Is a policy of free movement of workers sustainable?*

Pierre M. Picard
Département Economie, Université du Luxembourg, 162A avenue de la Faiencerie, L-1511, Luxembourg, & CORE, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
pierre.pircard@uni.lu

Tim Worrall
School of Economics, University of Edinburgh, 30 Buccleuch Place, Edinburgh, EH8 9JT, UK.
tim.worrall@ed.ac.uk

November 2013 (Revised March 2015)

Abstract

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

Keywords: Labor market flexibility, migration, sustainable plan

JEL: F22, J61.

*We thank the editor and two referees for many useful comments and suggestions. We are also thankful to participants at the 2011 NORFACE Conference on migration and seminars at Erasmus University, Oxford University, and University Carlos III, and University of Lille for comments. We are grateful to M. Beine, C. Détang-Dessendre, G. Fachini, A. Margherita, P. Neary, T. Mueller, E. Toulemonde, and A. Venables for helpful comments and suggestions. The first author gratefully acknowledges the support of the grant F2R-CRE-PUL-10EGQH, University of Luxembourg. The second author gratefully acknowledges the support of the Hallsworth Research Fellowship Fund at the University of Manchester. The usual caveats apply.

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 to implement this policy, implementing it in stages, 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 the United States of America and Canada. The main reason for this reluctance lies in the fear that inflows of migrant workers may depress local labor market conditions and lower the welfare of the host country's 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 (see, for example, Klein and Ventura, 2009; Clemens et al., 2010) but imposes short run costs to local workers, particularly to the low skilled ones (see, for example, 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 a simple stochastic structure. 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. Although our main aim is to consider the implications of migration and trade together, we proceed in stages to develop the model. Initially, we consider a model where immigration imposes a cost on locals without specifying how

[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).

these costs arise (Sections 2 and 3). Taking the source of this cost as given, we examine how it might be offset by a long run benefit stemming from a policy of free movement, in which labor markets are integrated and workers (and their descendants) are able to relocate to more productive countries, allowing them to smooth their consumption. In considering these costs and benefits, it is important to take into account a possible externality created by free movement of labor. Free movement will lead to equalization of utilities in equilibrium, but there is no specific price for the migration decision. Migrants do not internalize the effect of their move on the productivity and consumption of local workers. *If* free movement creates a greater cost in the receiving country than gain in the sending country, then reducing the extent of migration could increase aggregate welfare: free movement leads to excessive agglomeration of labor in the receiving country. On the other hand, *if* free movement creates a lower cost in the receiving country than gain in the sending country, then increasing the extent of migration could increase aggregate welfare: free movement leads to under-agglomeration of labor in the receiving country. Under- or excessive agglomeration decrease the long-run benefits of a policy of free movement of labor. But they have opposite effects in the short run: excessive agglomeration increases the short run costs while under-agglomeration ameliorates the short run costs.

It is therefore important to know more about the the relative costs and benefits of a policy of free movement of labor. We consider these costs and benefits by more precisely specifying the model in Sections 4 and 5. In Section 4, we examine a standard migration model without trade but with productivity shocks. If production is iso-elastic, then there is no excessive agglomeration. However, weaker congestion enhances countries' incentives to adopt a policy of free movement of workers because it mitigates the short-run costs of immigration. Similarly, stronger risk aversion enhances countries' incentives to adopt such a policy because it raises the benefits from income smoothing that the policy brings. We also show how labor market frictions, in terms of minimum wages, impact on the sustainability of policies of free movement of labor.

Finally, in Section 5, we discuss the costs and benefits of adopting a policy of free movement of workers in the presence of international trade. In the Hecksher-Ohlin benchmark, factor prices are equalized across countries and workers have no incentives to migrate. To discuss migration issues, we therefore depart from this model and present a simple Ricardo-Viner trade model with production of both tradeable and non-tradeable goods. Because there is a non-tradeable good and differences in technologies across countries, factor price equalization does not hold and there will still be an incentive for migration. As has been recognized by Davis and Weinstein (2002) (see also, Felbermayr and Kohler, 2007), in the presence of trade, migration induces a terms of trade effect. With both tradeable and non-tradeable goods, the change in the terms of trade caused by migration has an adverse effect on the relative import prices and the consumption basket of domestic workers. However, this terms of trade effect also attenuates the impact of productivity shocks on consumption and reduces workers' incentives to move. We show that the presence of trade leads to excessive agglomeration even with iso-elastic production functions and Cobb-Douglas preferences because migrants do not take into account their effects on terms of trade. This excess agglomeration is higher for weaker congestion factors and for lower trade shares. An implication is that there is a trade-off between the trade share and congestion: weak congestion is good for sustainability of free movement in the standard migration model, but the presence of trade causes excessive agglomeration, which is bad for sustaining free movement. We present a parameterized example to illustrate this trade-off. An interesting special case is where there are no non-tradeable goods, as in Davis and Weinstein (2002). In this case, we show that changes in the terms of trade fully absorb the shocks and eliminate any migration incentives when countries and shocks are symmetric. As in Section 4, we also consider the implications of labor market frictions on the sustainability of policies of free movement of labor.

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 modeling 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. Second, 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. (see, for example, Borjas 2003, Facchini and Testa 2009, Wahba and Zenou 2012, Ottaviano and Peri 2012). It nevertheless puts at the forefront uncertainty and intertemporal welfare that are important factors in assessing migration policy. In particular, it complements the migration literature by studying free migration policies as a reciprocal mechanism based on the potential benefit of labor market flexibility and income smoothing. It goes some way to explaining why some of the large gains from migration (identified by Klein and Ventura 2009 and Clemens et al. 2010) may be difficult to realize. Third, the paper contributes to our understanding of the interaction of trade and migration and extends in an interesting way the results of Davis and Weinstein (2002) and Felbermayr and Kohler (2007).

Our analysis combines a standard dynamic framework with limited commitment and standard trade models to tackle the issue of the adoption of a free movement policy. Yet, this combination is novel and does improve our understanding of the linkages between migration, trade and sustainability of a policy of free movement of workers. The model we adopt is stylized but tractable. Although the model could be extended and improved, we do think it highlights some reasons why it sometimes proves difficult to sustain a policy of free movement of workers.

2 The general model

We consider a model with continuum of workers, of mass $\bar{L}$, 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 suppose there is a set of states 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 in 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 short-run equilibrium migration decisions for a given state s. Assuming an interior solution when not all labor flows to one country, free movement of workers implies that utility is equal in each country. Since all workers are alike and have the same preferences, consumption is equalized too.[3]Hence, given that the consumption functions are decreasing, there will be a unique equilibrium

[3]We assume away moving costs in this text. However, our analysis holds provided that a share of workers face no moving costs and the average moving cost is not to high. Then, the equalization condition will apply to a marginal migrant who incurs no mobility cost.

labor allocation in each state, L_s, that satisfies:

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

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 favored 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 labor of the home country is given by

$$\frac{\mathrm{d}W_s}{\mathrm{d}L_s} = U\left(C_s\left(L_s\right)\right) - U\left(C_s^*\left(L_s^*\right)\right) + L_s U'\left(C_s\left(L_s\right)\right) C_s'\left(L_s\right) - L_s^* U'\left(C_s^*\left(L_s^*\right)\right) C_s^{*\prime}\left(L_s^*\right). \tag{1}$$

Assuming an interior solution, the free movement of workers is efficient if this expression is zero at the equilibrium. At the equilibrium allocation,there is excess agglomeration in the home country if this expression is negative and under agglomeration if the expression is positive. Since $C_s(L_s) = C_s^*(L_s^*)$ in equilibrium, the two first terms in (1) cancel out. Hence, there is excess agglomeration in the home country if and only if

$$\frac{L_s^* C_s^{*\prime}\left(L_s^*\right)}{L_s C_s'\left(L_s\right)} < 1. \tag{2}$$

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 $\lambda_s > 0$ or the iso-elastic case where $C_s(l) = \lambda_s l^\alpha$ for $\alpha \in (0, 1)$ and $\lambda_s > 0$. 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 $\delta \in (0,1)$.[4] 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 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 follows 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.[5] 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 of nature, that a utilitarian planner will choose the labor allocation in each state where the expression in (1) is equal to zero.[6]

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 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

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

[5] 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.

[6] 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.

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 (1) 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 (1) 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 current consumption (given our assumption that $C_s(l)$ is strictly decreasing). In contrast, provided there are states where workers can 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 and control the local labor supply.[7]

[7] As a very practical case, suppose that a E.U. country decides to leave the E.U. and breach Article 45 on the free movement of workers. Then, its natives working in the E.U. would be subject to the E.U. third-country association standards and would face the "E.U. preference of labor market access" that allows positive discrimination for E.U. workers. After some time, those natives and their descendants (if not naturalized) will be expelled.

This setting is clearly special, but we claim that it is not wholly 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 targeted on designated professional occupations, limited to three years but renewable 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 currently 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. H-1B 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 second 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 the dominant strategy of each country is to allow no entry to, and not renew any work permit of, 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]

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

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: $V_s^0 = u_s(L^0) + \delta \mathrm{E}_q V_q^0$. Hence,

$$V_s^0 = u_s(L^0) + \frac{\delta}{1-\delta}\mathrm{E}_s u_s(L^0),$$

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

$$u_s(L_s) - u_s(L^0) + \frac{\delta}{1-\delta}\mathrm{E}_q[u_q(L_q) - u_q(L^0)] \geq 0 \qquad \forall s \in \mathcal{S}, \tag{3}$$

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 (3) is most stringent for the state(s) with the highest short run cost, $\overline{s} \in \arg\max_s\{u_s(L^0) - u_s(L_s)\}$. Similarly, the equivalent of Condition (3) for the foreign country is more stringent in the state(s) $\overline{s}^* \in \arg\max_s\{u_s^*(\bar{L} - L^0) - u_s^*(\bar{L} - L_s)\}$. That is, if Condition (3) is satisfied in state $\overline{s}$, then it is satisfied for all states, with a similar statement applying to the foreign country at state $\overline{s}^*$. Thus, the policy of free movement of workers is sustainable if and only if

$$\begin{aligned} &u_{\overline{s}}(L_{\overline{s}}) - u_{\overline{s}}(L^0) + \tfrac{\delta}{1-\delta}\mathrm{E}_q[u_q(L_q) - u_q(L^0)] \geq 0, \\ &u^*_{\overline{s}^*}(\bar{L} - L_{\overline{s}^*}) - u_{\overline{s}^*}(\bar{L} - L^{0*}) + \tfrac{\delta}{1-\delta}\mathrm{E}_q[u_q^*(\bar{L} - L_q) - u_q^*(\bar{L} - L^0)] \geq 0. \end{aligned} \tag{4}$$

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 (4) is satisfied for $\delta \geq \delta^c$.[9]

We can also ask how sustainability relates to the agglomeration issue outlined in the previous

[9]It should be noted that when Condition (4) is satisfied, there is no breakdown in equilibrium. Since these constraints are forward-looking, the probabilities of states $\overline{s}$ and $\overline{s}^*$ matter only in so far as they affect the long term benefits/costs.

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 equation (1). 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 $\overline{s}$, then the short run costs of immigration, $u_{\overline{s}}(L_{\overline{s}}) - u_{\overline{s}}(L^0)$ 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 a policy of 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, higher risk aversion is associated with greater 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 and the foreign country's consumption functions. 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 $\lambda \in (0, 1)$.[10] 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 (3) becomes 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 $\mathrm{E}_q[u_q(L_q) - u_q(L^0)]$ becomes non-positive, so that free movement of workers is never a sustainable policy. We summarize this discussion in the following proposition.

Proposition 1 (Freedom of movement of workers) *The policy of free movement of workers is never sustainable if it brings no long-run benefit (*$\mathrm{E}_q[u_q(L_q) - u_q(L^0)] < 0$*). Otherwise, it is sustainable*

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

for high enough discount factor δ ($\delta > \delta^c$). The policy is more likely to be sustainable if workers are more risk averse. 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 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, improve analytical tractability and 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 the state $s = 1$ 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 $u_s(L_s) = u_s^*(L_s^*)$. Consequently, 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 (4) collapse to the single condition

$$\frac{\delta}{2-\delta} \geq G(\theta) \equiv \frac{u_1(\bar{L}/2) - u_1(L_1)}{u_2(L_2) - u_2(\bar{L}/2)}, \tag{5}$$

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 $G(\theta) \geq 1$. Using (5), it follows that $G(\theta) < 1$ if and only if $u_1(L_1) > (1/2)(u_1(\bar{L}/2) + u_2(\bar{L}/2))$. We summarize this discussion in the following lemma.

Lemma 1 (Two state model) *Consider a symmetric model with two equiprobable, anti-correlated states. Then, the policy of free movement of workers is never sustainable if* $u_1(L_1) < \frac{1}{2}[u_1(\bar{L}/2)+u_2(\bar{L}/2)]$. *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 labor used to produce a non-tradeable good (in Section 5 we will add tradeable goods). 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 $\alpha^* > 0$. For our discussion we will focus on three cases: constant returns to scale where $f(l) = l$, decreasing returns to scale where $f'' < 0$ and the fixed output case where $f(l) = 1$. In the constant returns to scale case, each worker's marginal productivity remains constant irrespective of home country production and labor force. By contrast, when there are decreasing returns, aggregate production exhibits congestion as the marginal productivity falls with inward migration. Finally, when $f(l) = 1$, output is equal to α_s, independently of the size of the labor force. This latter case can be interpreted as a purely agrarian economy with a random crop of fixed size α_s or, as we show below in section 4.4, where both economies have severe labor market frictions. In what follows it will also be instructive to discuss the specification where the production function is iso-elastic: $f(l) = l^\beta$, $\beta \in (0, 1]$. In this specification, the above three cases correspond to $\beta = 1$, $\beta < 1$ and the limit $\beta = 0$ respectively. Finally, for simplicity we assume 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 specification, β is inversely related to a measure of congestion. If $\beta = 1$, there is no congestion; increasing labor supply does not impact on the marginal product of existing workers. If $\beta = 0$, 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 may 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

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: $C_s(L_s) = \alpha_s g(L_s)$ where $g(l) \equiv f(l)/l$ is average productivity (a strictly decreasing function when there are decreasing returns to scale). With constant returns to scale all workers flow to the country with the higher productivity shock in each period and state. With decreasing returns under free movement of workers, and assuming an interior solution, workers migrate until there is equality of purchasing power:

$$\frac{C_s(L_s)}{C_s^*(L_s^*)} = \frac{\alpha_s g(L_s)}{\alpha_s^* g(L_s^*)} = 1. \tag{6}$$

Hence, $L_s \geq L_s^*$ if and only if $\alpha_s \geq \alpha_s^*$. Workers move into the more productive country because it offers higher purchasing power. The equilibrium population increases with the productivity ratio α_s/α_s^*.[11] Therefore, using equation (6),

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

There is excess agglomeration if and only if this term is less than one. As a special case presented in Section 2, the allocation of labor is efficient 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 ($U(C) = C^{1-\rho}/(1-\rho)$, $\rho \geq 0$, $\rho \neq 1$).[12]

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 $G(\theta)$, defined in equation (5), is equal to zero. 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

[11] That is, $\mathrm{d}L_s/\mathrm{d}(\alpha_s/\alpha_s^*) = -g(L_s)/(g'(L_s) + g'(L_s^*)) > 0$.

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

of migration policies on local citizens. It will be qualified in the next section.

As argued above, full congestion ($f(l) = 1$ and $g(l) = 1/l$) can be interpreted as an agrarian economy with random crops. The equilibrium 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 $G(\theta) = 1$ so that a policy of free movement of workers is never sustainable (for $\delta < 1$). Workers have a short run cost in the good state of nature that is exactly equal to 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 $\lim_{z\to 1} G(\theta) = 1$ for any production functions f. 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 specification where the production function is iso-elastic ($f(l) = l^\beta$) and $\beta \in (0,1)$. The equilibrium wage is equal to $w_s = \alpha_s \beta L_s^{\beta-1}$ and the consumption is given by $C_s(L_s) = \alpha_s L_s^{\beta-1} = w_s/\beta$. Without free movement, $L_s = \bar{L}/2$ and the wages are $w_s^0 \equiv \alpha_s\beta(\bar{L}/2)^{\beta-1}$. With free movement of workers, consumption is equalized across countries so that wages become equal. By symmetry, the same wage, w^e, applies in both states. Let $\overline{\alpha} \equiv (\frac{1}{2}\alpha_1^{1/(1-\beta)} + \frac{1}{2}\alpha_2^{1/(1-\beta)})^{1-\beta}$ be the 'average' of the productivity shocks.[13] It is easily checked that the equilibrium labor allocation and wage satisfy:

$$L_s^e = \left(\frac{\alpha_s}{\overline{\alpha}}\right)^{\frac{1}{1-\beta}} \left(\frac{\bar{L}}{2}\right); \quad \text{and} \quad w^e = \overline{\alpha}\beta\left(\frac{\bar{L}}{2}\right)^{\beta-1}, \tag{7}$$

where $L_1^e > \bar{L}/2 > L_2^e$ and $w^e \in (w_2^0, w_1^0]$. Using $z = \alpha_1/\alpha_2 > 1$ and substituting into (5), we can write

$$G(z,\beta) = \frac{w_1^0 - w^e}{w^e - w_2^0} = \frac{z - \overline{z}}{\overline{z} - 1}.$$

where the 'average' of relative productivity shocks $\overline{z} \equiv (\frac{1}{2} + \frac{1}{2}z^{1/(1-\beta)})^{1-\beta}$ increases with β and z. It can be shown that both the differences $z - \overline{z}$ and $\overline{z} - 1$ increase with z, but that the latter increases at a faster rate. Hence, $G(z,\beta)$ is strictly decreasing in β and z and $0 = G(z,1) \le G(z,\beta) \le G(z,0) = 1$. The overall effect is determined by two components: the short run cost given by $z - \overline{z}$ and the overall

[13]The average here is the weighted power mean of α_1 and α_2 with power $1/(1-\beta)$.

long term benefit that is proportion to $\overline{z} - 1/2(1+z)$. When there is less congestion (larger β), the short run cost falls and the longer term benefit rises. This yields a smaller critical discount factor for which the policy of free movement of workers can be sustained. The effect of the relative productivity shock is more complex because an increase in z raises both the short-term cost and long-term benefit. Nevertheless, the latter benefit dominates: the larger the relative shock z, the lower the discount factor that can sustain free movement of labor.

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 strongest. This is a long run benefit that must be balanced against the short run cost of migration. For a given discount factor, if the congestion factor or the shocks are small enough, then 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 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 larger are the shocks or the weaker is the congestion factor (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 workers' productivity, earnings or consumption. However, for full congestion ($f(l) = 1$), one can show that $G(\theta) < 1$ in the presence of risk aversion. Hence, in a random crop economy, risk aversion is a motive for workers to accept a policy of free movement of workers, which they would not do under risk neutrality. The insurance benefit is additional to the labor market flexibility benefit and offsets the impatience of workers.

Assuming iso-elastic production and preferences with constant relative risk aversion ρ, we can write

$$G(z, \beta, \rho) = \frac{z^{1-\rho} - \overline{z}^{(1-\rho)}}{\overline{z}^{(1-\rho)} - 1}.$$

This expression 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.*

4.4 Wage rigidities

It has been argued that many countries have been reluctant to allow uncontrolled inflows of workers in times of high unemployment. Boeri and Brücker (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 can be a rationale against the adoption of free movement of workers only when minimum wages have a permanent impact of economic outcomes.

We consider a simple situation where unemployment stems from downward wage rigidities. For simplicity, the wages w_s and w_s^* cannot fall below minimum wages $\underline{w}$ and $\underline{w}^*$ that are set exogenously and independently of whether there is free movement of labor or not. To shorten our discussion, we focus on identical minimum wages ($\underline{w} \equiv \underline{w}^*$), risk neutral workers and the symmetric two-state model where $z = \alpha_1/\alpha_2 = \alpha_2^*/\alpha_1^* > 1$ and $L^0 = \bar{L}/2$. Let l_s denote the number of employed workers and L_s denote the population including migrants, so that $L_s - l_s$ is unemployment.[14] With an iso-elastic production function, output is $\alpha_s f(l_s) = \alpha_s l_s^\beta$. Let $\widetilde{l}_s(\underline{w}) = (\underline{w}/(\alpha_s\beta))^{1/(\beta-1)}$ denote the level of employment at which the minimum wage just binds. Then, the employment level is equal to $l_s = \min\{L_s, \widetilde{l}_s(\underline{w})\}$ while the equilibrium wage is given by $w_s = \max\{\alpha_s\beta L_s^{\beta-1}, \underline{w}\}$, with symmetric expressions for foreign variables l_s^* and w_s^*. We also 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. Therefore, irrespective of whether the minimum wage

[14] The reader can interpret l_s as worked hours and $L_s - l_s$ as underemployment.

constraint binds or not, individual consumption can be written as

$$\widetilde{C}_s(L_s) = \frac{w_s}{\beta}\frac{l_s}{L_s},$$

where l_s/L_s can be interpreted as the employment rate. Recalling that $C_s(L_s) = \alpha_s L_s^{\beta-1}$ is the consumption where there is no minimum wage, it is easily checked that $\widetilde{C}_s(L_s) \leq C_s(L_s)$ with strict inequality where the minimum wage binds and with equality where it does not.

First, consider the short run equilibrium. In the absence of free movement of workers, minimum wages bind if $\widetilde{l}_s(\underline{w}) \leq \bar{L}/2$, that is, if $\underline{w} \geq w_s^0 = \alpha_s\beta(\bar{L}/2)^{\beta-1}$. For low values of $\underline{w} \leq w_2^0$, the minimum wage will not bind in either state; for intermediate values, $w_2^0 < \underline{w} \leq w_1^0$, the minimum wage will bind only in the low productivity state; and for $w_1^0 < \underline{w}$ it will bind in both states. In the presence of free movement of workers, consumptions are equalized across countries and also across states because states and countries are symmetric: $\widetilde{C}_1(L_1) = \widetilde{C}_2(\bar{L} - L_1)$.[15] When the minimum wage is sufficiently low ($\underline{w} < w^e$), the equilibrium migration and wage (L_s^e, w^e) are given by equation (7) where $w^e \in (w_2^0, w_1^0)$ and, as in previous sub-sections, there is no unemployment in either country. By contrast, if $\underline{w} > w^e$, the minimum wage binds, $w_s = \underline{w}$, and there is unemployment in both countries with employment levels $l_s = \widetilde{l}_s(\underline{w})$. Since consumptions are equalized in equilibrium, it follows that migration is determined by equation (7) and employment rates $\widetilde{l}_s(\underline{w})/L_s$ are also equalized across countries. Since employment is fixed irrespective of migration, there are fixed production levels in each country and the equilibrium with binding minimum wages has similar properties to the random crop model (with fixed crop: $\beta = 1$) discussed in section 4.2.

To assess the sustainability of the policy of free movement of workers as the minimum wage varies, we compute the value of

$$G(\underline{w}) = \frac{\widetilde{C}_1(\bar{L}/2) - \widetilde{C}_1(L_1)}{\widetilde{C}_2(L_2) - \widetilde{C}_2(\bar{L}/2)}.$$

There are four cases to consider according to whether the minimum wage lies above or below the thresholds w_2^0, w^e and w_1^0. First, for $\underline{w} \leq w_2^0$, the minimum wage does not bind in either state, either in the presence or the absence of migration, so that Proposition 2 applies unchanged. We know from equation (4.2) in section 4.2 that $G(\underline{w}) < 1$ for $\beta < 1$ and $z > 1$. That is, free movement of workers is sustainable for a high enough discount factor.

Second, for $w_2^0 < \underline{w} \leq w^e$, the minimum wage binds only in the low productivity country and in the absence of a free migration policy. That is, absent migration in the low productivity state,

[15] By symmetry, $\widetilde{C}_1(L_1) = \widetilde{C}_2^*(L_1)$ and $\widetilde{C}_2(\bar{L} - L_1) = \widetilde{C}_1^*(\bar{L} - L_1)$.

$w_2 = \underline{w} > w_0^2$ and $l_2 = \widetilde{l}_2(\underline{w}) < \bar{L}/2$. Recalling that $C_2(\bar{L}/2) = w_2^0/\beta$, we have

$$\widetilde{C}_2(\bar{L}/2) = C_2(\bar{L}/2)\left(\frac{\underline{w}\widetilde{l}_2(\underline{w})}{w_2^0\left(\bar{L}/2\right)}\right) < C_2(\bar{L}/2), \tag{8}$$

where the inequality follows because the elasticity of labor demand is greater than one and hence, $\underline{w}\widetilde{l}_2(\underline{w}) < w_2^0(\bar{L}/2)$. Equally, because the wage bill, $\underline{w}\widetilde{l}_2(\underline{w})$, is inversely related to $\underline{w}$, it follows that consumption $\widetilde{C}_2(\bar{L}/2)$ falls as the minimum wage rises.[16] Hence, in this case

$$G(\underline{w}) = \frac{z - \overline{z}}{\overline{z} - \left(\frac{\underline{w}\widetilde{l}_2(\underline{w})}{w_2^0\left(\bar{L}/2\right)}\right)}.$$

The binding minimum wage makes migration more attractive in the low productivity state (denominator), whereas the short run cost to the receiving country does not depend on the minimum wage (numerator). Equally, it can be seen that $G(\underline{w})$ is decreasing in $\underline{w}$ in this range, meaning that a higher minimum wage can help sustain the policy of free movement of workers. Intuitively, a binding minimum wage creates a distortion that, absent free movement, leads to underemployment in the low productivity country. A policy of free movement of workers allows workers to leave the underemployment country, which reduces labor supply and raises the wage above the minimum wage. This eliminates the labor market distortions and inefficiencies caused by downward rigidities in the wage. The larger the distortion, the greater are the benefits of avoiding it through migration.[17]

Next, consider a high minimum wage such that $\underline{w} > w_1^0$. In this case, the minimum wage binds in both the high and low productivity countries both with and without free movement of labor. Labor is determined by its demand and production is fixed irrespective of labor movements. As described above, employment rates are equalized and the equilibrium shares the same properties as the random crop model. As a consequence, $G(\underline{w}) = 1$ and the policy of free movement is not sustainable. Intuitively, the short run cost of immigration is exactly matched by the short run benefit of emigration, so that in the high productivity state, impatient individuals are not willing to bear the short run cost for a future uncertain gain.

Finally, for $w_e < \underline{w} \leq w_1^0$, the minimum wage binds when there is free movement while, absent free movement, it binds only in the low productivity country. In this case, an increase in the minimum wage

[16]The elasticity of labor demand is $1/(1-\beta) \geq 1$ with strict inequality for $\beta > 0$.

[17]It follows that a minimum wage in this range improves welfare for some discount factors. For example, without the minimum wage, there may be no discount factor that sustains the policy of free movement even though it is desirable because free movement improves average consumption. With the minimum wage, $G(\underline{w})$ may be lowered sufficiently that the policy does become sustainable, improving long run welfare.

increases the short run cost of immigration in the high productivity state because consumption in the free movement equilibrium falls with an increase in the minimum wage. In contrast, the consumption without free movement is unaffected by a change in the minimum wage. Equally, the net benefit of emigration in the low productivity state also falls: the net benefit of emigration in the low productivity state is directly proportional $\underline{w}\widetilde{l}_2(\underline{w})$. Thus, $G(\underline{w})$ is increasing in this range. In particular, it can be shown that for $w_e < \underline{w} \leq w_1^0$,

$$G(\underline{w}) = \frac{\left(\frac{w_1^0(\bar{L}/2)}{\underline{w}\widetilde{l}_1(\underline{w})}\right) L_1^e - (\bar{L}/2)}{L_1^e - (\bar{L}/2)}.$$

In this case, $w_1^0 \geq \underline{w}$ and $\bar{L}/2 \leq \widetilde{l}_1(\underline{w})$ with equality only if $\underline{w} = w_1^0$. Because the elasticity of labor demand is greater than one, $w_1^0(\bar{L}/2) \leq \underline{w}\widetilde{l}_1(\underline{w})$, again with equality only if $\underline{w} = w_1^0$. Thus, it is easily seen that $G(\underline{w}) \leq 1$ with equality only if $\underline{w} = w_1^0$. To sum up, a higher minimum wage makes the policy of free movement less sustainable. Intuitively, free movement transmits the labor market distortion into the high productivity country and does not eliminate it in the low productivity country. So, workers are less willing to accept the policy of free movements when they face a good shock.

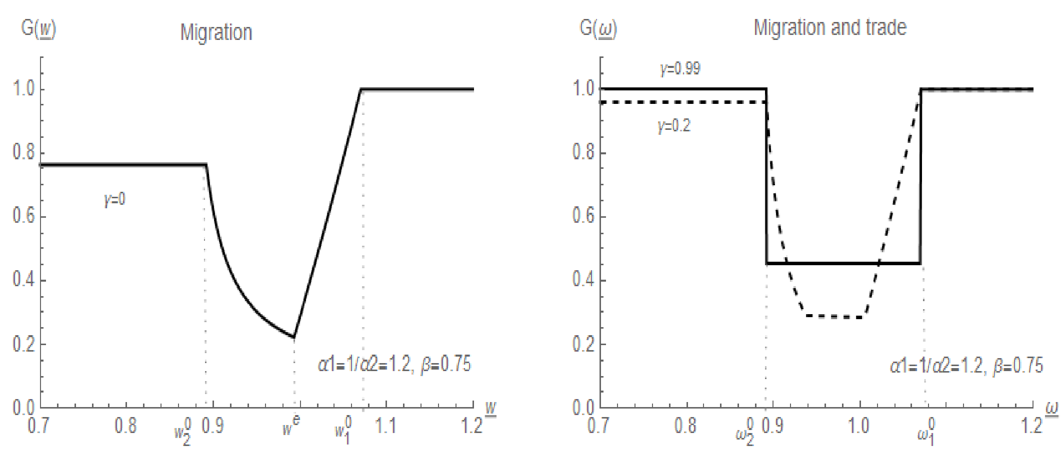

Figure 1: **Wage rigidities and relative cost of adopting a policy of free movement of workers**
Comment: The relative cost of adopting the policy of free movement of workers, G, is displayed as a function of a exogenous common minimum wage. The left panel shows the case of pure migration ($\gamma = 0$). The right panel depicts the economy with almost pure trade ($\gamma = 0.99$, solid line) and an economy with 20% trade ($\gamma = 0.2$, dashed line). The policy is sustained for lower discount factors if G is low. This occurs when only the low productivity country has a binding minimum wage (at the lowest G).

The left panel of Figure 1 presents a numerical example of the relative cost of adopting the policy of free movement of workers, $G(\underline{w})$, as a function of the exogenous common minimum wage $\underline{w}$ (the right panel will be discussed in section 5.4 when trade is added to the model). As argued above, as the minimum wage rises, this relative cost $G(\underline{w})$ is firstly constant and below 1, then falls and increases again to reach one. In particular, (i) for high values of the minimum wage ($\underline{w} > w_1^0$), a policy of free movement of workers is not sustainable; (ii) for low values of the minimum wage ($\underline{w} < w_2^0$), a policy of free movement of workers is sustainable for high enough discount factors; (iii) for values of the minimum wage $w_2^0 < \underline{w} \leq w^e$, an increases in the minimum wage enhances sustainability ($G(\underline{w})$ is decreasing); (iv) for values of the minimum wage $w_e < \underline{w} \leq w_1^0$, an increases in the minimum wage is detrimental to sustainability ($G(\underline{w})$ is increasing with $G(\underline{w}) \to 1$ as $\underline{w} \to w_2^0$). This pattern might help explain why some countries with enduring labor market frictions and large unemployment rates have seen a hardening of migration conditions within the E.U.

We summarize this finding in the following proposition:

Proposition 4 (Effect of a minimum wage) *Consider the two anti-correlated state model with risk-neutral workers. The effect of a common minimum wage on the policy of free movement of workers is non-monotone. As the minimum wage rises from zero, this policy firstly becomes more sustainable, then less sustainable and finally unsustainable.*

5 Migration and trade model

We now extend our discussion to consider the impact of trade on the sustainability of a policy of free movement of workers. 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 $F_s(L_s^i) = \alpha_s (L_s^i)^\beta$, $i \in \{X, Z\}$ 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.[18] 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 the Appendix. 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,

$$\frac{w_s}{w_s^*} = \frac{L_s^*}{L_s}. \tag{9}$$

Thus, the terms of trade moves in inverse proportion to the flows of labor. Equilibrium consumption of the composite good is $C_s = (P_s^X)^{-\gamma/2} (P_s^{X*})^{-\gamma/2} (P_s^Z)^{\gamma-1} Y_s / L_s$.[19] 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

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

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

consumption is

$$\frac{C_s}{C_s^*} = \frac{(P_s^Z)^{\gamma-1} w_s}{(P_s^{Z*})^{\gamma-1} w_s^*}. \tag{10}$$

Substituting for prices gives home consumption as a function of population:

$$C_s(L_s) = \alpha_s \left(\frac{\alpha_s^*}{\alpha_s}\right)^{\gamma/2} \left(\frac{\bar{L} - L_s}{L_s}\right)^{\beta\gamma/2} L_s^{\beta-1}. \tag{11}$$

where $\bar{L} - L_s = L_s^*$. Equation (11) provides a specific form for the function described in Section 2.[20] 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 (2), under which there is excess agglomeration of workers, becomes

$$\frac{L_s^* C_s^{*\prime}(L_s^*)}{L_s C_s'(L_s)} = \frac{1 - \beta + \beta\gamma/2 + (\beta\gamma/2)(L_s^*/L_s)}{1 - \beta + \beta\gamma/2 + (\beta\gamma/2)(L_s/L_s^*)} \leq 1. \tag{12}$$

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 $L_s = L_s^*$.[21]

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 constant returns to scale) and then analyze the general case 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 analyze 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 (9), 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 (10), free movement of workers equalizes

[20] With no traded good, $\gamma = 0$, equation (11) reduces to $C_s(L_s) = \alpha_s L_s^{\beta-1}$, which was the case studied in section 4.2.

[21] For $\beta\gamma = 0$, (12) holds with equality and any allocation of labor across countries is efficient.

consumption, so 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 purchasing power of local workers.

Secondly, changes in the terms of trade absorb the effect of productivity differences on consumption and therefore mitigate migration incentives. From (11), 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, which implies that $\alpha_s^* \alpha_s$ and $L_s L_s^*$ are constant. Then, the consumption levels do not depend on the states of nature. The terms of trade *fully* absorb the productivity shocks 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. Thus, the short run equilibrium in the symmetric case has no migration: $L_1 = L_2 = \bar{L}/2$. Applying this labor allocation in (12), one can see that the free movement equilibrium is trivially efficient. The same analysis applies in the presence of congested factors ($\beta < 1$) as long as $\gamma = 1$. Finally, because consumption is perfectly smoothed across states, there are no cost and no gain from migration. Countries are therefore indifferent to the adoption of a policy of free movement of workers. We summarize these results in the following proposition:

Proposition 5 (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 countries are indifferent to the adoption of a policy of free movement of workers.*

In the two anti-correlated 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 workers produce a different good at a different productivity when they move across border. Also, movements of workers and commodities are complements in the sense that positive local productivity shocks increase both export and immigration (Markusen, 1983; Neary, 1995).

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

$$\frac{L_s^*}{L_s} = \left(\frac{\alpha_s^*}{\alpha_s}\right)^{\frac{1-\gamma}{1-\beta(1-\gamma)}}. \tag{13}$$

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 (12) and (13), 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. As mentioned in the introduction, 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 6 *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 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 2 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 share on tradeable goods of $\gamma = 0.2$ and a weak congestion factor of $\beta = 0.8$, the domestic country hosts 69% of the total population compared to 61% in the socially optimal allocation. Larger shocks yield more excessive agglomeration.

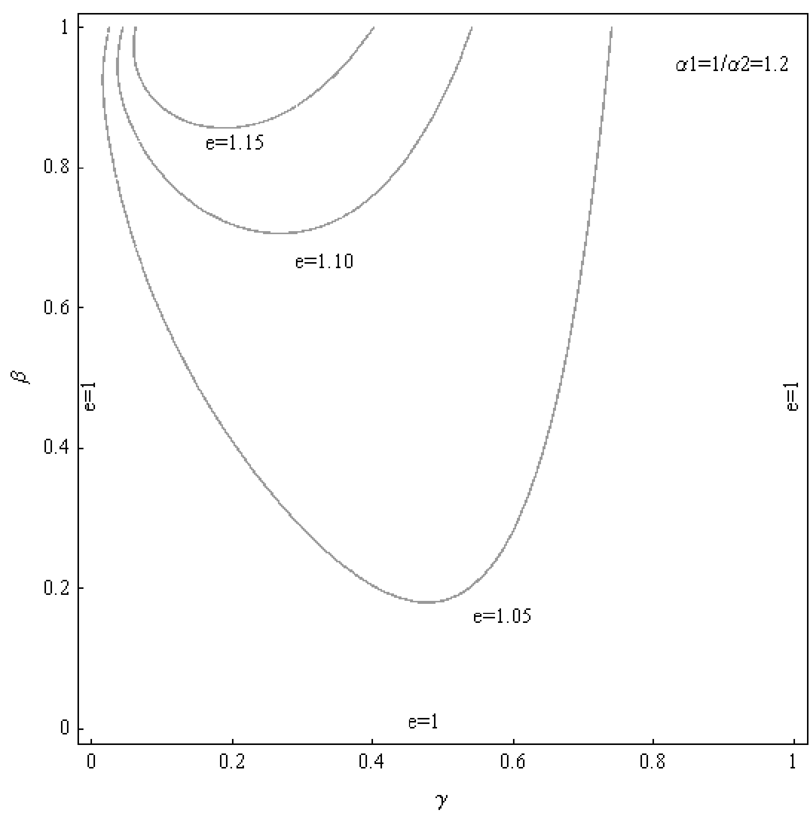

Figure 2: **Excessive agglomeration of workers**
Comment: The ratio between the equilibrium and first best numbers of workers, e, is larger for intermediate trade (γ) and low congestion (high β).

Figure 2 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 migrants when there exists no congestion (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 and there are incentives for migration. Lower congestion exacerbates these incentives. At the same time, immigration changes the terms of trade, which harms 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 wages 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 3 plots the locus of the relative cost of adopting the policy of free movement of workers for the parameters $\theta = (\gamma, \beta)$. More precisely, it plots the loci of $G(\gamma, \beta) = 0.50$, 0.75 and 1. These values respectively correspond to critical discount factors $\delta = 0.66$, 0.85 and 1.[22] The shock structure is the same as in Figure 2. The area (a) corresponds to $G(\theta) > 1$, where a policy of free movement is not sustainable, 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.

Figure 3 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 3 (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 separated by a dashed curve that corresponds to the locus

[22]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 4, 2 and 0 percent ($r = \delta^{-1/10} - 1$).

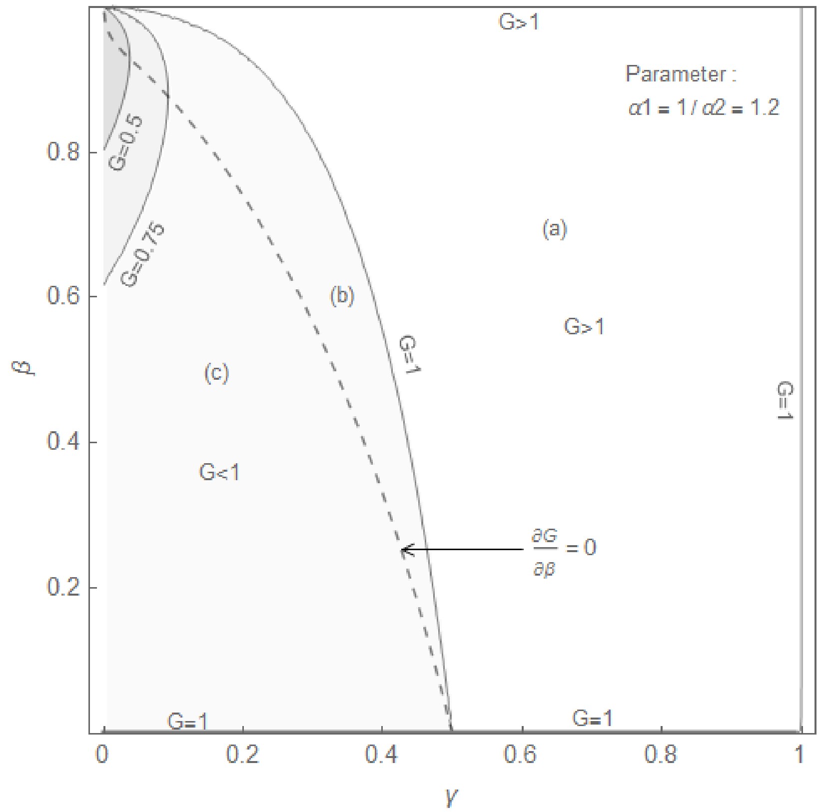

Figure 3: **Relative cost of adopting the policy of free movement of workers,** $G(\theta)$**.** Comment: The relative cost of adopting the policy of free movement of workers, G, is related to the cost of sustaining the policy of free movement of workers. Large trade sectors (γ) makes the policy unsustainable. Sustainability is not monotone with respect to factor congestion (β).

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 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 results in the following proposition.

Proposition 7 *Suppose that workers are risk neutral and that countries face two 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.*

5.4 Wage rigidities

In this section we reconsider the relationship between the sustainability of free movement of workers and downward wage rigidities when countries trade. For the sake of simplicity, we again assume an exogenous common minimum wage $\underline{\omega}$ expressed in each local currency. We assume that the wages w_s and w_s^* cannot fall below $\underline{\omega}P_s^X$ and $\underline{\omega}^* P_s^{X*}$. As before, let l_s and L_s denote the number of employed workers and the number of residents including migrants. To illustrate the effects of labor frictions, we consider the simplest example with iso-elastic production function, local income redistribution, risk neutral individuals and two symmetric states.

The analysis of the minimum wage is very similar to that given in section 4.4. Let $\widetilde{l}_s(\underline{\omega}) = (\underline{\omega}/(\alpha_s\beta))^{1/(\beta-1)}$. In the presence of a minimum wage $\underline{\omega}$, the number of domestic employed workers is $l_s = \min\{L_s, \widetilde{l}_s(\underline{\omega})\}$ while the equilibrium wage is $w_s/P_s^X = \max\{\alpha_s\beta L_s^{\beta-1}, \underline{\omega}\}$, with symmetric expressions holding for foreign variables. The trade balance condition maintains equality of wage bills, so that $w_s l_s = w_s^* l_s^*$. With the appropriate choice of the constant κ, individual consumption can be written as:

$$\widetilde{C}_s\left(L_s\right) = \frac{1}{L_s}\left(\frac{w_s l_s}{P_s^X}\right)^{1-\frac{\gamma}{2}}\left(\frac{w_s^* l_s^*}{P_s^{X*}}\right)^{\frac{\gamma}{2}}. \tag{14}$$

These consumption functions can be used to evaluate $G(\underline{\omega})$, which is defined in the same way as given in equation (8). Let $\omega_s^0 = \alpha_s\beta(\bar{L}/2)^{\beta-1}$ denote the thresholds above which the minimum wage binds in state s when there is no migration. Unlike the case without trade, analyzed in section 4.4, where the wage is equalized in a free movement equilibrium, here there are two threshold wages with free movement of labor: below the lower threshold, there is no unemployment in either country; above the higher threshold, there is unemployment in both countries. Let L_s^e and L_s^ϵ denote the respective free movement equilibrium allocations when the minimum wage does not bind in either country and when it binds in both countries.[23] Furthermore, let $\omega_s^e = \alpha_s\beta(L_s^e)^{\beta-1}$ and $\omega_s^\epsilon = \alpha_s\beta(L_s^\epsilon)^{\beta-1}$ denote the associated wages. In the two state symmetric case with $\gamma \in (0,1)$ and $\beta \in (0,1)$, we have $\omega_2^0 < \omega_2^e < \omega_2^\epsilon$ and $\omega_1^\epsilon < \omega_1^e < \omega_1^0$. It is also easily checked that $\omega_2^\epsilon \leq \omega_1^\epsilon$ with strict inequality for $\gamma > 0$. With free movement of labor, we have that the minimum wage does not bind in either country for $\underline{\omega} < \omega_2^e$ and it binds for both countries for $\underline{\omega} > \omega_1^\epsilon$. This means there are five possible regimes. For $\underline{\omega} < \omega_2^0$, the

[23] It can be shown that

$$L_s^e = \left(\frac{\alpha_s^{\frac{1-\gamma}{1-\beta(1-\gamma)}}}{\alpha_1^{\frac{1-\gamma}{1-\beta(1-\gamma)}} + \alpha_2^{\frac{1-\gamma}{1-\beta(1-\gamma)}}}\right)\bar{L}; \quad \text{and} \quad L_s^\epsilon = \left(\frac{\alpha_s^{\frac{1-\gamma}{1-\beta}}}{\alpha_1^{\frac{1-\gamma}{1-\beta}} + \alpha_2^{\frac{1-\gamma}{1-\beta}}}\right)\bar{L}$$

so that $L_1^\epsilon > L_1^e > \overline{L}/2 > L_2^e > L_2^\epsilon$ for $\beta \in (0,1)$ and $\gamma \in (0,1)$.

minimum wage never binds; for $\underline{\omega} \in (\omega_2^0, \omega_2^e)$ it binds only in the low productivity country and absent free movement of workers; for $\underline{\omega} \in (\omega_2^e, \omega_1^\epsilon)$, it binds only in the low productivity country in both the absence and presence of free movement; for $\underline{\omega} \in (\omega_1^\epsilon, \omega_1^0)$, it binds in both countries when there is free movement, but binds only in the low productivity country when labor is immobile; and finally for $\underline{\omega} > \omega_1^0$, it binds in all situations.

First, consider the case where the common minimum wage $\underline{\omega} \leq \omega_2^0$. In this case, the minimum wage is low enough such that it never binds in either country whether there is free movement of workers or not. This is the case discussed in section 5.3. Whether free movement is sustainable or not will depend on the parameter values. In particular, if γ is large enough, the previous results show that free movement is not a sustainable policy, that is $G(\underline{\omega}) > 1$ for $\underline{\omega}$ in this range and for γ large enough. Secondly, consider the case where $\underline{\omega} \geq \omega_1^0$, so that the minimum wage binds in both countries whether there is free movement or not. In this case, production is determined by labor demand and the consumptions are similar to the random fixed crop model.[24] Identically to section 4.4, it can be shown that $G(\underline{\omega}) = 1$ for a minimum wage in this range.

Following an identical reasoning to that given in section 4.4, it can be shown that $G(\underline{\omega})$ is decreasing in $\underline{\omega}$ for $\underline{\omega} \in (\omega_2^0, \omega_2^e)$. In this case, the minimum wage binds only in the low productivity country and absent migration. The policy of free movement has the added advantage of allowing workers in the country with low productivity to emigrate and escape local unemployment. The larger the minimum wage the greater is this advantage.

Next, consider a minimum wage $\underline{\omega} \in (\omega_1^\epsilon, \omega_1^0)$. In this case, the minimum wage binds in both countries with free movement of labor; whereas it binds in the low productivity country but not in the high productivity country when labor is immobile. Again, following a similar reasoning to that given in section 4.4, it can be shown that $G(\underline{\omega})$ is increasing in $\underline{\omega}$ for $\underline{\omega} \in (\omega_1^\epsilon, \omega_1^0)$. This follows because an increase in the minimum wage increases the short run cost of immigration in the high productivity state: in contrast, the minimum wage does not affect consumption in the absence of immigration whereas an increase of minimum wage leads to more unemployment and a fall in consumption in the free movement equilibrium. In addition, the net benefit of emigration in the low productivity state also falls: although unemployment rises with a higher minimum wage, per capita consumption falls less in the presence of migration because some workers leave the low productivity country and reduce congestion of productive factors there. Thus, it follows that there is some minimum wage such that $G(\underline{\omega}) < 1$. Consequently, for some discount factors, a policy of free movement becomes sustainable

[24] It can be seen from (14), that the total domestic consumption in state s, $L_s \widetilde{C}_s(L_s)$ is equal to the (random) constant $(\alpha_s \beta)^{\frac{1-\gamma/2}{1-\beta}} (\alpha_s^* \beta)^{\frac{\gamma/2}{1-\beta}} \underline{\omega}^{-\frac{\beta}{1-\beta}}$, which is independent of L_s.

even it were not sustainable with no minimum wage: free movement of workers allows a reallocation of workers that diminishes the inefficiencies caused by wage rigidities in the low productivity country. Hence, we can conclude:

Proposition 8 *In the two-state economy with risk-neutral workers, traded goods and a common minimum wage, the effect of a common minimum wage on the policy of free movement of workers is non-monotone: there is a non-empty set of minimum wages such that a policy of free movement of workers is sustainable even if it would be unsustainable without the minimum wage.*

We still have to consider cases where $\underline{\omega} \in (\omega_2^e, \omega_1^\epsilon)$. In this interval, when there is free movement of labor, the minimum wage binds in the low productivity country but not in the high productivity country. The equilibrium allocation under free movement of workers, L_1, is given by the solution of:[25]

$$\frac{(L_1)^{1-\beta(1-\gamma)}}{\bar{L} - L_1} = \left(\frac{\alpha_1}{\alpha_2}\right)^{1-\gamma} \left(\frac{\underline{\omega}}{\alpha_2 \beta}\right)^{\frac{\beta(1-\gamma)}{1-\beta}}.$$

Although it is difficult to obtain the analytical properties of $G(\underline{\omega})$ for $\underline{\omega} \in (\omega_2^e, \omega_1^\epsilon)$, it is easy to compute the solution numerically. The right panel of Figure 1 plots $G(\underline{\omega})$ for two particular values of γ: $\gamma = 0.20$ and $\gamma = 0.99$. The panel confirms that $G(\underline{\omega})$ can initially be above or below one for a low minimum wage, but at first decreases to a value less than one, then increases in $\underline{\omega}$ and finally reaches one for $\underline{\omega} > \omega_1^0$. In particular, it should be noted that for γ close to one, an intermediate value of the minimum wage allows a policy of free movement to be sustained when such a policy would never be sustained otherwise. As in the standard migration model, this analysis suggests that labor market rigidities help to sustain a policy of free movement of workers provided that the rigidities are neither permanent nor widespread.

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

[25] It can be checked that the solution of this equation is L_1^e when $\underline{\omega} = \omega_2^e$ and is L_1^ϵ when $\underline{\omega} = \omega_1^\epsilon$.

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 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 provided these rigidities are not permanent.

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, the model is stylized and many other important aspects of the problem are left out of the account. We have taken as given that immigration always has short-run costs. This may not be true or only true for particular sectors or skill groups. We have also only considered the case where migrants remain citizens of the origin country. An earlier version of the paper, Picard and Worrall (2014), considered the case of a full right migration policy where migrants could acquire citizenship. It was shown that the full right migration policy is less likely to be sustained than the free labor mobility policy because it changes the default positions. An important omission is the lack of skill heterogeneity across workers. With skill heterogeneity, there may be offsetting worker flows in response to productivity shocks. We have also not considered controlled migration. Although, as we have argued such a policy is difficult to implement relative to a policy of free movement, a policy of controlled migration may have efficiency benefits as it may reduce excess agglomeration issues. Equally, although one of our motivations was to apply the analysis of sustainable trade policies to migration, we have assumed free trade. We have not considered how the adoption of trade and migration policies might be coordinated. All these are possible interesting directions for further research.

References

Acemoglu, D., Golosov, M., and Tsyvinski, A. (2010), Dynamic Mirrlees Taxation under Political Economy Constraints, *Review of Economic Studies*, 77, 841–881.

Bagwell, K. and Staiger, R. W. (1990), A Theory of Managed Trade, *American Economic Review*, 80, 779–95.

Bagwell, K. and 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. and Brücker, 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. and Kehoe, P. J. (1990), Sustainable Plans, *Journal of Political Economy*, 98, 783–802.

Clemens, M. A., Montenegro, C. E., and 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. and Weinstein, D. E. (2002), Technological Superiority and the Losses from Migration, NBER Working Papers 8971, National Bureau of Economic Research, Inc.

Facchini, G. and Testa, C. (2009), Who is against a Common Market?, *Journal of the European Economic Association*, 7, 1068–1100.

Felbermayr, G. J. and Kohler, W. (2007), Immigration and Native Welfare, *International Economic Review*, 48, 731–760.

Grossman, G. M. and Helpman, E. (1995), The Politics of Free-trade Agreements, *American Economic Review*, 85, 667–90.

Klein, P. and Ventura, G. (2009), Productivity Differences and the Dynamic Effects of Labor Movements, *Journal of Monetary Economics*, 56, 1059–1073.

Ligon, E., Thomas, J. P., and 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.

Ortega, F. (2010), Immigration, Citizenship, and the Size of Government. *The B.E. Journal of Economic Analysis & Policy*, 10.

Ottaviano, G. P. and Peri, G. (2012), Rethinking the Effect of Immigration on Wages. *Journal of the European Economic Association*, 10, 152–197.

Picard, P. M. and Worrall, T. (2014), Is a Policy of Free Movement of Workers Sustainable? IZA Discussion Papers 8035, Institute for the Study of Labor.

Staiger, R. W. and Bagwell, K. (1999), An Economic Theory of GATT, *American Economic Review*, 89, 215–248.

Thomas, J. P. and Worrall, T. (1988), Self-enforcing Wage Contracts, *Review of Economic Studies*, 55, 541–554.

Wahba, J. and Zenou, Y. (2012), 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.

Wellisch, D. and Walz, U. (1998), Why do Rich Countries Prefer Free Trade over Free Migration? The Role of the Modern Welfare State, *European Economic Review*, 42, 1595–1612.

Appendix

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: $P_s^X F_s(L_s^X) = P_s^{X*} F_s^*(L_s^{X*})$. 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 $w_s / w_s^* = L_s^* / L_s$.