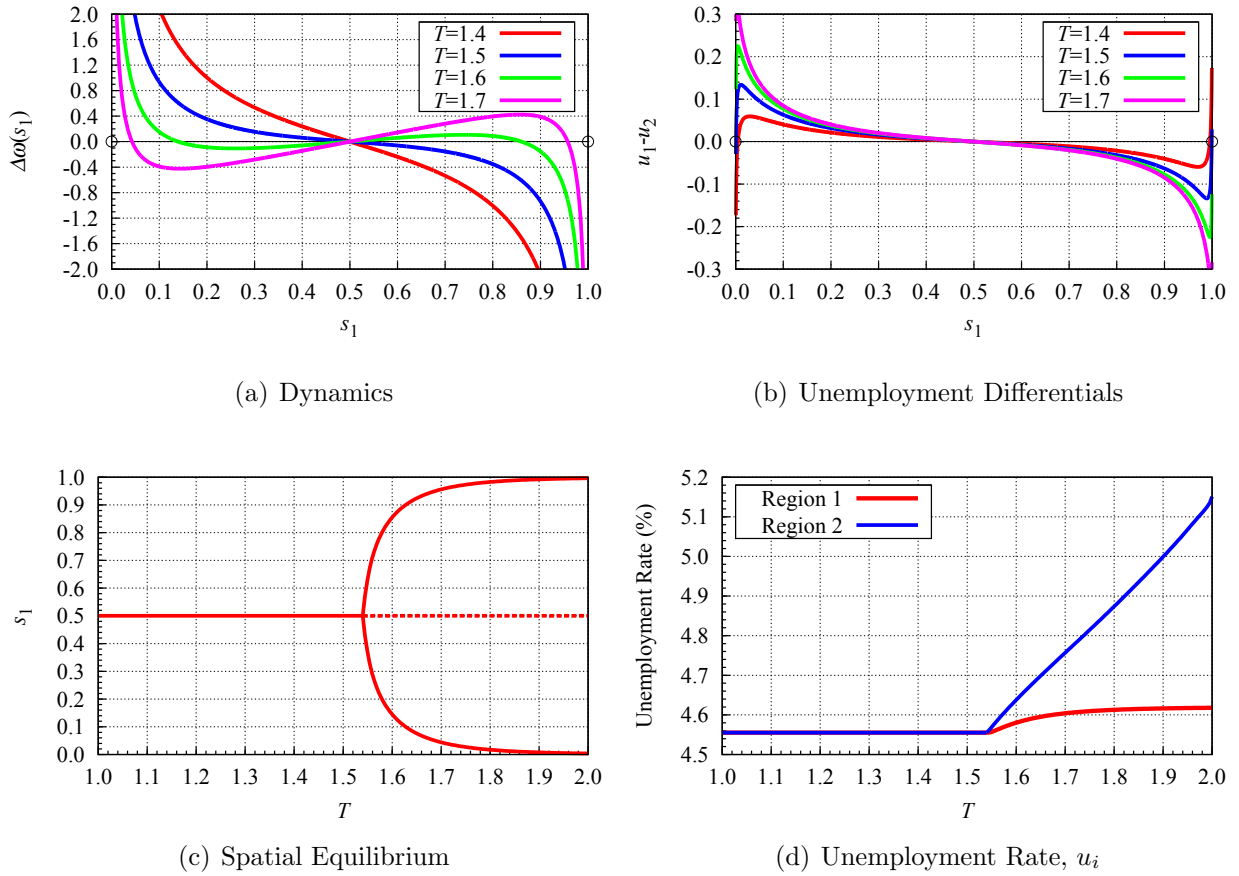


Figure 5: Numerical Simulation Results When Agglomeration Has Weak Negative Externalities on Matching Efficiency

Notes: The solid and dashed lines in Panel (c) denote stable and unstable equilibria, respectively. The parameters used in this numerical analysis are shown in Table 1.

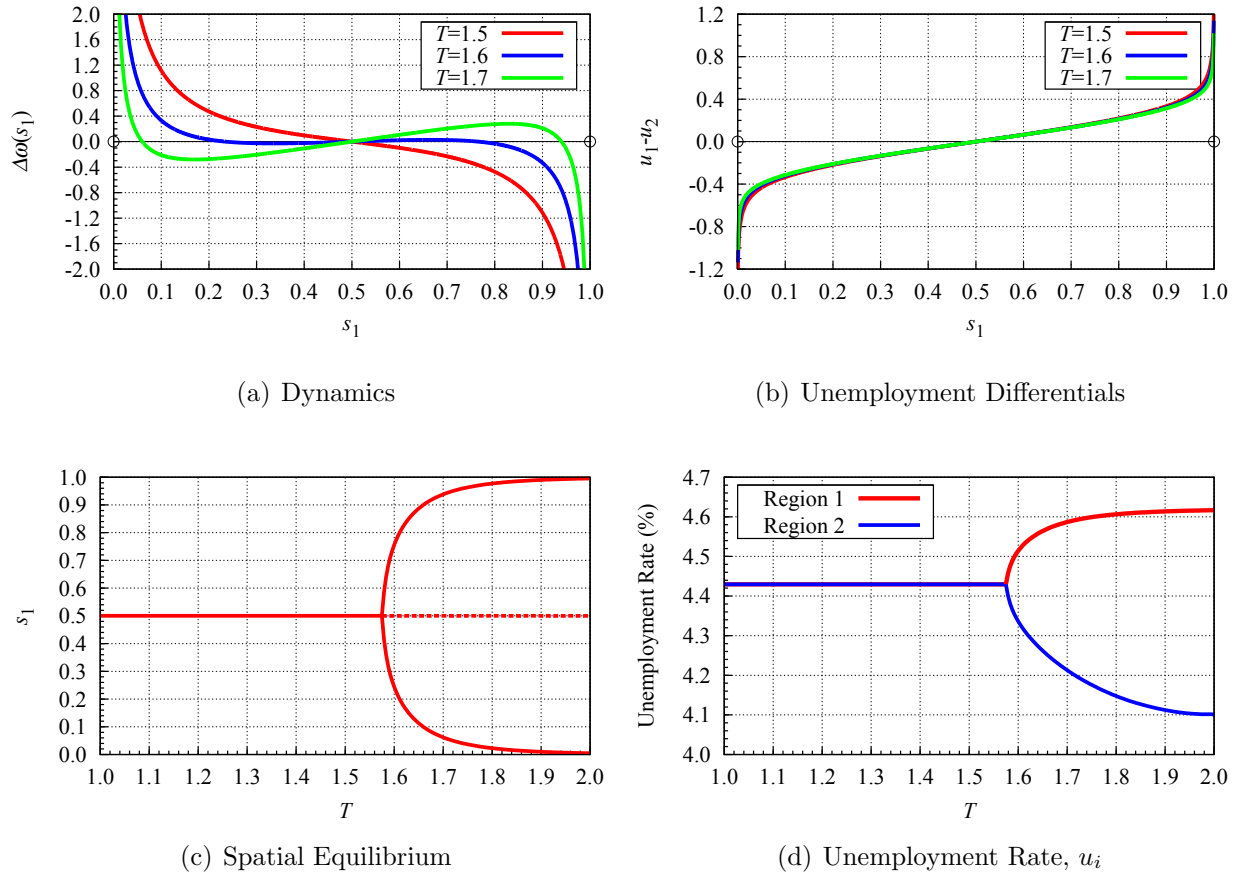


Figure 6: Numerical Simulation Results When Agglomeration Has Strong Negative Externalities on Matching Efficiency

Notes: The solid and dashed lines in Panel (c) denote stable and unstable equilibria, respectively. The parameters used in this numerical analysis are shown in Table 1.

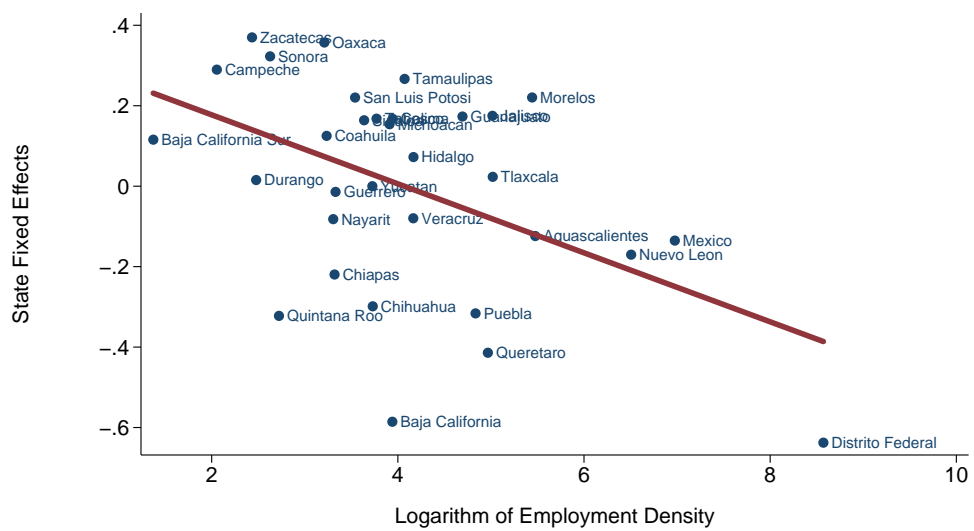


Figure 7: Matching Efficiency and Employment Density