IZA DP No. 8035

Is a Policy of Free Movement of Workers Sustainable?

Pierre M. Picard
Tim Worrall

March 2014

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Is a Policy of Free Movement of Workers Sustainable?

Pierre M. Picard
University of Luxembourg,
CORE, Université Catholique de Louvain and IZA

Tim Worrall
University of Edinburgh

ABSTRACT

Is a Policy of Free Movement of Workers Sustainable?[*]

This paper studies the costs and benefits of the adoption of the policy of free movement for workers. For the countries to agree on uncontrolled movement of workers, the short run costs must be outweighed by the long term benefits that result from better labor market flexibility and income smoothing. We show that such policies are less likely to be adopted when workers are impatient and less risk averse workers, when production technologies display decreasing returns and when countries trade a share of their products.

JEL Classification: F22, J61

Keywords: labor market flexibility, migration, sustainable plan

Corresponding author:

Pierre M. Picard
Département Economie
University of Luxembourg
162A avenue de la Faiencerie
L-1511 Luxembourg
Luxembourg
E-mail: pierre.picard@uni.lu

1 Introduction

Since its inception, the European Union has aimed at implementing a policy of free movement of workers between member states (Article 45, E.U. Lisbon Treaty). However, some E.U. member states have been reluctant implement this policy, either implementing it in stages or applying different standards of implementation, or in some cases applying policies as restrictive as for non-E.U. immigrants. Similar projects and difficulties have emerged in other country associations like the North American Free Trade Area (N.A.F.T.A.) in which the TN-status currently allows the free movement of workers only between United States of America and Canada. The main reason of this reluctance lies in the fear that inflows of migrant workers may depress local labor market conditions and the welfare of the host country workers.[1]

It is not difficult to find evidence for why country associations might aim for a policy of free movement of workers and fail to agree on it. Migration offers large benefits to the immigrants (e.g. Klein and Ventura, 2009; Clemens et al., 2010) but imposes short run costs to local workers, particularly, to the low skilled ones (e.g. Borjas, 2003). In this paper, we consider the balance of short-term costs with long-term benefits and examine the decision of countries to open their borders and adopt policies of unconditional and uncontrolled movements of workers. The approach we adopt is similar to the literature on sustaining free trade (see, for example, Bagwell and Staiger, 1990; Staiger and Bagwell, 1999; Bagwell and Staiger, 2005; Grossman and Helpman, 1995).[2] We consider a repeated model with uncertainty. To sustain a policy of free movement of workers requires that no country finds it unilaterally beneficial to breach the policy. As in the sustainable trade literature, we suppose that a breach leads to reversion to repetition of a short-run Nash equilibrium where labor migration is completely controlled. Each country therefore weighs the short-run cost of immigration against the long-term benefits and the policy of free movement of workers is sustainable if at each date and state the long-term benefits exceed the short-run costs for each country.

To examine this issue, we consider a simple, two-country, dynamic model where, under free movement of labor, workers freely choose their location in each time period. We initially consider a model where

---

[1] In 2005, the referenda rejecting the European Constitution by the Dutch and French publics occurred in the context of the debate over the accession of Turkey to the E.U. In 2011, the fear of uncontrolled immigration waves has enticed France to threaten to suspend its obligation to the E.U. freedom of movement (Schengen Treaty) (Waterfield, 2011).

[2] More generally, our approach relates to the discussion about sustainable government policies (see, for example, Chari and Kehoe, 1990; Acemoglu et al., 2010) and self-enforcing insurance mechanisms (see, for example, Thomas and Worrall, 1988; Ligon et al., 2002).

immigration imposes a cost on locals (Sections 2 and 3) and then justify these costs by examining a standard migration model without trade but with productivity shocks (Section 4) and add a simple Ricardo-Viner trade model with production of both tradeable and non-tradeable goods (Section 5). Section 5 also considers the implications of labor market rigidities.

In Section 3, we show how a policy of free movement of workers integrates labor markets and allows workers (and their descendants) to relocate to more productive countries and allows them to smooth their consumption. Risk averse workers benefit from better income smoothing under the policy of free movement of labor. Section 3 highlights that permanent differences in consumption levels between countries make it more difficult to sustain a policy of free movement of labor. This suggests a reason why uncontrolled or unconditional migration policies seem more likely to be adopted amongst developed countries rather than between developed and less developed countries. To our knowledge, this view that migration fosters labor market flexibility and insurance has not been discussed in either the labor or migration literature.

Section 4 presents a particular case where production is subject to country specific productivity shocks and constant or decreasing returns to scale. It is shown that, absent trade effects, weaker congestion enhances countries' incentives to adopt a policy of free movement of workers because it mitigates the short-run costs of immigration. Likewise, increased risk aversion also enhances countries' incentives to adopt such a policy because it raises the benefits from income smoothing that the policy brings.

Section 5 discusses the sustainability of the policy of free movement of labor amongst two trade partners. The presence of trade alters the costs and benefits of a policy of free movement of workers through a terms of trade effect. This terms of trade effect is discussed by Davis and Weinstein (2002) and Felbermayr and Kohler (2007). Our analysis with only tradeable goods is similar to Davis and Weinstein (2002) who show that immigration have an negative welfare impact on local workers through the changes in the terms of trade. However, the terms of trade also attenuate the impact of productivity shocks on consumption and mitigate workers' incentives to move. We show that they fully absorb the shocks and eliminate any migration incentives when countries and shocks are symmetric. As a result, we are able to show that the terms of trade mitigate or even eliminate the incentives to move.[3] In the presence of both tradeable and

---

[3]In this case, free trade policies are substitutes for labor movement policies from the viewpoint of consumption smoothing. This resembles the statement in Mundell (1957) statement about the substitutability between movements of factors and goods

non-tradeable goods, changes in the terms of trade have an adverse effect on the relative import prices and the consumption basket of domestic workers. Thus, free movement of workers is less likely to be sustained.

The main contribution of our paper is threefold. First, as far as we know, this is the first paper to apply the methods used in modelling sustainable trade, assessing short run costs against long run benefits (as in Bagwell and Staiger, 1990), to the issue of migration. Unlike that literature that considers optimal tariffs, here we consider only a policy of free movement of labor and not a policy of optimal controlled migration. We concentrate on the case of free movement of labor of because we believe this is the relevant case for many countries and because many country associations operate or are discussing the operation of policies of free movement of labor. Although our model abstracts from many important factors considered in the migration literature (such as, skill heterogeneity, skill complementary in production, physical capital adjustment, voter sentiment, etc.), it puts at the forefront uncertainty and intertemporal welfare that are important factors in assessing migration policy. Second, the paper contributes to our understanding of the interaction of trade and migration and we extends the results of Davis and Weinstein (2002) and Felbermayr and Kohler (2007). We show how the terms of trade effects interact with congestion factors and with the share of tradeable and non-tradable goods in determining if a policy of free movement of workers is sustainable. Finally the paper complements to the migration literature by studying free migration policies as a reciprocal mechanism based on the potential benefit of labor market flexibility and income smoothing. In contrast to our approach, that literature mainly focuses on the (often static) welfare assessment of migration in the presence of skill heterogeneity, skill complementary in production, physical capital adjustment, political economy, social capital, etc. (e.g. Borjas, 2003; Facchini and Testa 2009, Wahba and Zenou 2009, Ottaviano and Peri 2012). Calibration studies have concluded the existence of large gains from migration (Klein and Ventura, 2009; Clemens et al., 2010). This papers shows that, although gains may be high, free movement of workers may not be adopted on the ground of the reciprocity and labor market flexibility.

---

in Hecksher-Ohlin frameworks. However, it is not equivalent because workers do not produce the same goods after they move. Also, movements of workers and commodities are complements in the sense that local productivity shocks increase both export and immigration (Markusen, 1983; Neary, 1995).

2 The general model

We consider a model with continuum of workers, of size $\bar{L}$, distributed over two countries. The home country has L workers and the foreign country L^* where $L + L^* = \bar{L}$ (an asterisk * denotes foreign variables). We let L^0 denote the initial population of workers in the home country (with $\bar{L} - L^0$ in the foreign country). We assume that all workers derive utility from a composite consumption good C according to an increasing, differentiable and concave utility function $U(C)$. Labor is mobile between countries and we let l denote the amount of labor working in the local country, whether home or foreign. As well as mobility of labor between countries, there is uncertainty and we describe this uncertainty by supposing there is a set of state of nature $\mathcal{S} \equiv \{1, ..., S\}$. In Sections 4 and 5 we model this uncertainty as deriving from a productivity shock. For the moment it is useful to think of consumption in the two countries depending on both the state s and the labor in each country l. In particular, we suppose that the consumption of domestic and foreign workers is determined by strictly decreasing differentiable functions $C_s(l)$ and $C_s^*(l)$, $l \in [0, \bar{L}]$. Furthermore, we assume a no-crossing property such that for any $s \in \mathcal{S}$ either $C_s(l) \geq C_s^*(l)$ for all $l \in [0, \bar{L}]$ or $C_s(l) \leq C_s^*(l)$ for all $l \in [0, \bar{L}]$. Microeconomic foundations for this formulation will be given the subsequent sections and will include both a labor market and a tradeable goods market. We first discuss the implications of free labor mobility between countries in a short-run equilibrium and in a long-run dynamic equilibrium with participation constraints.

2.1 Short-run Equilibrium

We assume that labor movement occurs in response to the observed state of nature. Thus, we can consider the equilibrium migration decisions for a given state s. Free movement of workers implies that utility is equal in each country. Since workers have same preferences, this implies consumption is equalized too.[4] Hence, given that the consumption functions are decreasing, there will be a unique equilibrium labor allocation in each state, L_s, that satisfies:

$$C_s(L_s) = C_s^*(\bar{L} - L_s). \tag{1}$$

---

[4]Appendix A shows that this equilibrium condition remains valid in the presence of heterogeneous migration costs.

It follows from the no crossing property that $L_s \geq L_s^*$ if and only if $C_s(l) \geq C_s^*(l)$ for any $l \in [0, \bar{L}]$. That is, population is higher in the country that is favoured by the state of nature. It is worth noting that this equilibrium is Pareto efficient in the sense that there is no possibility to relocate an infinitely small mass of workers without negatively affecting the utility of some other workers. However, an interesting question to ask is whether there is too much or too little movement of labor when there is scope for compensation. To address this question consider the utilitarian aggregate welfare in state s, $W_s \equiv L_s U(C_s(L_s)) + L_s^* U(C_s^*(L_s^*))$. The effect on welfare in state s of a marginal increase in population of the home country is given by

$$\frac{\mathrm{d}W_s}{\mathrm{d}L_s} = U\left(C_s\left(L_s\right)\right) - U\left(C_s^*\left(L_s^*\right)\right) + L_s U'\left(C_s\left(L_s\right)\right) C_s'\left(L_s\right) - L_s^* U'\left(C_s^*\left(L_s^*\right)\right) C_s^{*\prime}\left(L_s^*\right) \tag{2}$$

The free movement of workers is efficient if this expression is zero at the equilibrium. There is excess agglomeration in the home country if this expression is negative, Since $C_s\left(L_s\right) = C_s^*\left(L_s^*\right)$, the two first terms in (2) cancel out. Hence, there is excess agglomeration in the home country if and only if

$$\frac{L_s^* C_s^{*\prime}\left(L_s^*\right)}{L_s C_s'\left(L_s\right)} < 1 \tag{3}$$

The free movement equilibrium yields the socially optimal spatial distribution of workers if and only if the marginal migrant causes a fall in the total domestic consumption (denominator) that is equal to the increase in the total foreign consumption (numerator). If the former is larger than the latter, there is excess agglomeration of workers. The marginal welfare valuation of the policy of free movement of workers does not consider the marginal migrant because the latter has only a second order gain. The first order gains and losses accrue respectively to the populations that the marginal migrant quits and joins. There are some specifications of the consumption functions such that free labor movement does maximize W_s. These include the logarithmic case where $C_s(l) = \lambda_s \ln l$ for $\lambda_s \in \mathbb{R}_0^+$ or the iso-elastic case where $C_s(l) = \lambda_s l^\alpha$ for for $\alpha \in (0, 1)$ and $\lambda_s \in \mathbb{R}_0^+$. As will be shown in Sections 4 and 5, even with Cobb-Douglas preferences, the consumption function $C_s(l)$ may not satisfy these conditions when there are both tradeable and non-tradeable goods.

2.2 Long-run Equilibrium

We now consider a dynamic version of this model. We focus on a discrete time model with an infinite horizon. We assume that workers are infinitely lived and have the same discount factor $\delta \in (0,1)$.[5] We assume that states of nature are i.i.d. across periods and that state s occurs with probability $p_s > 0$ where $\sum_s p_s = 1$. With this assumption and an infinite time horizon, agents' decisions depend only on the current state, so that we can analyze all decisions in the current time period and to drop the reference to time. We denote the contemporaneous utility by $u_s(l) \equiv U(C_s(l))$ and $u_s^*(l) \equiv U(C_s^*(l))$.

Let V_s denote the expected discounted utility of a worker in the home country in state s with V_s^* similarly defined for a worker in the Foreign country. With free mobility of labor, workers can choose where to work and we have

$$V_s = \max\{\, u_s(L_s) + \delta \mathrm{E}_q V_q, u_s^*(L_s^*) + \delta \mathrm{E}_q V_q^* \,\}$$

The same expression applies to the Foreign worker. Because free movement of workers implies $V_s = V_s^*$, it follow that $u_s(L_s) = u_s^*(L_s^*)$ and $C_s(L_s) = C_s^*(L_s^*)$. Consequently, the equilibrium in the dynamic model coincides with the sequence of the short-run equilibria.[6] Therefore, $V_s = V_s^* = u_s(L_s) + (\delta/(1-\delta))E_s u_s(L_s)$ where L_s are the short-run equilibrium labor allocations.

It is also possible to consider the welfare properties of this dynamic equilibrium. It is clear, given our assumptions on mobility and the i.i.d. assumption on states, that a utilitarian planner will choose the labor allocation in each state where the expression in (2) is equal to zero.[7]

Thus, our earlier discussion of excess agglomeration applies equally in the static and dynamic cases. Only for certain specifications of the consumption functions does the free movement of workers coincide with the socially optimal allocation. Equally, since all workers are alike, if we assume that workers are randomly allocated across countries, then the ex ante welfare of an individual worker under free movement

---

[5]Under this assumption agents can also be interpreted as dynasties where each generation has an altruism coefficient δ.

[6]This property is valid only under 'uncontrolled' movement of workers and greatly eases the nature and the exposition of our subsequent analysis. It is not valid under 'controlled' movement of workers because future utility levels then differ across countries.

[7]The utilitarian planner chooses the set of labor allocations $\{L_s^t\}$ with time $t = 0,...,\infty$ that maximizes $\sum_{t=0}^{\infty} \delta^t \mathrm{E}_s[L_s^t U(C_s(L_s^t)) + L_s^{*t} U(C_s^*(\bar{L} - L_s^t))]$, which is equivalent to maximizing W_s w.r.t. to L_s and yields the same labor allocation as in the static model.

will be the same as the utilitarian planner's objective. To see this, let $\omega_s \equiv L_s/\bar{L}$ denote the probability that a home worker is allocated to the home country and let $\omega_s^* \equiv L_s^*/\bar{L} = 1 - L_s/\bar{L}$ be the probability of locating to the foreign country. Then a worker's ex ante expected utility is $\mathcal{U} \equiv \sum_s p_s \left[\omega_s u_s(L_s) + \omega_s^* u_s^*(L_s^*) \right]$. Since the states of nature are independent, the preferred allocation maximizes $\left[L_s u_s(L_s) + L_s^* u_s^*(L_s^*) \right] / \bar{L}$, in each state s. This is proportional to the short-run welfare objective W_s. Hence, the preferred allocation is where the expression in (2) is equal to zero. By contrast, the equilibrium expected utility is $\mathrm{E}_s[u_s(L_s)] = \sum_s p_s u_s(L_s)$. If the expression in (2) is not equal to zero at the equilibrium allocation, then a policy of controlling worker movement could improve a worker's ex ante utility.

3 Sustainable policy

We now study whether policies of free movement of workers will be adopted by the two countries. Here we assume that the objective of the country coincides with the objective of its (homogenous) workers. This will be the case if the country takes into account the welfare of its initial L^0 citizens and does not take into account the welfare of migrants into the country. With this objective, immigration has a short-term cost because it lowers consumption (given our assumption that $C_s(l)$ is strictly decreasing). In contrast, however, provided there are states where workers can earn achieve higher consumption by moving to the other country, then there will be long-term benefits from a policy of free movement of workers.

We define a policy of *free movement of workers* as the removal of any control over the movement of workers between countries. More precisely, it is a *common* policy in which both countries *unconditionally* grant *non-permanent* work permits to workers who obtain a job in their jurisdiction. As is typical of many actual migration policies, these non-permanent work permits are automatically associated with non-permanent residence permits. In this paper, we keep a distinction between, on the one hand, work permits and, on the other hand, citizenship and the socioeconomic and political rights that are associated with citizenship. This distinction is important for two reasons. First, it fixes the group of individuals that each government considers as its nationals wherever they work and reside. When workers do not change citizenship or nationality, this group is invariant to the possible relocation of labor between countries. Second, this

distinction determines the alternative policy when countries do not adopt a policy of free movement of workers or when a country decides to breach such a policy. In such cases, we assume that the opting out and breaching countries are able to exert a control on the issue of work permits by putting restrictions and conditions on the number of non-permanent work permits. As a result they can stop renewing existing work permits granted and control the local labor supply.

This setting is not unrealistic. Common policies allowing non-nationals access to local labor markets are often embedded in third-country association agreements or guest worker programs. These agreements and programs permit the economic immigration of third-country nationals into a host country under the control of quotas or individualized labor certifications. For example, the E.U. had such agreements with many Eastern European countries during the 1990s and still has such agreements with some neighboring countries including Turkey and Morocco. Hence, our discussion relates to the E.U. decision to adopt a policy of free movement of workers with Eastern European countries in the 1990s or to the current debate about Turkey's access to the E.U. labor market. Our discussion may also be relevant for the popular concerns about migration issues during the 2005 French referendum about the European Constitution. In North America, the N.A.F.T.A. includes a policy in favor of free movement of workers. In particular, the TN-status gives the right to Canadian, Mexican and citizens of the U.S. to work in each other's countries. The TN-status is limited to three years and to designated professional occupations and can be renewed indefinitely. In practice, the U.S. has implemented a differential treatment for Canadians and Mexicans. Whereas the TN-status is granted to any Canadian at the U.S. border without control and quotas, it is not offered to Mexican nationals. Thus, our discussion also relates to the U.S. and Canadian decision to adopt a common uncontrolled mobility of their nationals within the N.A.F.T.A.; it relates to the U.S. and Mexican decision to remove the present controls and quotas on Mexicans; and it is finally applicable to the extension of the TN-status to other professional occupations and other countries and to the U.S. H1B visa or to the U.S. employment-based green cards, etc.

We suppose that each country has two options: either to adopt the policy of free movement of workers or to independently control the inflow of workers. Consider the option where each country sets immigration controls. Under our assumptions, a country cannot change the welfare of its citizens working in the other

country and puts no weight on the immigrants residing in its own jurisdiction. It then follows from the fact that the consumption function $C_s(l)$ is strictly decreasing that each the dominant strategy of each country is to allow no entry to foreign workers. Thus, the second option reduces to no movement of workers. We assume that the revocation of the right to work is instantaneous, so that any breach from a policy of free movement of labor means that the allocation of labor reverts straightaway to the initial distribution $(L^0, \bar{L} - L^0)$. Furthermore, and for simplicity, we assume that once the agreement about the policy of free movement of workers is breached, the countries play their dominant strategy in all subsequent periods so that there is no movement of labor after a breach.[8]

We say that a policy of free movement of workers is *sustainable* if and only if no country has an incentive to deviate from the policy at any date or state given that deviation means an instantaneous and irrevocable transition to no movement of labor between countries. Let V_s^0 denote the expected discounted utility of a worker in the home country in a given state s when workers are not allowed to cross borders. Analogously to V_s (the expected discounted utility of a worker when there is a policy of free movement of workers defined in the previous section) it is defined recursively by: $V_s^0 = u_s(L^0) + \delta E_q V_q^0$. Hence,

$$V_s^0 = u_s(L^0) + \frac{\delta}{1-\delta} E_s u_s(L^0)$$

with a similar expression holding for foreign workers. Since a country's welfare is identified with that of its representative workers, the policy of free movement of workers is sustainable if and only if $V_s \geq V_s^0$; that is, if

$$u_s(L_s) - u_s(L^0) + \frac{\delta}{1-\delta} E_q[u_q(L_q) - u_q(L^0)] \geq 0 \qquad \forall s \in \mathcal{S}. \tag{4}$$

where L_s and L_q are the equilibrium labor supplies under free movement of workers, with similar expressions holding for the foreign country. We refer to these conditions as *participation* or *self-enforcement* constraints. These conditions compare the short run cost of immigration (first two terms) with the long run benefit of the free movement policy (last term). Condition (4) is most stringent for the state(s) with the highest short

---

[8]This latter assumption may be relaxed without qualitatively altering the results.

run cost, $\bar{s} \in \arg\max_s\{u_s(L^0) - u_s(L_s)\}$. Similarly the equivalent of Condition (4) for the foreign country is more stringent in the state(s) $\bar{s}^* \in \arg\max_s\{u_s^*(\bar{L} - L^0) - u_s^*(\bar{L} - L_s)\}$. Rewriting Condition (4), the policy of free movement of workers is sustainable if and only if

$$
\begin{aligned}
&u_{\bar{s}}(L_{\bar{s}}) - u_{\bar{s}}(L^0) + \tfrac{\delta}{1-\delta}\mathrm{E}_q[u_q(L_q) - u_q(L^0)] \geq 0 \\
&u_{\bar{s}^*}^*(\bar{L} - L_{\bar{s}^*}) - u_{\bar{s}^*}^*(\bar{L} - L^{0*}) + \tfrac{\delta}{1-\delta}\mathrm{E}_q[u_q^*(\bar{L} - L_q) - u_q^*(\bar{L} - L^0)] \geq 0
\end{aligned}
\tag{5}
$$

These conditions lead to the following conclusions. First, sustainability is possible only if there are positive future expected gains. This means that countries must incur shocks in the future that generate emigration to the country with higher consumption. Second, if there are positive future expected gains, then there is a critical discount factor $\delta^c \in (0,1)$ such that Condition (5) is satisfied for $\delta \geq \delta^c$.

We can also ask how sustainability relates to the agglomeration issue outlined in the previous section, to risk aversion and to permanent differences in the consumption functions of the two countries. First recall that excessive agglomeration depends on the sign of $\mathrm{d}W_s/\mathrm{d}L_s$ in (2). If $\mathrm{d}W_s/\mathrm{d}L_s < 0$, then the equilibrium labor allocations exhibit excessive agglomeration, and if $\mathrm{d}W_s/\mathrm{d}L_s > 0$, then the equilibrium labor allocations exhibit under-agglomeration. Either excessive or under-agglomeration mean that the future expected gain $\mathrm{E}_q[u_q(L_q) - u_q(L^0)]$ is lower than it would be if the labor allocation could be chosen optimally. The effect in the short-run is however, different. Under-agglomeration will tend to lower the short-run costs of the policy of free movement of workers, whereas excessive agglomeration will increase the short-run cost. In particular, if there is excessive agglomeration in state $\bar{s}$, then the short run costs of immigration, $u_{\bar{s}}(L_{\bar{s}}) - u_{\bar{s}}(L^0)$, may be increased sufficiently to offset any long-term future gains. Thus, factors that cause agglomeration to be excessive tend to reduce the extent to which a policy of free movement of workers is sustainable. Such factors will be considered in more detail in the next two sections.

Secondly, the impact of increasing risk aversion is a priori ambiguous. Since it is assumed that all workers have the same utility function U, risk aversion does not affect the equilibrium labor allocation, L_s. To the extent to which free movement of workers reduces the variability in consumption and the expected labor population $\mathrm{E}_s[L_s]$ coincides with the initial population L^0, increased risk aversion increases the long-term benefits of the policy. The effect on the short run cost $u_s(L_s) - u_s(L_0)$ is ambiguous, but typically it will

decline with increasing risk aversion, so that the overall impact on sustainability of increasing risk aversion is unclear. However, if workers are infinitely risk averse, then utility is evaluated by the consumption in the worst state and this is always improved by a policy of free mobility. Hence, if risk aversion is sufficiently large, there exists a large enough discount factor δ^c above which free movement of workers becomes a sustainable policy.

Thirdly, consider a case where there are permanent differences between the home country consumption function and the foreign country. For example, suppose the home country consumption function is $C_s(l)$ and the corresponding foreign country consumption is $\lambda C_s^*(l)$ for some parameter $\lambda \in (0, 1)$.[9] Consequently, a fall in λ increases the equilibrium labor allocation $L_s = L_s(\lambda)$ in all states. Since it reduces the domestic instantaneous utility $u_s(L_s)$ in all states, condition (4) become more stringent and the critical discount factor above which the free movement of workers is sustained will be higher. For low enough λ, the long run benefit $E_q[u_q(L_q) - u_q(L^0)]$ becomes non-positive, so that free movement of workers is never a sustainable policy.

Proposition 1 (Freedom of movement of workers) *The policy of free movement of workers is never sustainable if it brings no long-run benefit ($E_q[u_q(L_q) - u_q(L^0)] < 0$). Otherwise, it is sustainable for high enough discount factor δ ($\delta > \delta^c$). The policy is unsustainable for large enough permanent differences in country consumption functions (λ small).*

Proposition 1 shows that a policy of free movement may be adopted provided there are long term benefits that outweigh any short term costs. However, it also shows that where there are no, or small long term benefits, or permanent differences between countries, then such a policy won't be sustainable. The latter observation is particularly true for labor flows between developing and developed countries. As a case in point, European institutions have repeatedly been concerned with immigration consequence of the policy of free movement of labor. For example, the E.U. has discussed Italian migration at its inception stage in the 1950s and the migration of Eastern European workers during its enlargement phase at the beginning of this century. In North America, the TN-status, which offers permission to work within the U.S. under the N.A.F.T.A., has been subject to huge restrictions for Mexican natives whereas it has included very few

---

[9]Such a parametrization can be justified by the models presented in Sections 4 and 5.

restrictions for Canadians. As in Wellisch and Walz (1998) and Ortega (2010), this argument shows that developed countries have greater reluctance to accept migration flows from less developed countries because of a permanent redistribution towards immigrants.

The conditions for sustainability depend on the distribution of the states of nature and their associated consumption levels. To aid understanding and improve analytical tractability and allows us to discuss the impact of uncertainty and dynamics on migration incentives, consider then the simplest symmetric example with two equiprobable and anti-correlated states (similar to the example considered in Bagwell and Staiger, 1990)). That is, suppose $s \in S = \{1, 2\}$, $C_1(l) = C_2^*(l)$ and $C_2(l) = C_1^*(l) \, \forall \, l \in [0, \bar{L}]$ while $L^0 = \bar{L}/2$. This yields the contemporaneous utility levels $u_1(l) = u_2^*(l)$ and $u_2(l) = u_1^*(l)$. Furthermore, order states so that $s = 1$ is more favorable to the home country and state $s = 2$ to the foreign country. Then, $C_1(l) > C_1^*(l)$ and $C_2(l) < C_2^*(l)$ so that $u_1(l) > u_1^*(l)$ and $u_2(l) < u_2^*(l)$. Free movement of workers implies that $u_s(L_s) = u_s^*(L_s^*)$. As a result, the favored country attracts migrants: $L_1 > L^0 > L_2$ where $L_2 = \bar{L} - L_1$. Because states are symmetric, the utility level is the same in all states under free movement of workers: $u_1(L_1) = u_2(L_2)$. Furthermore, because $u_s' < 0$, we have $u_1(L_1) < u_1(L_2)$ and $u_2^*(L_2^*) < u_2^*(L_1^*)$. In the short run, countries with a good state can achieve higher utility levels if they restrict immigration. In this two-state case, the participation constraints (5) collapse to the single condition

$$\frac{\delta}{2 - \delta} \geq G(\theta) \equiv \frac{u_1(\bar{L}/2) - u_1(L_1)}{u_2(L_2) - u_2(\bar{L}/2)} \tag{6}$$

where $G(\theta) > 0$ measures the *relative cost of adopting the policy of free movement of workers* and θ is a vector of parameters of the utility and consumption functions. The value of $G(\theta)$ increases with the fall in domestic consumption after immigration in the good state (numerator) and decreases with the rise in domestic consumption after emigration in the bad state (denominator). Because $\delta/(2 - \delta)$ is an increasing function with a supremum of 1, the policy of free movement of workers is not sustainable when $G(\theta) > 1$. Using (6), it follows that $G(\theta) \leq 1$ if and only if $u_1(L_1) \geq (1/2)(u_1(\bar{L}/2) + u_2(\bar{L}/2))$. We summarize this discussion in the following lemma.

Lemma 1 (Two state model) *Suppose symmetric native populations and two equiprobable anti-correlated*

states. Then, the policy of free movement of workers is never sustainable if $u_1(L_1) \leq \frac{1}{2}[u_1(\bar{L}/2) + u_2(\bar{L}/2)]$. Otherwise, it is sustainable if the discount factor is high enough.

Next we present two models with and without trade that rationalize the properties of the consumption functions we have assumed. We examine the parameter values such that $G(\theta) < 1$ where a policy of free movement of workers is sustainable if the discount factor is high enough.

4 Standard migration model

In the previous section we have been ambivalent about the source of any long-run benefit from a policy of free movement of workers. In this section we discuss two principal sources: labor market flexibility and consumption smoothing. Toward this aim, we embed the above analysis in a simple model with production and labor markets. With decreasing returns to scale at the country or firm level, labor demand schedules are decreasing and therefore local wages and consumption fall with net immigration. In the remainder of the paper, uncertainty is represented by country productivity shocks that shift labor demand.

We consider a two-country model in which the home country produces a local non-tradeable good Z_s. The foreign country produces another local non-tradeable good Z_s^*. The price of these non-tradeable goods can be normalized to one. For this section, there is no trade and workers consume only local goods, so that $C_s = Z_s$. Each worker inelastically supplies one unit of homogeneous labor to the production sector. In the home (foreign) country, L_s (L_s^*) workers are employed at a wage w_s (w_s^*). Each country has a unit mass of firms that produce according to the production function $F_s(l) = \alpha_s f(l)$, where $f(l)$ is increasing and concave and $\alpha_s > 0$ measures country productivity. The foreign country has the same production $f(l)$ but has a productivity parameter $\alpha^* > 0$. Production displays constant returns to scale if $f(l) = l$. In this case, each worker's marginal productivity remains constant irrespective of home country production and labor force. By contrast, production displays decreasing returns to scale or congestion if $f'' < 0$. When $f(l) = 1$, output is equal to α_s, independently of the size of the labor force. This case can be interpreted as a purely agrarian economy with a random crop of fixed size α_s. In what follows it will also be convenient to discuss the case where the production function is iso-elastic: $f(l) = l^\beta$, $\beta \in (0, 1]$. Finally, for simplicity we assume

14

that local profits are redistributed to local individuals.

As is standard, it is possible to interpret the extent of decreasing returns as a measure of congestion. Thus, in the iso-elastic case, β is inversely related to a measure of congestion. If $\beta = 1$, then $f(l) = l$ and there is no congestion; increasing labour supply does not impact on the marginal product of existing workers. If $\beta = 0$, then $f(l) = 1$ and production cannot be increased even with extra workers. It is important to understand this congestion measure. First, the congestion force can be interpreted either at a firm or sector level. At a firm level, each firm, which hires l workers, can be thought of holding a unit of local indivisible capital, which embeds either natural resources, such as land or water, or local human resources, such as local human capital, entrepreneurial skills, etc. At the sector level, decreasing returns to scale can be interpreted as the sharing of common infrastructures, resources and land. In this case, the production function $F_s(L_s)$ applies to the production sector with L_s being the sector employment. Then, each firm can be interpreted as experiencing a sector specific productivity that is equal to $F'_s(L_s) = \alpha_s f'(L_s)$. Second, the reader may interpret the no-congestion case ($f(l) = l$) as a case where production involves capital and labor and where capital is instantaneously and elastically supplied. Third, in a dynamic setting, if capital is chosen before the realization of productivity shocks, the production function has decreasing returns to scale in the short-run and labor demand is downward sloping. The fact that international labor movements impact negatively on local wages is a possible explanation of the observed reluctance amongst the public to accept uncontrolled movements of workers.

4.1 Short run equilibrium and sustainability

In the short run equilibrium, workers consider only current payoffs when deciding whether to move to another country. For the sake of conciseness, we characterize the variables for the home country, those for the foreign country being symmetric. Since workers earn the local wage and receive the local profit, the value of their consumption is equal to the value of their production: $C_s = \alpha_s g(L_s)$ where $g(l) \equiv f(l)/l$ is average productivity (a decreasing function). Under free movement of workers, workers migrate until there

15

is equality of purchasing power:

$$\frac{C_s}{C_s^*} = \frac{\alpha_s g(L_s)}{\alpha_s^* g(L_s^*)} = 1. \tag{7}$$

Hence, $L_s \geq L_s^*$ if and only if $\alpha_s \geq \alpha_s^*$. Workers move into the more productive country because the more productive country offers higher purchasing power. The equilibrium population increases with the productivity ratio α_s/α_s^* (that is, $\mathrm{d}L_s/\mathrm{d}(\alpha_s/\alpha_s^*) = -g(L_s)/(g'(L_s) + g'(L_s^*)) > 0$). There is excess agglomeration of workers if

$$\frac{L_s^* C_s^{*\prime}(L_s^*)}{L_s C_s'(L_s)} = \frac{\alpha_s^* L_s^* g'(L_s^*)}{\alpha_s L_s g'(L_s)} = \frac{L_s^* g'(L_s^*)/g(L_s^*)}{L_s g'(L_s)/g(L_s)} \leq 1,$$

where we use (7) in the second equality. As mentioned in Section 2, the allocation of labor is efficient only if the production function is iso-elastic (i.e. $lg'(l)/g(l)$ is constant).

In the two state case, we let $z = \alpha_1/\alpha_2 > 1$ measure the relative shock. At the end of this section, we will also suppose that workers' preferences exhibit constant relative risk aversion ($U(C) = C^{1-\rho}/(1 - \rho)$, $\rho \geq 0$, $\rho \neq 1$).[10]

4.2 Labor market flexibility

First we eliminate any insurance motive for free movement of workers and suppose that all workers are risk neutral ($\rho = 0$). Thus, any benefits from free movement of workers is derived from the additional labor market flexibility that allows labor supply to locate where demand for labor is strongest. First, note that under constant returns to scale ($f(l) = l$ and $g(l) = 1$), it is easy to check that $G(\theta) = 0$. So, the policy of free movement of workers is always sustainable for any positive discount factor. In each period and state, all workers flow to the country with the higher productivity shock. Local citizens keep the same productivity, so that they incur no short run cost. Immigrants consume what they produce and do not affect the productivity and consumption of locals. Such a benchmark result is regularly used in the literature to argue about the limited impact of migration policies on local citizens. It will be qualified in the next section.

---

[10]For the sake of conciseness, we do not report the case where $U(C) = \ln C$ ($\rho = 1$), although the analysis is essentially similar.

As argued above, full congestion ($f(l) = 1$ and $g(l) = 1/l$) can be interpreted as an agrarian economy with random crops. The allocation of labor depends only on the relative productivities: $L_1/L_2 = z$ or equivalently, $L_1 = (1/(1+z))\bar{L}$. From this it can be checked that $G(\theta) = 1$. Hence, a policy of free movement of workers is never sustainable (given the assumption that $\delta < 1$). Workers have a short run cost in the good state of nature that is exactly equal to the their short run benefit in the bad state. Risk-neutral but impatient workers will not be prepared to incur this short run cost for an equal but uncertain benefit in the future.

Similarly, for small productivity shocks, it can be checked that $\lim_{z \to 1} G(\theta) = 1$. In the limit as $z \to 1$, the (infinitely) small short run cost is equal to the (infinitely) small short run benefit, so that workers have no incentive to relocate. These two cases highlight the fact that free movement of workers may not be a sustainable policy simply because the benefits of a policy of free movement of workers is delayed and uncertain when compared to the costs that are immediate and known. This occurs even if there is no excess agglomeration of workers.

Next consider the case where the production function is iso-elastic ($f(l) = l^\beta$). In this case, $C_s(L_s) = \alpha_s L_s^{\beta-1}$ and $L_1 = (1 + z^{1/(\beta-1)})^{-1}\bar{L}$. Hence,

$$G(\theta) \equiv G(z, \beta) = z \frac{\left(\frac{\bar{L}}{2}\right)^{\beta-1} - L_1^{\beta-1}}{\left(\bar{L} - L_1\right)^{\beta-1} - \left(\frac{\bar{L}}{2}\right)^{\beta-1}} = \frac{z - \left(\frac{1}{2} + \frac{1}{2}z^{\frac{1}{1-\beta}}\right)^{1-\beta}}{\left(\frac{1}{2} + \frac{1}{2}z^{\frac{1}{1-\beta}}\right)^{1-\beta} - 1},$$

where the second equality follows from substituting for L_1 and using the equilibrium condition that $zL_1^{(\beta-1)} = (\bar{L} - L_1)^{(\beta-1)}$. It can be shown that $G(z, \beta) \in (0, 1)$ and is strictly increasing in z and strictly decreasing in β.[11] The smaller the relative shock $z > 1$ and the weaker is the congestion factor (larger is β), the smaller is the critical discount factor for which the policy of free movement of workers can be sustained. This is because smaller shocks and weaker congestion factors reduce the short run costs.

To sum up, when workers are risk neutral, the only benefit from free movement of workers derives from improved labor market flexibility. Labor supply moves to where the demand is stronger. This is a long run

---

[11] Define $\psi(z, \beta) = \left(\frac{1}{2} + \frac{1}{2}z^{\frac{1}{1-\beta}}\right)^{(1-\beta)}$. It is easy to check that that $\partial\psi(z,\beta)/\partial\beta > 0$ and that $\partial\psi(z,\beta)/\partial z \in (0,1)$. Hence, $G(z,\beta)$ is strictly increasing in z and strictly decreasing in β. Since $L_1 > 1/2$, it follows that $G(z,\beta) > 0$. It is also easy to check that $2\psi(z,\beta)/(1+z)$ is increasing in z and that $\lim_{z \to 1} 2\psi(z,\beta)/(1+z) = 1$, so that $G(z,\beta) < 1$.

benefit that must be balanced against the short run cost of migration. This short run cost increases with the relative productivity shock and is higher the greater is the congestion factor. If the congestion factor or the shocks are large enough, the policy of free movement of workers will not be sustainable.

Proposition 2 (Labor market flexibility) *Suppose there is no trade and workers are risk neutral. Then, the agglomeration of workers is efficient only for iso-elastic production functions. The policy of free movement of workers is always sustainable for constant returns to scale and never sustainable in a random fixed crop economy. The smaller are the shocks or the weaker is the congestion factors (larger β), the smaller is the discount factor for which the policy of free movement of workers can be sustained.*

4.3   Income smoothing

We now turn to the issue of insurance. The policy of free movement of workers also benefits risk averse workers because it allows them to smooth their consumption by moving across borders. In the previous section we have established that the policy of free movement of workers is sustainable when workers are infinitely risk averse. However, one can see that risk aversion has no impact on workers location and consumption when the production technology has constant returns to scale ($f(l) = l$). This is because immigration has no impact on local citizens' productivity, earnings or consumption. However, for full congestion ($f(l) = 1$), one can show that $G(\theta) < 1$. Hence, in a random crop economy, risk aversion entices workers to accept a policy of free movement of workers although they would not under risk neutrality. The insurance benefit is additional to the labor market flexibility benefit and offsets the impatience of workers.

Assuming that worker preferences exhibit constant relative risk aversion with coefficient ρ, it can be shown that

$$G(\theta) \equiv G(z, \beta, \rho) = \frac{z^{1-\rho} - \left(\frac{1}{2} + \frac{1}{2} z^{\frac{1}{1-\beta}}\right)^{(1-\beta)(1-\rho)}}{\left(\frac{1}{2} + \frac{1}{2} z^{\frac{1}{1-\beta}}\right)^{(1-\beta)(1-\rho)} - 1},$$

which can be shown to fall as ρ increases for all values of z and β. As the coefficient of risk aversion increases, the cost of the free movement policy decreases because workers have lower marginal utility from consumption in the good state of nature. Conversely, their benefit increases because they have higher marginal utility

of consumption in the bad state. Hence, increases in risk aversion reduce the critical discount factor above which the policy of free movement of workers can be sustained. The insurance motive reinforces the labor market flexibility motivation for the free movement of workers. We summarize this discussion in the following proposition:

Proposition 3 (Consumption smoothing) *Under constant relative risk aversion, the more risk averse are workers, the lower is the critical discount factor above which the policy of free movement of workers can be sustained.*

5 Migration and trade model

The adoption of a policy of free movement of workers has often been associated with agreements on free trade. The E.U. and N.A.F.T.A. both had high levels of internal trade before adopting policies on the free movement of workers. It is important therefore, to examine a model that includes trade as well as migration. As has been noted by a number of authors (see, for example, Davis and Weinstein, 2002; Felbermayr and Kohler, 2007), migration can have an impact on the terms of trade and this will feedback to the desirability and sustainability of a policy of free movement of workers. We therefore extend our model to incorporate trade in a simple and analytically tractable way.

We consider a Ricardo-Viner model that has two non-tradeable goods (Z_s, Z_s^*) and two tradeable goods (X_s, X_s^*). The home country specializes in the production of X_s and the foreign country in the production of X_s^*. Individuals consume a Cobb-Douglas composite good $C_s \equiv \kappa\ X_s^{\gamma/2}(X_s^*)^{\gamma/2}Z_s^{1-\gamma}$ where κ is a constant (with a similar expression holding for composite consumption in the foreign country). The parameter $\gamma \in [0,1]$ reflects the preferences for tradeable goods and, given the Cobb-Douglas structure, the expenditure share on domestic traded goods is $\gamma/2$. The standard migration model of the previous section corresponds to the case where $\gamma = 0$. When $\gamma = 1$, all goods are traded.

As before, each worker inelastically supplies one unit of homogeneous labor. In the home country, L_s^X individuals work in the tradeable good sector while L_s^Z are employed in the local non-tradeable good sector. Workers move freely between sectors and therefore, are paid the same wage w_s in each sector. Each

tradeable and non-tradeable sector includes a unit mass of firms that produce according to a production function $F_s(L_s^i) = \alpha_s \left(L_s^i\right)^\beta$, $i \in \{X, Z\}$ where $\alpha_s > 0$ is the country productivity and $\beta \in (0, 1]$ measures the decreasing returns in the production sector. The productivity parameter α_s is assumed to be identical across tradeable and non-tradeable sectors.[12] The foreign country has the same production structure (in particular, β is the same across countries), but has a productivity parameter $\alpha^* > 0$. As in Davis and Weinstein (2002), allowing free movement of workers expands the feasible world production set when technologies across countries are not identical. World production increases with migration.

5.1 Short run equilibrium

The short run equilibrium consists of a set of prices, wages, income and sectorial labor distribution that satisfy both profit maximization and market clearing conditions for labor and goods. The solution of the model is standard and details are given in Appendix B. Firms hire workers so that the marginal product of labor equals the wage: $P_s^i F_s'(L_s^i) = w_s$. Firms' sales and profits are proportional to the wage bill: $P_s^i F_s(L_s^i) = w_s L_s^i / \beta$. Because production functions are the same across sectors, labor allocates across the tradeable and non-tradeable sectors according to the product demands: $L_s^X = \gamma L_s$ and $L_s^Z = (1 - \gamma)L_s$. The terms of trade (w_s/w_s^*) adjust to equalize the values of exports and imports and consequently,

$$\frac{w_s}{w_s^*} = \frac{L_s^*}{L_s}. \tag{8}$$

Thus, the terms of trade moves in inverse proportion to the flows of labor. Equilibrium consumption of the composite good is $C_s = (P_s^X)^{-\gamma/2}(P_s^{X*})^{-\gamma/2}(P_s^Z)^{\gamma-1}Y_s/L_s$.[13] National income is $Y_s/L_s = \beta^{-1}w_s$ and the non-tradeable good price is $P_s^Z = (\alpha_s\beta)^{-1}[(1-\gamma)L]^{1-\beta}w_s$. The ratio of home and foreign consumption is

$$\frac{C_s}{C_s^*} = \frac{(P_s^Z)^{\gamma-1}w_s}{(P_s^{Z*})^{\gamma-1}w_s^*}. \tag{9}$$

---

[12]This eliminates any bias when we compare the economies with and without trade.

[13]The constant κ is normalized to cancel out other constants in this expression.

Substituting for prices, and using (8), gives home consumption as a function of population:

$$C_s(L_s) = \alpha_s \left(\frac{\alpha_s^*}{\alpha_s}\right)^{\gamma/2} \left(\frac{L_s^*}{L_s}\right)^{\beta\gamma/2} L_s^{\beta-1}. \tag{10}$$

This provides a specific form for the function described in Section 2. Note that, as mentioned in Section 2, unless $\beta\gamma = 0$, consumption is not an iso-elastic function of L_s and the free labor movement equilibrium is inefficient.

In the equilibrium with free movement of workers, $C_s(L_s) = C_s^*(L_s^*)$. For $\alpha_s > \alpha_s^*$, condition (3), under which there is excess agglomeration of workers, becomes

$$\frac{L_s^* C_s^{*\prime}(L_s^*)}{L_s C_s'(L_s)} = \frac{1 - \beta + \beta\gamma/2 + (\beta\gamma/2)(L_s^*/L_s)}{1 - \beta + \beta\gamma/2 + (\beta\gamma/2)(L_s/L_s^*)} \leq 1 \tag{11}$$

For $\beta\gamma > 0$, this is true if and only if $L_s \geq L_s^*$. Thus, the allocation of labor across countries is efficient only if $L_s = L_s^*$.[14]

To highlight the differences to the standard model of migration considered in the previous section, we first discuss an economy that includes only tradeable goods and then analyze an economy with both tradeable and non-tradeable goods.

5.2 Pure trade economy

Following Davis and Weinstein (2002), suppose that the economy includes only tradeable goods and production displays constant returns to scale ($\gamma = \beta = 1$). Even with constant returns to scale, migration affects consumption through a terms of trade effect. Changes in the terms of trade absorb productivity differentials between countries, reducing the incentives to migrate. We analyse each of these effects in turn.

First, immigration has a negative effect on the consumption of home workers through its impact on the terms of trade. To see this, consider a case where the home country initially has a smaller population $L_s < L_s^*$. By (8), the terms of trade satisfy $w_s/w_s^* > 1$ before any migration. Suppose that the home country has a good shock: $\alpha_s > \alpha_s^*$. Then, by (9), free movement of workers equalizes consumption, so

---

[14]For $\beta\gamma = 0$, (11) never holds and any allocation of labor across countries is efficient.

that $C_s/C_s^* = w_s/w_s^* = 1$ and the terms of trade fall. Given constant returns to scale, immigration does not affect the home country's purchasing power for the domestic good, $w_s/P_s^X = \alpha_s$, but it reduces that for the foreign good, $w_s/P_s^{X*} = (w_s/w_s^*)\alpha_s^*$. Thus, home workers lose from immigration through its adverse effect on the terms of trade. The same argument shows that foreign workers gain. This stands in contrast to the conclusion of the previous section where immigration had no impact on welfare when firms produce under constant returns to scale. In other words, immigration increases the production of local goods, which deflates the price of those goods and consequently the income of local workers.

Secondly, changes in the terms of trade absorb the effect of productivity differences on consumption and therefore mitigate migration incentives. From (10), it can be seen that home consumption is $L_s C_s(L_s) = \sqrt{\alpha_s^* \alpha_s}\sqrt{L_s^* L_s}$. Consumption depends on the geometric mean of the productivity parameters and population sizes. Consider again the symmetric case with two anti-symmetric shocks. Then, because $\alpha_s^* \alpha_s$ and $L_s L_s^*$ are constant, the consumption levels do not depend on the states of nature. The terms of trade *fully* absorb the productivity shocks. This is because a positive productivity shock stimulates both local production and income. This leads to an increase in exports and fall in the export price and at the same time the rise in income leads to higher imports and import prices, which rebalances foreign production and income. With Cobb-Douglas preferences, the terms of trade effect completely offset consumption differences. The short run equilibrium has no migration: $L_1 = L_2 = \bar{L}/2$. Applying this labor allocation in (11), one can see that the free movement equilibrium is trivially efficient. Each country achieves the same consumption level, whether the policy of free movement of workers is adopted or not. It can be also shown that the relative cost of adopting the policy of free movement of workers, $G(\theta) = 1$ as $\gamma = 1$. Therefore, free movement of workers is never a sustainable policy. The same analysis applies in the presence of congested factors ($\beta < 1$) as long as $\gamma = 1$. We summarize those results in the following proposition:

Proposition 4 (Pure trade economy) *Consider an economy with only tradeable goods ($\gamma = 1$). Then, immigration leads to changes in the terms of trade that adversely affect local workers. Changes in the terms of trade also attenuates the consumption discrepancies caused by productivity shocks. In the two anti-correlated state model, the changes in the terms of trade fully absorb productivity shocks and eliminate consumption fluctuations altogether: there is no migration in equilibrium and the free movement of workers is never a*

sustainable policy.

In the two anti-corretated state model, a trade policy is a substitute for labor movement policy *for the purpose of income smoothing.* Trade is nevertheless not a substitute for migration in the Hecksher-Ohlin sense where the movement of factor is equivalent to the movement of goods (Mundell 1957). In this model, each country specializes in the production of its tradeable goods so that a worker produces a different good (at a different productivity) when she moves across border. In Hecksher-Ohlin model the workers is able to produce the same goods when she moves.

To sum up, whereas Davis and Weinstein (2002) highlight the short run cost of migration through terms of trade between the U.S. and Mexico, this model gives a case for President Salinas' claim that emigration may not occur because of trade. Things, however, are different in the presence of non-tradeable goods and we now turn to that case.

5.3 Economy with tradeable and non-tradeable goods

We now study the more general situation where there are both tradeable and non-tradeable goods ($\gamma < 1$). We focus on the effect of terms of trade and labor market flexibility by assuming that workers are risk neutral ($U(C) = C$). In this case, the allocation of workers across countries is given by

$$\frac{L_s^*}{L_s} = \left(\frac{\alpha_s^*}{\alpha_s} \right)^{\frac{1-\gamma}{1-\beta(1-\gamma)}} \tag{12}$$

One can check that $d(L_s^*/L_s)/d(\alpha_s^*/\alpha_s) > 0$, while $d(L_s^*/L_s)/d\beta < 0$ and $d(L_s^*/L_s)/d\gamma > 0$ if $\alpha_s > \alpha_s^*$. As expected, for any $\gamma < 1$, workers move into the most productive country because the latter offers higher consumption levels. However, in equilibrium, the labor allocation in the most productive country decreases with the intensity of local congestion and the share of tradeable goods.

From (11) and (12), it can be seen that the labor allocation across countries is efficient only if $\beta\gamma = 0$, or $\gamma = 1$. Otherwise, there is excess agglomeration. This excess agglomeration occurs because migrants do not internalize the effect of their move on the productivity and consumption of local workers. There is a missing price: while wages and product prices give the signals for production and consumption, there is no

specific price for the migration decision.

Proposition 5 *The policy of free movement of workers yields excessive agglomeration of workers in the high productivity country in the presence of decreasing returns to scale and both tradeable and non-tradeable goods.*

Consider again the symmetric country two anti-correlated state model and let $\tilde{L}$ denote the efficient allocation when the home country has the high productivity shock. Let $e = L/\tilde{L}$ denote the ratio of the equilibrium to efficient allocation, so that $e = 1$ corresponds to an efficient allocation and $e > 1$ corresponds to excess agglomeration. It can be shown that the equilibrium labor level L_s increases faster than the efficient level as β rises. When local factor congestion is weaker, agglomeration in the higher productivity country is more pronounced both in the free labor movement equilibrium and in the efficient allocation. The externality in the location decisions however exacerbates the agglomeration process at the cost of reducing aggregate consumption. This is because, as β increases, equilibrium wages become less elastic to the relocation of workers and do not give appropriate location incentives to workers. Therefore, *the agglomeration of workers becomes increasingly excessive for weaker local factor congestion*. Figure 1 plots the contours of e in (γ, β)-space for shock values $\alpha_1 = 1/\alpha_2 = 1.2$. It shows that the excessive agglomeration of workers can be significant. For example, with a moderate expenditure on tradeable goods of $\gamma = 0.2$ and a weak congestion factor of $\beta = 0.8$, the domestic country host 69% of the total population compared to 61% in the socially optimal allocation. Stronger shock amplitudes yield more excessive agglomeration.

<center>INSERT FIGURE 1 HERE</center>

Figure 1 also shows that the impact of trade on excess of agglomeration is non-monotonic with respect to the size of the tradeable sector. Excessive agglomeration increases with γ for small γ while it decreases with γ for large γ. Therefore, *the agglomeration of workers is most excessive for intermediate shares of trade*. At the two extremes $\gamma = 0$ and $\gamma = 1$, the welfare optimum and the equilibrium allocation coincide.

Agglomeration of workers is most excessive when production has weak congestion and each country trades a small share of its production. In particular, the more productive country attracts too many

<center>24</center>

migrants when there exists no congestion or constant returns to scale. This runs counter to the standard argument that migration is innocuous under constant returns to scale because workers move with both their constant productivity and consumption to the hosting country. It is explained by the previous arguments about the effects of the terms of trade. When some goods are not traded, the terms of trade cannot fully absorb the productivity shocks so that migration incentives are present. Lower congestion exacerbates these incentives. At the same time, immigration change the terms of trade, which harms the local workers. Again, as explained before, immigrants increase their productivity and produce more of the good of the destination country, increase congestion and depress its price and local wages. They also demand more of the good produced in the low productivity country and increase its price. Local workers in the more productive country therefore see their wage fall and the price of imports rise. A planner would prefer to reduce labor movements to partly restore the wages and consumption levels of those in the more productive country.

We now return to the issue of the sustainability of the policy of free movement of workers. Figure 2 plots the locus of the relative cost of adopting the policy of free movement of workers, $G(\theta)$ in (γ, β)-space for $G(\theta)$ equal to 0.25, 0.50, 0.75 and 1. These values respectively correspond to critical discount factors $\delta = 0.40, 0.66, 0.85$ and $1.^{15}$ The shock structure is the same as in Figure 1. The area (a) corresponds to $G(\theta) > 1$ and the areas (b) and (c) to $G(\theta) < 1$. The relative cost of adopting free movement of workers, $G(\theta)$, falls as we move to the North-West of the figure. As a result, free movement of workers is more likely to become a sustainable policy in economies with lower trade and weaker congestion.

INSERT FIGURE 2 HERE

Figure 2 also shows that $G(\theta)$ increases as more goods are traded (larger γ). Because trade is a substitute for labor movement for the purpose of income smoothing, free movement of workers is less useful when trade is important. On the other hand, $G(\theta)$ is not monotone with respect to the intensity of congestion parameter β. Indeed, as we move downward in Figure 2 (β falls), $G(\theta)$ firstly decreases when the parameters (β, γ) lie in the area (b) but it increases when those parameters lie in the area (c). In the figure, areas (b) and (c) are

---

[15]For example, if one considers a time period of ten years between the shocks, these values correspond to annual opportunity costs of time of respectively 9, 4, 2 and 0 percents ($r = \delta^{-1/10} - 1$).

separated by a dashed curve that corresponds to the locus where the partial derivative $\partial G(\theta)/\partial\beta = 0$. This locus shows, for a given γ, the value of β for which free movement of workers can be supported for the lowest discount factor. Whereas lower congestion implies that domestic workers' productivity and wages are less affected by the inflow of international workers, it also implies that the incentive for migration is not offset enough by the downward pressure on wages in the high productivity country. Excessive agglomeration of workers occurs and can be so inefficient that the domestic country does not find it desirable to opt for free movement of workers. In this case, the short run cost of accepting an excessive inflow of foreign workers in good states of nature does not outweigh the benefit of the migration option in bad states of nature. We summarize our result in the following proposition.

Proposition 6 *Suppose that individuals are risk neutral and that countries face two-state anti-correlated shocks. Then, free movement of workers is not a sustainable policy when trade is important and congestion is weak. Otherwise, there exists a discount factor δ such that free movement of workers is sustainable. The critical discount factor is lower when fewer goods are traded and for intermediate values of local factor congestion.*

We finally extend the above model to investigate the effects of unemployment on the sustainability of the policy of free movement of workers.

5.4 Wage rigidities

The reluctance to opt for free movement of workers is often based on a claim about local labor market frictions. Many countries have been reluctant to allow uncontrolled inflows of workers in times of high unemployment. Boeri and Brcker (2005) present evidence of the hardening of migration conditions within the E.U., particularly for richer countries with large unemployment levels, such as France and Belgium. We demonstrate that the existence of unemployment stemming from labor market rigidities is not a rationale against the adoption of free movement of workers in our model.

We suppose that unemployment stems from downward nominal wage rigidities. For the sake of simplicity, suppose that the wage w_s cannot fall below some exogenous minimum wage $\underline{w}$. Let l_s denote the number of

26

employed workers (or hours worked) and L_s denote the population including migrants. Then, $L_s - l_s$ can be interpreted as either unemployment or under-employment. Assume that the foreign country is subject to the same wage rigidity. We make the simplifying assumption that governments implement lump sum redistribution to the unemployed so that employed and unemployed workers residing in a same country get the same utility. The analysis of the short run equilibrium is the same as in Section 5 except that L_s must be replaced by l_s. The trade balance condition (8) now gives the employment ratio: $l_s/l_s^* = w_s^*/w_s$. Consumption is given by:

$$C_s = \alpha_s^{1-\gamma/2} \, (\alpha_s^*)^{\gamma/2} \, (l_s)^{\beta(1-\gamma/2)} \, (l_s^*)^{\beta\gamma/2} \, L_s^{-1} = \alpha_s^{1-\gamma/2} \, (\alpha_s^*)^{\gamma/2} \, (w_s^*/w_s)^{\beta\gamma/2} \, l_s^\beta L_s^{-1}$$

with a similar expression holding for the foreign country.

Suppose that the domestic country faces a good productivity shock: $\alpha_s > \alpha_s^*$. Then, if labor is immobile and if the minimum wage $\underline{w}$ is high enough, downward wage rigidities imply that the foreign country faces unemployment ($w_s^* = \underline{w}$ and $l_s^* < L_s^*$) whereas the domestic country has full employment ($w_s \geq \underline{w}$ and $l_s = L_s$). The consumption levels are given by:

$$\widetilde{C}_s (L_s) = \alpha_s^{1-\gamma/2} \, (\alpha_s^*)^{\gamma/2} \, (\underline{w}/w_s)^{\beta\gamma/2} \, L_s^{\beta-1}$$

$$\widetilde{C}_s^* (L_s^*) = (\alpha_s^*)^{1-\gamma/2} \, (\alpha_s)^{\gamma/2} \, (w_s/\underline{w})^{\beta\gamma/2} \, (l_s^*/L_s^*)^\beta \, (L_s^*)^{\beta-1}$$

By contrast, when labor is allowed to move across countries, foreign workers move to the home country as long as $\widetilde{C}_s/\widetilde{C}_s^* > 1$: that is, if $(\alpha_s/\alpha_s^*)^{1-\gamma}(\underline{w}/w_s)^{\beta\gamma}(l_s^*/L_s^*)^{-\beta}(L_s/L_s^*)^{\beta-1} > 1$. So, L_s increases whereas L_s^* decreases to l_s^*. At this point, the foreign country reaches full employment and the allocation of labor across countries achieves the short run equilibrium described in (12). Therefore, free movement of workers eliminates unemployment. Free movement of workers implies a better use of productive resources in terms of both labor time and allocation of labor across countries. Workers get the same consumption, $C_s(L_s)$, and the same utility level, $u_s(L_s) = U[C_s(L_s)]$, as derived in the absence of rigidities. Note that labor market rigidities imply that $\widetilde{C}_s(L_s) < C_s(L_s)$ for all L_s, $s \in \mathcal{S}$. This means that $\widetilde{u}_s(L_s) \equiv U\left[\widetilde{C}_s(L_s)\right] < u_s(L_s)$ for

all states $s \in \mathcal{S}$.

Free movement of workers is a sustainable policy for the domestic country if and only if the participation constraint

$$u_s(L_s) - \widetilde{u}_s(L^0) + \frac{\delta}{1-\delta} \left[E_s u_s(L_s) - E_s \widetilde{u}_s(L^0) \right] \geq 0$$

is satisfied. Since $\widetilde{u}_s(L_s) < u_s(L_s)$, this condition is less stringent than condition (4). This allows us to conclude the following:

Proposition 7 *The free labor mobility is a sustainable policy for lower discount factors when countries face downward wages rigidities and unemployment.*

This analysis suggests that the gains from labor market flexibility obtained from a free labor mobility policy are strengthened by labor market rigidities. Presumedly, so are the gains from insurance or income smoothing. The resistance to the policy is therefore probably grounded on (the perception of) permanent productivity differences or on other issues that are outside of the scope of this paper (e.g. skill differentials, rent-seeking, mis-perception of immigration effects, fiscal issues, etc.).

6   Conclusion

In this paper we have studied the factors that help countries mutually agree on common policies of unconditional and uncontrolled movement of workers. For the countries to agree on such a policy, short run costs must be outweighed by long term benefits. While the costs lie in the congestion and adverse effects of term of trade, the benefits considered here stem from labor market flexibility and consumption smoothing. Countries facing good productivity shocks incur short run costs because they allow foreign workers to participate in their local labor markets, which reduces local wages and/or purchasing power. By contrast, countries facing bad productivity shocks benefit from free movement of workers because they are able to invite their citizens to work temporarily or permanently in more prosperous countries. We show that such policies are less likely to be adopted when workers are impatient and less risk averse, when production technologies

28

display decreasing returns and when countries trade a greater share of their products. Large permanent differences in consumption levels prohibits the adoption of the policy. The presence of trade increases the cost of immigration cost but at the same time it mitigates the migration incentives. In the presence of both tradeable and non-tradeable goods, the policy of free movement of workers is more likely to be sustainable as fewer goods are traded and for intermediate values of local factor congestion. Finally, wage rigidities improve the sustainability of the policy.

Our analysis is designed to highlight the role of labor market flexibility and insurance in the adoption of uncontrolled and unconditional migration. It shows the importance of time discounting, risk aversion, factor congestion and trade. We consider these to be important ingredients in the decision of countries to adopt a policy of free movement of labor and believe that the interplay of trade and congestion offers new insights into the these decisions. Admittedly, many other important aspects of the problem are left out of the account: the political decision process, the skill heterogeneity across workers, the public finance issues, the possibility to control movements of workers, etc. These issues are left for further research.

References

Acemoglu, D., Golosov, M., Tsyvinski, A., 2010. Dynamic Mirrlees taxation under political economy constraints. Review of Economic Studies 77, 841–881.

Bagwell, K., Staiger, R.W., 1990. A theory of managed trade. American Economic Review 80, 779–95.

Bagwell, K., Staiger, R.W., 2005. Enforcement, private political pressure, and the General Agreement on Tariffs and Trade/World Trade Organization escape clause. The Journal of Legal Studies 34, 471–513.

Boeri, T., Brcker, H., 2005. Why are Europeans so tough on migrants? Economic Policy 20, 629–703.

Borjas, G.J., 2003. The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. The Quarterly Journal of Economics 118, 1335–1374.

Chari, V.V., Kehoe, P.J., 1990. Sustainable plans. Journal of Political Economy 98, 783–802.

Clemens, M.A., Montenegro, C.E., Pritchett, L., 2010. The Place Premium: Wage Differences for Identical Workers across the US Border. Working Paper wp321. University of Chile, Department of Economics.

Davis, D.R., Weinstein, D.E., 2002. Technological Superiority and the Losses from Migration. NBER Working Papers 8971. National Bureau of Economic Research, Inc.

Facchini, G., Testa, C., 2009. Who is against a common market. Journal of the European Economic Association 09, 1068–1100.

Felbermayr, G.J., Kohler, W., 2007. Immigration and native welfare. International Economic Review 48, 731–760.

Grossman, G.M., Helpman, E., 1995. The politics of free-trade agreements. American Economic Review 85, 667–90.

Klein, P., Ventura, G., 2009. Productivity differences and the dynamic effects of labor movements. Journal of Monetary Economics 56, 1059–1073.

Ligon, E., Thomas, J.P., Worrall, T., 2002. Informal insurance arrangements with limited commitment: Theory and evidence from village economies. Review of Economic Studies 69, 209–244.

Markusen, J.R., 1983. Factor movements and commodity trade as complements. Journal of International Economics 14, 341–356.

Mundell, R.A., 1957. International trade and factor mobility. American Economic Review 47, 321–335.

Neary, P., 1995. Factor mobility and international trade. Canadian Journal of Economics 28, 4–23.

Ottaviano, G., Peri, G., 2012. Rethinking the gains from immigration: Theory and evidence from the us. Journal of the European Economic Association 10, 1068–1100.

Staiger, R.W., Bagwell, K., 1999. An economic theory of GATT. American Economic Review 89, 215–248.

Thomas, J.P., Worrall, T., 1988. Self-enforcing wage contracts. Review of Economic Studies 55, 541–554.

Wahba, J., Zenou, Y., 2009. Out of sight, out of mind: Migration, entrepreneurship and social capital. Regional Science and Urban Economics 42, 890–903.

Waterfield, B., 2011. France threatens to 'suspend' Schengen treaty. The Telegraph, 22nd April.

Appendix A

We consider a rudimentary model of moving costs. The aim is to show that the assumption of zero moving costs in the main body of the paper can be relaxed provided moving costs are not too high on average. To do this requires two elements. First we assume that a moving cost of μ is equivalent to a reduction in consumption. Suppose that consumption in the home country is higher than consumption in the foreign country when there are $L_s \geq L^0$, $C(L_s) > C^*(L_s)$. A foreign worker will move if $U\left[C(L_s) - \mu\right] \geq U\left[C(L_s)\right]$ or $\mu \leq C(L) - C^*(L)$. Second assume that moving costs are heterogeneous across workers and that the distribution of moving costs across workers is a negative exponential distribution with an average moving cost of $\bar{\mu}$. That is the distribution of moving costs is given by the cumulative distribution function $\Phi(\mu) = 1 - \exp(-\mu/\bar{\mu})$. This distribution is convenient as there are always some workers with low moving costs and because the distribution is described simply by its mean. Other assumptions on the distribution of moving costs could be made but the exponential is particularly analytically convenient. Using the exponential distribution, the number of foreign workers with moving costs less than or equal to $C_s(L_s) - C_s^*(L_s)$ is

$$L^0 \Phi\left[C_s(L_s) - C_s^*(L_s)\right] = L^0 \left[1 - \exp(-(C(L_s) - C^*(L_s))/\bar{\mu})\right]$$

For there to be an equilibrium in which $C_s(L_s)$ and $C_s^*(L_s)$ are equated, this number of workers should be no less that $L_s - l$ where L_s is the equilibrium labor force and $l \in [L^0, L_s)$. That is

$$L^0 \left(1 - \exp(-(C_s(l) - C_s^*(l))/\bar{\mu})\right) \geq L_s - l$$

31

for $l \in [L^0, L_s)$. If this inequality is satisfied then for any $l \in [\bar{L}/2, L_s)$ there are always enough workers prepared to move such that consumption is equated in the two countries. Rewriting this inequality gives

$$\bar{\mu} \le \frac{C_s(L_s) - C_s^*(L_s)}{-\log\left(1 - \left(\frac{L_s - l}{L^0}\right)\right)}$$

for $l \in [L^0, L_s)$. Note that the term on the right-hand-side is positive and bounded above. To see it is positive, note that $C_s(L_s) > C_s^*(L_s)$ by assumption and the term inside the logarithm is less than one. It is bounded above for $l \in [L^0, L_s)$. In the limit as $l \to L_s$ both the numerator and denominator vanish, but using L'Hospital's rule it can be shown that this limit is positive and finite. Thus provided the average moving cost is not too high, and assuming a similar condition holds for Home workers considering migration to the Foreign country, the equilibrium allocation with moving costs will be the same as when moving costs are zero. The welfare properties will, however, be different.

Appendix B

This appendix derives the short run market equilibrium of Section 5. We proceed in four steps. First, because profits are redistributed locally we have that national income Y_s is equal to the value of domestic production $P_s^X X_s + P_s^Z Z_s$ where P_s^i is the price of good in sector i. Second we calculate labor demand from the condition that the value of the marginal product equals the wage rate, $P_s^i F_s'(L_s^i) = w_s$, or equivalently, $P_s^i \alpha_i \beta \left(L_s^i\right)^{\beta - 1} = w_s$. This implies that the value of production in each sector is proportional to the wage bill: $P_s^i F_s(L_s^i) = \beta^{-1} w_s L_s^i$. The national income in wage units is then equal to $Y_s = \beta^{-1} w_s L_s$. Third, given the Cobb-Douglas preference individuals spend a share $\gamma/2$ of their income on each of the tradeable goods and a share $1 - \gamma$ on the local non-tradeable good. So, the goods market clearing condition in the non-tradeable sector gives $\beta^{-1} w_s L_s^Z = (1 - \gamma)Y_s$ and hence $L_s^Z = (1 - \gamma)L_s$ since $Y_s = \beta^{-1} w_s L_s$. Then using the labor market clearing condition in the domestic market we have that $L_s^X = \gamma L_s$. We can further use these conditions to compute the price of tradeable and non-tradeable goods in wage units as $P_s^X = (\alpha_s \beta)^{-1}(\gamma L_s)^{1-\beta} w_s$ and $P_s^Z = (\alpha_s \beta)^{-1}[(1-\gamma)L_s]^{1-\beta} w_s$. Finally, we consider the market clearing conditions for the tradeable good sectors in the domestic and foreign countries. With the Cobb-Douglas preference

the value of production is equal to the consumers' expenditure shares: $P_s^X F_s(L_s^X) = (\gamma/2)(Y_s + Y_s^*)$ and

$P_s^{X*} F_s^*(L_X^*) = (\gamma/2)(Y_s^* + Y_s)$. Therefore, the value of production of the tradeable good is the same in

both countries: $P_s F_s(L_s^X) = P_s^{X*} F_s^*(L_s^{X*})$. Because the value of production in each sector is proportional

to the wage bill (with proportion β) the wage bills in each country in the tradeable sectors must be equal:

$w_s L_s^i = w_s^* L_s^{i*}$. This then further applies to the non-tradeable sector and hence the equilibrium ratio of
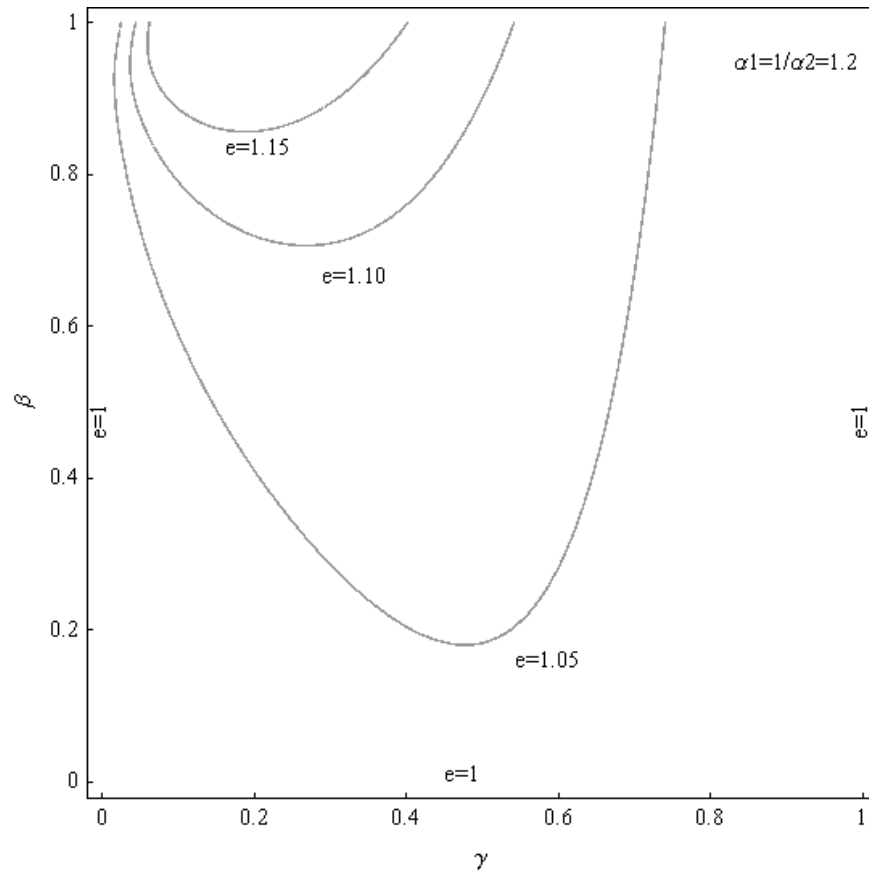
wages is $w_s/w_s^* = L_s^*/L_s$.

Figure 1: Excessive agglomeration of workers

Comment: The ratio between the equilibrium and first best numbers of workers, e, is larger for intermediate trade (γ) and low congestion (high β).
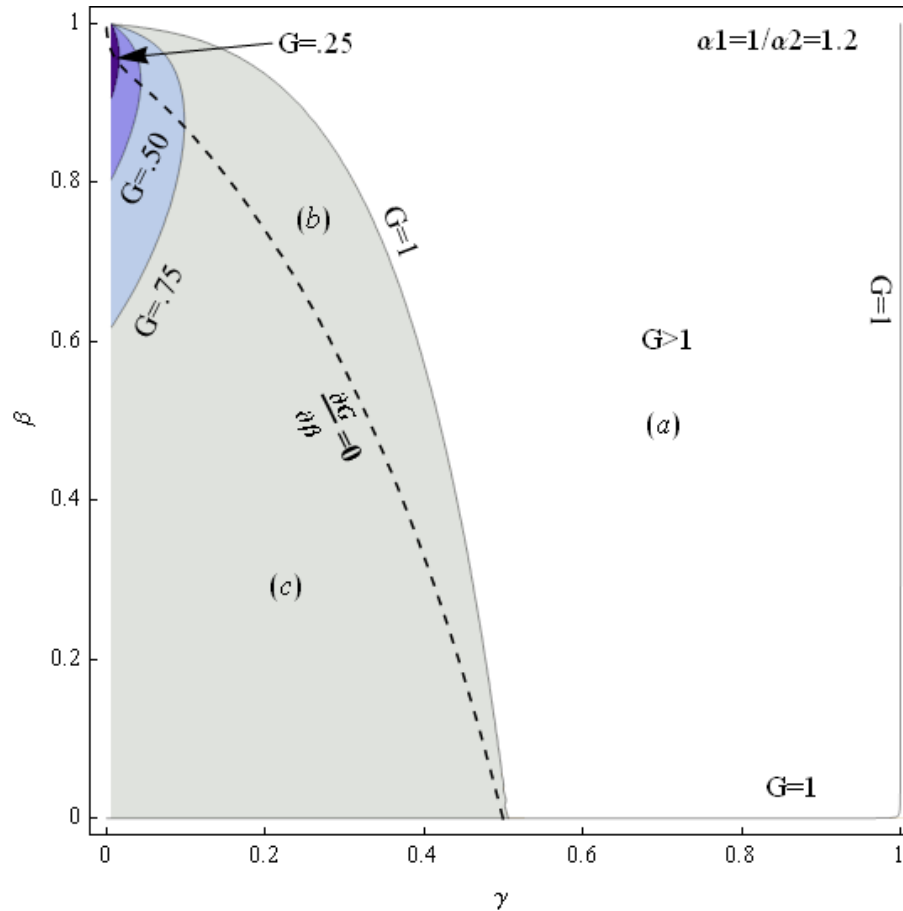
Figure 2: Relative cost of adopting the policy of free movement of workers, G(θ).

Comment: The relative cost of adopting the policy of free movement of workers, G, is related to the cost of sustaining the policy of free movement of workers. Large trade sectors (γ) makes the policy unsustainable. Sustainability is not monotone with respect to factor congestion (β).