CIRPÉE
Centre interuniversitaire sur le risque, les politiques économiques et l'emploi

Cahier de recherche/Working Paper **03-24**

Agglomeration Effects and the Competition for Firms

Robin Boadway
Katherine Cuff
Nicolas Marceau

Avril/April 2003

Boadway: Department of Economics and CIRPÉE, Queen's University, Kingston, Ontario, Canada, K7L 3N6. E-mail: boadwayr@qed.econ.queensu.ca

Cuff (corresponding author): Department of Economics and CIRPÉE, McMaster University, 1280 Main St.West, Hamilton, Ontario, Canada, L8S 4M4. E-mail: cuffk@mcmaster.ca

Marceau: Département des sciences économiques and CIRPÉE, UQÀM, Case Postale 8888, Succ.Centre-Ville, Montréal, Québec, Canada, H3C 3P8. E-mail: marceau.nicolas@uqam.ca

An earlier (June 2002) version of this paper was presented at the Canadian Economics Association Meetings at McGill University, Montréal and the Japan-Canada Conference on Fiscal Federalism and Local Public Finance at Meiji Gakuin University, Tokyo as well as in seminars at McMaster University, Ryerson University and SUNY at Buffalo. We are grateful for comments by the various participants as well as for suggestions by John Burbidge and Toshihiro Ihori. Comments made by two anonymous referees and an editor were also useful. Financial support from the Social Sciences and Humanities Research Council of Canada is also acknowledged.

Résumé: Nous étudions un monde à deux régions, chacune d'elles étant habitée par un nombre donné mais différent de travailleurs immobiles. Un nombre donné de firmes mobiles choisissent d'opérer dans l'une ou l'autre région. Les firmes créent des emplois là où elles se localisent mais un certain niveau de chômage frictionnel persiste malgré tout. Des effets d'agglomération de deux types sont envisagés, soient les économies d'échelle dans l'appariement des travailleurs aux firmes et les externalité de production découlant de la présence d'autres firmes. Nous étudions le cas où les régions font partie d'un État unitaire dans lequel le gouvernement central décide des politiques régionales. Nous étudions aussi le cas d'un État fédéral dans lequel les gouvernements régionaux sont responsables des politiques régionales. Pour chacun des types d'État, nous caractérisons l'allocation des ressources et nous identifions les instruments de politique économique nécessaires à l'atteinte de l'optimum social.

Abstract: A two-region economy consists of a given but different number of immobile workers in each region, and a given number of mobile firms. Firms create jobs where they locate, but there is frictional unemployment. Two sorts of agglomeration effects arise: those from economies of scale in matching, and those from production economies external to the firm. Regions may either be part of a unitary state in which case all regional policies are decided by the central government, or they may be part of a federal state in which case some policies are determined by the regional governments. We characterize the resource allocations in both a unitary and a federal state, and identify the set of instruments that are required to replicate the social optimum in each state.

Key Words: Agglomeration, Inter-Jurisdictional Competition, Unemployment

JEL Classification: H2, H7, J6, R3

1. INTRODUCTION

Economic activity tends to cluster. In Canada, for example, a disproportionately large share of the manufacturing sector is located in Ontario while in other provinces, in particular in the Maritime provinces, this sector is disproportionately small. In the USA, the clustering of several industries in very few locations is identified, documented, and explained in Krugman (1993). The most prevalent explanation for the clustering, or agglomeration, of economic activity revolves around the existence of increasing returns of some sort. Suppose, for example, that firms are more profitable when they are located near other firms. Then, a cluster of firms in some location makes it attractive for other firms to locate there, thereby making the cluster more attractive and reinforcing the phenomenon.

This clustering of economic activity has important real world implications. If there are individuals who are not willing or simply unable to move out of a region in which there is a limited amount of economic activity, then these individuals are likely to have fewer opportunities than those located in regions where economic activity is flourishing. This, of course, is problematic if society has some aversion to inequality. In this case, a difficult policy problem must be solved. One possibility is for policy to encourage firms to locate in regions in which they have not clustered in the past, so that economic activity and opportunities will be more equally spread out. Alternatively, policy can take the clusters as given and simply redistribute income from the citizens of well-endowed regions to those of less-endowed regions. This paper attempts to identify which of these policies should be used.

Specifically, this paper models the process of job creation explicitly and studies the optimal allocation of firms between regions and the policies that have to be implemented to achieve it. We also study the equilibrium allocation(s) of firms when the policies are decentralized to regional governments which compete for firms, and the role of a central government in inducing the regions to choose optimal policies. We assume a two-region federation with immobile workers in each region, and a given number of perfectly mobile firms. Jobs in each region are created by firms that locate there, but there is imperfect matching of workers to jobs which gives rise to frictional unemployment.[1] Adding more firms to a region not only brings more job offers and employment directly, but can also have two types of scale or agglomeration effects. First, there may be economies of scale in the matching technology—or *matching agglomeration effects*—which implies that the proportion of job offers filled increases with the number of firms in the region. Second, there may be production economies external to the firm—or *production agglomeration effects*—which cause output per firm to rise with the total number of firms in the region. Both sorts of agglomeration give rise to inefficiencies that require corrective policy action. In the literature, these agglomeration forces fall into the category of *localization economies* — the general phenomenon which causes the production costs of firms in a particular industry to decrease when the total output of the industry increases. Localization economies are perceived as being important and relevant for policy. In his discussion of the empirical literature on agglomeration economies, O'Sullivan (1993, pp. 31–32) reports estimates for the elasticity of output per worker with respect to industry output — a measure of the size of localization economies — ranging between 0.02 and 0.11.[2] Note that the

---

[1] Unemployment may arise for other reasons. For example, Fuest and Huber (1999) examine the implication of unemployment resulting from wage bargaining for fiscal competition and coordination.

[2] Thus, depending on the industry, a 10 percent increase in industrial output increases the productivity of a worker by 0.20 to 1.10 percent.

particular agglomeration forces we focus on differ from those generally studied in the economic geography literature. The latter models usually focus on increasing returns internal to imperfectly competitive firms, pecuniary externalities, and transportation costs (see Kind *et al.* (2000) and Ludema and Wooton (2000)), which nevertheless also all fall into the category of localization economies. In addition, full employment is usually assumed in the standard economic geography literature.

To focus on the role of firms, we adopt a model with relatively few elements. Both firms and workers are identical across regions. Workers are entirely passive: all participate in the labour market, and are either employed or involuntarily unemployed. Firms choose where to locate and how many job offers to make. All agents are well-informed. Firms anticipate market outcomes, knowing how many job offers will be filled as well as market prices and government policies, and take them as given. The government can observe who is involuntarily unemployed, and offers full unemployment insurance. It is, however, constrained by the matching process that determines the number of job offers that are filled. The government is also able to impose lump-sum taxes and transfers on firms and workers, and does so to pursue a Pareto optimum involving the welfare of workers in each region and the owners of firms. The only difference between the two regions is the number of immobile workers, although we treat as a special case that in which the regions are the same size.

To characterize the social optimum, we allow a central planner to choose all of the relevant variables such as consumption, employment, vacancies, the amount of involuntary unemployment, and the location of firms. We then examine the resulting allocations under the special cases of no agglomeration, production agglomeration, and matching agglomeration. With no agglomeration, the optimal allocation is symmetric in the sense that per worker variables are identical in both regions whether or not they are the same size. With matching or production agglomeration, the optimal allocation may be symmetric if the two regions are the same size, but necessarily asymmetric if they are not. If the matching function has increasing returns to scale, but diminishing marginal products, there will be a unique local and global optimum. In this case of matching agglomeration, if regions have the same population, this optimum will again be symmetric. However, if the regions differ in size, the larger region will also have a higher proportion of firms per worker. With production agglomeration, things are much more complicated. There will generally be multiple local optima, even in the case of same-sized regions. In the global optimum, either region can have the higher ratio of firms to workers.

The policies that are needed to internalize the externalities of agglomeration are straightforward to specify. The matching technology gives rise to hiring externalities whether or not agglomeration effects are present. These can be corrected by an employment subsidy or tax. Production agglomeration effects, which induce a non-optimal allocation of firms across regions, require differential firm taxation. In a decentralized federation, regional governments will in fact choose the correct employment subsidy. However, they will not choose the correct firm tax. This is because, although their choice of tax on firms will properly take account of the production externality, it will include an element of inefficient tax competition. We show how a central government can intervene with its own firm tax to ensure that the optimal allocation of firms across regions is achieved.

The remainder of the paper is organized as follows. In the next section, we study the behaviour of the private sector, given government policies. In Section 3, we characterize the socially optimal allocation of resources and then determine, in Section 4, the set of unitary state government tax and transfer policies that are necessary to replicate the social optimum in a market economy. In Section 5, the replication of the social optimum in a market economy is again examined, but in the context of a federal state. Finally, in Section 6, we conclude and offer a brief discussion about possible extensions.

2. PRIVATE SECTOR BEHAVIOUR

There are two regions indexed by $i = 1, 2$. Regions may be part of a unitary state in which case regional policies are decided at the central government level as examined in Section IV, or they may be part of a federal state in which case some regional policies are determined at the regional level, and others at the central level as examined in Section V. The private sector in each region is composed of both workers and firms. When dealing with a representative region, we suppress the regional subscript. We also adopt the convention of using uppercase letters to refer to region-wide variables and lowercase letters to individual firm or worker variables.

Region i has L_i immobile workers, with $L_1 \geq L_2$. All workers supply one unit of labour and have identical utility functions $u(c)$ with $u' > 0 > u''$, where c is a composite consumption good whose price is normalized to unity. Utility is the same whether individuals are working or not.[3] Our analysis focuses on the behaviour of firms and so, to simplify, we assume workers are passive. In all equilibria characterized below, they will remain attached to the labour market — employed or involuntarily unemployed.

Firms make two types of decisions. First, they choose the region in which to operate. Second, they post vacancies, succeed at filling some of them, and then produce and earn profits. The proportion of vacancies filled, and therefore the unemployment rate in each region, will be determined by a simple matching technology described below. When making their decisions, firms take as given government policies, which include both employment and firm taxes. In a market economy, the sequence of events is therefore as follows:

Stage 1: Choice of policies by government(s);

Stage 2: Location decision by firms;

Stage 3: Hiring decision by firms, and determination of employment and output.

The rest of this section describes firm decision-making at Stages 2 and 3 for a representative region, and solves for the resulting regional labour market equilibrium. The governments' policy choices in Stage 1 are examined in Sections 4 and 5. The relevant equilibrium concept for the various games studied is that of subgame perfection. Thus, we begin by first solving for Stage 3 and then for Stage 2.

2.1. Stage 3: Regional Labour Market Equilibrium

At this stage, the number of firms in the representative region has been determined. We begin by looking at how individual firms behave, and then turn to equilibrium in the regional labour market as a whole.

[3] Incorporating some disutility of work (or utility of leisure) would only complicate the analysis and not affect the basic results.

Firm Equilibrium

There are $\overline{N}$ identical firms which are costlessly mobile between the two regions, so $\overline{N} = N_1 + N_2$. In the representative region, each firm produces output according to the production function $a(N)f(e)$, where e is per firm employment, $a(N) \geq 0$, $a' \geq 0$, $f' > 0 > f''$ and $f(0) = 0$. If $a' > 0$, there is said to be *production agglomeration*: the more firms there are in a given region, the higher output per firm will be in that region for a given level of employment.[4] If there is no production agglomeration, then $a' = 0$. In this case, we assume for simplicity that $a(N) = 1$.

Frictional unemployment arises in this economy as a result of the imperfect matching of workers to jobs.[5] In order to hire workers, each firm posts v vacancies. Labour markets are regional in nature. The total number of unemployed workers who are matched to posted vacancies in a region, denoted by M, depends on the number of involuntary unemployed I and the total number of job vacancies $V = Nv$ in the region. The number of matches is determined by a matching technology $M(I, V)$, which is assumed to satisfy $M_I, M_V > 0$ and $M_{II}, M_{VV} < 0$. This matching function is taken to be homogeneous of degree $\rho \geq 1$ in (I, V). If $\rho > 1$, there is *matching agglomeration.* Increasing I and V by the same factor increases the number of matches in the region by that factor to the power ρ. Given the homogeneity of $M(I, V)$, it is useful to rewrite it in the intensive form. Define $s \equiv I/V$. This ratio represents the slackness in the labour market from a potential employer's perspective. By homogeneity, we have:

$$m(s) \equiv M\left(\frac{I}{V}, 1\right) = V^{-\rho}M(I, V)$$

It is straightforward to show that, given our assumptions about the function $M(I, V)$, $m' > 0 > m''$, $sm'(s)/m(s) \equiv \eta \in (0, 1)$, and $\rho - \eta < 1$.[6] Note for future reference that the total number of matches is $M(I, V) = V^{\rho}m(s)$, and the proportion of vacancies filled is $M/V = V^{\rho-1}m(s)$. To simplify our analysis in what follows, we assume that η is a constant. This will be the case if $M(I, V)$ is of the Cobb-Douglas form.

The proportion of a given firm's vacancies v that are filled depends both on the average proportion of vacancies filled region-wide, M/V, as well as on the wage the firm offers to pay. The higher the firm's wage offer relative to other firms in the economy, the more likely its vacancies will be filled. If the number of job matches at the firm level were independent of wages, then firms would have no incentive to offer positive wages.[7] Let r be the firm's relative wage, so its wage

---

[4] The assumption that production agglomeration depends on the number of firms rather than, say, total employment or output is not innocuous, as we shall see later. Our presumption is that each firm is associated with an entrepreneur or management team whose ideas or production methods might spill over to other firms independent of the number of workers it employs. Agglomeration associated with total employment is captured by our matching externality discussed below.

[5] We assume that there is imperfect matching even though the heterogeneity of firms and workers is not explicitly modeled. This simplifying assumption is standard in the search/matching literature (e.g. see Diamond, 1982).

[6] Differentiating $M(I, V) = V^{\rho}m(s)$, we obtain: $M_I = V^{\rho-1}m'(s) > 0$, $M_V = V^{\rho-1}m(s)(\rho - \eta) > 0$, $M_{II} = V^{\rho-2}m''(s) < 0$, and $M_{VV} = V^{\rho-2}[\rho - \eta - 1]m(s)(\rho - \eta) < 0$, which implies the stated properties.

[7] This is no longer the case if firms bargain with workers to determine a firm-specific wage once a

offer is rw where w is the market wage. For simplicity, we assume that the proportion of firm vacancies filled takes the simple multiplicative form, rM/V. In each 'period', a proportion b of those employed at any given firm will exogenously separate from their job.[8] In the steady-state, each firm's new hires will equal its separations, so $(rM/V)v = be$. This equality will be referred to as the *firm matching constraint*.[9] It is also assumed that there is a fixed cost per retained worker $(1-b)e$ and per vacancy created v (equivalently, the number of employees e plus the number of vacancies created but not filled, $v - be$), assumed for simplicity to be ψ for both.[10]

Below, we assume that governments use two policy instruments that affect firms: an employment tax σ for each worker the firm employs, and a lump-sum tax per firm T. It will be shown that these policy instruments are sufficient for achieving efficient production outcomes in a federal state. When making their choices, firms take these policies as given. They also treat the region-wide variables I and V as fixed, that is, they take the ratio of these two variables s and the proportion of vacancies filled $M/V = V^{\rho-1}m(s)$ as given. Firms also take as given the wage w workers can receive elsewhere in the region. The representative firm's problem can therefore be written as:

$$\max_{\{r,e\}} \quad \pi - T = a(N)f(e) - (rw+\sigma)e - \psi\left[(1-b)e + \frac{beV^{1-\rho}}{rm(s)}\right] - T \tag{1}$$

where we have substituted out v using the firm's matching constraint.

In equilibrium, all firms will behave identically so $r = 1$. The first-order conditions of the representative's firm problem can then be written as:[11]

$$-w + \frac{\psi b V^{1-\rho}}{m(s)} = 0 \tag{r}$$

$$a(N)f'(e) - (w+\sigma) - \psi\left[1 - b + \frac{bV^{1-\rho}}{m(s)}\right] = 0 \tag{e}$$

match has been formed. Such an alternative wage determination process is standard in the matching literature (for example, see Pissarides, 2000).

[8] Allowing for the number of firm vacancies filled to depend directly on the firm's actual wage would not change any of our qualitative results. Allowing the separation rate b to be endogenous would complicate the model considerably.

[9] The matching model used here is a simplified version of that found in Layard and Johnson (1986). Our model is in fact the steady-state version of a fully dynamic model presented in the Appendix. As shown there, the firm's matching constraint used here corresponds to the transition equation in the fully dynamic problem. When evaluated in the steady state, the Euler equations of the problem in the Appendix correspond to the first-order conditions below. In this paper, we are interested in steady-state outcomes and how government policies affect them.

[10] The fixed cost of posting and maintaining vacancies reflects advertising costs, human resources staff and other associated search costs. The fixed cost per retained worker reflects the operating costs of the firm such as maintaining office spaces and providing office supplies. Assuming the latter fixed cost to be zero would not affect our results. We include it here to reflect a commitment of the firm to filling its posted vacanices. However, assuming a positive cost of creating vacanices is necessary for our story of equilibrium unemployment. Therefore, we can think of the fixed cost, ψ, as reflecting matching frictions in the labour market.

[11] The second-order conditions are satisfied.

The condition (r) determines the market wage rate that all firms are satisfied with (i.e., such that $r = 1$ is optimal), $w(V,s) = \psi b V^{1-\rho}/m(s)$ where $w_V \leq 0$ and $w_s < 0$. Substituting this into (e), we obtain

$$a(N)f'(e) = \sigma + \psi \left[1 - b + \frac{2bV^{1-\rho}}{m(s)}\right] \tag{2}$$

which yields $e(N,\sigma,V,s)$, where $e_N \geq 0$, $e_\sigma < 0$, $e_V \geq 0$, and $e_s > 0$.[12] Firm profits can then be written as:

$$\pi(N,\sigma,V,s) = a(N)f(e(\cdot)) - \sigma e(\cdot) - \psi \left[(1-b)e(\cdot) + \frac{2be(\cdot)V^{1-\rho}}{m(s)}\right]$$

where $\pi_N \geq 0$, $\pi_\sigma < 0$, $\pi_V \geq 0$, and $\pi_s > 0$.[13] For later reference, it is useful to rewrite the firm's profits in a more conventional form by substituting in its first-order conditions (2) to give:

$$\pi(N,\sigma,V,s) = a(N)f(e(\cdot)) - e(\cdot)a(N)f'(e(\cdot)) \tag{3}$$

Regional Labour Market Equilibrium

Equilibrium in a representative region requires that those in the regional labour force are either employed or involuntarily unemployed and that the aggregate matching constraint be satisfied. Thus, the following two conditions must be satisfied:[14]

$$L = Ne(N,\sigma,V,s) + sV; \qquad bNe(N,\sigma,V,s) = V^\rho m(s) \tag{4}$$

These can be solved for $V(N,\sigma)$ and $s(N,\sigma)$. Total differentiation of the equations in (4) yields:

$$V_N = \frac{(Ne_N + e)(V^\rho m'(s) + bV)}{\Delta} > 0 \qquad V_\sigma = \frac{Ne_\sigma(V^\rho m'(s) + bV)}{\Delta} < 0$$

$$s_N = -\frac{(Ne_N + e)(\rho V^{\rho-1} m(s) + bs)}{\Delta} < 0 \qquad s_\sigma = -\frac{Ne_\sigma(\rho V^{\rho-1} m(s) + bs)}{\Delta} > 0$$

where, by our assumption that $\rho - \eta < 1$,

$$\Delta = V^\rho m(s)(\rho - \eta) - 2Ne_\sigma \frac{\psi b}{m(s)}\left(m'(s) + bV^{1-\rho}(1-\rho+\eta)\right) > 0$$

which implies that the labour market equilibrium is stable.

Using these, we obtain equilibrium wages, employment and profits:

$$w(V(N,\sigma), s(N,\sigma)), \quad e(N,\sigma,V(N,\sigma),s(N,\sigma)), \quad \pi(N,\sigma,V(N,\sigma),s(N,\sigma))$$

---

[12] Differentiating (2), we obtain $e_N = -a'(N)f'(e)/[a(N)f''(e)] \geq 0$, $e_\sigma = 1/[a(N)f''(e)] < 0$, $e_V = 2w_V/[a(N)f''(e)] \geq 0$, and $e_s = 2w_s/[a(N)f''(e)] > 0$.

[13] By the Envelope Theorem, we have $\pi_N = a'(N)f(e) \geq 0$, $\pi_\sigma = -e < 0$, $\pi_V = -2ew_V \geq 0$, and $\pi_s = -2ew_s > 0$.

[14] Recall, L is the size of the regional labour force. Therefore, the first condition states that $L = Ne + I$ where we have substituted out for I using the definition of $s = I/V$.

where:

$$
\begin{aligned}
\frac{dw}{dN} &= w_V V_N + w_s s_N > 0; & \frac{dw}{d\sigma} &= w_V V_\sigma + w_s s_\sigma < 0 \\
\frac{de}{dN} &= e_N + e_V V_N + e_s s_N; & \frac{de}{d\sigma} &= e_\sigma + e_V V_\sigma + e_s s_\sigma < 0 \\
\frac{d\pi}{dN} &= \pi_N + \pi_V V_N + \pi_s s_N; & \frac{d\pi}{d\sigma} &= \pi_\sigma + \pi_V V_\sigma + \pi_s s_\sigma < 0
\end{aligned}
\tag{5}
$$

The signs of $d\pi/dN$ and de/dN are ambiguous when there are production agglomeration effects. If there are none ($a' = 0$), these expressions are necessarily negative. In other words, increasing the number of firms in a region reduces per firm employment and profits unless there are sufficiently large production externalities.

2.2. Stage 2: Location Equilibrium

Since firms are choosing where to locate, it is necessary to distinguish between the two regions in what follows. Given regional policies $(\sigma_1, \sigma_2, T_1, T_2)$, firms fully anticipate the after-tax profits that they can earn in each region. In equilibrium, after-tax profits must be equalized in both regions:

$$
\pi_1(N_1, \sigma_1, V_1(N_1, \sigma_1), s_1(N_1, \sigma_1)) - T_1 = \pi_2(N_2, \sigma_2, V_2(N_2, \sigma_2), s_2(N_2, \sigma_2)) - T_2 \tag{6}
$$

where $N_2 = \overline{N} - N_1$. This equation determines N_1 as a function of regional policies $(\sigma_1, \sigma_2, T_1, T_2)$. For stability of the location equilibrium, the difference in per firm profits between region 1 and 2 must be decreasing in the number of firms locating in region 1, so $d\pi_1/dN_1 + d\pi_2/dN_2 < 0$. As mentioned, in the absence of production agglomeration, $d\pi/dN < 0$, and the stability condition is met. However, it is possible that $d\pi/dN$ becomes positive when there is production agglomeration. In what follows, we assume that the location equilibrium is stable.

This completes our analysis of private sector behaviour. What remains to be considered is Stage 1 of the sequence of events. We return to that in Sections 4 and 5 for the cases of a unitary state and a federation. First, we present the social optimum benchmark against which the outcomes of Section 4 and 5 will be compared.

3. THE SOCIAL OPTIMUM

To be as general as possible, we characterize the social optimum in this economy by assuming there exists a central planner who solves the Pareto problem of maximizing the per capita utility of workers in one region subject to some minimum level of per capita utility $\overline{U}$ in the other region. This implies that only workers count. In the next section, we allow for the fact that there may be a group of non-workers who own firms, in which case a minimum profit constraint has to be added to the problem to account for their utility. For now, all output will be assumed to accrue to workers. All workers face some risk of unemployment. Since they are risk-averse and the governments can observe who is involuntarily unemployed, we assume from the start that there is full unemployment insurance, so individuals receive the same level of consumption whether or not they are working. Restricting the ability of governments to identify the involuntarily unemployed would imply that only second-best policies with less than full unemployment insurance could be used. This would complicate the problem without adding any new insight.[15]

[15] Optimal policies under imperfect information in the labour market are analyzed in Boadway, Cuff and Marceau (2003).

The planner can choose the location of firms, as well as firm employment, firm vacancies, the ratio of regional involuntary unemployment to regional vacancies, and individual consumption in each region. However, it is constrained by the imperfect matching between workers and jobs in each region, and by the fixed number of firms in the economy as a whole. The planner's problem can be written as:

$$\max_{\{c_1,c_2,e_1,e_2,v_1,v_2,s_1,s_2,N_1\}} \quad u(c_1)$$

subject to a minimum level of per capita utility in region 2,

$$u(c_2) \geq \overline{U} \tag{φ}$$

the resource constraint,

$$N_1[a(N_1)f(e_1) - \psi((1-b)e_1 + v_1)] + (\overline{N} - N_1)[a(\overline{N} - N_1)f(e_2) - \psi((1-b)e_2 + v_2)] - L_1c_1 - L_2c_2 = 0 \tag{λ}$$

the labour market equilibrium conditions,

$$L_1 - N_1e_1 - s_1N_1v_1 = 0 \tag{δ_1}$$

$$L_2 - (\overline{N} - N_1)e_2 - s_2(\overline{N} - N_1)v_2 = 0 \tag{δ_2}$$

and the regional matching constraints,

$$(N_1v_1)^{\rho}m(s_1) - bN_1e_1 = 0 \tag{γ_1}$$

$$((\overline{N} - N_1)v_2)^{\rho}m(s_2) - b(\overline{N} - N_1)e_2 = 0 \tag{γ_2}$$

where the constraints are labeled by the Lagrange multipliers used in the constrained maximization problem.

The social optimum is the solution to the following five conditions obtained from the first-order conditions and the matching constraints (γ_1) and (γ_2) after routine substitution, where s has been substituted out using the labour market equilibrium condition, $s = (L/N - e)/v$:[16]

Firm Allocation by Region:

$$a(N_1)f(e_1) - e_1a(N_1)f'(e_1) + N_1a'(N_1)f(e_1) = a(N - N_1)f(e_2) - e_2a(N - N_1)f'(e_2) + (N - N_1)a'(N - N_1)f(e_2) \tag{7}$$

Employment per Firm:

$$a(N_i)f'(e_i) = \psi\left[1 - b + \frac{m'((L_i/N_i - e_i)/v_i) + b\,(N_iv_i)^{1-\rho}}{m((L_i/N_i - e_i)/v_i)(\rho - \eta)}\right] \qquad i = 1, 2 \tag{8}$$

[16] A technical appendix available on request from the authors presents the details of the derivations of the results in this section.

Matching Constraints:

$$bN_ie_i = (N_iv_i)^\rho \, m((L_i/N_i - e_i)/v_i) \qquad i = 1, 2 \qquad (9)$$

These five equations determine the socially optimal values of $\{e_1, e_2, v_1, v_2, N_1\}$. Their values will depend critically on what type of agglomeration there is, if any, in the economy. With the existence of agglomeration effects of either sort, the size of the region enters explicitly into the condition determining optimal firm employment, while with production agglomeration alone size enters explicitly into the condition determining the optimal firm allocation.

Note that none of the five equations depends on the relative utility levels in the two regions, that is, on the point achieved on the national utility possibilities frontier (UPF). In effect, the solution to the above problem maximizes the total product in the economy. The amount of resources available for total consumption, $L_1c_1 + L_2c_2$, is determined from the resource constraint (λ) evaluated at the above solution. How these resources are distributed between workers in the economy will depend on the value of $\overline{U}$, that is, which point on the economy-wide UPF the planner wants to achieve. For example, under a utilitarian objective, the planner would want to equalize the marginal utility of consumption of all individuals which, given our assumption on the utility function, translates into equal per capita consumption, $c_1 = c_2 = c$, where c would be the maximized total product in the economy divided by total population $L_1 + L_2$.

Using (7)–(9), we next characterize the possible outcomes in the social optimum by considering the following three special cases: no agglomeration, matching agglomeration only, and production agglomeration only.

3.1. Characterization of the Social Optimum

Recall that $L_1 \geq L_2$, so regions may be the same size, with $L_1 = L_2$, or they may differ in size, in which case Region 1 is assumed to be the larger one, so $L_1 > L_2$. Whether regions differ in size affects the nature of the social optimum when there are agglomeration effects.

Case 1: No Agglomeration

With no agglomeration, $a = 1$ and $\rho = 1$. In this case, the optimal allocation is symmetric in terms of per capita variables. Thus, the ratio of regional population to number of firms is equalized across regions ($L_1/N_1 = L_2/N_2$), and all per firm and per worker variables are identical in both regions ($e_1 = e_2$, $v_1 = v_2$). If regions have the same population, then the outcome is fully symmetric since there are equal number of firms in both regions. The second-order conditions of the social optimum are satisfied everywhere when there are no agglomeration effects. Therefore, the solution to the above conditions (7)–(9) is unique.

Case 2: Matching Agglomeration

With only matching agglomeration, $a = 1$ and $\rho > 1$. In this case, if regions are the same size ($L_1 = L_2$) the ratio of regional population to number of firms will be the same ($L_1/N_1 = L_2/N_2$). But if $L_1 > L_2$, then optimality requires that $L_1/N_1 < L_2/N_2$: disproportionately more firms should be allocated to the larger region to exploit the matching agglomeration effect. As in the no-agglomeration case, the second-order conditions are satisfied everywhere, so the solution to (7)–(9) is again unique. With same-sized regions, the fully symmetric outcome again satisfies all the necessary conditions and therefore it is the unique optimum.

Case 3: Production Agglomeration

When there is only production agglomeration, $a' > 0$ and $\rho = 1$. In this case, the second-order conditions are not necessarily satisfied, so there can be multiple local optima with sufficiently large agglomeration effects. If the regions are the same size, then the equilibrium can be the symmetric one with $L_1/N_1 = L_2/N_2$. But, even if regions are the same size and the second-order conditions are not satisfied everywhere, then the social optimum might involve $L_1/N_1 \gtrless L_2/N_2$. If one region is larger, then the equilibrium must be asymmetric: $L_1/N_1 \neq L_2/N_2$. Therefore, with regions that differ in size, there are two possible types of equilibria: $L_1/N_1 < L_2/N_2$ and $L_1/N_1 > L_2/N_2$ with $N_1 \neq N_2$. Unlike with matching agglomeration, it may be optimal to put disproportionately more firms in the smaller region in the presence of production agglomeration. If the production agglomeration is particularly strong, then the social optimum may be characterized by all firms being concentrated in a single region.

4. THE UNITARY STATE

Decentralizing the social optimum to a market economy setting turns out to be relatively straightforward, as long as we retain the assumption of a unitary state, i.e. a single level of government. Going back to the sequence of events described in Section 2, a unitary state government is assumed to decide on all regional policies in Stage 1. Then, firms choose where to locate in Stage 2, and how many workers to hire and the number of vacancies to create in Stage 3. The instruments needed by the unitary state government to decentralize the social optimum are regional-specific taxes/transfers on employment, firms, and households.

We first examine firm hiring (Stage 3), and characterize the employment tax needed to internalize the externalities arising from imperfect matching. Then, turning to firm location in Stage 2, we determine the necessary taxes on firms required to ensure the efficient distribution of firms between regions. We then discuss how these taxes together with lump-sum taxes on working households and transfers to the unemployed can ensure that the desired point on the socially optimal UPF is achieved. The allocation of resources achieved by this three-stage procedure is the result of a sub-game perfect Nash equilibrium (SPNE).

4.1. Employment Taxes

Recall that the marginal productivity condition determining the choice of e in the representative region in the social optimum is given by (8), while that determining the choice of e by competitive firms is given by (2). To ensure that the choice of e by firms corresponds to the socially optimal e, the employment tax must be set such that the right-hand side of (2) is equal to the right-hand side of (8), that is,

$$\sigma = \psi \left[\frac{m'(s) + bV^{1-\rho}}{m(s)(\rho - \eta)} - \frac{2bV^{1-\rho}}{m(s)} \right] \tag{10}$$

Note that this rule for the employment tax σ is not affected by $a(N)$, but does depend on matching frictions in the labour market. In other words, *σ corrects for an externality arising from the matching process (including matching agglomeration effects through ρ), but not the production agglomeration externality.*[17]

[17] The tax/subsidy σ would also depend on production agglomeration if the latter were a function of

To gain further insight into nature of the optimal employment tax, we can characterize explicitly the matching externality. It arises in this economy because each firm j of the N firms in a representative region ignores the full effects its choice of employment e_j (or equivalently, job vacancies v_j) has on total output in that region. To see this, note that total output in a region with N firms, denoted by Y, is:

$$Y = \sum_{k=1}^{N}[a(N)f(e_k) - \psi((1-b)e_k + v_k)] = \sum_{k=1}^{N}\left[a(N)f(e_k) - \psi((1-b)e_k + \frac{be_kV^{1-\rho}}{m(s)})\right]$$

where we have used the firm matching constraint with $r_k = 1$ in equilibrium. The economy-wide variables V and s depend on the employment decisions of all firms in the economy. In particular, the regional labour market equilibrium conditions (4) can be written as $L = E + sV$ and $bE = V^\rho m(s)$, where aggregate employment E is given by $E = \sum_{k=1}^{N} e_k$. These conditions determine V and s as functions of E, where:

$$\frac{\partial V}{\partial E} = \frac{bV + V^\rho m'(s)}{V^\rho m(s)(\rho - \eta)} > 0 \qquad \frac{\partial s}{\partial E} = \frac{-\rho V^{\rho-1}m(s) - bs}{V^\rho m(s)(\rho - \eta)} < 0$$

Consider now the effect of an increase in firm j's employment on total output Y. Noting that $\partial E/\partial e_j = 1$, differentiating the above expression for Y with respect to e_j and evaluating it at a symmetric equilibrium where $e_j = e$ for all firms, we obtain:

$$\frac{\partial Y}{\partial e_j} = \left\{a(N)f'(e) - w - \psi\left[1 - b + \frac{bV^{1-\rho}}{m(s)}\right]\right\} + \left\{w + \psi\frac{bNeV^{1-\rho}m'(s)}{m(s)^2}\frac{\partial s}{\partial E} - \psi\frac{(1-\rho)bNeV^{-\rho}}{m(s)}\frac{\partial V}{\partial E}\right\} \tag{11}$$

where we have added and subtracted the equilibrium wage rate, w. The two terms in braces reflect the effect of an increase in employment perceived by the firm and the external effect imposed on the rest of the regional economy. From the firm's optimization problem, the first term is equal to zero. The externality term has the following interpretation. The wage rate w appears because the firm in its choice of employment treats the wage rate as a cost to itself along with the fixed costs of employment ψ. However, only the latter is a true social cost. Involuntarily unemployed workers have no opportunity cost. Therefore, hiring one more worker from the pool of involuntarily unemployed does not involve any social cost other than the additional resources needed to pay for the increase in the fixed costs arising from hiring that additional worker. The last two components of the externality term reflect the additional hiring costs that are imposed on other firms when one firm increases its employment: a change in e_j will increase total employment E, which will affect the economy-wide variables V and s, in turn tightening all other firms' matching constraints. This is apparent when substituting the firm matching constraint, $v = beV(E)^{1-\rho}/m(s(E))$, in the externality term in (11). The externality term can then be written as:

$$\left\{w - \psi N\frac{dv}{dE}\right\} \tag{12}$$

total regional employment rather than the total number of firms in the region, i.e. if production were given by $a(Ne)f(e)$ with $a'(Ne) \geq 0$. In this case, the production agglomeration effect term $(Na'(Ne)f(e))$ in the social optimum would appear in the condition determining the optimal firm employment rather than the optimal firm location.

where $dv/dE > 0$.[18] To maintain a given level of employment, firms must post more vacancies when total employment goes up.

More relevant for our purpose is the fact that the externality term in (11) reduces to the negative of the right-hand side of (10) when the expressions for $w = (\psi b V^{1-\rho})/m(s)$, $\partial s/\partial E$ and $\partial V/\partial E$ are substituted into (11). Thus, setting σ as in equation (10) corrects for the externalities arising from imperfect matching in the labour market. The sign of this externality effect is indeterminate: *either an employment subsidy or an employment tax will generally be needed to decentralize the social optimum to a unitary state setting.*

For an indication of the sign and magnitude of this externality, we can re-write the externality component of (11) as follows by using the expression for the optimal choice of the firm's wage given by (r) earlier and the firm's matching constraint:

$$\psi \left\{ \frac{v}{e} - N \frac{dv}{dE} \right\} = \psi \frac{v}{e} \left\{ 1 - \frac{dv}{dE} \frac{E}{v} \right\}$$

The sign of this expression is positive if the elasticity of per firm vacancies with respect to total employment is less than one, and negative if the elasticity is greater than one. In the former case, per firm employment is too low from a social point of view and an employment subsidy is required to achieve the social optimum. In the latter case, per firm employment is too high and an employment tax is necessary to achieve the social optimum. In terms of the fundamental parameters of this economy, using (11), it can be shown that in the case of no matching agglomeration ($\rho = 1$), the externality is negative if the following condition holds:

$$b(1 - 2\eta) < m'(s)$$

A sufficient condition for the above to be satisfied is if $\eta \geq 1/2$. Therefore, if the matching function is sufficiently elastic then the government should be taxing employment. A necessary condition for an employment subsidy is that $\eta < 1/2$ and the sufficient condition is given by $b(1 - 2\eta) > m'(s)$.

4.2. Taxes on Firms and Households

In order to fully decentralize the social optimum, the allocation of firms between regions must also be efficient. Ensuring that the socially optimal decision rule for e is used is not enough to do this. In particular, condition (7) determining the optimal distribution of firms must also be satisfied. To ensure this, the equilibrium firm location condition (6) must give rise to the efficient distribution of firms. Let T_1 and T_2 be regional-specific taxes on firms. These taxes potentially play two roles. First, they provide an incentive for firms to locate efficiently and second, they divert profits from firms to households in the event that firms are not owned by the households that make up the working population. Consider these two roles in turn.

First, firms locate such that net profits are equalized or, in other words until the location equilibrium condition, $\pi_1 - T_1 = \pi_2 - T_2$, is satisfied. Using the expression (3) for firm profits, we can rewrite this equilibrium condition as:

$$a(N_1)f(e_1) - e_1 a(N_1) f'(e_1) - T_1 = a(N_2) f(e_2) - e_2 a(N_2) f'(e_2) - T_2 \tag{13}$$

[18] Using the firm's matching constraint, we have $dv/dE = (\partial v/\partial V)(\partial V/\partial E) + (\partial v/\partial s)(\partial s/\partial E) = be[bV^{1-\rho}(1 - \rho + \eta) + m'(s)]/[m(s)^2 V^{\rho}(\rho - \eta)] > 0$.

Condition (13) will correspond to the optimal firm allocation rule (6) provided the following relationship between T_1 and T_2 holds:

$$T_2 - T_1 = N_1 a'(N_1) f(e_1) - N_2 a'(N_2) f(e_2) \tag{14}$$

This condition ensures the efficient distribution of firms. It is worth noting that if there were no production agglomeration, then $a' = 0$ and the tax on firms is not needed to correct for an inefficient distribution of firms. In this case, $T_1 = T_2$ and profits before tax give the correct signal of the value of the firm to the region, so an efficient distribution of firms will result without intervention even if matching agglomeration existed.

The above instruments ensure that the total product in the economy is maximized. The next issue is how to ensure the optimal distribution of this product between workers in the two regions. As in the previous section, we assume that the government has enough information to provide full insurance. At the same time, the level of per capita consumption can differ across regions. By varying the latter, the unitary state government can choose the most preferred point on the nation-wide UPF. If the firms are owned by households in the two regions, all wages and after-tax profits will accrue to them. It is then sufficient for the government to impose lump-sum region-specific taxes and transfers on households to allocate total output optimally among households. In this case the absolute level of the taxes on firms T_i is not important, only their relative values across regions matters. On the other hand, if firms are owned by non-workers who receive no welfare weight, then the unitary government simply sets the taxes on firms so that the net return to a firm is at its minimum level which is assumed to be zero. Either way, the unitary state government is able to distribute the total resources produced in the economy to households.

We now turn to the decentralization of government policies to a federal setting. We show how allowing regional governments to choose regional policies will result in an inefficient outcome in the absence of a central government. Regional governments acting as Nash competitors will distort their policies to attract firms to their regions. As a result, national product will not be maximized. Corrective central government policies will be required to achieve the social optimum.

5. THE FEDERAL STATE

We focus first on the case where all regional fiscal decisions are taken by regional governments. Later we consider how a central government might intervene to ensure that the social optimum is attained. We consider the general case in which there can be both types of agglomeration and assume that the regional government policies include an employment tax σ, a tax per firm of t (rather than T as in the unitary state case above), and lump-sum taxes and transfers on households. Regional governments maximize the per capita utility of households in their region, which is equivalent to maximizing the sum of utilities since households are immobile. We assume initially that firms are owned by the working population in each region, and that households in region i own an equal per capita share of $\widehat{N}_i$ firms where $\widehat{N}_1 + \widehat{N}_2 = \overline{N}$. Regional governments are first-movers and set their policies in Stage 1. Once regional policies are set, firms decide in which region to locate in Stage 2, and then the regional labour market comes to an equilibrium in Stage 3. Again, the equilibrium is an SPNE. Having solved Stages 2 and 3 in Section 2, we immediately turn to the setting of regional policies in Stage 1.

5.1. Stage 1: Regional Government Policies

Regions anticipate the effect of their policies both on the location equilibrium and on their own labour markets, and take as given the policy choices of the other region. Since the government can impose lump-sum taxes on households, choosing those taxes is equivalent to choosing per capita consumption. For analytical convenience, we treat c as the choice variable of the representative regional government (which we take to be region 1). The problem of this government can be written as follows, where we have suppressed the regional subscripts:

$$\max_{\{c,t,\sigma\}} \quad u(c)$$

subject to its resource constraint,

$$\begin{aligned}[w(V(N,\sigma),s(N,\sigma))+\sigma]Ne(N,\sigma,V(N,\sigma),s(N,\sigma))&\\ +Nt+\widehat{N}[\pi(N,\sigma,V(N,\sigma),s(N,\sigma))-t]-Lc=0&\end{aligned} \qquad (\lambda)$$

and the location equilibrium condition,

$$\begin{aligned}\pi(N,\sigma,V(N,\sigma),&s(N,\sigma))-t\\ &-\pi_2(\overline{N}-N,\sigma_2,V_2(\overline{N}-N,\sigma_2),s_2(\overline{N}-N,\sigma_2))+t_2=0\end{aligned} \qquad (\mu)$$

where again the constraints are labeled by the Lagrange multipliers used in the constrained maximization problem. Notice that we are treating the number of firms N as an artificial control variable, while adding the location equilibrium condition as a constraint. We could equivalently have simply taken the number of firms to be determined endogenously by the solution to the location equilibrium constraint $N(\sigma,\sigma_2,t,t_2)$. As well, we are implicitly assuming that a non-negative after-tax profit constraint on firm profits is not binding.

The first-order conditions are:

$$u'(c)-\lambda L=0 \qquad (c)$$

$$\lambda(N-\widehat{N})-\mu=0 \qquad (t)$$

$$\lambda\left[(w+\sigma)N\frac{de}{d\sigma}+Ne\frac{dw}{d\sigma}+Ne+\widehat{N}\frac{d\pi}{d\sigma}\right]+\mu\frac{d\pi}{d\sigma}=0 \qquad (\sigma)$$

$$\lambda\left[(w+\sigma)N\frac{de}{dN}+Ne\frac{dw}{dN}+(w+\sigma)e+t+\widehat{N}\frac{d\pi}{dN}\right]+\mu\left[\frac{d\pi}{dN}-\frac{d\pi_2}{dN}\right]=0 \qquad (N)$$

Substituting (t) along with the relevant expressions in (5) for $dw/d\sigma$, $de/d\sigma$ and $d\pi/d\sigma$ into (σ) and simplifying, we can solve for σ. This yields (10), the same expression as in the unitary state case. As shown before, this level of employment tax will ensure that the regional firms are following the efficient decision rule for e. In other words, regions will choose their employment tax policies so that the matching externality is correctly taken into account. However, since the allocation of firms may differ between the SPNE and the social optimum, this alone does not ensure that firm employment will be socially optimal.

Next, consider the regional government's choice of the tax on firms. We can derive an expression for t by substituting (t) into (N):[19]

$$t=-(w+\sigma)N\frac{de}{dN}-Ne\frac{dw}{dN}+(w+\sigma)e-N\frac{d\pi}{dN}+(N-\widehat{N})\frac{d\pi_2}{dN}$$

[19] The details of this analysis are available on request in the Technical Appendix.

Then, substituting in for w, σ, $d\pi/dN$, de/dN, and dw/dN from (5), and simplifying:

$$t = (\widehat{N} - N)\frac{d\pi_2}{dN_2} - Na'(N)f(e) \tag{15}$$

An analogous expression is obtained for region 2. Recall from (14) above that to ensure a socially optimal allocation in the unitary state, regional firm taxes should be set such that $t_2 - t_1 = N_1a'(N_1)f(e_1) - N_2a'(N_2)f(e_2)$. Thus, since the first term on the right-hand side of (15) is not expected to be zero, regional governments generally set their taxes on firms non-optimally. Although the regions take proper account of the production agglomeration effect (the second term in (15), they use t to attract firms to their regions.[20] Given that $d\pi_2/dN_2 < 0$ by the stability of the location equilibrium, the sign of the tax competition effect depends on the number of firms locating in the region relative to the number of firms owned by individuals in that region.

If, as we have assumed above, the firms are owned by the L members of the workforce, all profits will end up being consumed by them in the two regions. The allocation of resources will differ from the social optimum analyzed above for two reasons. First, national output will not be maximized because firms locate inefficiently between the two regions. Second, the relative levels of consumption between households in the two regions will differ from the social optimum since there is no mechanism to redistribute consumption between regions. If firms are not owned by households, but by owners who do not count in regional social welfare functions, then there will be a third problem. Indeed, to the extent that the tax on firms is competed down, there will be positive after-tax profits that are not available for consumption.

To gain further insight into the consequences of tax competition, it is useful to characterize the SPNE in the special case of no agglomeration. Recall that without agglomeration ($a' = 0, \rho = 1$), the social optimum is symmetric in per firm terms, which implies that firm employment and profits are the same in both regions and $L_1/N_1 = L_2/N_2$. There is no need for differential firm tax rates between regions. Consider first the case in which firms are owned by members of the workforce. Then, by (15), regions choose tax rates according to the following tax rules in the SPNE:

$$t_1 = (\widehat{N}_1 - N_1)\frac{d\pi_2}{dN_2} \qquad t_2 = (N_1 - \widehat{N}_1)\frac{d\pi_1}{dN_1}$$

where we have used the fact that $\widehat{N}_1 + \widehat{N}_2 = \overline{N} = N_1 + N_2$. Clearly, the region which is a net gainer of firms will want to tax them, and the net loser will want to subsidize them. This is analogous to the standard terms-of-trade effect in the capital tax competition literature. Whether the social optimum can be supported as an SPNE depends on the difference between the distribution of firm ownership $\widehat{N}_i$ and the location of firms within each region in the efficient allocation, say, N_i^o. If it is the case that $\widehat{N}_i = N_i^o$ for $i = 1, 2$, then both regions will set their tax rates to zero and the SPNE will be efficient. For example, suppose as assumed in Bucovetsky (1991) and Wilson (1991) that individuals in both regions own an equal share of all the firms in the economy, so $\widehat{N}_i = L_i\overline{N}/(L_1 + L_2)$. Since in the social optimum $L_1/N_1 = L_2/N_2 = (L_1 + L_2)/\overline{N}$ for $L_1 \geq L_2$ we

[20] Note regions do not use employment taxes or subsidies to compete for firms: σ is chosen so as to internalize the matching externality. If regions were for some reason precluded from using a tax on firms, say because it was assigned to the central government, then regions would be induced to choose non-optimal employment policies (see Boadway, Cuff, and Marceau, 2002).

have $\widehat{N}_i = N_i^o$. The efficient allocation of firms can be supported as a SPNE.[21] On the other hand, if $\widehat{N}_i \neq N_i^o$, then starting from the social optimum allocation of firms, regions will have an incentive to set differential firm taxes and the SPNE will be inefficient. Depending on the distribution of firm ownership, one region will tax firms and the other will subsidize. Or, if firms are owned by non-workers whose utility does not count, then both regions will tax firms.[22]

The upshot is that in the absence of agglomeration effects, firm allocation can be efficient in the SPNE provided households in a given region own the same number of firms as are located in that region in the social optimum. This continues to be true in the presence of matching externalities.[23]

Allowing instead for a production externality gives both regions an incentive to subsidize firms regardless of the distribution of firm ownership, and therefore will generally lead to an inefficient distribution of firms.

5.2. Central Government Intervention

It is clear that in a fully decentralized regional setting, an inefficient allocation of resources occurs primarily because firms are allocated inefficiently between regions. In addition, the relative levels of consumption in the two regions may differ from what might be regarded as socially optimal. We now examine how a first-mover central government could decentralize the social optimum in a federal setting. We examine this problem assuming that the central government's policy instruments include taxes on firms in each region, denoted by (T_1, T_2), and an inter-regional transfer from region 1 to region 2, denoted by B. It can select its policies before regional governments choose theirs.

The location equilibrium condition will now be given by,

$$\pi_1(N_1, \sigma_1, V_1(N_1, \sigma_1), s_1(N_1, \sigma_1)) - t_1 - T_1 - \pi_2(\bar{N} - N_1, \sigma_2, V_2(\bar{N} - N_1, \sigma_2), s_2(\bar{N} - N_1, \sigma_2)) + t_2 + T_2 = 0 \tag{16}$$

which determines $N_1(\sigma_1, t_1, T_1, \sigma_2, t_2, T_2)$ where $\partial N_1/\partial t_i = \partial N_1/\partial T_i$ for $i = 1, 2$. As before, the regional governments must select (c_i, t_i, σ_i) to maximize regional welfare. We envision Stage 1 (Government Policies) as being decomposed into two sub-stages:

Stage 1a: Choice of (T_1, T_2, B) by the central government;

Stage 1b: Choice of $(c_1, t_1, \sigma_1, c_2, t_2, \sigma_2)$ by the regional governments.

---

[21] In Bucovetsky (1991) and Wilson (1991), the SPNE's are inefficient. These inefficiencies arise because regional governments are restricted to use only capital/firm taxes rather than capital/firm *and head taxes*, where we interpret capital in their models as firms in our's.

[22] If regions are identical, then the distribution of firms will be efficient in the SPNE since firm taxes will be the same. If regions have different populations, then firm taxes will differ and there will be an inefficient allocation of firms in the SPNE. In both cases, profits will be diverted from the workers given the minimum profit constraint is not binding.

[23] Recall, with same-sized regions, $N_1 = N_2$ in the social optimum. Therefore, if $\widehat{N}_i = \overline{N}/2$, the distribution of firms in the SPNE with matching agglomeration will be efficient. With regions that differ in size, $L_1/N_1 < L_2/N_2$ in the social optimum. The number of firms owned by workers in region 1 must be disproportionately larger than the number owned in region 2 for the SPNE to have an efficient allocation of firms.

At Stage 1b, central taxes (T_1, T_2) and the inter-regional transfer B have already been set and are taken as given by regional governments. As well, each regional government takes as given the policy choices of the other regional government. The problem of the representative region's government (taken again as region 1) is as above except that the inter-regional transfer B enters negatively into its revenue constraint and the location equilibrium is given by (16). From the first-order conditions of this problem, we obtain the reaction functions of region 1. These will be independent of the inter-regional transfer B: $t_1 = R_1^t(t_2,\sigma_2; T_1, T_2)$ and $\sigma_1 = R_1^\sigma(t_2, \sigma_2; T_1, T_2)$. By a similar approach, we can solve for the reaction functions for region 2: $t_2 = R_2^t(t_1, \sigma_1; T_1, T_2)$ and $\sigma_2 = R_2^\sigma(t_1, \sigma_1; T_1, T_2)$. A Nash equilibrium of Stage 1b is then a 4-tuple $\{t_1^N(T_1,T_2)$, $\sigma_1^N(T_1,T_2)$, $t_2^N(T_1,T_2)$, $\sigma_2^N(T_1,T_2)\}$ such that $t_i^N(T_1,T_2)=R_i^t(t_j^N(T_1,T_2)$, $\sigma_j^N(T_1,T_2); T_1, T_2)$ and $\sigma_i^N = R_i^\sigma(t_j^N(T_1,$ $T_2)$, $\sigma_j^N(T_1, T_2); T_1, T_2)$, for $i = 1, 2$, $i \neq j$. Thus, depending on (T_1, T_2), different Nash equilibria will arise in Stage 1b.

In Stage 1a, the central (federal) government anticipates how its policies will affect the Nash equilibrium solution for the regions' policies. Recall that the regions will use the efficient rule in choosing employment taxes σ_i. The central government can then achieve a social optimum if it can induce the regions to choose policies that result in the efficient allocation of firms between regions, and if all profits are diverted appropriately to households in the two regions. The instruments available for achieving these two objectives are the choice of taxes on firms (T_1, T_2) and inter-regional transfers B. Consider these in turn.

Denote by N_1^o the number of firms located in region 1 in the social optimum. Then, the federal government must choose (T_1, T_2) such that the following two equations are simultaneously satisfied:

$$N_1^o = N_1\bigg(\sigma_1^N(T_1, T_2), t_1^N(T_1, T_2), T_1, \sigma_2^N(T_1, T_2), t_2^N(T_1, T_2), T_2\bigg)$$
$$\pi_1(\cdot) - t_1^N(T_1, T_2) - T_1 = \pi_2(\cdot) - t_2^N(T_1, T_2) - T_2$$

where the latter equation must be equal to zero when the households do not own the firms. If this system of two equations has a unique solution, then the federal government will clearly be capable of replicating the social optimum. However, uniqueness of the solution to this system of equations itself rests on the uniqueness of the Stage 1b Nash equilibrium, and it is not possible to guarantee it at the level of generality of our analysis.[24]

If the Stage 1b Nash equilibrium is unique, then once the optimal number of firms has been implemented, and all profits are either in the hand of households or taken as taxes by the two levels of government, the federal government could simply choose its interregional redistribution scheme B to achieve whatever relative levels of consumption c_1 and c_2 it desires. Changes in B induce movements along the national UPF as in the unitary state case

Finally, it is worth noting the consequences of assuming that the regions do not have access to both σ and t as policy instruments. Without using σ, the regions would not internalize the matching externality on labour markets. However, this does not affect the ability of the central government to ensure that the social optimum is achieved, provided the regional governments' game (Stage

[24] Totally differentiating the first-order conditions of each region's problem (four equations), yields a 4×4 Jacobian matrix. If this Jacobian matrix is negative quasi-definite, then the solution to the four first-order conditions is the unique Nash equilibrium. See Friedman (1986) for more details on Nash equilibrium uniqueness.

1b) yields a unique Nash equilibrium. Assuming that the central government can set differential employment subsidy rates in the two regions, it can set them at their socially optimal values. On the other hand, if regions do not have access to taxes on firms, it is straightforward to show that they will choose their employment subsidies non-optimally.[25] That is, they will use employment subsidies as a device for attracting firms, setting the subsidy rate at too low a level. The social optimum can only be achieved if the central government also has access to employment subsidies in the two regions that it can use to ensure that the overall employment subsidy is optimal. If employment subsidies are not available to the central government, it will not be possible to achieve the social optimum.

6. CONCLUDING REMARKS

There are several natural extensions to the above analysis. First, in our model, all workers are identical and participate in the labour force. In the real world, job creation is often thought of as an instrument for inducing marginal workers to participate in the labour force, and can be viewed as a form of redistribution. To examine this aspect of policy, there must be some worker heterogeneity. A standard approach is to allow different worker abilities. We would then require a more general production function which uses different skilled individuals as different inputs or allows for separate production processes. Alternatively, we could allow workers to have differing preferences. Second, the process giving rise to unemployment could be modeled differently. For example, we could have allowed the matching technology to assign job matches randomly across firms rather than the same number per firm. Alternatively, an efficiency wage argument could have been used to explain unemployment. Third, we have adopted the extreme assumption of immobile labour. Allowing for labour mobility would raise similar issues as those we have discussed in this paper. Fourth, we have assumed a fixed number of firms in the federation. If the number of firms were endogenous, then the social optimum would have some optimal number of firms being created. Obviously, when there is fiscal decentralization there is no reason to expect that the number of firms in equilibrium will be optimum. Some additional central instrument will no doubt be required to ensure a socially optimal allocation.

Finally, we have adopted the assumption of perfect information on the part of governments for simplicity. If governments could not monitor whether individuals were looking for work, then full unemployment insurance would not result and the full social optimum could not be decentralized even to a unitary state. In the case, the analysis could be carried out using the benchmark of an informationally constrained second-best social optimum, and the same sorts of qualitative results could be derived.

[25] This is analyzed in some detail in Boadway, Cuff and Marceau (2002).

7. APPENDIX

7.1 Dynamic Problem of the Firm

In the basic model presented here, we abstract from agglomeration effects and ignore government policies. Incorporating them is straightforward.

The basic notation we use is as follows:

$n_t =$ number of workers at beginning of period t
$v_t =$ number of vacancies posted in period t
$e_t =$ number of of workers used in production in period t

<u>Timing during period t</u>

- ⋄ Firm starts period t with n_t workers.
- ⋄ Separation occurs: $(1-b)n_t$ workers stay with the firm; bn_t workers leave the firm.
- ⋄ Firm chooses hiring strategy (v_t, r_t).
- ⋄ Aggregate variables (w_t, m_t) are determined.
- ⋄ Fixed cost $\psi[(1-b)n_t + v_t]$ is incurred.
- ⋄ Matching takes place: $r_t m_t v_t$ vacancies are filled; $(1 - r_t m_t)v_t$ vacancies remain unfilled.
- ⋄ Production takes place with $e_t = (1-b)n_t + r_t m_t v_t$ workers.

The transition equation is simply: $n_{t+1} = e_t$

<u>The problem of the firm</u>

We have: $e_t = (1-b)n_t + r_t m_t v_t$. It follows that $v_t = [e_t - (1-b)n_t]/r_t m_t$ and that the transition equation can be written as $n_{t+1} = e_t$.

The problem of the firm can be written as:

$$V(n_t) = \max_{e_t, r_t} \quad f(e_t) - r_t w_t e_t - \psi\left((1-b)n_t + \frac{e_t - (1-b)n_t}{r_t m_t}\right) + \beta V(n_{t+1})$$

$$\text{s.t.} \quad n_{t+1} = e_t \qquad n_0 \quad \text{given}$$

The Euler equations are:

$$e_t: \quad f'(e_t) - r_t w_t - \frac{\psi}{r_t m_t} + \beta V'(n_{t+1}) = 0$$

$$r_t: \quad -w_t e_t + \frac{\psi}{r_t}\frac{[e_t - (1-b)n_t]}{r_t m_t} = 0$$

Using the Benveniste-Scheinkman formula, we can re-write the Euler equations as:

$$e_t: \quad f'(e_t) - r_t w_t - \frac{\psi}{r_t m_t} + \beta(1-b)\psi\frac{[1 - r_{t+1}m_{t+1}]}{r_{t+1}m_{t+1}} = 0 \tag{A1}$$

$$r_t: \quad -w_t e_t + \frac{\psi}{r_t}\frac{[e_t - (1-b)n_t]}{r_t m_t} = 0 \tag{A2}$$

Steady state analysis

Imposing $r = 1$ and eliminating time subscripts, we obtain:

$$\begin{aligned} e &= (1-b)n + mv \\ n &= e \end{aligned}$$

which in turn leads to:

$$be = mv \tag{A3}$$

The Euler equations in the steady state when $r = 1$ become:

$$f'(e) - w - \frac{\psi}{m} + \beta(1-b)\psi\frac{(1-m)}{m} = 0 \tag{A1$'$}$$

$$-we + \psi\frac{[e-(1-b)n]}{m} = 0 \tag{A2$'$}$$

Combining (A3) and (A2′), we obtain:

$$w = \frac{\psi b}{m}$$

Now combining (A1′) when $\beta = 1$ with (A2′) yields:

$$f'(e) = \psi\left[(1-b) + \frac{2b}{m}\right]$$

which is (2) in the text in the absence of agglomeration effects and employment subsidies.

REFERENCES

Boadway, R., K. Cuff, and N. Marceau. (2003). "Redistribution and Employment Policies with Endogenous Unemployment." *Journal of Public Economics*, forthcoming.

Boadway, R., K. Cuff, and N. Marceau. (2002). "Inter-Jurisdictional Competition for Firms." *International Economic Review* **43**, 761–782.

Bucovetsky, S. (1991). "Asymmetric Tax Competition." *Journal of Urban Economics* **30**, 167–81.

Diamond, Peter A. (1982). "Aggregate Demand Management in Search Equilibrium." *Journal of Political Economy* **90**, 881–894.

Friedman, J.W. (1986). *Game Theory with Applications to Economics*. New York: Oxford University Press.

Haufler, A. and I. Wooton. (1999). "Country Size and Tax Competition for Foreign Direct Investment." *Journal of Public Economics* **71**, 121–139.

Johnson, G.E. and P.R.G. Layard. (1986). "The Natural Rate of Unemployment: Explanation and Policy." In O. Ashenfelter and P.R.G. Layard (eds.), *Handbook of Labor Economics*. Amsterdam: North-Holland, 921–999.

Kind, H., K. Knarvik, and G. Schjelderup. (2000). "Competing for Capital in a 'Lumpy' World." *Journal of Public Economics* **78**, 253–274.

Krugman, P. (1993). *Geography and Trade*. Cambridge: MIT Press.

Ludema, R. and I. Wooton. (2000). "Economic Geography and the Fiscal Effects of Regional Integration." *Journal of International Economics* **52**, 331–357.

O'Sullivan, A. (1993). *Urban Economics*, Second Edition. Boston: Irwin.

Pissarides, C.A. (2000). *Equilibrium Unemployment Theory*, Second Edition. Oxford: Basil Blackwell.

Wilson, J. D. (1991). "Tax Competition with Interregional Differences in Factor Endowments." *Regional Science and Urban Economics* **21**, 423–51.

Wilson, J. D. (1999). "Theories of Tax Competition." *National Tax Journal* **52**, 269–304.

8. TECHNICAL APPENDIX

8.1 The Social Optimum

The first-order conditions are:

$$(c_1) \quad u'(c_1) - \lambda L_1 = 0$$
$$(c_2) \quad \varphi u'(c_2) - \lambda L_2 = 0$$
$$(e_i) \quad \lambda[N_i a(N_i) f'(e_i) - \psi(1-b)] - \delta_i N_i - \gamma_i b N_i = 0 \quad i = 1,2$$
$$(v_i) \quad -\lambda\psi - \delta_i s_i + \gamma_i \rho (N_i v_i)^{\rho-1} m(s_i) = 0 \quad i = 1,2$$
$$(s_i) \quad -\delta_i N_i v_i + \gamma_i (N_i v_i)^{\rho} m'(s_i) = 0 \quad i = 1,2$$
$$\begin{aligned}(N_1) \quad \lambda\Big[& a(N_1)f(e_1) - \psi((1-b)e_1 + v_1) + N_1 a'(N_1) f(e_1) - a(N_2) f(e_2) + \psi((1-b)e_2 + v_2) \\ & - N_2 a'(N_2) f(e_2)\Big] - \delta_1(e_1 + s_1 v_1) + \delta_2(e_2 + s_2 v_2) \\ & + \gamma_1(\rho N_1^{\rho-1} v_1^{\rho} m(s_1) - b e_1) - \gamma_2(\rho N_2^{\rho-1} v_2^{\rho} m(s_2) - b e_2) = 0\end{aligned}$$

Using (s_i), (v_i), and (e_i), we obtain:

$$a(N_i) f'(e_i) = \psi\left[1 - b + \frac{m'(s_i) + b(N_i v_i)^{1-\rho}}{m(s_i)(\rho - \eta)}\right] \qquad \text{(TA1)}$$

Using (e_i), (N_1), and (v_i), we get:

$$a(N_1)(f(e_1) - e_1 f'(e_1)) + N_1 a'(N_1) f(e_1) = a(N_2)(f(e_2) - e_2 f'(e_2)) + N_2 a'(N_2) f(e_2) \qquad \text{(TA2)}$$

Second-Order Conditions of the Social Optimum

These conditions are relevant since agglomeration effects can give rise to non-convexities that might entail corner solutions or multiple local optima. To characterize the second-order conditions of the social optimum, it turns out to be easier to suppress c from the above problem and assume that the objective is to maximize total product rather than utility. This gives the same solution as above without specifying a particular point on the economy-wide UPF. The problem can be further simplified by disaggregating the planner's problem into two steps. In the first, the planner assigns firms to the two regions and in the second, each regional planner optimizes with respect to employment and vacancies given the matching technology and the labour market-clearing condition. The analysis proceeds in reverse order starting with the second step.

Step 2: Regional Planning Optimum

In this second step, the representative regional planner given the size of its population L takes the number of firms in its region N as given and solves the following problem:

$$\max_{\{e,v\}} \quad N[a(N)f(e) - \psi((1-b)e + v)] \qquad \text{subject to} \qquad (Nv)^{\rho} m(s(e,v,N)) - bNe = 0 \quad (\gamma)$$

where $s(e, v, N, L) = (L - Ne)/Nv$.

The first-order necessary conditions are:

$$N[a(N)f'(e) - \psi(1-b)] + \gamma\left[(Nv)^{\rho}m'(s)s_e - bN\right] = 0 \tag{e}$$
$$-N\psi + \gamma[\rho(Nv)^{\rho-1}Nm(s) + (Nv)^{\rho}m'(s)s_v] = 0 \tag{v}$$

Define $z = G(e,v) = N[a(N)f(e) - \psi((1-b)e + v)]$ and $g(e,v) = bNe - (Nv)^{\rho}m(s(e,v,N))$. The second-order sufficient condition for a relative maximum of z is that d^2z is negative definite subject to $dg = 0$ which holds if and only if the determinant of the bordered Hessian matrix evaluated at the solution to the necessary conditions is positive. This determinant is given by the following expression:

$$|H| = -g_e[g_e(G_{vv} - \gamma g_{vv}) - g_v(G_{ev} - \gamma g_{ev})] + g_v[g_e(G_{ve} - \gamma g_{ve}) - g_v(G_{ee} - \gamma g_{ee})]$$

Differentiating g and F:

$$g_e = N[(Nv)^{\rho-1}m'(s) + b] > 0$$
$$g_v = -(Nv)^{\rho}\frac{m(s)}{v}(\rho - \eta) < 0$$
$$g_{ee} = -(Nv)^{\rho}m''(s) > 0$$
$$g_{vv} = -(Nv)^{\rho}\frac{m(s)}{v^2}(\rho - \eta)(\rho - \eta - 1) > 0$$
$$g_{ev} = g_{ve} = (Nv)^{\rho}\frac{m(s)}{v^2}(\rho - \eta) > 0$$
$$G_{ee} = Na(N)f''(e) < 0$$
$$G_{vv} = G_{ve} = G_{ev} = 0$$

$|H|$ can then be rewritten as $|H| = g_e^2\gamma g_{vv} - 2g_e g_v \gamma g_{ev} - g_v^2 G_{ee} + g_v^2\gamma g_{ee} > 0$. Therefore, the second-order sufficient conditions are satisfied for a relative maximum.

Next, substituting the expression for γ into (e) and re-organizing terms, we obtain (TA1). Condition (TA1) together with the matching constraint (γ) determines $e(N,L)$ and $v(N,L)$. Totally differentiating these two expressions and using Cramer's Rule and the expressions for s_e, s_v, s_N, and s_L, we get:

$$\frac{de}{dN} = \frac{C_N F_v - C_v F_N}{C_e F_v - C_v F_e} \qquad \frac{dv}{dN} = \frac{C_e F_N - C_N F_e}{C_e F_v - C_v F_e} \tag{TA3}$$
$$\frac{de}{dL} = \frac{C_L F_v - C_v F_L}{C_e F_v - C_v F_e} \qquad \frac{dv}{dL} = \frac{C_e F_L - C_L F_e}{C_e F_v - C_v F_e} \tag{TA4}$$

where

$$C_e = -N(m'(s)(Nv)^{\rho-1} + b) < 0 \qquad C_N = be\left(1 - \rho + \eta + \frac{m'(s)}{m(s)}\frac{e}{v}\right) > 0$$
$$C_v = (Nv)^{\rho}\frac{m(s)}{v}(\rho - \eta) > 0 \qquad C_L = -(Nv)^{\rho-1}m'(s) < 0$$
$$F_e = -a(N)f''(e) - \psi\frac{m''(s)}{vm(s)(\rho - \eta)} + \psi m'(s)\frac{m'(s) + b(Nv)^{1-\rho}}{vm(s)^2(\rho - \eta)} > 0$$
$$F_v = -\psi\frac{sm''(s)}{vm(s)(\rho - \eta)} + \psi\frac{\eta m'(s) + b(Nv)^{1-\rho}(1 - \rho + \eta)}{vm(s)(\rho - \eta)} > 0$$
$$F_L = -\psi\left[\frac{m''(s)}{m(s)(\rho - \eta)} + m'(s)\frac{m'(s) + b(Nv)^{1-\rho}}{m(s)^2(\rho - \eta)}\right]\frac{1}{Nv} > 0$$
$$F_N = a'(N)f'(e) + \frac{L}{N^2v(\rho - \eta)}\left[\frac{\psi m''(s)}{m(s)} - \frac{m'(s)^2}{m(s)^2}\right] - \frac{\psi b(Nv)^{1-\rho}}{Nm(s)(\rho - \eta)}\left[(1 - \rho + \eta) + \frac{m'(s)e}{m(s)v}\right] \gtrless 0$$

The determinant is negative, $C_eF_v - C_vF_e < 0$. However, as the sign of F_N is ambiguous so too is the sign of de/dN. If there is *no production agglomeration*, then $a' = 0$ and $F_N < 0$ in which case $de/dN < 0$. However, if there is *production agglomeration* e may be increasing or decreasing in N. Regardless of the type of agglomeration, per firm employment is increasing in L, $de/dL > 0$.

Applying the Envelope Theorem to the maximum value function of the regional planning problem, $W(N)$, and combining with (TA1), we get:

$$\frac{dW}{dN} = a(N)f(e) + Na'(N)f(e) - ea(N)f'(e) \tag{TA5}$$

Step 1: Central Planning Optimum

The problem of the central planner is to maximize:

$$\max_{\{N_1\}} \quad W_1(N_1) + W_2(\overline{N} - N_1)$$

The first-order condition is:

$$\frac{dW_1}{dN_1} - \frac{dW_2}{dN_2} = 0$$

Using (TA5), this may be written as (TA2).

The second-order condition of the central planner's problem is then:

$$\frac{dW_1/dN_1 - dW_2/dN_2}{dN_1} < 0 \tag{TA6}$$

Using (TA2), the left-hand side of this condition can be written as:

$$\begin{aligned} &N_1a''(N_1)f(e_1) + a'(N_1)[2f(e_1) - e_1f'(e_1)] + [N_1a'(N_1)f'(e_1) - e_1a(N_1)f''(e_1)]\frac{de_1}{dN_1} \\ &+ N_2a''(N_2)f(e_2) + a'(N_2)[2f(e_2) - e_2f'(e_2)] + [N_2a'(N_2)f'(e_2) - e_2a(N_2)f''(e_2)]\frac{de_2}{dN_2} \end{aligned} \tag{TA7}$$

The sign of (TA7) is generally ambiguous.

<u>Characterizing the Social Optimum</u>

Case 1: No Agglomeration

Suppose there are no production agglomeration effects: $a' = 0$. Then, as shown above, $de/dN < 0$. The second-order condition of the social optimum is given by:

$$-e_1a(N_1)f''(e_1)\frac{de_1}{dN_1} - e_2a(N_2)f''(e_2)\frac{de_2}{dN_2} < 0 \tag{TA8}$$

which is satisfied everywhere. Therefore, the solution which satisfies the first-order conditions, (TA1) and (TA2), and the regional matching constraints, (γ_i) for $i = 1, 2$, is the unique global optimum.

In this case, the first-order necessary condition reduces to:

$$f(e_1) - e_1f'(e_1) = f(e_2) - e_2f'(e_2) \tag{TA2$'$}$$

Recall, that e is determined at Step 2 and depends on both N and L. Since $f(e) - ef'(e)$ is strictly increasing in e, (TA2′) implies that N_1 is set such that:

$$e(N_1, L_1) = e(\overline{N} - N_1, L_2)$$

Now, suppose $L_1 = L_2$. Then this condition can only be satisfied if $N_1 = \overline{N}/2$. As the second-order conditions are satisfied, this symmetric optimum (same e, v, and s) is unique and holds regardless of whether or not there are matching agglomeration effects. Note, however, that with matching agglomeration effects, $\rho > 1$, the value of e at the social optimum will differ from that in the case of no matching agglomeration.

Next, suppose that $L_1 > L_2$. Then, at $N_1 = N_2$, $e_1 > e_2$ since $de/dL > 0$. In order, to equalize the e's more firms must be located in Region 1 since $de/dN < 0$. Therefore, $N_1 > \overline{N}/2$. In this case, when $\rho = 1$ the symmetric outcome in terms of e, v and s is still the social optimum, but it no longer is if $\rho > 1$. To see this, first suppose that $\rho = 1$. Then, the first-order conditions from Step 2 and the regional matching constraints for both regions can be written as follows, using the fact that $m'(s)/m(s) = \eta/s$:

$$f'(e) = \psi \left[1 - b + \frac{\eta}{s_i(1-\eta)}\right] \quad i = 1, 2 \tag{TA1′}$$

$$v_i m(s_i) - be = 0 \quad i = 1, 2 \tag{γ_i}$$

Note that neither of these conditions depend on N_i explicitly. Given $e_1 = e_2 = e$, it follows from (TA1′) that $s_1 = s_2$ which then implies from (γ_i) that $v_1 = v_2$. From the definition of $s_i = (L_i/N_i - e_i)/v_i$ and the fact that $s_1 = s_2$ and $v_1 = v_2$, it must be that $L_1/N_1 = L_2/N_2$.

Case 2: Matching Agglomeration

Suppose $\rho > 1$, then the first-order conditions from Step 2 and the regional matching constraints for both regions can be written as follows, using the fact that $m'(s)/m(s) = \eta/s$:

$$f'(e) = \psi \left[1 - b + \frac{\eta}{s_i(\rho-\eta)} + \frac{b(N_i v_i)^{1-\rho}}{m(s_i)(\rho-\eta)}\right] \quad i = 1, 2 \tag{TA1″}$$

$$(N_i v_i)^{\rho} m(s_i) - bN_i e = 0 \quad i = 1, 2 \tag{γ_i}$$

Suppose $s_1 = s_2$ and $v_1 = v_2$ which implies from the definition of s that $L_1/N_1 = L_2/N_2$. This assumption violates the matching constraints since they can only both be satisfied if $N_1 = N_2$ which contradicts the fact that $N_1 > N_2$. Therefore, it cannot be that $s_1 = s_2$ and $v_1 = v_2$ which implies that $L_1/N_2 \neq L_2/N_2$ in the social optimum with matching agglomeration and regions that differ in size.

Substituting out s_i and using the matching constraint (TA1″), we obtain:

$$f'(e) = \psi \left[1 - b + \frac{\eta v_i}{((L_i/N_i) - e)(\rho-\eta)} + \frac{v_i}{e(\rho-\eta)}\right] \quad i = 1, 2 \tag{TA1‴}$$

The right-hand side of the above condition is increasing in v_i and decreasing in L_i/N_i. Therefore, since $e_1 = e_2 = e$ it must be that $v_1 > (<) v_2$ and $L_1/N_1 > (<) L_2/N_2$ (equality of these variables violates the matching constraint). Then, substituting out s_i using its definition and dividing by N^ρ, the matching constraint becomes:

$$v_i^\rho m((L_i/N_i - e)/v_i) - bN_i^{1-\rho}e = 0 \quad i = 1, 2 \qquad (\gamma_i)$$

The left-hand side is increasing in both v_i and L_i/N_i. Given $N_1 > N_2$, it follows from the matching constraint together with (TA1$'''$) that $v_1 < v_2$ and $L_1/N_1 < L_2/N_2$. More firms are located in the larger region relative to the case when there are no matching agglomeration. If instead we substituted v_i out of the conditions, then it would be the case that both the right-hand side of (TA1$'''$) and the matching constraint are increasing in s and L/N and therefore, $s_1 < (>) s_2$ and $L_1/N_1 > (<) L_2/N_2$. The fact that $N_1 > N_2$ does not seem to eliminate one of these possibilities. However, we know that $s_1 \neq s_2$ since if it did then L/N would be the same from (TA1$'''$) which violates the matching constraint. Therefore, given $L_1/N_1 < L_2/N_2$ it must be that $s_1 > s_2$.

Case 3: Production Agglomeration

Suppose $a' > 0$. Then, as shown above de/dN may be increasing or decreasing in N. In this case, the second-order condition of the social optimum is given by (TA7). Here, whether the above condition is satisfied or not will depend on the form of $a(N)$ and whether de/dN is positive or negative. Unlike the case with no production agglomeration, the sign of de/dN depends on the point it is evaluated at. Therefore, there may be multiple local optima for the social planner's problem when there is production agglomeration.

8.2 Federal State

From (t), we have $\mu = \lambda(N - \widehat{N})$. Substituting this and (5) into (σ), and solving for σ, we obtain:

$$\sigma = -\frac{\psi b V^{1-\rho}}{m(s)} + \frac{\psi}{m(s)(\rho-\eta)}(m'(s) + bV^{1-\rho}(1-\rho+\eta)) = \psi\frac{m'(s) + bV^{1-\rho}}{m(s)(\rho-\eta)} - 2\frac{\psi b V^{1-\rho}}{m(s)}$$

Next, substituting out for μ in (N) and solving for t, we have

$$t = -(w+\sigma)e - (w+\sigma)N\frac{de}{dN} - Ne\frac{dw}{dN} - N\frac{d\pi}{dN} + (N - \widehat{N})\frac{d\pi_2}{dN}$$

Noting that $de/dN = e_N + 2e_\sigma dw/dN$, $d\pi/dN = a'(N)f(e) - (2e)dw/dN$, $(w+\sigma) = \psi(V^\rho m'(s) + bV(1-\rho+\eta)/[V^\rho m(s)(\rho-\eta)]$, and substituting in for dw/dN, we can rewrite the above as:

$$\begin{aligned} & t + Na'(N)f(e) - (N-\widehat{N})\frac{d\pi_2}{dN} = \\ & -\psi\frac{X}{V^\rho m(s)(\rho-\eta)}\left[e + Ne_N + N2e_\sigma\frac{e+Ne_N}{\Delta}\frac{\psi b}{V^\rho m(s)}X\right] + Ne\frac{e+Ne_N}{\Delta}\frac{\psi b}{V^\rho m(s)}X \end{aligned}$$

where $X = V^\rho m'(s) + Vb(1-\rho+\eta)$.

Using the matching constraint $bNe = V^\rho m(s)$ and the definition of Δ, it can be shown that the right-hand side of the above is zero. Therefore:

$$t = (N-\widehat{N})\frac{d\pi_2}{dN} - Na'(N)f(e)$$