Queen's Economics Department Working Paper No. 941

Ex ante Free Mobility, Ex post Immobility, and Time-Consistent Policy in a Federal System

Kiyoshi Mitsui
Motohiro Sato

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

11-1996

Ex ante Free Mobility, Ex post Immobility, and Time-Consistent Policy in a Federal System*

Kiyoshi Mitsui** and Motohiro Sato***

preliminary

Nov. 19,1996

* We wish to thank Professor R. Boadway and the participants of the workshop at Queen's University for helpful and thoughtful comments. We also thank J. Spicer for comments and corrections. All remaining errors are the responsibility of the authors.
** Department of Economics, Queen's University, Kingston, Ontario, Canada and Department of Economics of Meijigakuin University, Tokyo, Japan. (e-mail address: mitsuik@qed.queensu.ca)
*** Department of Economics, Queen's University, Kingston, Ontario, Canada and Department of Economics of Hitotsubashi University, Tokyo, Japan

Abstract: This paper examines regional population distribution when there is an interregional transfer policy without commitment. We introduce explicitly the following time structure of actions. Individuals make decisions on locational choices freely *ex ante*, but are immobile *ex post*. The interregional transfer policies by regional governments and the central government are implemented after individuals' migration decisions. We obtain the following results. First, locally stable time-consistent equilibria are single-community equilibria when there is a pure local public good. When we extend the basic model by taking account of capital, congestion, and spillovers in the provision of a public good, it is shown that whether or not central government intervention enhances the efficiency of the population distribution depends upon several economic factors.

1. Introduction

There is a large literature on the efficiency of a federal system. Flatters, Henderson, and Mieszkowski (1974), Boadway (1982), and Boadway and Flatters (1982) focus on the fiscal externality or fiscal migration effects of a federal system. They examine how inefficient population distribution is when there is no central government intervention and then they show how the central government could make transfers from one region to another to achieve an efficient distribution of population. Myers (1990) takes into account the possibility of voluntary transfers by regional governments. He finds that competition among regional governments with voluntary transfers leads to an efficient regional population distribution. Central government is not necessary to improve the efficiency of the regional population distribution.

Traditional discussion has been based upon static models. Time structure is not explicitly introduced. Although such a framework provides a useful tool to capture the nature of a federal system, some attention must be paid to whether or not the outcomes derived from this setting can be extended to a dynamic analysis. In the present paper, we attempt this extension, focusing on a time-consistency issue, a topic of a growing concern in the public finance literature.

When individuals can move across regions, a time-consistency problem can arise due to a difference in timing between individuals' decision makings of location and fiscal decisions made at the national and local levels.[1] An individual's choice of residence is usually made based on a long-run perspective. When individuals decide where to live, they choose their occupations at the same time. There are large sunk costs to changing jobs. Individuals spend much time and energy searching for new jobs and, after getting jobs, they invest in human capital, a part of which is firm specific.[2]

[1] The time-consistency problem was introduced by macroeconomists interested in the issues of capital taxation and default of public debts. Kydland and Prescott (1977) and Fisher (1980) investigate time-consistency problems using representative consumer models. The development of the literature on the time-consistency problem is well surveyed in Chari, Kehoe, and Prescott (1989) and Persson and Tabellini (1990).

[2] Mansoorian and Myers (1993) and Wellisch (1994) investigate another sort of immobility. They examine the possibility that individuals have an attachment to a particular region for cultural or nationalistic reasons. Mansoorian and Myers (1993) show that it is not possible for the central

Given these consideration, we assume that individual mobility is imperfect *ex post*. This kind of *ex post* imperfect mobility is examined, for example, by Boadway and Wildasin (1990). They analyze occupational choice when industries are subject to technological uncertainty: workers can choose occupations freely *ex ante*, but must pay some cost to change their jobs once the uncertainty is removed. In the present paper, to simplify analysis, we make the stronger assumption that individuals are perfectly immobile *ex post*.

We assume the following time structure. First, the central government and local governments announce their policies. Then, individuals decide where to live. The central government, then, puts into practice its policy on interregional transfers and, finally, regional governments provide a public good in their regions and make voluntary transfers to other regions.

Placing our model in the literature, if regional governments can commit to their policies, our model is the same one analyzed by Myers (1990).[3] However, if regional governments cannot commit to their policies, voluntary transfers no longer occur. With respect to the central government, the central government, if it has the ability to commit, our model reduces to Boadway and Flatters (1982). The inter-regional transfers are made from an *ex ante* view point and thus net social products among regions will be equalized.[4] We study the case that neither level of government can make a credible commitment.

We assume that all individuals are homogeneous in all aspects except residence and provide a fixed unit of labour in their residence. Each local jurisdiction provides a single (local) public good financed by a head tax and / or property tax. We show that

---

government to enhance the efficiency of population distribution, even if individuals attachment to home is heterogeneous. The efficiency result remains even when there is an attachment to home. Wellisch (1994) takes into account the interregional spillovers in the provision of the public good. He concludes that (i) if individuals are perfectly mobile across regions, then decentralized public good provision is socially efficient, and (ii) if individuals are attached to particular regions, then central government intervention can improve efficiency.

[3] When local governments can commit to their policies and design voluntary transfers, the Pareto efficient population distribution is achieved by voluntary transfers because of the strong incentive equivalency among local governments. Even in the presence of central government intervention, local governments can redesign their transfer scheme to offset its policy.

[4] The equalization transfer is practiced from an efficiency concern, since *ex ante* all households (workers) are homogeneous in all respects.

in such a circumstance, the centrally organized inter-regional transfers (or equalization grants) can further distort population distribution and, in particular, are likely to induce the population concentration in a specific region. Although the non-intervention equilibrium is inefficient, national policy can only make the situation worse. This contradicts with the traditional normative view which favours central government intervention when the inefficiency arises in a federal system. Myers (1990) casts a doubt on such a view in a sense that the central intervention is not necessary to achieve the efficient population distribution. Our conclusion casts a further doubt on such a view: the time consistent issue can make the judgment of the desirability of national policy substantially difficult.

One of the keys to our conclusion stems from the difference in the purpose of interregional transfers between *ex ante* and *ex post*. All individuals are identical before they decide residence and thus there is no equity concern *ex ante*. However, all people may not be identical *ex post*. Since population distribution is already fixed, any transfer is neutral and purely redistributive from an *ex post* view point. In contrast to the recommendation by Boadway and Flatters (1982), the equalization grants will not be based on efficiency concerns, but used instead to promote *ex post* equity. Since the central authority faces immobility of individuals *ex post*, its optimization is simply characterized by a social welfare maximization subject to resource constraints in the economy as a whole.

Another key to obtaining our conclusion is the scale merit of consumption in a local public good. The benefit of a public good cannot be redistributed across regions. Then the consumption level of a public good in a region is always the same as the provision level in the region. Moreover, since a public good has publicness in its consumption, a relatively large community can spread the cost of the public good provision more widely, resulting in a cheaper per capita tax price for a given level of public good provision This scale merit of consumption in a public good prevents full equalization of utility levels across regions when there are population gaps among regions. In order to maximize social welfare, the central government will make inter-regional transfers so that the utility level of a relatively high populated region remains higher than that of a relatively less populated region. As supposed in the literature of time-consistency, individuals, being rational, anticipate *ex post* redistributive policy correctly and thus move to a congested region *ex ante*, resulting in greater population concentration.

Then, we examine some extensions of our model. We investigate how the introduction of capital, congestion cost, and spillovers in the provision of a public good affects the results. In particularly, we consider the effect on our result when we allow spillovers into neighboring regions. Wellisch (1994) argues that if individuals are perfectly mobile, decentralized provision of a public good is socially efficient regardless of the spillover effect and that if individuals are attached to particular regions, central government intervention can enhance efficiency. In our model, if we assume that the central government can commit to its policy, the results are be the same. However, this is no longer necessarily true if the central government has no commitment power and the spillover effect is weak.

Section 2 presents the setup of the basic model. We describe first the time structure of our model. The local government problem and the central government's problem are discussed when they cannot commit to policy. Then we provide the definition of a migration equilibrium and characterize one. Section 3 is devoted to extending our model to take into account capital as an input, congestion costs, and spillovers in the provision of a public good. In section 4, we compare the utility when there is and when there is not central government intervention. We investigate an example with a log-linear utility function and Cobb-Douglas production functions. Section 5 provides the concluding remarks.

2. Basic Model

A nation consists of M regions, indexed $i = 1, \cdots, M$. We assume a continuum of individuals. The population (and labour force) in region i is denoted by n_i and the aggregate population is fixed at $N = \sum_{i=1}^{M} n_i$. Individuals are assumed to be endowed with one unit of labour and to have homogeneous tastes. Their preference are defined by a strictly increasing and strictly concave utility function, $u_i = u(x_i, G_i)$, where x_i is the consumption of a private good and G_i is the quantity of a public good provided in region *i*. It is assumed that $u_{xG} \geq 0$ which implies x_i and G_i are normal goods and that the public good is a pure public good within the region. We assume a concave production function, $f_i(n_i)$ with $f_i' > 0$ and $f_i(0) = 0$. The marginal rate of transformation between the private good and public good is assumed to be unity.

There are two levels of governments. Local government *i* chooses its public good provision, G_i, and non-negative (voluntary) transfers to other regions. The central government determines a lump-sum transfer, S_i $\left(\sum_{i=1}^{M} S_i = 0\right)$, for each region. We assume governments are benevolent. Local governments maximize utility for their residents and the central government maximizes social welfare. Moreover, we postulate the rent sharing case in the sense that all the output value of a region accrues to its own residents and the local public good is financed from its citizens by means of a lump-sum tax.

To capture the *ex ante* mobility and *ex post* immobility, we suppose the following time structure, illustrated in figure 1. At stage 0, the central government announces its transfer policy $\tilde{S} = (S_1 \cdots, S_M)$ to regions and regional governments announce their policies which consist of the level of local public good G_i and any voluntary transfers they will make to other regions. At stage 1, individuals decide where to reside, producing a population distribution $\tilde{n} = (n_1, \cdots, n_M)$. At stage 2, the central government enacts its transfer policy and at stage 3, regional governments execute their policy.

Figure 1 is around here

We consider the case in which there is no commitment technology for either the central government or local governments.[5] We assume sequential rationality for both private and public agents. Backward optimization, therefore, produces a consistent time path for policy and private agent decisions. First, we begin with the optimization problem facing local governments. They take as given the decisions of both the central government and private agents. Second, the problem of the central government is solved, taking as given the decisions of individuals and the optimal behavior of local governments. Finally, the individual's problem is examined with the assumption of rational expectation in the sense that individuals forecast future policies as being sequentially rational for the central government and local governments.

2.1. Local Government Problem

Since local governments cannot commit to their policy, we can preclude the possibility of voluntary transfers. For given n_i and $\widetilde{S}$, local government i chooses x_i and G_i to maximize $u(x_i, G_i)$ subject to the budget constraint:

$$n_i \cdot x_i + G_i = f_i(n_i) + S_i . \qquad (1)$$

Local governments act on behalf of their residents. By substituting (1) for x_i into the utility function, the maximization problem becomes

$$\text{(I)} \quad \max_{G_i} u\left(\frac{f_i(n_i) + S_i - G_i}{n_i}, G_i\right)$$

The optimal mix of private good and public good is obtained by the following first order condition for (I):

$$n_i \cdot \frac{u_G(x_i, G_i)}{u_x(x_i, G_i)} = 1 \qquad (2)$$

for any $i \in \Lambda^+(\widetilde{n})$ where $\Lambda^+(\widetilde{n})$ denotes the index set of regions which have positive population. This is a familiar Samuelson condition. We denote by $x_i(S_i, n_i)$ and

[5] As mentioned in the introduction, our model can be placed in existing works when either the central or local governments can commit credibly to policy. The relationship among Boadway and Flatters (1982), Myers (1990), and ours are examined in appendix A.

$G_i(S_i, n_i)$ the solutions to (1) and (2). For any $i \in \Lambda^+(\tilde{n})$, the indirect utility function is

$$V_i(S_i, n_i) = u(x_i(S_i, n_i), G_i(S_i, n_i)). \tag{3}$$

2.2 Central Government Problem

Now we move to the optimization problem of the central government. A social welfare function is defined as

$$SW = \sum_{i \in \Lambda^+(\tilde{n})} n_i \cdot W(u_i) \tag{4}$$

where $W' > 0$ and $W'' \leq 0$. The welfare weights are assumed to be the same for all individuals. When $W'' = 0$, the social welfare function (4) reduces to the utilitarian social welfare function. Although the formulation of the social welfare function (4) is fairly general, it does not include the Rawlsian social welfare function, which exhibits infinite elasticity of inequality aversion. Introducing such a form of social welfare will change our argument. We turn to it later.

The central government maximizes social welfare with respect to its transfer policy:

$$(\text{II}) \quad \max_{S_i} \sum_{i \in \Lambda^+(\tilde{n})} n_i \cdot W(V_i(S_i, n_i))$$

subject to $\sum_{i=1}^{M} S_i = 0$ and $S_i + f_i(n_i) \geq 0$. It follows that $S_i = 0$ for any $i \in \Lambda^0(\tilde{n})$, where $\Lambda^0(\tilde{n})$ denotes the index set of empty regions. The first order condition associated with (II), applying the envelope theorem, implies

$$W'(V_j(S_j, n_j)) \cdot u_x(x_j^*, G_j^*) = W'(V_k(S_k, n_k)) \cdot u_x(x_k^*, G_k^*). \tag{5}$$

for any $i, k \in \Lambda^+(\tilde{n})$, where $x_i^* = x_i(S_i, n_i)$ and $G_i^* = G_i(S_i, n_i)$. Since the allocation of private and public goods is determined efficiently within each region according to the Samuelson rule, the marginal social welfare of a region induced by the increase of transfer to the region can be measured in terms of the marginal utility of the private good. Denote the solution of (II) by $S_i(\tilde{n})$ and let $V_i(\tilde{n})$ be the maximum value function associated with problem (II): $V_i(\tilde{n}) = V_i(S_i(\tilde{n}), n_i)$.

The following proposition states that individuals in a highly populated region have a higher utility than those in a lowly populated region.

Proposition 1. For a given $\tilde{n} = (n_1, \cdots, n_M)$,

(i) if $n_j > n_k > 0$, then $V_j(\tilde{n}) > V_k(\tilde{n})$, and

(ii) if $n_j = n_k > 0$, then $V_j(\tilde{n}) = V_k(\tilde{n})$

PROOF: See appendix B.

To consider the intuition for this result, let $W'' = 0$ and $u_{xG} = 0$. In this case, the consumption of the private good is fully equalized from (5) and the transfers are made so as to equalize private good consumption among regions. The pureness of the public good implies economies of scale: per capita cost of providing one unit of the public good is less in more populated regions. Thus, the consumption level of the public good is higher in the relatively highly populated regions. The interregional transfer policy, therefore, leads to the higher utility in relatively highly populated regions.

Explanation of this proposition can be presented for two region case $(M = 2)$. The social welfare function is

$$SW = n_1 \cdot W(V_1(S_1, n_1)) + n_2 \cdot W(V_2(-S_1, n_2)) \quad . \tag{6}$$

From the strict concavity of the utility function and the social welfare function, it follows

$$\frac{d^2 SW}{d S_1^2} < 0 . \tag{7}$$

Figure 2 is around here

Define $\hat{S}_1$ such that $V_1(\hat{S}_1, n_1) = V_2(-\hat{S}_1, n_2)$ $[\equiv \hat{V}]$. Figure 2 illustrates the indifference curve $u(x_i, G_i) = \hat{V}$. Suppose $n_1 > n_2$. Then, under the interregional transfer $\hat{S}_1$, region i chooses the private and public goods mix (x_i, G_i) according to the Samuelson rule. Since $n_1 > n_2$, the relative price of the private good in region 1 is higher than that in region 2. Thus, we have $\hat{x}_1 < \hat{x}_2$ and $\hat{G}_1 > \hat{G}_2$ from the strict concavity of the utility function. Individuals in region 1, therefore, consume less private goods but consume more public good than those in region 2 if the utility levels of both residents are the same. Therefore, we have

$$u_x(\hat{x}_1, \hat{G}_1) > u_x(\hat{x}_2, \hat{G}_2). \tag{8}$$

from $u_{xx} < 0$ and $u_{xG} \geq 0$. In other terms, if the utility levels in both regions were the same, the marginal utility of the private good would be higher in the more populated region than in the less populated region.

Evaluating the derivative dSW / dS_1 at $S_1 = \hat{S}_1$ yields

$$\frac{dSW}{dS_1} = W'(\hat{V}) \cdot \{u_x(\hat{x}_1, \hat{G}_1) - u_x(\hat{x}_2, \hat{G}_2)\} \tag{9}$$

from the envelope theorem. Combining (8) and (9) implies $dSW / dS_1 > 0$ at $S_1 = \hat{S}_1$. Thus, the optimal transfer, $S_1(\tilde{n})$, is greater than $\hat{S}_1$ as a result of the concavity of social welfare function (see (7)). We derive, therefore, that $V_1(\tilde{n}) > V_2(\tilde{n})$ since $V_i(S_i, n_i)$ is a strictly increasing function with respect to S_i (see figure 3). Consequently, we conclude that the benevolent government should implement the interregional transfer policy to make the utility of the high populated region higher than that of the low populated region.[6]

Figure 3 is around here

It would be worth investigating the implication of proposition 1 in more detail. Let $x_i(\tilde{n})$ and $G_i(\tilde{n})$ be respectively consumption of a private good and a public good under the optimal policy on interregional transfers. Consider the case where $n_1 > n_2$. Then,

---

[6] Atkinson and Stiglitz (1980) show that the social optimal involves the asymmetric treatment of identical individual in the context that the central government can choose the population distribution. See also Hartwick (1980).

$$u(x_1(\tilde{n}), G_1(\tilde{n})) > u(x_2(\tilde{n}), G_2(\tilde{n})) \tag{10}$$

follows from proposition 1. Therefore, (5) and $W'' \leq 0$ imply

$$u_x(x_1(\tilde{n}), G_1(\tilde{n})) \geq u_x(x_2(\tilde{n}), G_2(\tilde{n})). \tag{11}$$

We can argue by contradiction that

$$G_1(\tilde{n}) > G_2(\tilde{n}). \tag{12}$$

Suppose not. From (10), it follows $x_1(\tilde{n}) > x_2(\tilde{n})$. On the other hand, the inequality (11) implies, from $u_{xG} \geq 0$, that $u_x(x_1(\tilde{n}), G_1(\tilde{n})) \geq u_x(x_2(\tilde{n}), G_1(\tilde{n}))$. Combining this inequality and $u_{xx} < 0$ implies $x_1(\tilde{n}) \leq x_2(\tilde{n})$. This is a contradiction. Consequently, we establish inequality (12).

In order to examine the nature of this central authority's policy making, consider the following examples. First, examine the case in which the central government maximizes a utilitarian social welfare ($W'' = 0$). In this case, $u_x(x_1(\tilde{n}), G_1(\tilde{n})) = u_x(x_2(\tilde{n}), G_2(\tilde{n}))$ follows from (5). Therefore, combining this and (12) yields $x_1(\tilde{n}) \geq x_2(\tilde{n})$. If the central government does not concern much about equalization of utility between the regions, the central authority makes transfer in order to make use of the complementarity between private goods and a public good. This transfer policy, therefore, leads to the higher consumption level of private goods in the high populated region.

Second, examine the case where $u_{xG} = 0$. In this case, $x_1(\tilde{n}) \leq x_2(\tilde{n})$ follows from (11) and (12). Since the degree of complimentarity between private goods and a public good is small in this example, the central authority makes transfer so that the consumption level in the low populated region is higher than in the high populated region in order to equalize the social marginal utility between the regions.

It seems to be worth considering that the direction of the interregional transfer under a simple production function. Per capita transfer is given by

$$\frac{S_i(\tilde{n})}{n_i} = x_i(\tilde{n}) - \frac{f_i(n_i)}{n_i} + \frac{G_i(\tilde{n})}{n_i}. \tag{13}$$

Suppose $n_1 > n_2$. If $f_i(n_i)$ has a concavity, it follows $f_1(n_1)/n_1 < f_2(n_2)/n_2$. If $W'' = 0$, we have already seen $x_1(\tilde{n}) \geq x_2(\tilde{n})$. Moreover, if demand for G_i is sufficiently elastic with respect to a tax price $(1/n_i)$, we can expect $G_1(\tilde{n})/n_1 > G_2(\tilde{n})/n_2$.[7] Thus we can conclude $S_1(\tilde{n})/n_1 > 0 > S_2(\tilde{n})/n_2$. We will examine the direction of the transfer again in an extended model (see section 3.1).

2.3 *Individual Problem and Migration Equilibrium*

Now we examine the locational decision making of individuals who anticipate the time-consistent policies described above. A continuum of individuals means that when an individual decides where to reside, he takes as given population distribution and, therefore, utility in each region.

Because individuals move to the region which maximizes utility, the only sustainable allocations under free mobility are those which equalize utility levels across regions which have residents. In addition, the utility levels in depopulated regions are no greater than in positive population regions. Therefore, we can define the time-consistent migration equilibrium (hereafter referred to as TCE) as follows.[8] A population distribution $\tilde{n}^* = (n_1^*, \cdots, n_M^*)$ is a TCE if it satisfies

(i) if n_j^*, $n_k^* > 0$, $V_j(\tilde{n}^*) = V_k(\tilde{n}^*)$,

(ii) if $n_i^* = 0$, $V_i(\tilde{n}^*) \leq V_j(\tilde{n}^*)$ for any $j \in \Lambda^+(\tilde{n}^*)$,

(iii) $\sum_{i=1}^{M} n_i^* = N$.

---

[7] One of the sufficient condition for this inequality is $u(x,G) = u(x) + \frac{G^{1-\mu}}{1-\mu}$ where $u' > 0$, $u'' < 0$, and $\mu > 1$. When $n_1 > n_2$, $x_1(\tilde{n}) \leq x_2(\tilde{n})$ follows from $u_{xG} = 0$. Moreover, combining (2) and (11) yields $n_1 \cdot G_1(\tilde{n})^{-\mu} \geq n_2 \cdot G_2(\tilde{n})^{-\mu}$. Thus, it follows $G_1(\tilde{n})/n_1 > G_2(\tilde{n})/n_2$ from $\mu > 1$ and $G_1(\tilde{n}) > G_2(\tilde{n})$.

[8] An "optimal" TCE is often defined as a TCE which leads to the highest social welfare level among those achieved by TCEs. This optimal TCE is sometimes referred to as a TCE briefly. If the central government has an opportunity to announce its policy before individuals make their locational choices, this optimal TCE has a reality. In contrast, there need not exist the announcement stage in order to define a TCE.

$V_i(\tilde{n})$ is defined by $V_i(\tilde{n}) = \lim_{\|\tilde{\varepsilon}\| \to 0} V_i(\tilde{n} + \tilde{\varepsilon})$ for any $i \in \Lambda^0(\tilde{n})$, where $n_i + \varepsilon_i > 0$, and $\sum_{i=1}^{M} \varepsilon_i = 0$.[9] [10] Then, we can derive the following proposition.

Proposition 2 A population distribution $\tilde{n} = (n_1, \cdots, n_M)$ is a TCE if, and only if, each populated region has the same population.

PROOF: Suppose $n_j, n_k > 0$. Then, from proposition 1, $V_j(\tilde{n}) = V_k(\tilde{n})$ holds if and only if $n_j = n_k$. Suppose region i is a no-resident region. From proposition1, if $\varepsilon_i < n_j + \varepsilon_j$, it follow that $V_i(\tilde{n} + \tilde{\varepsilon}) < V_j(\tilde{n} + \tilde{\varepsilon})$ for any $j \in \Lambda^+(\tilde{n})$. When $\|\tilde{\varepsilon}\|$ decreases to zero, $\varepsilon_i < n_j + \varepsilon_j$ holds for any $j \in \Lambda^+(\tilde{n})$ since $n_j > 0$. As a result, if $n_i = 0$, $V_i(\tilde{n})$ satisfies the requirement of (ii) in the definition of TCE. □

This time-consistent migration equilibrium has a multiplicity. Yet, some of them are not robust in the sense that if the population is disturbed from the equilibrium, it will diverge.[11] In order to examine this sustainability, define the local stability of a TCE. When the population distribution is disturbed from an equilibrium, if individuals have incentives to move to a region where the population increases, the equilibrium is not locally stable. That is, a TCE $\tilde{n}^*$ is locally stable if the following condition is satisfied: for any $\tilde{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_M)$ such that $\sum_{i=1}^{M} \varepsilon_i = 0$ and $\|\tilde{\varepsilon}\|$ is small, if $\varepsilon_i < 0 < \varepsilon_j$, then $V_i(\tilde{n}^* + \tilde{\varepsilon}) > V_j(\tilde{n}^* + \tilde{\varepsilon})$ for any $i, j = 1, \cdots, M$ $(j \neq i)$. Then, we obtain the following proposition.

---

9 $\|\tilde{\varepsilon}\|$ represents a norm of $\tilde{\varepsilon}$. If we choose a Euclidean norm, $\|\tilde{\varepsilon}\|$ is defined as $\|\tilde{\varepsilon}\| = \sqrt{\varepsilon_1^2 + \cdots + \varepsilon_M^2}$.

10 The basic reason why we do not define the utility in depopulated regions without taking a limit is that budget constraints for the regions do not constrain the level of private consumption. The rationalization of this definition of utility in no-resident regions is discussed in appendix C.

11 As we assumed that individuals are immobile *ex post*, the process of divergence can be thought to proceed through generations.

Proposition 3. A population distribution $\tilde{n} = (n_1, \cdots, n_M)$ is a locally stable TCE if, and only if, it is a single-community equilibrium.

PROOF: Suppose that there are *m* communities $(m \geq 2)$ with *N/m* residents. We can assume that $\tilde{n} = (N/m, \cdots, N/m, 0, \cdots, 0)$ without loss of generality. Choose δ such that $\delta < 2(N/m)^2$. Then, define $\tilde{\varepsilon}$ such that $\tilde{\varepsilon} = (-\sqrt{\delta/2}, \sqrt{\delta/2}, 0, \cdots, 0)$. It is easy to see that $\|\tilde{\varepsilon}\| = \delta$ and

$$V_1(\tilde{n} + \tilde{\varepsilon}) < V_2(\tilde{n} + \tilde{\varepsilon}) \qquad (14)$$

for any $\delta > 0$ from proposition1. This contradicts the definition of local stability.

Suppose that there is a community where all individuals reside. We can assume that $\tilde{n} = (N, 0, \cdots, 0)$ without loss of generality. When $\|\tilde{\varepsilon}\| < N/2$, $n_1 + \varepsilon_1 > n_j + \varepsilon_j = \varepsilon_j$ for any $j = 2, \cdots, M$. Then, if we choose δ such that $\delta < N/2$, we have $V_1(\tilde{n} + \tilde{\varepsilon}) > V_j(\tilde{n} + \tilde{\varepsilon})$ for any $\tilde{\varepsilon}$ $(\|\tilde{\varepsilon}\| \leq \delta)$, for $j = 2, \cdots, M$ from proposition1. □

Proposition 2 and 3 are predictable consequences of proposition 1, which argues that the *ex-post* policy always favours relatively highly populated regions. Time-consistency requires that the individuals correctly anticipate the outcome of the *ex-post* optimal policy: they can achieve the highest utility if they successfully reside in one of the highest populated regions (proposition 2).

When one region has the highest population, every individual will move to that region. This locational choice pattern means all regions are completely depopulated except for one (proposition 3). In so far as an efficient population allocation is interior, the *ex-post* policy does not realize Pareto efficient outcome given the restriction that the utility of all populated regions are the same. Although the decentralized migration equilibrium is not efficient when local governments cannot commit to their policies, intervention may worsen the efficiency of population distribution.[12]

[12] We define formally a decentralized (or no-intervention) equilibrium when local governments have no commitment technology in section 4.

Note that no assumption about technology was required to derive propositions 2 and 3. In contrast, the stability of a decentralized equilibrium depends upon the nature of production functions. The strong diminishing returns to labour in production results in the stability of interior decentralized equilibrium.[13] The optimal population distribution may not be single communities when each region faces strong diminishing returns to scale. Hence, the single-community equilibria may entail inefficiency.

Our analysis has not so far included the Rawlsian social welfare function. Our results are not relevant for the Rawlsian case. The maximization of a Rawlsian welfare function implies equalization of utility among regions rather than that of social marginal utility of private good. Proposition 1 should be revised as follows: for a given $\tilde{n} = (n_1, \cdots, n_M)$, if $n_j, n_k > 0$, then $V_j(\tilde{n}) = V_k(\tilde{n})$. In contrast to proposition 2, this implies that any population distribution, including an efficient one, becomes a TCE. The resulting distribution may be the first best one.[14] However, any TCE in this case is not locally stable; once a perturbation takes place, there is no incentive for any individual to move from their new residence.

3. Some Extensions: Capital, Congestion, and Spillover Effect

We now consider some extensions of our model. First, we investigate how the introduction of capital into our model affects the results. Secondly, congestion effects are examined. Finally, we investigate the effect of spillovers in the provision of a public good on the previous results.

3.1 Capital Introduced as a Resource of Production[15]

[13] We will examine the stability of the decentralized migration equilibrium without commitment, using a simple example in section 4.

[14] In contrast, an optimal TCE leads to a first best population distribution under the Rawlsian social welfare function.

[15] The importance of interregional capital mobility as well as residential mobility has been emphasized in the literature of tax competition. See, for example, Zodrow and Mieszkowski (1986) and Wildasin (1988).

Capital is assumed to be perfectly mobile even *ex post* and each person is assumed to own the same proportion of capital (K/N) where K is the total amount of capital in the economy. The production function is assumed to be common among regions and is given by $f(n_i, K_i)$ where K_i is the capital input in region i. We assume, for simplicity, constant returns to scale.[16]

Here we do not consider capital tax. The perfect mobility of capital implies that the rate of returns to capital, r, is common for all populated regions:

$$f_K(n_i, K_i) = [\equiv r].^{17} \tag{15}$$

for any $i \in \Lambda^+(\tilde{n})$. It is assumed that $f_{nn} < 0$, $f_{KK} < 0$, and $f_{nK} > 0$. Let w_i denote the wage (rate) in region i. Then w_i is assumed to coincide with the marginal product of labour $f_n(n_i, K_i)$. Since the production function is homogeneous of degree one, it follows that $n_i / K_i = n_j / K_j$. Therefore, $n_i > n_j$ implies that $K_i > K_j$. Capital follows labour in order to obtain higher returns. Thus proposition 1 can be applied to this case which leads to the concentration of population in a locally stable TCE.

Consider the direction of the interregional transfer under a given population distribution (see section 2.2). Since the production function is homogeneous of degree one, the marginal productivity of labour $f_n(n_i, K_i)$ is homogeneous of degree zero with respect to n_i and K_i. Thus, per capita labour incomes are the same for all regions. Moreover, per capita capital income becomes $r \cdot K/N$, since the rate of return on capital is independent of regions and each individual is assumed to have the same portion of capital. Therefore, the per capita income, before interregional transfers, $w + r \cdot K/N$ is the same in all regions where w is the common wage rate in all regions. Focus on a simple example: $M = 2$, $n_1 > n_2$. As discussed in section 2.2, if

---

[16] Though, for simplicity, we ignore the existence of other inputs, we can take into account, for instance, land as a fixed input. In that case, if the production function is assumed to be constant returns to scale with respect to all inputs, we can proceed the analysis as the same manner as in the text.

[17] When there exist proportional taxes on capital returns and they are exogenous, the requirement for capital market equilibrium (15) is replaced by

$$(1 - z_i) \cdot f_K(n_i, K_i) = (1 - z_j) \cdot f_K(n_j, K_j)$$

for any $i, j = 1, \cdots, M$ where z_i is the tax rate on capital returns for region i. The argument in this case can proceed similarly as the case in the text.

$u_{xG} = 0$, $x_1(\tilde{n}) \le x_2(\tilde{n})$ and if the demand for G_i is sufficiently inelastic with respect to tax price, we can expect $G_1(\tilde{n})/n_1 < G_2(\tilde{n})/n_2$, which implies $x_1(\tilde{n}) + G_1(\tilde{n})/n_1$ $< x_2(\tilde{n}) + G_2(\tilde{n})/n_2$. As shown above, the per capita income before interregional transfer is the same, $w + r \cdot K/N$, for all regions. Thus, the authority transfers incomes from region 1 to region 2. Yet, workers desire to move to a higher populated region to obtain a higher utility level, even though there is a transfer from the higher populated region to the lower populated region. Though we examine a possibility that transfer is made from a low populated region to a high populated region in the basic model (see section 2.2), the authority makes transfer in the opposite direction in this extended model.

3.2 Congestion

The introduction of capital does not affect the previous population concentration results. However, this concentration doesn't necessary occur when we assume congestion in consumption of a public good. Rewrite the utility function $u(x_i, G_i, n_i)$ in order to capture the congestion effects.[18]

As an example, suppose the utility function is

$$u_i = u(x_i, G_i / n_i^\theta) \qquad (16)$$

where $0 \le \theta \le 1$. The good G_i which is provided by the public sector becomes a pure public good when $\theta = 0$ and a publicly provided private good when $\theta = 1$.

Let g_i denote G_i / n_i^θ. Call g_i the service of the public good. The tax price of g_i is $n_i^{\theta - 1}$ since the budget constraint for region i is $n_i \cdot x_i + n_i^\theta \cdot g_i = f_i(n_i) + S_i$. When $\theta = 1$, it follows that distribution of population is indeterminate because the prices of the public good in both regions are the same or there is no scale merit in consumption of the public good. Otherwise, the result of population concentration still holds since the tax price of the public good service is lower in a high populated region

[18] The problem of congestion in public goods is analyzed, for example, by Bergstrom and Goodman (1973).

than in a low populated region when $0 \le \theta < 1$. Consequently, proposition 2 and proposition 3 are still relevant.

Now consider another type of congestion; not directly related to the public good:

$$u_i = u(x_i, G_i) - \phi(n_i) \tag{17}$$

where $\phi' > 0$ and $\phi'' > 0$. Denote by $V_i(\tilde{n})$ the utility, $u(x_i, G_i)$, after it is optimized by the interregional transfer.[19] For example, if the authority maximizes a utilitarian social welfare, then the optimization problem faced by the central government is independent of $\phi(n_i)$. Therefore, if $n_j > n_k$, then $V_j(\tilde{n}) > V_k(\tilde{n})$ from proposition 1. However, this does not necessarily imply that $V_j(\tilde{n}) - \phi(n_i)$ $> V_k(\tilde{n}) - \phi(n_k)$ since $\phi(n_i) > \phi(n_k)$. If the congestion cost, $\phi(n_i)$, increases drastically as n_i approaches *N*, the concentration of population is stopped before everyone resides in the same region.

3.3 Spillover

The extensions examined above do not affect the role of local governments. The role is to allocate the consumption of private good and public good according to the Samuelson rule by means of a lump-sum tax program. Therefore, even though the distribution of population is inefficient, the allocation of the private good and public good is efficient under a given population distribution. Yet, when we consider the spillover effect or interregional externality in a public good, this is not the case.

We assume that there are two regions in the economy to consider the effects. The utility function is,

$$u_i = u(x_i, G_i, G_j) \tag{18}$$

For example, if an individual wants to travel across regions, it is important to him that transportation systems are good not only in his own region but also in neighbouring regions.

---

[19] $V_i(\tilde{n}) = V_i(S_i(\tilde{n}), n_i)$ and $V_i(S_i, n_i) = u(x_i(S_i, n_i), G_i(S_i, n_i))$.

When we consider the spillover effect, the strategy of a local government i depends on the strategy of the other region j. The first order condition (2) is modified to become

$$n_i \cdot \frac{u_{G_i}(x_i, G_i, G_j)}{u_x(x_i, G_i, G_j)} = 1 \tag{19}$$

for any $j \neq i, i = 1,2$. The budget constraint of region i is

$$n_i \cdot x_i + G_i = f_i(n_i) + S_i, \tag{20}$$

for $i = 1,2$. From (19) and (20), we obtain a Nash equilibrium.

In this framework, local governments choose their provision of the public good strategically. Since each local government takes no account of the contribution of its public good provision to other regions, the provision level becomes lower than the optimal level. This is the well-known free-rider problem.[20]

4. Welfare Comparison with Time-Consistent Equilibrium and Equilibrium without Central Government Intervention

We have already seen from proposition 3 that the time-consistent policy substantially disturbs the population distribution unless the scale merit in forming a community is significant. The welfare comparison in this section ascertains this prospect.[21] As a benchmark, we employ the same equilibrium without national intervention as Boadway and Flatters (1982). There is no voluntary transfer at the local level since local governments cannot credibly commit. We call the equilibrium BFE hereafter.

---

[20] We will examine the case with the externality in a simple example in section 4.

[21] This welfare comparison also has the following policy implication which is closely related to the issue of discretionary policy versus a rule. Choosing "policy as a rule" has the benefit that the authority can commit to its policy, but the cost is that the rule need to be simple. On the other hand, choosing a discretionary policy has the advantage that the authority can design a policy which depends on the varieties of situations and the cost is difficulty for the authority to commit to its policy.

That the central government does not intervene in the transfer policy at all is a simple rule concerning an interregional transfer policy. It is easy for private agents to verify whether the authority obeys this simple rule or not. Though there may be other simple policies which are easy for the authority to commit to, no-intervention policy is thought of as one of the simplest policies among them. We adopt the equilibrium without central government intervention as a reference point in order to characterize TCE.

For the purpose of comparing social welfare, we examine two cases. One of them is a basic case in which the consumption of the public good has no congestion or no spillovers. The second case is a situation in which there are spillover effects leading to multiple-community equilibrium.[22]

4.1 Welfare Comparison in the Basic Model

We consider the case of $M = 2$ to allow for a diagrammatic exposition. Using maximum value function $V_i\ (0, n_i)$ (see (3)), the population distribution in BFE is obtained by $V_1\ (0, n_1^B) = V_2\ (0, N - n_1^B)$ if $0 < n_1^B < N$. Moreover, $n_1^B = 0$ if $V_1(0,0) \le V_2(0,N)$ and $n_1^B = N$ if $V_1(0,N) \le V_2(0,0)$. We denote by V_{BFE} and V_{TCE} the maximum utility level among those which are realized respectively at locally stable BFEs and at locally stable TCEs. It follows that $V_{TCE} = \max(V_1(0,N), V_2(0,N))$, since the population distribution in a locally stable equilibrium is a single-community one.

Figure 4 is around here

The relationship between V_{BFE} and V_{TCE} is illustrated in figure 4. It depends on the scale merit in consumption of a public good and the degree of diminishing returns to scale in the production technology. Figure 4(a) illustrates a situation in which diminishing returns are large for both regions. In this case, a locally stable BFE is an internal equilibrium and the utility is higher than that realized if all individuals reside in a single community: the utility of the locally stable BFE is higher than that of a locally stable TCE. Figure 4(b) illustrates a case where the scale merit of public good consumption is large for both regions. In this case, the locally stable BFE is a single-community and the utility of the locally stable BFE and that of the locally stable TCE are the same. Figure 4(c) illustrates a case where the scale merit in consumption of the public good is greater than the demerit of diminishing returns to scale in region 1 but the demerit is larger than the scale merit in region 2. In this case, the locally stable

[22] If we take into account congestion in the consumption of public good, there also exists a

BFE is an internal equilibrium and the utility is lower than that realized if all individuals reside in a single community. Only in this case, national intervention improves welfare: the utility of the locally stable BFE is lower than that of the locally stable TCE.

4.2 Welfare Comparison in A Simple Example with Spillover Effect

Consider a simple example to compare the TCE with the BFE. We assume that the central government maximizes a utilitarian social welfare and that there are only two regions. In addition, in order to capture the spillovers in the provision of a public good, we assume the following utility function:

$$u_i = \alpha \cdot \log x_i + \beta \cdot \log G_i + \gamma \cdot \log G_j, \quad (21)$$

for $i = 1,2$, $j \neq i$, where $\alpha + \beta + \gamma = 1$ and $\beta \geq \gamma \geq 0$. The parameter γ captures the degree of the externalities of the public good. If $\gamma = 0$, then this utility function reduces to the case of no-externality. If $\gamma > 0$, then the marginal rate of substitution between G_i and G_j diminishes.[23] When $\beta = \gamma = 1$, the second and third terms capture the "weaker link" as named by Cornes (1993).[24] Using the Samuelson condition, utility in each region becomes

$$V_i\,(\tilde{S},\tilde{n}) = A + (1-\gamma)\cdot \log(f_i(n_i) + S_i) + \gamma \cdot \log(f_j(n_j) + S_j) - \alpha \cdot \log n_i \quad (22)$$

for $i = 1,2$, $j \neq i$, where $\tilde{S} = (S_1, S_2)$, $S_1 + S_2 = 0$, and $A = \alpha \cdot \log \alpha + (\beta+\gamma)\cdot \log \beta - \log(\alpha+\beta)$. The derivation of the following equations are included in appendix D.

Under the utilitarian social welfare function ($SW = n_1 \cdot V_1 + n_2 \cdot V_2$), the utility level that an individual anticipates to obtain when he resides in region *i* is given by

$$V_i\,(\tilde{n}) = \mathrm{A} - \log(n_1 + n_2) + \log(f_1(n_1) + f_2(n_2))$$
$$+ (1-\gamma)\log((1-\gamma)\cdot n_i + \gamma \cdot n_j)$$

multiple-community equilibrium.

[23] Another important specification which capture the spillovers in public good is that the argument of utility function is a weighted average of public goods. See Ihori (1991).

[24] Hirshleifer (1983,1985) introduces the notion of "weakest-link" and " best-shot". Weakest-link is a case where the utility of individuals depends on the minimum level of each individual's contribution to public good. Best-shot is a case where utility of individual depends on the maximum level of each individual's contribution to public good.

$$+\gamma \cdot \log((1-\gamma)\cdot n_j + \gamma \cdot n_i) - \alpha \cdot \log n_i \tag{23}$$

for $i = 1,2$, $j \neq i$.

Denote by (n_1^*, n_2^*) a population distribution in a time-consistent migration equilibrium. From $V_1(\tilde{n}) = V_2(\tilde{n})$, the symmetric population distribution $(N/2, N/2)$ is a time-consistent migration equilibrium. When $\gamma = 0$, $(n_1^*, n_2^*) = (N,0)$ and $(0,N)$ are also TCE from Proposition 2. When $\gamma > 0$, $\lim_{\substack{n_1 \to N \\ n_1+n_2=N}} (V_1(\tilde{n}) - V_2(\tilde{n}))$ $= \lim_{\substack{n_2 \to N \\ n_1+n_2=N}} (V_2(\tilde{n}) - V_1(\tilde{n})) = -\infty$. Then it follows that $0 < n_1^*, n_2^* < N$ when $\gamma > 0$.

We denote by V_{TCE}^{SM} the utility level of a symmetric TCE. From (23), it follows

$$V_{TCE}^{SM} = A + \log \frac{f_1(N/2) + f_2(N/2)}{2} - \alpha \cdot \log(N/2). \tag{24}$$

Consider the local stability of the symmetric TCE. Under the constraint $n_1 + n_2 = N$, differentiating $V_1(\tilde{n}) - V_2(\tilde{n})$ with respect to n_1 and evaluating the derivative at $n_1 = n_2$ implies that the symmetric TCE is locally stable if

$$(1-2\gamma)^2 < \alpha. \tag{25}$$

When the degree of the externality of a public good γ is large, the central authority has an incentive to make the utility level of the less populated region higher than that of the highly populated region in order to make use of the merit of the externality. This characteristic of the policy formation makes the symmetric TCE locally stable.

In order to investigate the characteristics of the BFE, we assume a Cobb-Douglas production function:

$$f_i(n_i) = k_i \cdot n_i^{\delta}. \tag{26}$$

Assume $k_1 \geq k_2$. If BFE is an interior solution, from the BFE condition $V_1(\tilde{0}, \tilde{n}^B) = V_2(\tilde{0}, \tilde{n}^B)$ $[\equiv V_{BEF}]$, it follows that

$$V_{BFE} = A + \frac{\log k_1 + \log k_2}{2} + (\delta - \alpha)\cdot\left(\log N - \log \frac{\sqrt{\Omega}}{1+\Omega}\right), \tag{27}$$

where $\Omega = \left(\frac{k_1}{k_2}\right)^{\frac{1-2\gamma}{\alpha-(1-2\gamma)\delta}}$.

Now differentiating $V_1(\widetilde{0},\widetilde{n}) - V_2(\widetilde{0},\widetilde{n})$ with respect to n_1 under the constraint $n_1 + n_2 = N$, the interior BFE is locally stable if

$$(1-2\gamma)\cdot\delta < \alpha. \tag{28}$$

When the scale merit of production δ is small, the demerit of choosing residence in the highly populated region is large. This is because the income earned in the highly populated region is lower than in a less populated region. When the benefit of a public good in a region to its own residents, β ($=1-\alpha-\gamma$), is small, the merit of choosing the low populated region is large. So for a given α, when δ is small and / or γ is large, the interior BFE is locally stable.

Figure 5 is around here

Figure 5 illustrates the combination of (γ,α) under which the symmetric TCE and the interior BFE are locally stable for a δ given (see (25) and (28)). For example, a given α, if the scale merit in production, δ, is small and spillovers in public good provision, γ, are small, the combination (γ,α) is in area "a". In this case, the interior BFE is locally stable and the symmetric TCE is unstable. When both the scale merit in production and the public good externality are large, the combination (γ,α) is in area "c". In this case, the interior BFE is locally unstable and the symmetric TCE is stable.

Compare the utility level of locally stable TCE with that of locally stable BFE. When a combination (γ,α) is in areas "b" or "c", locally stable BFEs are $(N,0)$ and $(0,N)$. Since $V_i(\widetilde{0},(N,0)) = -\infty$ $(i=1,2)$, V_{BFE} becomes minus infinity in areas "b" or "c". When $\gamma > 0$, (n_1^*,n_2^*) satisfies $0 < n_1^*, n_2^* < N$, which implies $V_{TCE} = V_1(\widetilde{n}^*) > -\infty$. Hence $V_{TCE} > V_{BFE}$ in areas "b" or "c", unless $\gamma = 0$.

In area "d", the locally stable TCE is a symmetric TCE. Then from (24) and (27), we have

$$V_{TCE} - V_{BFE} = \log\frac{k_1 + k_2}{2} - \frac{\log k_1 + \log k_2}{2} - (\delta - \alpha)\cdot\log\frac{2\sqrt{\Omega}}{1+\Omega}. \tag{29}$$

Therefore, if $\delta \geq \alpha$, $V_{TCE} \geq V_{BFE}$ follows (the equality holds when $k_1 = k_2$). When the scale economies of production are large (areas "b" , "c", or "d"), no interregional transfer results in an over-population in one region. Thus, central government intervention can improve the efficiency of the population distribution.

To investigate areas "a" and "e", we assume that the production functions are identical in both regions ($k_1 = k_2$). In area "a", BFE is Pareto efficient given the restriction that both regions' utilities are the same. On the contrary, since a symmetric TCE is not locally stable in area "a", V_{TCE} is not Pareto optimal. Consequently, $V_{TCE} < V_{BFE}$.

To examine area "e", we maintain the assumption that $k_1 = k_2$. In this case, $n_1^B = n_2^B$ holds and a symmetric TCE is locally stable and entails no interregional transfers. Therefore, $V_{BFE} = V_{TCE}$ in area "e" and $k_1 = k_2$. These results comparing V_{BFE} and V_{TCE} are summarized in Table1.

Table 1 is around here

Finally, to compare V_{BFE} with V_{TCE} in areas "a" or "e" when $k_1 \neq k_2$. We use a numerical example. We choose the following parameters values: $k_1 = 11$, $k_2 = 10$, $\alpha = 0.7$, $\gamma = 0$, 0.02, 0.05, 0.07, 0.1, and $\delta = 0.6$, 0.7. The technology is described by k_1, k_2 and δ.

Table 2 is around here

Table 2 shows the results of this comparison. When the elasticity of product with respect to labour is small (δ=0.6), an interior solution of BFE is locally stable. If $\gamma \leq 0.07$, then we are in area "a" and $V_{BFE} > V_{TCE}$. The same results hold when

$k_1 = 11$ and $k_2 = 10$ that hold when $k_1 = k_2$. In contrast, if γ=0.1, then we are in area "e" and $V_{BFE} < V_{TCE}$. Figure 6 and figure 7 illustrate $V_i(\tilde{n}) - A$ when δ=0.6. When γ=0, $V_i(\tilde{n}) - A$ approaches minus infinity as n_i tends to zero (figure 6), resulting in a single-community equilibrium. When γ=0.05, $V_i(\tilde{n}) - A$ increases sharply when n_i approaches zero owing to the externality of the public good (figure 7). Thus, the externality stops the concentration of population.

Figure 6 and 7 are around here

When $\delta = \alpha$ =0.7, an interior solution of BFE is locally stable. If $\gamma \leq 0.07$, then we are in area "a". In this case, we obtain $V_{BFE} > V_{TCE}$. The characteristics when $k_1 = k_2$ still holds when $k_1 = 11$ and $k_2 = 10$. From this simple numerical example, when both scale merit in production and externality of public good are small (area "a"), utility is greater without the central authority's intervention.

When the spillovers in public good provision are large, the central authority makes interregional transfers to a less populated region in order to increase its public good provision. Reducing a discrepancy between levels of public good provision in regions contributes to improved efficiency. Individuals move to a less populated region, expecting a transfer policy which makes the utility in the region higher than in the other region. The intervention, therefore, improves the efficiency of population distribution. When the scale merit in production is large, the concentration of population takes place even if there is no intervention of the national authority. The expectation of interregional transfer does not distort the population distribution so much. The intervention, therefore, enhances the efficiency without a large externality of public good provision. In the basic model, we have seen that the central organized intervention results in the concentration of population. Thus, the intervention can make the population distribution less efficient. In contrast, when we take into account the externality of public good provision, whether or not central government intervention enhances the efficiency of population distribution depends, for example, upon the degree of the scale merit in production and externality of public good provision.

5. Concluding Remarks

This paper has investigated the properties of a time-consistent migration equilibrium. We have derived the following results. First, it is shown that the locally stable time-consistent migration equilibria are single-community equilibria. When the central government makes transfers after individuals have made decisions on residence, the *ex post* transfer policy makes the utility of the high populated region higher than that of a low populated region. Private agents anticipate this transfer policy rationally so that the migration process results in a concentration of population.

Second, we extend this basic model in some aspects. We show that the basic population concentration result still holds even when we introduce capital as a production input. In addition, it is shown that the concentration of population is avoided when we assume congestion in consumption of the public good or spillovers in the provision of a public good.

Third, we have examined whether central government intervention enhances efficiency in the absence of a commitment technology. The utility in the time-consistent migration equilibrium is compared with that in equilibrium without central government intervention. We show that a central government's intervention without commitment can enhance efficiency when the externalities of a public good are large.

The model can be extended in several ways. We assume *ex post* immobility to simplify the analysis. It is important to examine what would happen if we assume *ex post* imperfect mobility. If we extend our model to a two-period one, we can consider other problems in the model. In the second period, interregional transfers have the same characteristics as our model. However, the transfer policy in the first period may be designed in order to offset the distortion caused by the second period transfers. Treating the public good as a production resource could be considered in a two-period model. The accumulation of a public capital may be under-provided because of the income equalization policy in the second period. A matching grants policy may improve the population distribution as well as the public capital stock allocation.

Appendix A

The budget constraint of a region is

$$n_i \cdot x_i + G_i = f_i(n_i) + S_i + \sum_{j \neq i} Z_{ji} - \sum_{j \neq i} Z_{ij}. \qquad \text{(A-1)}$$

where Z_{ij} is a (non-negative) voluntary transfer from region i to region j. $\sum_{j \neq i} Z_{ji}$ represents the total transfer to region i from the other regions and $\sum_{j \neq i} Z_{ij}$ is the total transfer from region i to the other regions. Thus $\sum_{j \neq i} Z_{ji} - \sum_{j \neq i} Z_{ij}$ captures the net transfer to region i.

If local governments can commit to their policies, the population of region i is determined by the local government's policy. The analysis begins with the problem of migration equilibrium for a given $(\widetilde{G}, \widetilde{\mathbf{Z}}, \widetilde{S})$, where $\widetilde{G} = (G_1, \cdots, G_M)$, $\widetilde{Z}_i = (Z_{i1}, \cdots, Z_{i(i-1)}, Z_{i(i+1)}, \cdots, Z_{iM})$, and $\widetilde{\mathbf{Z}} = (\widetilde{Z}_1, \cdots, \widetilde{Z}_M)$. *Ex ante* free mobility implies that utility is the same across regions, and this determines the population distribution $\widetilde{n}$ as a function of regional transfers and public good provisions: $\widetilde{n} = g(\widetilde{G}, \widetilde{\mathbf{Z}} \colon \widetilde{S})$. This function is referred to as a migration response function. An equilibrium of this federal system with commitment is given by a Nash equilibrium. Denote by $(\widetilde{G}^*, \widetilde{\mathbf{Z}}^*)$ a Nash equilibrium when the interregional transfers by the central government are $S_1 = \cdots = S_M = 0$, where $\widetilde{G}^* = (G_1^*, \cdots, G_M^*)$ and $\widetilde{\mathbf{Z}}^* = (\widetilde{Z}_1^*, \cdots, \widetilde{Z}_M^*)$. It is easy to see that $(\widetilde{G}^*, \widetilde{\mathbf{Z}}^*)$ coincides with a Nash equilibrium in the model of Myers (1990).[25] For a given $\widetilde{S} = (S_1, \cdots, S_M)$, define $Z_{ij}^*(S_j) = Z_{ij}^* - S_j / M$, $\widetilde{Z}_i^*(\widetilde{S}) = (Z_{i1}^*(S_1), \cdots, Z_{iM}^*(S_M))$, and $\widetilde{\mathbf{Z}}^*(\widetilde{S}) = (\widetilde{Z}_1^*(\widetilde{S}), \cdots, \widetilde{Z}_M^*(\widetilde{S}))$. Then we have

$$S_i + \sum_{j \neq i} Z_{ji}^*(\widetilde{S}) - \sum_{j \neq i} Z_{ij}^*(\widetilde{S}) = \sum_{j \neq i} Z_{ji}^* - \sum_{j \neq i} Z_{ij}^*. \qquad \text{(A-2)}$$

---

[25] Myers (1990) mainly investigates the situation in which an instrument for making interregional transfers is a source based per unit tax on land. Yet, his analysis can be applied to the rent sharing case in which local governments have direct instruments for interregional transfer. See the appendix of Myers (1990).

Thus, it follows that $(\widetilde{G}^*, \widetilde{\mathbf{Z}}^*(\widetilde{S}))$ is a Nash equilibrium. Moreover, the population distribution under $(\widetilde{G}^*, \widetilde{\mathbf{Z}}^*(\widetilde{S}))$ is the same for any $\widetilde{S}$. Even in the presence of central government intervention, local governments can redesign their transfer scheme to offset its policy. Therefore, when local governments can commit to their policies and design voluntary transfers dependent upon the interregional transfer policy of the central government, the Pareto efficient population distribution is achieved by voluntary transfers because of the strong incentive equivalency among local governments.

Consider a case in which local governments cannot commit to their policy. For given n_i, $\widetilde{S}$, and $\widetilde{Z}_j$ $(j \neq i)$, local government i chooses x_i, G_i, and $\widetilde{Z}_i$ to maximize $u(x_i, G_i)$ subject to the budget constraint (A-1). By substituting (A-1) for x_i into the utility function, the local government maximization problem becomes

$$\text{(I')} \quad \max_{G_i, \widetilde{Z}_i} u\left(\frac{f_i(n_i) + S_i - G_i - \sum_{j \neq i} Z_{ji} - \sum_{j \neq i} Z_{ij}}{n_i}, G_i \right).$$

Under the assumptions of *ex post* immobility and no policy commitment of local governments, if $\widetilde{Z}_i' \neq (0, \cdots, 0)$ and $\widetilde{Z}_i'' = (0, \cdots, 0)$, then the strategy $(G_i, \widetilde{Z}_i')$ is strictly dominated by strategy $(G_i, \widetilde{Z}_i'')$ for any G_i. Local governments, therefore, have no incentives to make voluntary transfer to other regions. The basic reason for this result is that voluntary transfers have no impact on migration because of *ex post* immobility. Thus we can preclude the possibility that $\widetilde{Z}_i \neq (0, \cdots, 0)$ in any Nash equilibria. As $\widetilde{Z}_i = (0, \cdots, 0)$ is a dominant strategy, we focus on the strategy G_i of region *i*.[26]

If local governments have no commitment technology but the central government has, the local governments implement their policies, taking population distribution as given but the central government makes its interregional transfer, taking into account the migration response. Thus, the problem reduces to that addressed in

[26] As local governments are assumed to choose G_i and $\widetilde{Z}_i$ at the same time, there is no voluntary transfer. Yet, if we postulate that local governments choose $\widetilde{Z}_i$ first and then they choose G_i simultaneously, there exist (strategic) transfers. See Buchholz and Konrad (1995).

Boadway and Flatters (1982). They derive a familiar formula where the appropriate interregional transfer ensures that net benefits in all regions are the same:

$$\frac{G_1 + S_1 - R_1}{n_1} = \cdots = \frac{G_M + S_M - R_M}{n_M}, \tag{A-3}$$

where $R_i = f_i(n_i) - n_i \cdot f_i'(n_i)$ represents the total rents in region i. The population distribution $\tilde{n}$ is determined by equalizing utility across regions i.e. $V_1(S_1, n_1) = \cdots = V_M(S_M, n_M)$. When $M = 2$, therefore, the transfer from region 1 to 2 is

$$S_2 = \frac{n_1 \cdot n_2}{n_1 + n_2}\left[\left(\frac{G_2}{n_2} - \frac{G_1}{n_1}\right) + \left(\frac{R_1}{n_1} - \frac{R_2}{n_2}\right)\right]. \tag{A-4}$$

This ideal equalization scheme improves the efficiency of the population distribution as a Pigouvian taxation by internalizing the fiscal externality. The role of the policy is to correct the inefficiency of population distribution induced by the free migration process.

Appendix B

PROOF of **Proposition 1** : Consider the differentiation of the social welfare function under the constraints

$$S_j + S_k = S_j(\tilde{n}) + S_k(\tilde{n}), \tag{B-1}$$

$$S_l' = S_l(\tilde{n}) \quad (l \neq j, k). \tag{B-2}$$

From (3) and (4), we obtain

$$\begin{aligned} \left.\frac{dSW}{dS_j}\right|_{(B-1)\ and\ (B\text{-}2)} &= n_j \cdot W'(V_j(S_j, n_j)) \cdot \frac{\partial}{\partial S_j} V_j(S_j, n_j) \\ &\quad - n_k \cdot W'(V_k(S_k, n_k)) \cdot \frac{\partial}{\partial S_k} V_k(S_k, n_k) \\ &= W'(V_j(S_j, n_j)) \cdot u_x(x_j(S_j, n_j), G_j(S_j, n_j)) \\ &\quad - W'(V_k(S_k, n_k)) \cdot u_x(x_k(S_k, n_k), G_k(S_k, n_k)) \end{aligned} \tag{B-3}$$

from the envelope theorem. The second derivative of social welfare function is

$$\left.\frac{d^2SW}{dS_j^2}\right|_{(B-1)\ and\ (B\text{-}2)} = W''(V_j)\cdot\frac{(u_x^j)^2}{n_j} + W'(V_j)\cdot\left(\frac{u_{xx}^j}{n_j} + \frac{dG_j}{dS_j}\left(u_{xG}^j - \frac{u_{xx}^j}{n_j}\right)\right)$$

$$+W''(V_k)\cdot\frac{(u_k^j)^2}{n_k} + W'(V_k)\cdot\left(\frac{u_{xx}^k}{n_k} + \frac{dG_k}{dS_k}\left(u_{xG}^k - \frac{u_{xx}^k}{n_k}\right)\right), \quad \text{(B-4)}$$

where

$V_i = V_i(S_i, n_i)$,

$u_x^i = u_x(x_i(S_i, n_i), G_i(S_i, n_i))$,

$u_{x\bullet}^i = u_{x\bullet}(x_i(S_i, n_i), G_i(S_i, n_i))$.

From the Samuelson condition (2), it follows that

$$\frac{dG_i}{dS_i} = \frac{u_{xx}^i / n_i - u_{xG}^i}{n_i u_{GG}^i - 2u_{xG}^i + u_{xx}^i / n_i}. \quad \text{(B-5)}$$

From the strict concavity of the utility function, we obtain $u_{xx}^i u_{GG}^i - (u_{xG}^i)^2 > 0$. Therefore, we derive the inequality

$$\left.\frac{d^2SW}{dS_j^2}\right|_{(B-1)\ and\ (B\text{-}2)} < 0. \quad \text{(B-6)}$$

Define $\hat{S}_i$ $(i = j, k)$ such that

$$V_j(\hat{S}_j, n_j) = V_k(\hat{S}_k, n_k) \ \ [\equiv \hat{V}]. \quad \text{(B-7)}$$

Using $\hat{S}_i$, we denote $\hat{x}_i = x_i(\hat{S}_i, n_i)$ and $\hat{G}_i = G_i(\hat{S}_i, n_i)$.

We show that (i) holds. When $n_j > n_k$, we have $\hat{x}_j < \hat{x}_k$ and $\hat{G}_j < \hat{G}_k$ from the strict concavity of the utility function (see figure 2 which corresponds a case where $(j,k) = (1,2)$). Therefore, it follows that

$$u_x(\hat{x}_j, \hat{G}_j) > u_x(\hat{x}_k, \hat{G}_k) \quad \text{(B-8)}$$

from $u_{xx} < 0$ and $u_{xG} \geq 0$. Combining (B-3) and (B-8) yields

$$\left.\frac{dSW}{dS_j}\right|_{\substack{(B-1)\ and\ (B\text{-}2) \\ S_j = \hat{S}_j}} = W'(\hat{V})\cdot\{u_x(\hat{x}_j, \hat{G}_j) - u_x(\hat{x}_k, \hat{G}_k)\} > 0. \quad \text{(B-9)}$$

Suppose $S_j(\tilde{n}) \leq \hat{S}_j$. From (B-6) and (B-9), there is a S_j' such that $S_j' > \hat{S}_j$ and

$$SW(\widetilde{S}',\widetilde{n}) > SW(\widetilde{S}(\widetilde{n}),\widetilde{n}) \quad \text{(B-10)}$$

where $\widetilde{S}' = (S_1',\cdots,S_M')$, $S_k' = S_j(\widetilde{n}) + S_k(\widetilde{n}) - S_j'$, and $S_l' = S_l(\widetilde{n})$. This is a contradiction. As a result, we derive that $S_j(\widetilde{n}) > \hat{S}_j$ which implies $V_j(\widetilde{n}) > V_k(\widetilde{n})$ from (B-7) and $\partial V_i(S_i,n_i)/\partial S_j > 0$.

When $n_j = n_k$, it follows that $\hat{x}_j = \hat{x}_k$ and $\hat{G}_j = \hat{G}_k$. Therefore, we have

$$\left.\frac{dSW}{dS_j}\right|_{\substack{(B-1)\text{ and }(B-2)\\ S_j=\hat{S}_j}} = 0, \quad \text{(B-11)}$$

which implies that $\hat{S}_j = S_j(\widetilde{n})$. As a result, we can conclude that $V_j(\widetilde{n}) = V_k(\widetilde{n})$. □

Appendix C

When a region has no residents, the migration equilibrium requires that there is no incentive for any individual to move to the no-resident region. To examine whether or not an individual has such an incentive, we need to know the level of utility in empty regions. Yet, the consumption levels of a private good in no-resident regions are not determined endogenously in equilibrium because any x_i can satisfy the budget constraint (1). Thus, the utility level $V_i(\widetilde{n})$ may not be determined endogenously for empty region *i*. In order to resolve this difficulty, if there is no-resident community, we define $V_i(\widetilde{n})$ by taking the limits.

Though we assume that there are a continuum of individuals, the policy toward no-resident regions are thought of as the limit in the following sense. If an individual moved to an empty region, the individual is assumed to predict that the policy toward its region would be the policy when it has an infinitesimally small population, though he has no weight in the economy. In other terms, the individual who will move to a no-resident region will predict the limiting policy as the population decreases to zero.

To predict the limits utility for an empty region, the policies towards regions need to have continuity with respect to population. If policies have continuity, the *ex* *ante* prediction will be correct *ex post*. We can show that $S_i(\tilde{n}) = \lim_{\|\tilde{\varepsilon}\|\to 0} S_i(\tilde{n}+\tilde{\varepsilon})$ for any $i = 1,\cdots,M$ (see lemma C), which implies that $V_i(\tilde{n}) = \lim_{\|\tilde{\varepsilon}\|\to 0} V_i(\tilde{n}+\tilde{\varepsilon})$ for any $i \in \Lambda^+(\tilde{n})$. This assures that the prediction of individuals about $\lim_{\|\tilde{\varepsilon}\|\to 0} V_i(\tilde{n}+\tilde{\varepsilon})$ are consistent with $V_i(\tilde{n})$ for any $i \in \Lambda^+(\tilde{n})$ after they have decided residence. Owing to this continuity, we define $V_i(\tilde{n})$ by taking the limits.

Lemma C. For a given $\tilde{n} = (n_1,\cdots,n_M)$,

$$S_i(\tilde{n}) = \lim_{\|\tilde{\varepsilon}\|\to 0} S_i(\tilde{n}+\tilde{\varepsilon}) \tag{C-1}$$

for any $i = 1,\cdots,M$.

PROOF :

(Step 1):

Suppose $n_i = 0$ or $i \in \Lambda^0(\tilde{n})$. First, $S_i(\tilde{n}+\tilde{\varepsilon}) \geq -f_i(\varepsilon_i)$ from $x_i \geq 0$, $G_i \geq 0$, and regional budget constraint (1). From $\lim_{\varepsilon_i \to 0} f_i(\varepsilon_i) = 0$, we obtain that

$$\lim_{\|\tilde{\varepsilon}\|\to 0} S_i(\tilde{n}+\tilde{\varepsilon}) \geq 0 \quad . \tag{C-2}$$

Next, we will show by contradiction that $\lim_{\|\tilde{\varepsilon}\|\to 0} S_i(\tilde{n}+\tilde{\varepsilon}) \leq 0$ when $i \in \Lambda^0(\tilde{n})$. Suppose that there is a $\delta > 0$ such that $S_i(\tilde{n}+\tilde{\varepsilon}) > \delta$ for any $\tilde{\varepsilon}$. Then, consider $\hat{G}_i$ such that $0 < \hat{G}_i < \delta$. From (1), it follows that

$$x_i = (f_i(\varepsilon_i) + S_i(\tilde{n}+\tilde{\varepsilon}) - \hat{G}_i)/\varepsilon_i > (\delta - \hat{G}_i)/\varepsilon_i . \tag{C-3}$$

Then, we have

$$V_i(\tilde{n}+\tilde{\varepsilon}) \geq V_i(S_i(\tilde{n}+\tilde{\varepsilon}),\varepsilon_i) \geq u((\delta-\hat{G}_i)/\varepsilon_i, \hat{G}_i) \tag{C-4}$$

for any $\tilde{\varepsilon}$.

Let k be the smallest region among positive population regions ($n_k \leq n_j$ for any $j \in \Lambda^+(\tilde{n})$). Define $V_{\max}$ as

$$V_{\max} = u(\textstyle\sum_{i=1}^{M} f_i(N)/n_k,\ \sum_{i=1}^{M} f_i(N)) . \tag{C-5}$$

Then, it follows that

$$V_{\max} > V_j(\tilde{n}) \tag{C-6}$$

for any $j \in \Lambda^+(\tilde{n})$. Since $\lim_{\|\varepsilon\| \to 0} V_j(\tilde{n}+\tilde{\varepsilon}) = V_j(\tilde{n})$ for any $j \in \Lambda^+(\tilde{n})$, we have

$$V_{\max} > V_j(\tilde{n}+\tilde{\varepsilon}) \tag{C-7}$$

for a $\tilde{\varepsilon}$ such that $\|\tilde{\varepsilon}\|$ is sufficiently small. Combining (C-4) and (C-7), for a $\tilde{\varepsilon}$ such that $\|\tilde{\varepsilon}\|$ is sufficiently small, it follows

$$V_i(\tilde{n}+\tilde{\varepsilon}) > V_{\max} > V_j(\tilde{n}+\tilde{\varepsilon}) \tag{C-8}$$

for any $i \in \Lambda^0(\tilde{n})$ and for any $j \in \Lambda^+(\tilde{n})$, since $(\delta - \hat{G}_i)/\varepsilon_i$ increases to infinity as ε_i decreases to zero. This contradicts Proposition 1. Thus, we conclude that

$$\lim_{\|\tilde{\varepsilon}\| \to 0} S_i(\tilde{n}+\tilde{\varepsilon}) \le 0. \tag{C-9}$$

It is shown that $\lim_{\|\tilde{\varepsilon}\| \to 0} S_i(\tilde{n}+\tilde{\varepsilon}) = 0$ for any $i \in \Lambda^0(\tilde{n})$ from (C-2) and (C-9).

(Step 2):

Suppose $\tilde{S}(\tilde{n}) \ne \lim_{\|\tilde{\varepsilon}\| \to 0} \tilde{S}(\tilde{n}+\tilde{\varepsilon})$. Since $\tilde{S}(\tilde{n})$ is the solution of maximization problem of (**II**), it follows there exists $\delta > 0$ such that

$$\sum_{i\in\Lambda^+(\tilde{n})} n_i \cdot W(V_i(S_i(\tilde{n}),n_i)) - \sum_{i\in\Lambda^+(\tilde{n})} n_i \cdot W(V_i(\lim_{\|\tilde{\varepsilon}\| \to 0} S_i(\tilde{n}+\tilde{\varepsilon}),n_i)) > \delta \tag{C-10}$$

for any $\tilde{\varepsilon}$ $(\|\tilde{\varepsilon}\| > 0)$. The maximized social welfare under $\tilde{n}+\tilde{\varepsilon}$ is

$$\sum_{i\in\Lambda(\tilde{n})} (n_i+\varepsilon_i) \cdot W(V_i(S_i(\tilde{n}+\tilde{\varepsilon}),n_i+\varepsilon_i))$$

$$= \sum_{i\in\Lambda^+(\tilde{n})} (n_i+\varepsilon_i) \cdot W(V_i(S_i(\tilde{n}+\tilde{\varepsilon}),n_i+\varepsilon_i))$$

$$+ \sum_{i\in\Lambda^0(\tilde{n})} \varepsilon_i \cdot W(V_i(S_i(\tilde{n}+\tilde{\varepsilon}),\varepsilon_i)). \tag{C-11}$$

Consider the following interregional transfer policy:

$$S_i'(\tilde{n},\tilde{\varepsilon}) = \begin{cases} S_i(\tilde{n}+\tilde{\varepsilon}) & [i \in \Lambda^0(\tilde{n})] \\ S_i(\tilde{n}) - \dfrac{\sum_{j\in\Lambda^0(\tilde{n})} S_j(\tilde{n}+\tilde{\varepsilon})}{\#\Lambda^+(\tilde{n})} & [i \in \Lambda^+(\tilde{n})] \end{cases} \tag{C-12}$$

where $\#\Lambda^+(\tilde{n})$ is the number of regions which have a positive population. The social welfare under $S_i'(\tilde{n},\tilde{\varepsilon})$ becomes

$$\sum_{i\in\Lambda(\tilde{n})}(n_i+\varepsilon_i)\cdot W(V_i(S_i'(\tilde{n},\tilde{\varepsilon}),n_i+\varepsilon_i))$$

$$=\sum_{i\in\Lambda^+(\tilde{n})}(n_i+\varepsilon_i)\cdot W(V_i(S_i'(\tilde{n},\tilde{\varepsilon}),n_i+\varepsilon_i))$$

$$+\sum_{i\in\Lambda^0(\tilde{n})}\varepsilon_i\cdot W(V_i(S_i(\tilde{n}+\tilde{\varepsilon}),\varepsilon_i)). \tag{C-13}$$

From step 1, it follows

$$\sum_{i\in\Lambda^+(\tilde{n})}S_i'(\tilde{n},\tilde{\varepsilon})\cong 0. \tag{C-14}$$

for any $\tilde{\varepsilon}$ such that $||\tilde{\varepsilon}||$ is small. Thus, combining (C-11) and (C-13) yields

$$\sum_{i\in\Lambda(\tilde{n})}(n_i+\varepsilon_i)\cdot W(V_i(S_i(\tilde{n}+\tilde{\varepsilon}),n_i+\varepsilon_i))$$

$$-\sum_{i\in\Lambda(\tilde{n})}(n_i+\varepsilon_i)\cdot W(V_i(S_i'(\tilde{n},\tilde{\varepsilon}),n_i+\varepsilon_i))$$

$$\cong\sum_{i\in\Lambda^+(\tilde{n})}(n_i+\varepsilon_i)\cdot W(V_i(S_i(\tilde{n}+\tilde{\varepsilon}),n_i+\varepsilon_i))$$

$$-\sum_{i\in\Lambda^+(\tilde{n})}(n_i+\varepsilon_i)\cdot W(V_i(S_i'(\tilde{n},\tilde{\varepsilon}),n_i+\varepsilon_i))$$

$$\cong\sum_{i\in\Lambda^+(\tilde{n})}(n_i+\varepsilon_i)\cdot W(V_i(S_i(\tilde{n}+\tilde{\varepsilon}),n_i+\varepsilon_i))$$

$$-\sum_{i\in\Lambda^+(\tilde{n})}(n_i+\varepsilon_i)\cdot W(V_i(S_i(\tilde{n}),n_i)) \tag{C-15}$$

for any $\tilde{\varepsilon}$ such that $||\tilde{\varepsilon}||$ is small, from the continuity of $V_i(S_i,n_i)$ with respect to S_i and n_i except $n_i=0$. Hence, it follows from (C-10) that

$$\sum_{i\in\Lambda(\tilde{n})}(n_i+\varepsilon_i)\cdot W(V_i(S_i(\tilde{n}+\tilde{\varepsilon}),n_i+\varepsilon_i))$$

$$-\sum_{i\in\Lambda(\tilde{n})}(n_i+\varepsilon_i)\cdot W(V_i(S_i'(\tilde{n},\tilde{\varepsilon}),n_i+\varepsilon_i))\le-\delta<0 \tag{C-16}$$

for any $\tilde{\varepsilon}$ such that $||\tilde{\varepsilon}||$ is small. This contradicts the optimality of $\tilde{S}(\tilde{n}+\tilde{\varepsilon})$. As a result, we can conclude that $\tilde{S}(\tilde{n}) = \lim_{\|\tilde{\varepsilon}\|\to 0}\tilde{S}(\tilde{n}+\tilde{\varepsilon})$. □

Appendix D

In this appendix, we derive some equations in section 4.2.

1. The derivation of equation (22):

From the Samuelson condition, it follows that

$$x_i = \frac{\alpha}{n_i \cdot (\alpha+\beta)}(f_i(n_i)+S_i), \quad \text{(D-1)}$$

$$G_i = \frac{\beta}{\alpha+\beta} \cdot (f_i(n_i)+S_i). \quad \text{(D-2)}$$

Plugging (D-1) and (D-2) into (21) yields (22).

2. The derivation of (23):

The social optimum level of transfer is

$$S_1(\tilde{n}) = \frac{1}{n_1+n_2}[\{(1-\gamma)\cdot n_1 + \gamma \cdot n_2\} \cdot f_2(n_2)$$
$$- \{(1-\gamma)\cdot n_2 + \gamma \cdot n_1\} \cdot f_1(n_1)]. \quad \text{(D-3)}$$

Combining (D-3) and (22) implies (23).

3. The derivation of (25):

The derivative of $V_1(\tilde{n}) - V_2(\tilde{n})$ evaluated at $n_1 = n_2$ is

$$\frac{d(V_1(\tilde{n}) - V_2(\tilde{n}))}{dn_1} = \frac{4}{N} \cdot \{(1-2\gamma)^2 - \alpha\} \quad \text{(D-4)}$$

which implies (25).

4. The derivation of (27):

If BFE is an interior solution, it follows that

$$n_1^B = N \cdot \left\{1 - \frac{1}{1+(k_1/k_2)^{(1-2\gamma)/(\alpha-(1-2\gamma)\delta)}}\right\}, \quad \text{(D-5)}$$

$$n_2^B = \frac{N}{1+(k_1/k_2)^{(1-2\gamma)/(\alpha-(1-2\gamma)\delta)}}, \quad \text{(D-6)}$$

Substituting (D-6) and (D-7) into $V_i\ (\tilde{0}, \tilde{n}^B)$ yields (27).

5. The derivation of (28):

Differentiating $V_1(\tilde{0},\tilde{n}) - V_2(\tilde{0},\tilde{n})$ with respect to n_1 under the constraint $n_1 + n_2 = N$, we have

$$\frac{d(V_1(\tilde{0},\tilde{n}) - V_2(\tilde{0},\tilde{n}))}{dn_1} = \frac{N}{n_1 \cdot n_2} \cdot \{(1-2\gamma) \cdot \delta - \alpha\} \qquad \text{(D-7)}$$

which implies (28).

References

Atkinson, A. B. and J. E. Stiglitz (1980), "Lectures on Public Economics," McGraw-Hill, New York.

Bergstrom, T. C. and R. P. Goodman (1973), "Private Demands for Public Goods," *American Economic Review* 63, 280-96.

Boadway, R. W. (1982), "On the Method of Taxation and the Provision of Local Public Goods: Comment," *American Economic Review* 72, 846-51.

Boadway, R. W. and F. R. Flatters (1982), "Efficiency and Equalization Payments in a Federal System of Governments: A Synthesis and Extension of Recent Results," *Canadian Journal of Economics* 15, 613-633.

Boadway, R. W. and D. E. Wildasin (1990), "Optimal Tax-Subsidy Policies for Industrial Adjustment to Uncertain Shocks," *Oxford Economic Papers* 42 105-134.

Buchholz, W. and K. A. Konrad (1995), "Strategic Transfers and Private Provision of Public Goods," *Journal of Public Economics* 57, 489-505.

Chari, V. V., P. J. Kehoe, and E. C. Prescott (1989), "Time Consistency and Policy," in R. Barro (ed.) Modern Business Cycle Theory (Cambridge: Harvard University Press), Chapter 7.

Cornes, R. (1993), "Dyke Maintenance and Other Stories: Some Neglected Types of Public Goods," *The Quarterly Journal of Economics* February, 259-271.

Fisher, S. (1980), "Dynamic Inconsistency, Cooperation, and the Benevolent Dissembling Government," *Journal of Economic Dynamics and Control* 2(1), 93-107.

Flatters, F. R., J. V. Henderson, and P. M. Mieszkowski (1974), "Public Goods, Efficiency, and Regional Fiscal Equalization," *Journal of Public Economics* 3, 99-112.

Hartwick, J. M. (1980), "The Henry George Rule, Optimal Population, and Interregional Equity," *Canadian Journal of Economics* 13, 99-112.

Hirshleifer, J. (1983), "From Weakest-Link to Best Shot: The Voluntary Provision of Public Goods," *Public Choice* XLI, 371-86.

Hirshleifer, J. (1985), "From Weakest-Link to Best Shot: Correction," *Public Choice* XIVI, 21-23.

Ihori, T. (1991), "Immiserizing Growth with Interregional Externalities of Public Goods," *Regional Science and Urban Economics* 24, 485-96.

Kydland, E. F. and E. C. Prescott (1977), "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy* 85(3), 473-91.

Mansoorian, A. and G. M. Myers (1993), "Attachment to Home and Efficiency Purchases of Population in a Fiscal Externality Economy," *Journal of Public Economics* 52, 117-32

Myers, G. M. (1990), "Optimality, Free Mobility, and the Regional Authority in a Federation," *Journal of Public Economics* 43, 107-21

Persson, T. and G. Tabellini (1990), "Macroeconomic Policy, Credibility, and Politics," (Hardwood Academic Publishers, New York).

Wellisch, D. (1994), "Interregional Spillovers in the Presence of Perfect and Imperfect Household Mobility," *Journal of Public Economics* 55, 167-84

Wildasin, D. E. (1988), "Nash Equilibrium in Models of Fiscal Competition," *Journal of Public Economics* 35, 229-40.

Zodrow, G. R. and P. Mieszkowski (1986), "Pigou, Tiebout, Property Taxation, and the Underprovision of Local Public Goods," *Journal of Public Economics* 19, 356-70.

area	Interior BFE	Symmetric TCE	$k_1 \neq k_2$	:	$k_1 = k_2$
a	LS*	LUS	$V_{BFE} > V_{TCE}$ when $\gamma=0$	:	$V_{BFE} > V_{TCE}$
b	LUS**	LUS	$V_{BFE} \leq V_{TCE}$ (equality holds when $\gamma=0$)		
c	LUS	LUS	$V_{BFE} < V_{TCE}$		
d	LS	LS	$V_{BFE} < V_{TCE}$	:	$V_{BFE} = V_{TCE}$
e	LS	LS	?	:	$V_{BFE} = V_{TCE}$

*) LS stands for locally stable.

**) LUS stands for locally unstable.

Table 1

$< N{=}2,\ k_1 = 11,\ k_2 = 10, \text{ and } \alpha = 0.7 >$

δ	γ	n_1^*	V_{KE}	S_1	$\frac{S_1}{f_1(n_1^*)}$	$\frac{S_2}{f_2(n_2^*)}$	n_1^B	V_{BE}
	0.00	2.000	2.329	0.000	0.0	*	1.443	2.361
	0.02	1.986	2.277	0.320	1.9	-40.5	1.353	2.357
0.6	0.05	1.886	2.274	0.812	5.0	-29.9	1.262	2.354
	0.07	1.652	2.316	0.877	5.9	-16.5	1.219	2.353
	0.10	1.000	2.351	-0.500	-4.5	5.0	1.172	2.352
	0.00	2.000	2.398	0.000	0.0	*	2.000	2.398
	0.02	1.986	2.328	0.023	0.1	-4.5	1.927	2.350
0.7	0.05	1.886	2.302	0.228	1.3	-10.4	1.546	2.350
	0.07	1.652	2.328	0.294	1.9	-6.2	1.395	2.350
	0.10	1.000	2.351	-0.500	-4.5	5.0	1.266	2.350

Table 2

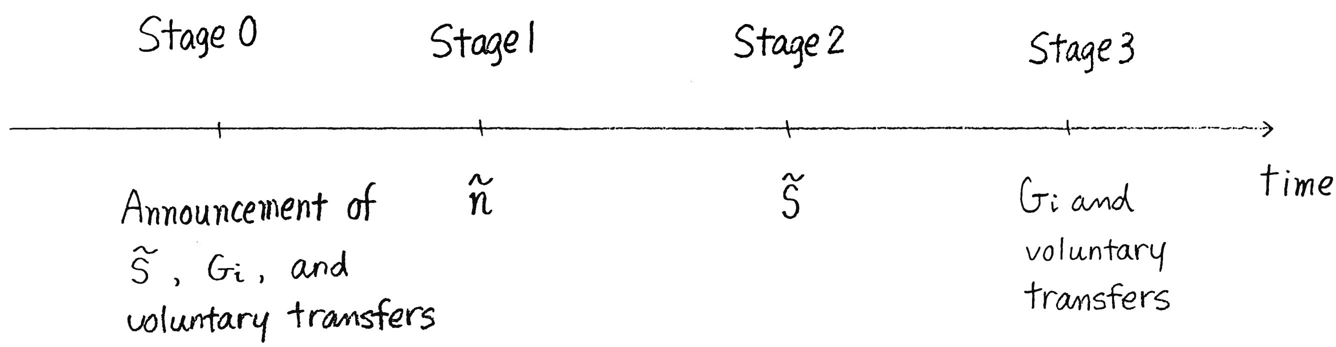

Stage 0
Stage 1
Stage 2
Stage 3
time
Announcement of
$\tilde{S}$, G_i, and
voluntary transfers
$\tilde{n}$
$\tilde{S}$
G_i and
voluntary
transfers

Figure 1

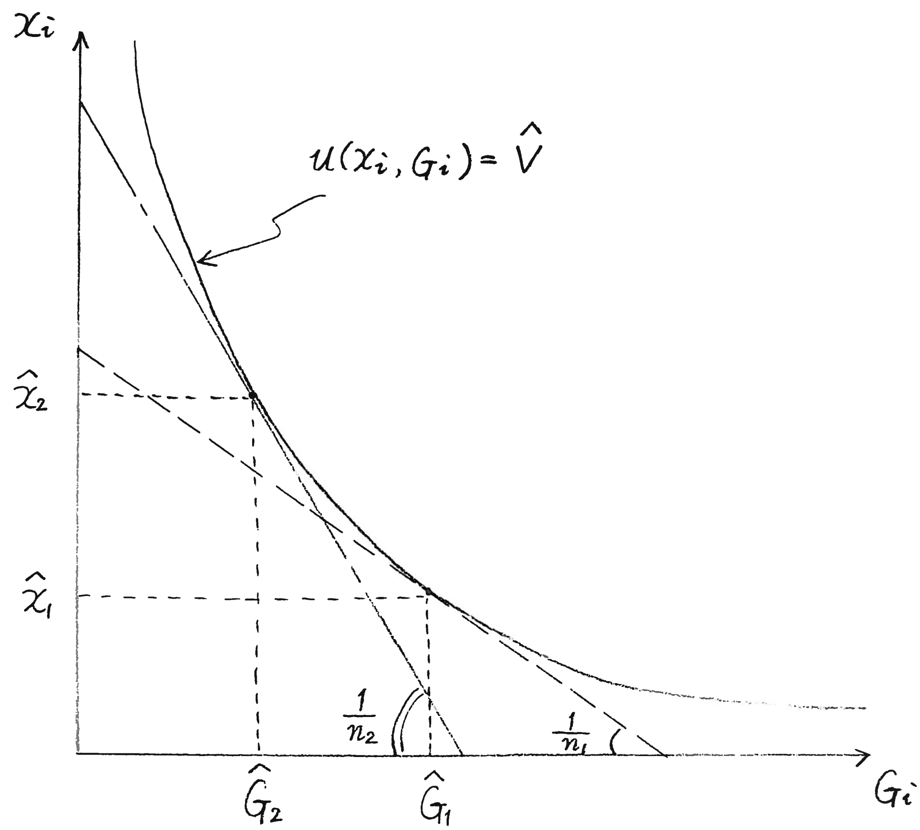


Figure 2

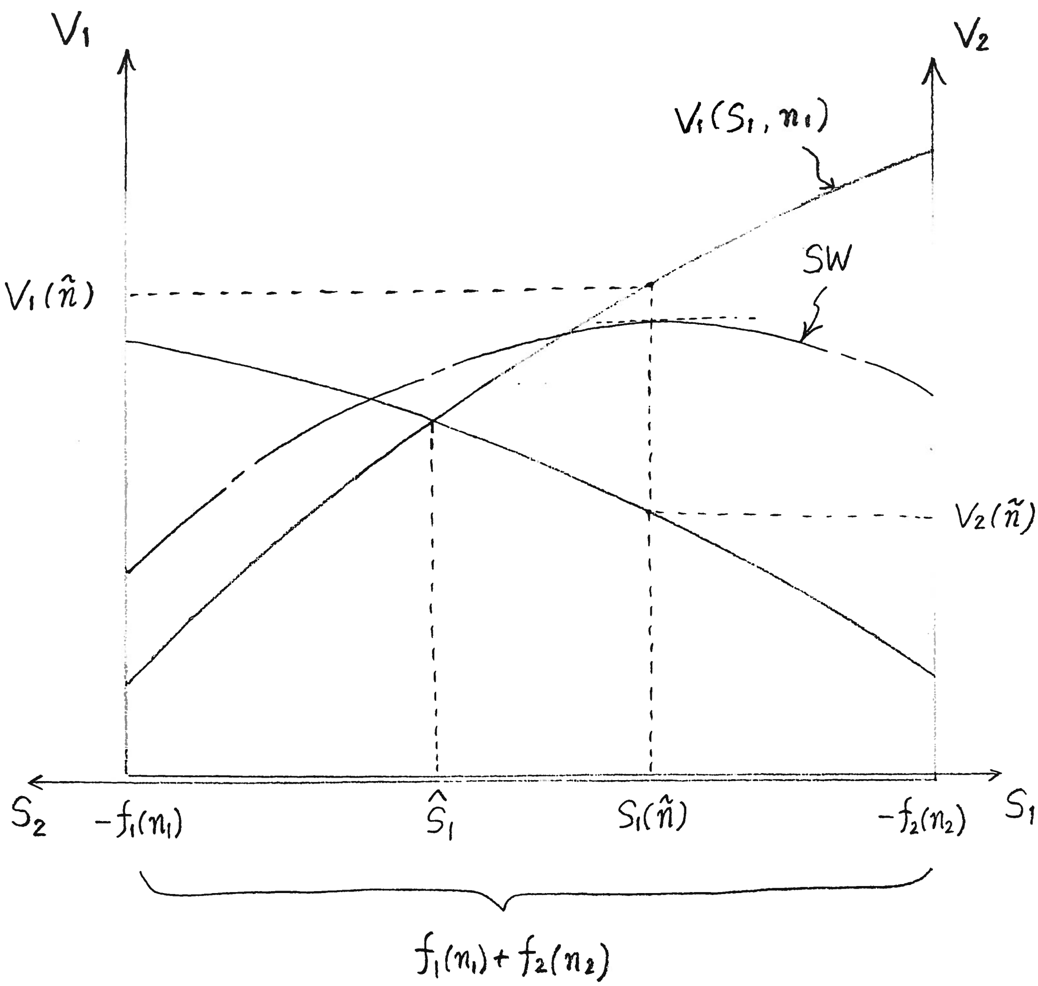


Figure 3

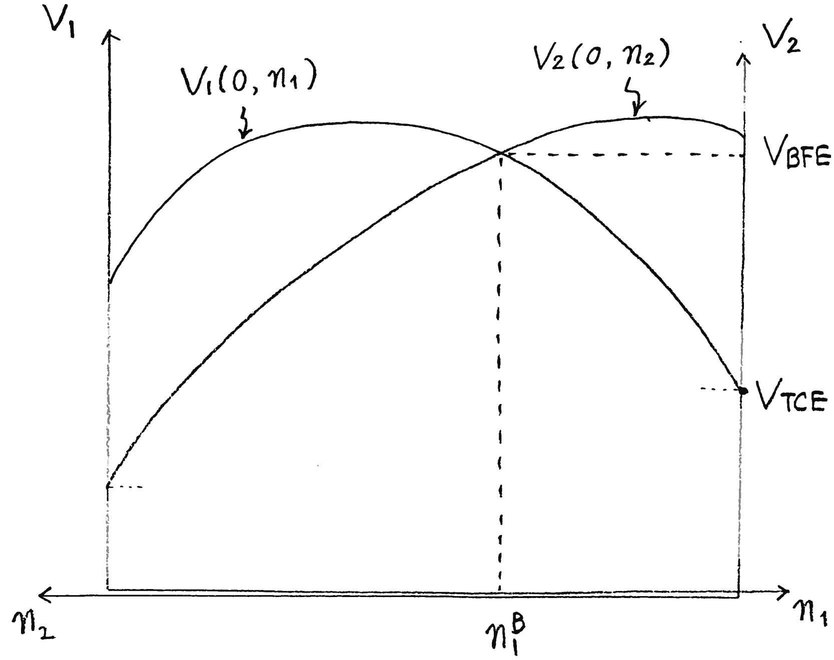


Figure 4 (a)

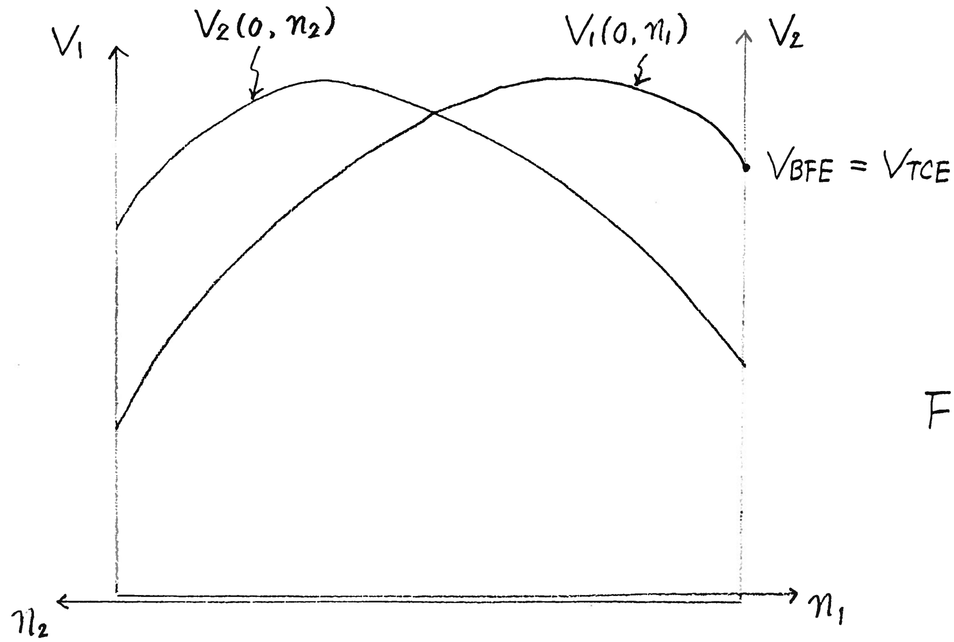


Figure 4 (b)

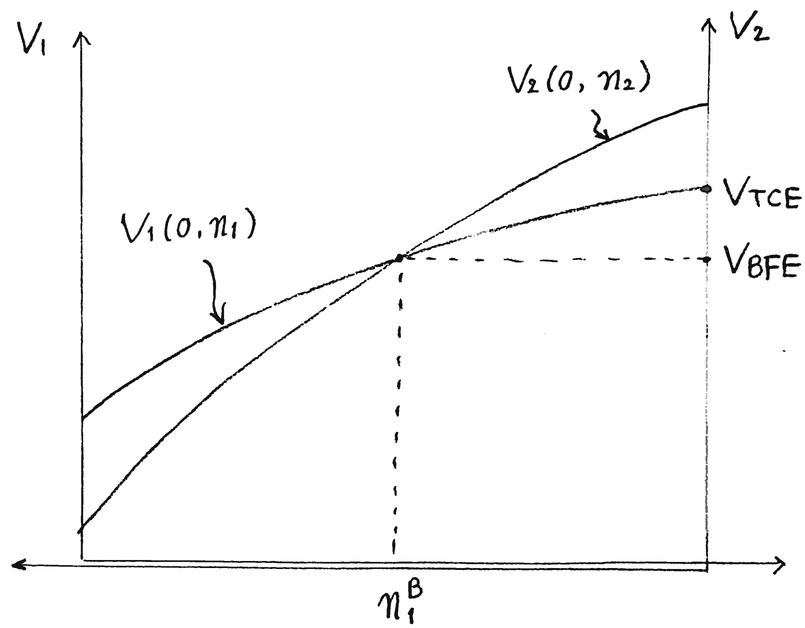
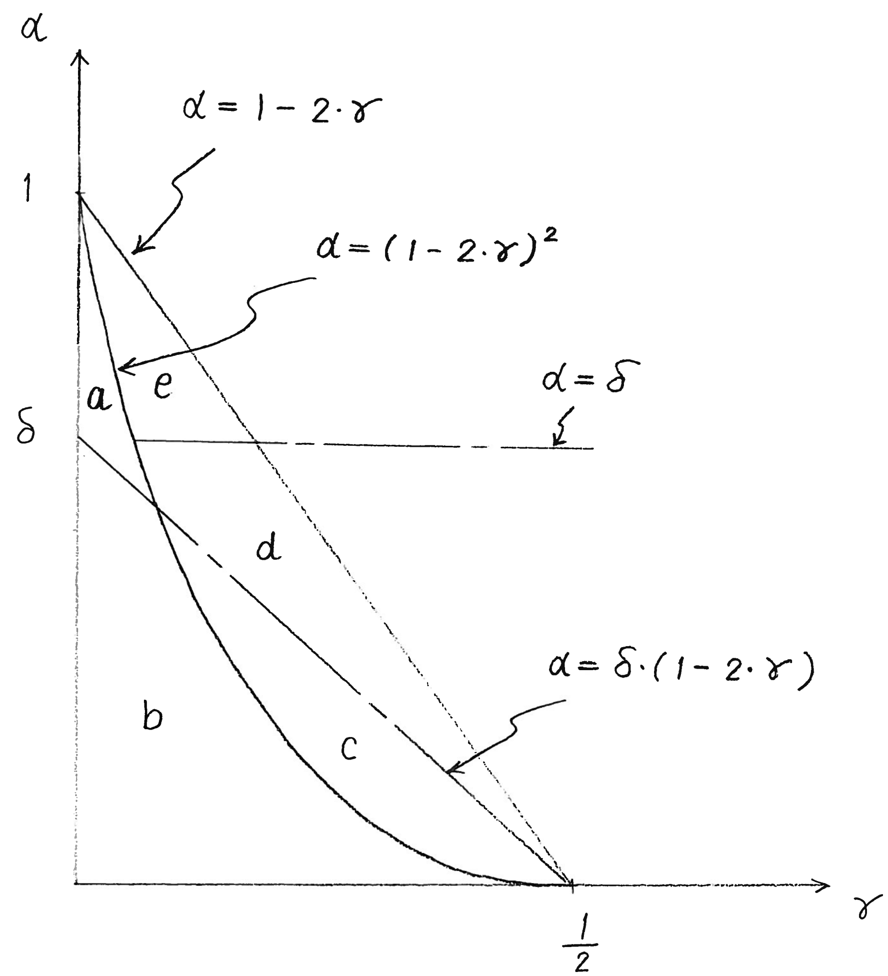


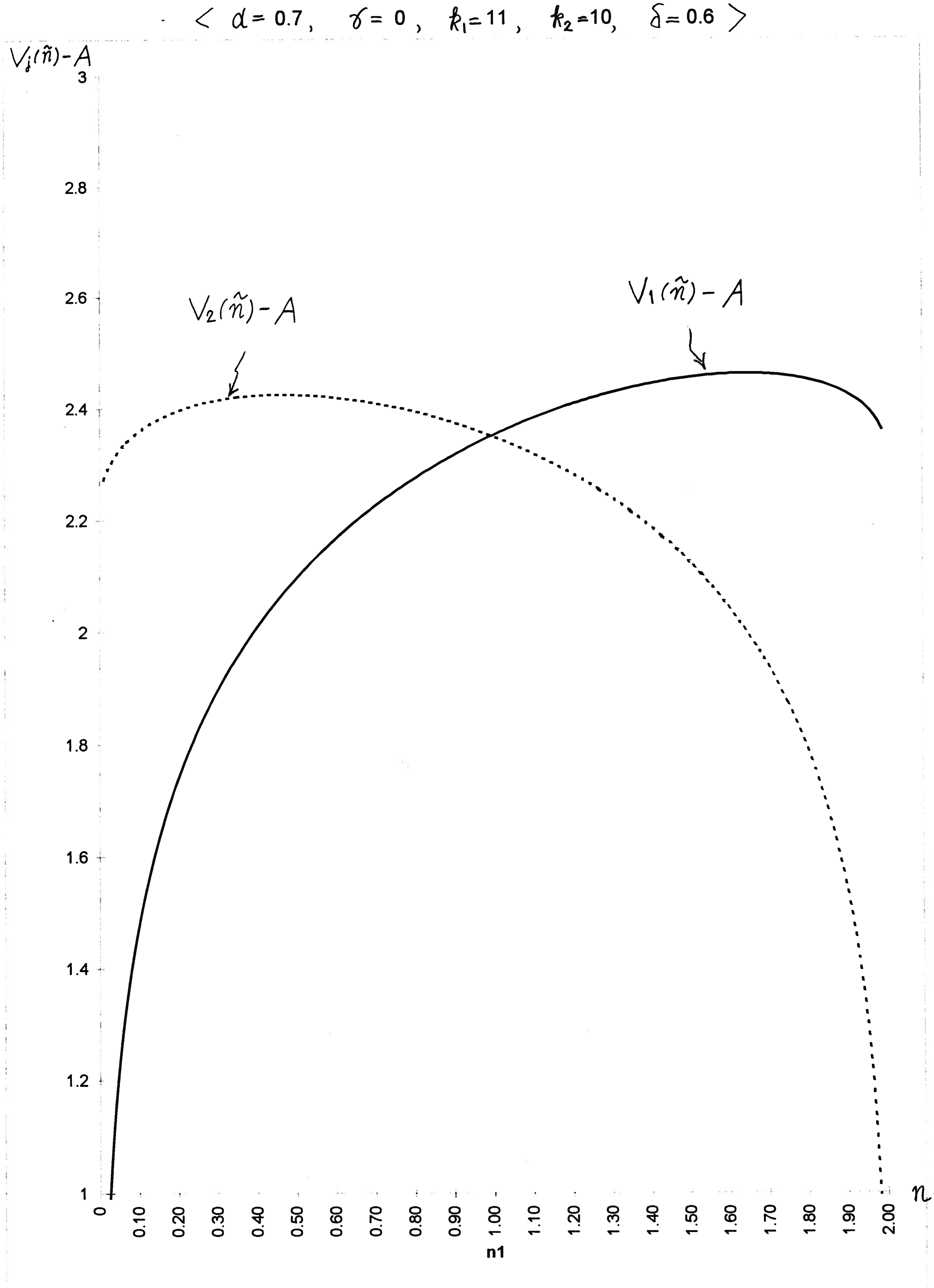
Figure 4(c)



*) Interior BFE is locally stable in area "a", "d", or "e".

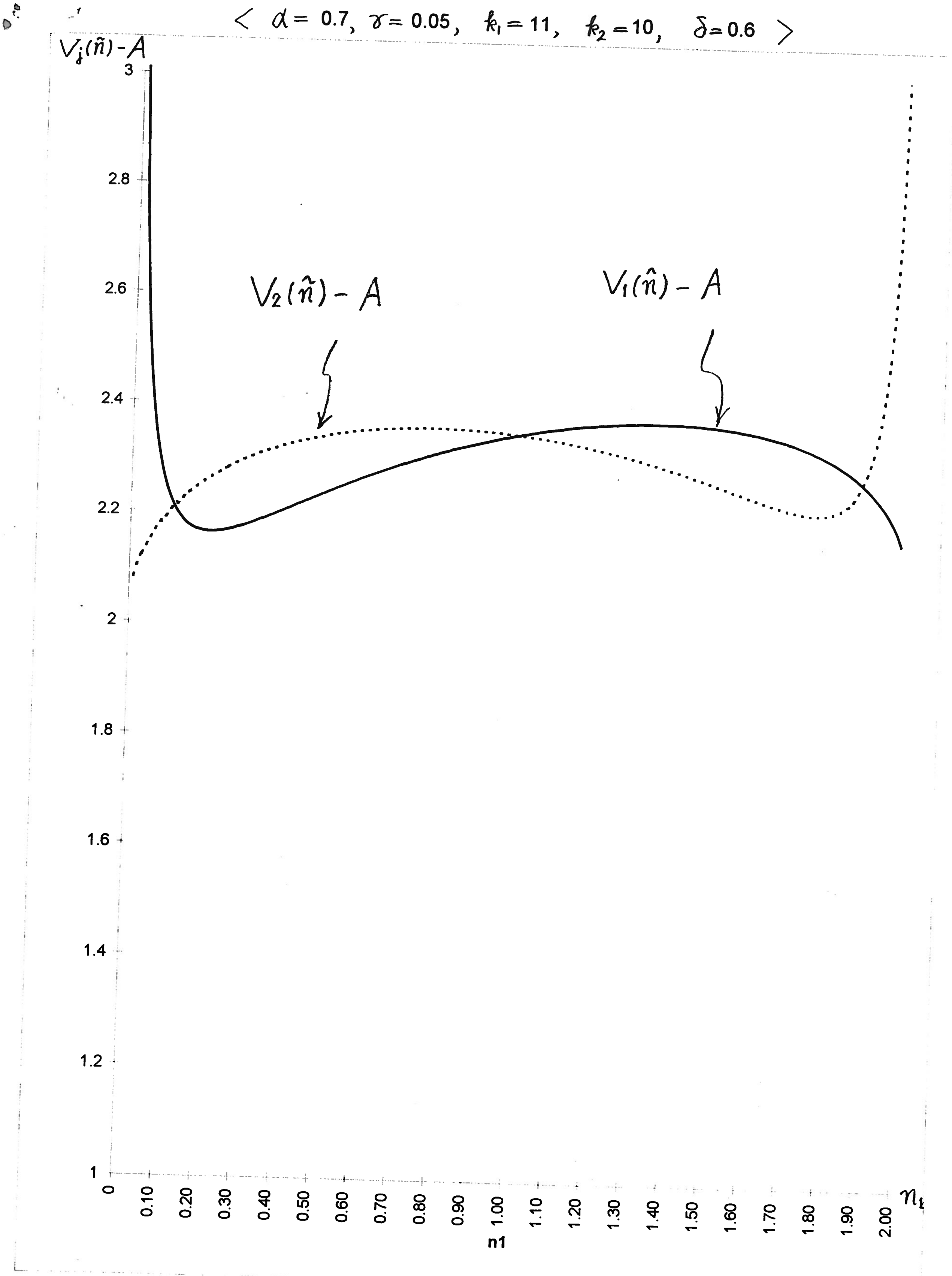
**) Symmetric TCE is locally stable in area "c", "d", or "e".

Figure 5



< α = 0.7, γ = 0, k1 = 11, k2 = 10, δ = 0.6 >
Vj(ñ) - A
V2(ñ) - A
V1(ñ) - A
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
1.10
1.20
1.30
1.40
1.50
1.60
1.70
1.80
1.90
2.00
n1
n1

Figure 6



< α = 0.7, γ = 0.05, k1 = 11, k2 = 10, δ = 0.6 >
Vj(ñ) - A
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
V2(n̂) - A
V1(n̂) - A
0
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
1.10
1.20
1.30
1.40
1.50
1.60
1.70
1.80
1.90
2.00
n1

Figure 7