Review of Economic Dynamics 18 (2015) 593–618

Contents lists available at ScienceDirect

Review of Economic Dynamics

www.elsevier.com/locate/red

An incentive mechanism to break the low-skill immigration deadlock ☆

David de la Croix [a,b,1], Frédéric Docquier [a,c,2]

[a] IRES, Université catholique de Louvain, Belgium
[b] CORE, Université catholique de Louvain, Belgium
[c] National Fund for Scientific Research, Belgium

ARTICLE INFO

Article history:
Received 26 March 2012
Received in revised form 15 September 2014
Available online 11 October 2014

JEL classification:
F22
F55
D58
D6
D7

Keywords:
Public good
Inequality aversion
Immigration policy

ABSTRACT

The international mobility of low-skilled workers is constrained by immigration restrictions set by rich countries, preventing extreme poverty from declining in the developing world. Under aversion to extreme poverty, one country's immigration policy affects the welfare of other countries: "nationalist" political decisions are subject to a prisoner's dilemma. Taking advantage of efficiency gains from coordination, we propose an alternative allocation of labor, which minimizes extreme poverty while keeping rich countries as well off as in the nationalist situation. This allocation can be decentralized through international taxes and subsidies. Numerical simulations show that the role of aversion to poverty is negligible when considering each country separately, but becomes important for the design of a coordinated migration policy.

1. Introduction

It is commonly admitted that low-skilled migration from poor to rich countries should be seen as an explicit component of the development policy of the rich world. By relaxing labor market constraints in developing countries and inducing large amounts of remittances, low-skilled migration generates gains for migrants, their families and, most importantly, the sending countries (see Pritchett, 2006; Clemens, 2011; Winters, 2001; Walmsley and Winters, 2003). These migration flows are currently restrained by highly restrictive, immovable immigration policies voted by the citizens of rich countries. Such regulations carry considerable economic costs for developing countries and prevent global inequality from declining (see Pritchett, 2006). Breaking the gridlock on labor migration requires solutions that are politically acceptable in rich countries.

☆ We wish to thank three anonymous referees, Gani Aldashev, Georg Kirchsteiger, Çaglar Özden, Maurice Schiff, and participants at conferences in Lille (SMYE 2008), Aix-en-Provence (Journées générations imbriquées 2008), CORE-ULg (2009), Paris (3rd Migration and Development Conference 2010) and at seminars at ETH Zurich and IEM-Nantes for their comments on previous drafts. We also wish to thank Axel Gosseries and Omer Moav for useful discussions at an early stage of this project. The authors acknowledge the financial support of the Belgian French-speaking Community (ARC conventions 09-14019 on "Geographical mobility of factors", and 09-14018 on "Sustainability").

E-mail addresses: david.delacroix@uclouvain.be (D. de la Croix), frederic.docquier@uclouvain.be (F. Docquier).
[1] Fax: +3210473945.
[2] Fax: +3210473945.

http://dx.doi.org/10.1016/j.red.2014.09.004
1094-2025/

Accounting for endogenous immigration restrictions (as in Facchini and Mayda, 2008), this paper proposes an incentive–compatible mechanism that could raise the number of South–North, low-skilled migrants, characterizes its properties, and quantifies its effectiveness.

We develop a political economy model describing an economy with a set of heterogeneous rich countries and one poor emigration country (representing the group of developing countries). The model formalizes natives' economic incentives to host low-skilled immigrants, i.e. the effects on wages (Borjas et al., 1997; Borjas, 2009), taxation (Facchini and Mayda, 2009), return to capital, as well as non-economic costs. In addition, it assumes that citizens from all rich countries are averse to extreme poverty, albeit to a small extent, and have internalized that welcoming low-skilled migrants from the South reduces global inequality. The key implication of altruism is that each developed country benefits from other states welcoming poor immigrants. Altruism confers a public good nature to immigration, implying that non-coordinated immigration policies are subject to a prisoner's dilemma problem: although households in rich countries suffer from the idea that people in developing countries are really poor, voting to welcome additional migrants is not optimal for them since they would be alone in bearing the cost, while the benefits would accrue to everybody in the world who cares about extreme poverty. We exploit the efficiency gains underlying this prisoner's dilemma to minimize extreme poverty subject to implementation constraints in rich countries.

Our theoretical and numerical analyses follow three steps. First, we construct an allocation of labor in which each rich country decides on its immigration level by voting so as to maximize the welfare of its citizens, taking the choices of other countries as given. This Nash equilibrium is labeled as the nationalist (or "Nashionalist") allocation. At that allocation, the resulting demand for low-skilled immigrants can reasonably be assumed to be lower than the supply. Assuming the nationalist allocation is the one observed, the model can be calibrated using statistics on bilateral migration, working-age population, skill premia, income shares, output, redistributive transfers, and before- and after-tax Gini coefficients. Second, we design an alternative allocation which leaves host countries indifferent compared with the nationalist scenario and maximizes the volume of South–North, low-skilled migration (or equivalently, maximizes the utility of the worst-off in the South). This allocation is labeled the no-regret allocation. Because there are positive externalities across rich countries when they jointly lower barriers to immigration, it is possible to increase global labor movements without any utility cost for the rich. Third, we show that the no-regret allocation can be decentralized through a tax-subsidy scheme which makes people internalize the social gains and costs of migration. For different degrees of aversion to extreme poverty, we can compute the global increase in low-skilled migration, and the country-specific tax and subsidy rates which allow the no-regret allocation to be decentralized. Our numerical simulations show that, even with small degrees of aversion, the coordination of immigration policies induces substantial increases in migration.

Clearly, aversion to extreme poverty is the key ingredient in our incentive–compatible mechanism. We use maximin altruistic preferences and consider that the utility of citizens in the rich world depends on their own level of consumption and the consumption of the worst-off in the world (i.e. low-skilled workers living in developing countries). Behavioral and experimental studies have revealed that individual choices are influenced by social preferences, including altruism, fairness, reciprocity and aversion to inequality (Fehr and Schmidt, 2006). The (quasi-)maximin model is related to the ideas presented by Yaari and Bar-Hillel (1984), according to which players want to help all other players, but are particularly keen to help someone who is worse off. It has been tested by Charness and Rabin (2002, 2005). Laboratory experiments reveal that this model does significantly better than altruism or pure self-interest, indicating that there is indeed much concern for those who are getting the lowest payoffs.

Is aversion to extreme poverty likely to affect development and immigration policies? Fehr and Schmidt (2006) examined the conditions under which the maximin motive plays a role in naturally occurring environments. In a competitive environment, or in an environment where the players view each other as agents behaving strategically, the maximin motive is not likely to be important. However, the maximin motive is highly relevant in the context of charitable giving or in the context of elections with a large number of people, in which strategic voting is unlikely to occur. Development assistance, donation to charities and NGOs, or the international community's efforts to reach the Millennium Development Goals are evidence of the rich world's concern for extreme poverty. It could be argued that altruism does not govern immigration policies, which are determined by the interests of the host populations only, except asylum policies which are adopted for humanitarian motives (Hatton and Williamson, 2006). We believe this view is changing as international organizations and scholars (through web sites, press releases, or academic studies among others) work on increasing public awareness of the link between migration and the development of poor countries.[3] Interestingly, our simulations show that the role of altruism is small when considering rich countries separately. However, when it comes to designing a coordinated migration policy, it becomes important.

As far as the existing literature is concerned, the benefits from coordinating immigration policies have been investigated in a few other studies. Based on humanitarian motives, hosting refugees raises the welfare of all potential host states. Hence, individuals might be expected to obtain some benefit from the knowledge that refugees also find safety in another country.

[3] In 2006, the International Organization for Migration initiated debates on EU immigration policy towards development-friendly measures. More recently, President Barack Obama announced that he would forbid the deportation of upstanding young people who are in the United States illegally, an attitude referred to as "an altruistic immigration policy" in the media. Commenting on this decision, David McKenzie, a lead economist for the World Bank, declared that "improved labor mobility is by far the greatest way to give a leg up to low-income people around the globe. High-skilled immigration is going to be useful for the high-skilled migrants, but allowing lower-skilled workers in is directly going to improve poverty in poor countries".

On this basis, Hatton (2004) and Hatton and Williamson (2006) argued that EU asylum policy can be viewed as the local provision of public goods, raising the issue of finding a cooperative outcome across countries. A similar issue is investigated in Bubb et al. (2011) who consider that distinguishing between refugees and economic migrants is costly. Host countries can then use restrictions (high standard of proof) to deter immigration. This choice exhibits strategic complementarity: "as more states increase their standard of proof, the best response to other states may be to increase the standard of proof". These complementarities give rise to multiple equilibria and coordination problems (high or low restrictions for all). The authors see the "1951 Convention Relating to the Status of Refugees" (basically, a convention imposing the "non-refoulement" of refugees) as a Pareto-improving contract that bound states to lower their standard of proof and provide a more efficient level of global public good. Unlike their paper, our model exhibits strategic substitutability: as more states decrease their level of low-skilled immigration, the best response from other states is to let more people come in. This excludes the possibility of multiple equilibria, although it does not solve the prisoner's dilemma problem. Finally, Skaksen et al. (2007) built a model with two types of spillover effects, a terms of trade externality (immigration causes wages and prices to decrease at destination) and a welfare policy externality (fiscal competition to attract immigrants). They show that non-cooperative behavior does not necessarily decrease the volume of immigration. It is only the case if the terms of trade externality dominates the welfare policy externality. None of these studies carried out a quantitative assessment of the welfare gain from moving from a non-cooperative equilibrium to a cooperative one.

Our paper is organized as follows. Section 2 presents our theoretical mechanism in a parsimonious framework in which the relaxation of labor market constraints is the only source of gain for those left in the South. The application of the theory using data on immigration, working-age population and output is proposed in Section 3. Then Section 4 introduces two ingredients which can affect our predictions quantitatively (i.e. migrants' remittances and development aid) and studies the properties of alternative utilitarian allocations. Finally, Section 5 concludes.

2. Theory

This section describes the framework used to endogenize the level of low-skilled immigration in rich receiving countries. We characterize the non-cooperative, nationalist allocation at which each host country votes on immigration, taking the immigration rates of other receiving states as given. Then, we describe the no-regret allocation in which the number of migrants is maximized subject to a set of participation constraints. Finally, we investigate how the no-regret alternative allocation can be decentralized through country-specific lump-sum taxes and subsidies per immigrant.

2.1. The environment

The world is made of $J+1$ countries, J developed ($j=1,\dots,J$) and 1 developing (indexed by 0) representing the whole developing world. Each country has two types of nationals, college graduates (denoted by $s=h$) and the less educated (denoted by $s=l$). The number of college graduates and low-skilled natives are denoted by n_j^h and n_j^l, respectively. The total national population is equal to $n_j = n_j^h + n_j^l$. The number of low-skilled migrants from country 0 to country j is denoted by m_j. As the international allocation of labor is determined by the votes of rich countries' citizens, the J developed states are the main actors in the game.

Inside each country, residents are either nationals or low-skilled migrants. Every person in the world has preferences defined over their own consumption and the consumption of the poorest individuals in the world, i.e. the less educated individuals left in developing countries. Migrants in country j consume c_j^m but their utility plays no part in our analysis because they have no right to vote.[4] Nationals of type s consume c_j^s and the less educated left in developing countries consume c_0^l. In addition, citizens' preferences depend on the migrant-to-citizen ratio in their country m_j/n_j through a convex quadratic cost parametrized by ε_j. This function aims at capturing non-economic factors in citizens' decision (such as the taste for an ethnically homogeneous society in South Korea and Japan), thereby allowing us to better isolate the economic factors underlying the immigration policy. The preferences of natives from country j are represented by the following utility function:

$$U_j^s = u(c_j^s) + \beta u(c_0^l) - \varepsilon_j \left(\frac{m_j}{n_j}\right)^2. \tag{1}$$

The utility from consumption $u(c)$ is increasing and concave. Individuals are altruistic with respect to the poorest people on in the world, i.e. $\beta \geq 0$.[5]

[4] Migrants are likely to be the main beneficiaries of the migration process, but their utility plays no part in our analysis because they have no right to vote. As argued in Hirschman (2001) for the United States, "Many immigrants tend to be fairly apolitical, are often slow to naturalize, and are more concerned with problems of day-to-day survival and their children's chances of upward mobility than with engagement in American politics". However the children and grandchildren of migrants might play a part in politics. In a dynamic setting, citizen voters would have to anticipate the effects of immigration on future domestic politics, as shown in Ortega (2010). In our static model, these effects will be very imperfectly captured through ε_j.

[5] See e.g. Azam and Laffont (2003) for a similar assumption.

We assume that each worker, high- or low-skilled, is endowed with one unit of raw labor. Moreover, high-skilled workers are endowed with one unit of human capital. As far as production is concerned, one physical good y_j is produced in each country using a constant return to scale technology $F_j(\cdot)$ combining high-skilled and low-skilled labor inputs, h_j and l_j respectively, as well as an unspecified fixed factor such as land or physical capital k_j. The choice of having two fixed factors is only for calibration purposes.

Notice that we rule out the possibility of capital movements across countries. Treating the number of high-skilled workers, total factor productivity and the amount of the fixed factor as exogenous, our static model focuses on the short- or medium-term effect of migration policies. We proceed that way for tractability reasons: assuming fixed capital spares us from modeling saving, foreign ownership of capital, external debt and so on. Nevertheless, we also believe that our assumptions are in line with the migration literature. The latter has scrutinized the effect of immigration on the returns to capital (Borjas, 1995; Ben-Gad, 2008) and, more recently, highlighted the effect of immigration on entrepreneurship and firm creation (di Giovanni et al., forthcoming). We are also in line with the literature on the labor market effects of immigration, which distinguishes between the short-run and long-run effects of immigration (Ottaviano and Peri, 2012). Finally, the long-run effect of immigration on capital movements is unclear when network effects operate. Although migration increases the marginal productivity of capital in destination countries, migration also reduces international transaction costs and facilitates capital outflows from receiving to sending countries (e.g. Kugler and Rapoport, 2007 or Javorcik et al., 2011) as well as technology transfers (e.g. Kerr, 2008; Agrawal et al., 2011). For the same reason, we do not account for trade.[6]

We treat the number of high-skilled workers, total factor productivity and the amount of the fixed factor as exogenous. We assume that the only variable input is the quantity of raw labor in efficiency units l_j:

$$y_j = F_j(k_j,\ h_j,\ l_j) = f_j(l_j)$$

where $f_j(\cdot)$ is increasing and concave, and the j subscript stands for differences in total factor productivity and factor shares. This specification implies that the productivity of a low-skilled workers depends crucially on the place where she works. A Mexican worker's low wage is related more to the place "Mexico" than to her competencies, as crossing the border multiplies her wage by four (see Clemens et al., 2008).

Raw labor used in the economy combines national raw labor and low-skilled immigrants through a function $l_j(n_j, m_j)$. This function is homogeneous of degree one and satisfies:

$$l_j(n_j, m_j) > 0, \qquad l'_{j,1}(n_j, m_j), l'_{j,2}(n_j, m_j) > 0, \qquad l''_{j,11}(n_j, m_j), l''_{j,22}(n_j, m_j) \le 0, \qquad l''_{j,12}(n_j, m_j) \ge 0.$$

The literature stresses that low-skilled immigrants and low-skilled natives are high substitutes (with an elasticity of substitution of about 20, see Section 3). To make the theoretical analysis simpler, we are going to prove some results under the assumption that they are perfect substitutes; these results remain however valid when the elasticity of substitution is high enough.

Assumption (A). $l'_{j,1}(n_j, m_j) = 1$. $l'_{j,2}(n_j, m_j) = \delta_j > 0$. $l''_{j,11}(n_j, m_j)$, $l''_{j,22}(n_j, m_j)$, $l''_{j,12}(n_j, m_j) = 0$.

The immigrants' productivity $l'_{j,2}(n_j, m_j)$ includes a discount δ_j to account for a possible lower efficiency because, for example, of language barriers.

The labor input in the poor country is made up of the citizens minus those who emigrate to each of the developed countries. Representing the number of low-skilled migrants from 0 to j by m_j, and the $j \times 1$ vector of the m_j's by $\mathbf{m}$, we have:

$$l_0 = l\big(n_0^l - \mathbf{m}'\, \mathbf{1_J}, 0\big), \tag{2}$$

where $\mathbf{1_J}$ is a column vector of ones of dimension J. We only model the equilibrium amount of low-skilled migration from 0 to j. Hence we disregard migration flows between developed countries (i.e. from j to $i \in J$) and the South–North brain drain.

High-skilled nationals receive the marginal product of their raw labor, $f'_j(l_j) l'_{j,1}(n_j, m_j)$, and some share $1 - \zeta_j$ of $s_j(l_j)$, the total income net of low-skilled labor wages:

$$s_j(l_j) = f_j(l_j) - f'_j(l_j) l_j.$$

This "surplus" includes human capital income, which is paid to high-skilled labor, and physical capital income, which is shared between natives in an exogenous way. The parameter ζ_j reflects both.

The government modifies the distribution of income by allocating a lump-sum net benefit to each low-skilled worker. Low-skilled nationals receive τ_j per person, while immigrants get $\tau_j \xi_j$. The consumption of nationals and immigrants in country j is given by:

[6] An analysis of the welfare consequences of migration with endogenous capital accumulation and trade can be found in Klein and Ventura (2007), and in di Giovanni et al. (forthcoming) respectively.

$$c_j^h = f_j'(l_j) l_{j,1}'(n_j, m_j) + \frac{(1-\zeta_j) s_j(l_j) - (n_j^l + \xi_j m_j)\tau_j}{n_j^h}, \tag{3}$$

$$c_j^l = f_j'(l_j) l_{j,1}'(n_j, m_j) + \frac{\zeta_j s_j(l_j)}{n_j^l} + \tau_j, \tag{4}$$

$$c_j^m = f_j'(l_j) l_{j,2}'(n_j, m_j) + \tau_j \xi_j. \tag{5}$$

It is reasonable to impose conditions on ζ_j such that $c_j^h > c_j^l$, at least when there is no redistribution ($\tau_j = 0$):

$$c_j^h > c_j^l \quad \Leftrightarrow \quad \frac{(1-\zeta_j) s_j(l_j)}{n_j^h} > \frac{\zeta_j s_j(l_j)}{n_j^l} \quad \Leftrightarrow \quad \frac{1-\zeta_j}{\zeta_j} > \frac{n_j^h}{n_j^l}.$$

The income for a low-skilled worker in the developing country (the low-skilled worker left behind) is equal to her marginal productivity:

$$c_0^l = f_0'\big(n_0^l - \mathbf{m}' \mathbf{1_J}\big) l_{0,1}'\big(n_0^l - \mathbf{m}' \mathbf{1_J}, 0\big), \tag{6}$$

which is a decreasing function of the quantity of labor l_0 left in developing countries. Eq. (6) will be modified later when aid and remittances are factored in.

We now define the level of m_j such that people in the developing country are indifferent between staying and leaving. Denote this level by $\breve{m}_j$. It is such that $c_0^l = c_j^m$ (assuming no migration cost). Such a level exists as c_0^l is increasing in m_j because of a decreasing return to labor and is unbounded, while c_j^m is decreasing in m_j. It allows to define the set of feasible m_j over which one can maximize the objective defined later.

Definition 1 *(Feasibility set).* Country's j migration level is feasible if it belongs to $[0, \breve{m}_j]$.

The immigration policy is determined through a vote by the citizen population. We assume a probabilistic voting set-up, which captures the idea that the political choice of individuals is not fully predictable: the probability that a given voter will vote for a party increases gradually as the party's platform becomes more attractive for her. As a consequence, each party's vote share varies continuously with the proposed policy platform, and probabilistic voting leads to the smooth aggregation of all voters' preferences, instead of depending solely on the median voter's preferences. With probabilistic voting, the outcome of the vote maximizes a weighted sum of utilities (this result was first derived by Coughlin and Nitzan, 1981):

$$\begin{aligned} W_j &\equiv \frac{n_j^h}{n_j} U_j^h + \frac{n_j^l}{n_j} \theta_j U_j^l \\ &= \frac{n_j^h}{n_j} u\big(c_j^h\big) + \frac{n_j^l}{n_j} \theta_j u\big(c_j^l\big) + \frac{n_j^h + \theta_j n_j^l}{n_j} \left(\beta u\big(c_0^l\big) - \varepsilon_j \left(\frac{m_j}{n_j}\right)^2 \right), \end{aligned} \tag{7}$$

where the parameter θ_j captures the extrinsic ratio of political power between low-skilled and college-educated nationals. It varies with each group's political participation and its response to the policy platform.

Definition 2 *(Country).* A country j is a sextuple $\Omega_j = \{n_j^h, n_j^l, \xi_j, \tau_j, \theta_j, \delta_j, \zeta_j\}$ representing the size and skill structure of the population, the intensity of redistribution, the degree of political participation, the productivity of immigrants relative to that of low-skilled nationals, the share of the surplus going to low-skilled natives, and functions $f_j(.)$, $l_j(.)$, which characterize the production technology.

2.2. Nationalist allocation

Only citizens have the right to vote and decide what is in their best interest. Remember that it is precisely in the context of elections with a large number of people, in which strategic voting is unlikely to occur, that maximin preferences are highly relevant (i.e. voters care about the worst-off). When voters in a given rich country decide on the migration level, they take the actions of other rich countries as given. Hence, the nationalist allocation is a "Nashionalist" equilibrium. Migrants are substitutable with low-skilled natives and generate a surplus mostly benefiting the highly skilled. The optimal immigration level depends on the political power of these two groups of nationals. Let us first define the value function of the political process (7) as a function of the migration vector by replacing c_j^h, c_j^l, c_0^l and l_0 by Eqs. (2), (3), (4) and (6):

$$\varphi_j(\mathbf{m}) = \frac{n_j^h}{n_j} u\left(f_j'(l_j) l_{j,1}'(n_j, m_j) + \frac{(1-\zeta_j)(f_j(l_j) - f_j'(l_j) l_j) - (n_j^l + \xi_j m_j)\tau_j}{n_j^h} \right)$$
$$+ \frac{n_j^l}{n_j}\theta_j\, u\left(f_j'(l_j) l_{j,1}'(n_j, m_j) + \frac{\zeta_j (f_j(l_j) - f_j'(l_j) l_j)}{n_j^l} + \tau_j \right)$$
$$+ \frac{n_j^h + \theta_j n_j^l}{n_j}\left(\beta u\left(f_0'\left(n_0^l - \mathbf{m}'\,\mathbf{1_J}\right) l_{0,1}'\left(n_0^l - \mathbf{m}'\,\mathbf{1_J}, 0\right)\right) - \varepsilon_j \left(\frac{m_j}{n_j}\right)^2 \right), \tag{8}$$

with $l_j = l_j(n_j, m_j)$. Note that $\partial\varphi_j(\mathbf{m})/\partial m_i > 0$, for $i \neq j$: any increase in the migration level to another country i increases the welfare of country j because of altruistic preferences.

The first-order condition for an interior maximum $\partial\varphi_j(\mathbf{m})/\partial m_j = 0$ defines an implicit function $\Psi_j(m_j, \Omega_j) = 0$, which can be decomposed as:

$$\Psi_j(.) = \frac{n_j^h}{n_j}\Psi_j^h(.) + \frac{n_j^l}{n_j}\theta_j \Psi_j^l(.) + \frac{n_j^h + \theta_j n_j^l}{n_j}\Psi_j^\beta(.) = 0, \tag{9}$$

where

$$\Psi_j^h(.) = u'\left(c_j^h\right)\left(f_j''(l_j) l_{j,2}'(n_j, m_j) l_{j,1}'(n_j, m_j) + f_j'(l_j) l_{j,12}''(n_j, m_j) - \frac{(1-\zeta_j) f_j''(l_j) l_{j,2}'(n_j, m_j) l_j + \xi_j \tau_j}{n_j^h} \right),$$

$$\Psi_j^l(.) = u'\left(c_j^l\right)\left(f_j''(l_j) l_{j,2}'(n_j, m_j) l_{j,1}'(n_j, m_j) + f_j'(l_j) l_{j,12}''(n_j, m_j) - \frac{\zeta_j f_j''(l_j) l_{j,2}'(n_j, m_j) l_j}{n_j^l} \right),$$

$$\Psi_j^\beta(.) = -\beta u'\left(c_0^l\right)\left(f_0''(l_0) l_{0,1}'(l_0, 0) + f'(l_0) l_{0,11}''(l_0)\right) - \frac{2\varepsilon_j m_j}{n_j}.$$

Assuming for a moment that the second-order condition for a maximum is verified, Eq. (9) describes the optimal number of migrants from the point of view of the receiving country, taking the choices by other countries as given. Immigration has several effects on the utility of citizens. At the nationalist allocation, these forces balance out at the margin:

- The term $\Psi_j^h(.)$ represents the income effect of the marginal migrant on high-skilled workers. Although increasing the labor force decreases the output per person (that is $f_j''(l_j) l_{j,2}'(n_j, m_j) l_{j,1}'(n_j, m_j) < 0$), it generates a higher income for college-educated nationals through the increase in profits ($-(1-\zeta_j) f_j''(l_j) l_{j,2}'(n_j, m_j) l_j > 0$). This second effect is weighted by $1 - \zeta_t$, the share of the surplus accruing to the high-skilled. The surplus is driven by the increased marginal productivity of the two fixed factors, physical and human capital.[7] In addition to these effects, low-skilled immigration reduces the income from raw labor ($f_j'(l_j) l_{j,12}''(n_j, m_j) \leq 0$) and increases the level of the redistributive tax ($-\xi_j \tau_j < 0$).
- The term $\Psi_j^l(.)$ represents the income effect of the marginal migrant on low-skilled workers. It has the same components as for the high-skilled. The weights attached to the effect on profits is now ζ_j. Borjas et al. (1997) identified a large and negative effect of immigration on low-skilled US citizens while Ottaviano and Peri (2012) and Manacorda et al. (2012) found limited effect due to complementarity between low-skilled immigrant and national workers.
- The term $\Psi_j^\beta(.)$ represents the non-economic part, including the altruistic part (β) and the other non-economic costs (ε_j). Welcoming an additional migrant raises the marginal productivity in the South and improves the well-being of the low-skilled left behind. If l_0 is large, this effect is likely to be small. For each developed country taken individually, the altruistic part plays a negligible part in determining the migration rate. On the contrary, the non-economic marginal cost is increasing linearly with m_j.

Let us now define a nationalist allocation:

Definition 3. A nationalist allocation is a feasible vector $\tilde{\mathbf{m}}$ such that the objective

$$\tilde{\Phi}_j(m_j) = \varphi_j\left([\tilde{m}_1..\tilde{m}_{j-1}, m_j, \tilde{m}_{j+1}.., \tilde{m}_J]'\right)$$

is maximized over $[0, \check{m}_j]$ with respect to m_j in each developed country $j = 1..J$ given the migration choices of other countries.

[7] There are debates about the size of the immigration surplus. Borjas (1995) provided rough estimates of the surplus for the United States. With one type of labor, he reached the pessimistic conclusion that a 10-percent increase in the workforce through migration affects citizens' income by about 0.1 percent. In a framework with two types of workers, the surplus varies with the educational structure of immigration. In addition, the size of the surplus is affected if immigration corrects distortions such as an excess level of home production created by income taxes (Kremer and Watt, 2009; Cortès and Tessada, 2011).

Such a maximizer $\tilde{m}_j$ can either be a corner ($\tilde{m}_j = 0$) or an interior solution. If it is an interior solution, it should satisfy: $\partial\tilde{\Phi}_j(\tilde{m}_j)/\partial\tilde{m}_j = 0$, $\partial^2\tilde{\Phi}_j(\tilde{m}_j)/\partial\tilde{m}_j^2 < 0$ and $\tilde{\Phi}_j(\tilde{m}_j) \geq \tilde{\Phi}_j(0)$ for all j.

The following condition on the characteristics of each country is sufficient to reject the corner solution $m_j = 0$ as a nationalist allocation. It defines two jointly determined thresholds $\bar{\theta}_j$ and $\bar{\xi}_j$, which can be read as an upper bound on the relative political power of low-skilled citizens θ_j together with an upper bound on the generosity of transfers ξ_j towards migrants.

Lemma 1 *(Interior solution). There exists thresholds*

$$\bar{\theta}_j = \frac{u'(c_j^h)}{u'(c_j^l)} < 1,$$

$$\bar{\xi}_j = \frac{f_j''(l(n_j,0))l'_{j,2}(n_j,0)}{\tau_j}\left(\left(u'\left(c_j^h\right)n_j^h + n_j^l\theta_j u'\left(c_j^l\right)\right)l'_{j,1}(n_j,0) - u'\left(c_j^h\right)(1-\zeta_j)l(n_j,0) - \theta_j u'\left(c_j^l\right)\zeta_j l_j\right),$$

where c_j^l and c_j^h are given by (3) and (4) with $m_j = 0$. For a given country j, $\theta_j < \bar{\theta}_j$ implies $\bar{\xi}_j > 0$. The corner solution $\tilde{m}_j = 0$ cannot be an optimum to country j's problem if $\theta_j < \bar{\theta}$ and $\xi_j < \bar{\xi}_j$.

Proof. See Appendix A.1. □

Evaluating the utility of each country taking as given the other countries' nationalist immigration level, we now derive the conditions under which there exists a level of migration $m_j^{\max} > \tilde{m}_j$ belonging to the feasibility set $[0, \breve{m}_j]$, such that country j is indifferent between taking no immigrants or a level $m_j^{\max}$ of immigration. Define $\breve{\varepsilon}$ the value of ε such that $\tilde{\Phi}_j(\breve{m}_j) = \tilde{\Phi}_j(0)$. $\tilde{\Phi}_j(\breve{m}_j)$ is evaluated at the vector of the nationalist allocation $\tilde{\mathbf{m}}$ except for country j where $\tilde{m}_j$ is replaced by $\breve{m}_j$. This value of ε exists as $\tilde{\Phi}(.)$ is linear in ε. Now, we have:

Lemma 2. *Under the condition of Lemma 1, there exists a threshold $\breve{\varepsilon}$ such that $\tilde{\Phi}_j(\breve{m}_j) = \tilde{\Phi}_j(0)$. If $\varepsilon > \breve{\varepsilon}$, there exists a migration level $m_j^{\max} < \breve{m}_j$ solving $\tilde{\Phi}_j(m_j^{\max}) = \tilde{\Phi}_j(0)$. This migration level is such that $\tilde{m}_j \in]0, m_j^{\max}[$.*

Proof. See Appendix A.2. □

Lemma 2 defines a number $m_j^{\max}$, depending solely on the parameters, which can be used as the upper bound on the optimal m_j.

The following proposition highlights the conflict of interest that arises between high- and low-skilled nationals when they have to decide on immigration policy.

Proposition 1. *Assume (A) and $f'''(x)x + f''(x) > 0$ for all $x > 0$.*[8] *When the nationalist allocation is an interior maximum for country j, there exists a low-skilled share of surplus*

$$\hat{\zeta} = \frac{n_j^l}{n_j + m_j^{\max}}$$

and a generosity of redistribution towards migrants

$$\hat{\xi} = \frac{-f_j''(n_j + m_j^{\max})\delta_j(n_j + m_j^{\max})\left(1 - \frac{\zeta_j n_j^h}{(1-\zeta_j)n_j^l}\right)}{\tau_j}$$

such that, for $\zeta < \hat{\zeta}$ and $\xi < \hat{\xi}$, each high-skilled national prefers to increase the level of immigration, whereas each low-skilled national prefers to decrease it.

Proof. See Appendix A.3. □

The intuition in this proposition is the following: at the nationalist allocation, the high-skilled prefer to increase migration, except when their share of the surplus is too low ($\zeta > \hat{\zeta}$), or when redistribution is high ($\xi_j > \hat{\xi}_j$). Fig. 1 illustrates

[8] This condition holds when $F(\cdot)$ is a Cobb–Douglas function or a CES with a high elasticity of substitution.

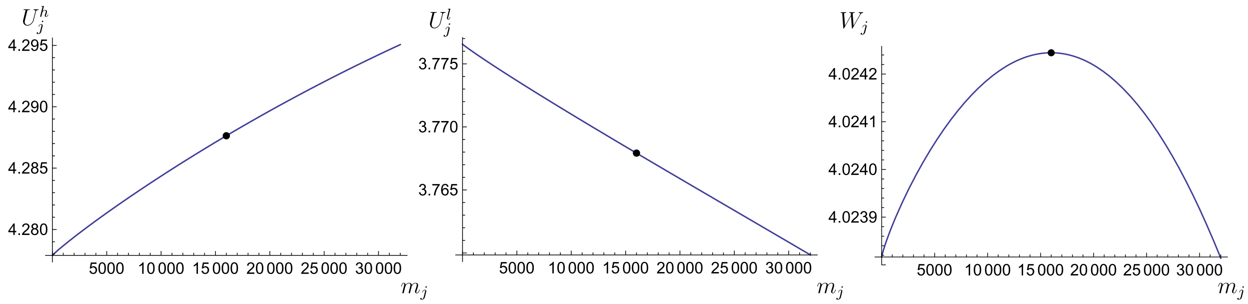


Fig. 1. Utility as a function of immigration.

Proposition 1, borrowing calibrated parameters for the United States from the next section. The left panel shows the utility of the high-skilled. The dot indicates the outcome of probabilistic voting. In the example, the high-skilled would like to allow more migrants to enter the country than what is effectively chosen. Their utility reaches a maximum for even higher numbers than what is displayed on the graph. The low-skilled, on the contrary, have a strongly decreasing utility in migration. W_j is the objective function that is maximized by the vote.

Henceforth, the conditions necessary for Proposition 1 are assumed to hold. The next proposition focuses on two key economic determinants of migration policy: the distribution of political power among nationals θ_j, and the strength of the redistribution implied by the transfer system ξ_j.

Proposition 2. *The immigration level $\tilde{m}_j$ is decreasing in the political power of the low-skilled θ_j. It is decreasing in redistribution ξ_j if $\zeta_j > l_j(n_j, \tilde{m}_j)/n_j^l$, or if the utility function is linear. The sign of the effect of ξ_j is ambiguous otherwise.*

Proof. See Appendix A.4. □

The interpretation is the following. First, countries with high θ_j will unsurprisingly be more reluctant to welcome immigrants as low-skilled nationals prefer to have fewer migrants. Second, the impact of redistribution is ambiguous. On the one hand, redistribution discourages high-skilled nationals from hosting migrants because they will have to pay more taxes to finance social benefits. On the other hand, redistribution reduces (resp. increases) the consumption of high-skilled (resp. low-skilled) nationals and increases (resp. reduces) their marginal utility. This tends to give more incentives to high-skilled nationals to increase immigration and fewer incentives to low-skilled nationals to combat immigration.

2.3. No-regret allocation

Due to altruism, the nationalist allocation of labor is subject to a prisoner's dilemma and leaves all countries worse off than in a cooperative equilibrium. An international agency which maximizes a social welfare function aggregating the utility levels of rich countries' citizens could reach a better allocation. As will be shown in Section 4, such a utilitarian allocation is detrimental to some rich countries. Here, we consider an alternative allocation which leaves host countries indifferent, compared with the nationalist scenario, and maximizes the volume of South–North low-skilled migration. In other words, the efficiency gains from coordinating immigration policies are totally used to improve the welfare of the worst-off in the world (i.e. low-skilled workers living in developing countries) subject to implementation constraints. If such an allocation existed, it would improve the level of income in developing countries (given Eq. (6)) and improve the utility of new migrants. Our no-regret allocation can be defined as follows[9]:

Definition 4. A no-regret allocation is a vector $\bar{\mathbf{m}}$ such that $\mathbf{m}'\,\mathbf{1_J}$ is maximized subject to:

$$\varphi_j(\tilde{\mathbf{m}}) \leq \varphi_j(\bar{\mathbf{m}}) \quad \text{for } j = 1..J. \tag{10}$$

Since utility in the South decreases with l_0, maximizing $\mathbf{m}'\,\mathbf{1_J}$ is identical to maximizing the average utility of low-skilled residents of developing countries. Inequality (10) is the participation constraint.

If citizens in the North had to make a choice between the nationalist allocation and the no-regret allocation, no country would reject the no-regret allocation. Considering the various types of households, high-skilled citizens would be better off with the no-regret allocation, and the low-skilled would be worse off.[10] Obviously, each rich country might be tempted to

[9] We implicitly assume that the supply of migrants is very large, and is not a binding constraint.

[10] It is however possible to design a transfer scheme such that no citizen is worse off.

free ride, i.e. let the others sign the no-regret treaty. Financial or political sanctions (e.g. exclusion from the GATT/WTO) could be implemented to avoid such free-riding practices.

Assuming that the solution $\bar{\mathbf{m}} > 0$, and denoting by λ_j the Lagrange multipliers associated with constraint (10) in country j, the first-order conditions associated with the maximization problem are:

$$1 + \lambda_j \frac{\partial \varphi_j(\bar{\mathbf{m}})}{\partial m_j} + \sum_{k \neq j} \lambda_k \frac{\partial \varphi_k(\bar{\mathbf{m}})}{\partial m_j} = 0 \tag{11}$$

$$\lambda_j\big(\varphi_j(\bar{\mathbf{m}}) - \varphi_j(\tilde{\mathbf{m}})\big) = 0, \qquad \lambda_j \geq 0, \qquad \varphi_j(\bar{\mathbf{m}}) \geq \varphi_j(\tilde{\mathbf{m}}),$$

for $j = 1..J$, where $\partial\varphi_j(\mathbf{m})/\partial m_j$ is given in Eq. (9) and $\partial\varphi_k(\mathbf{m})/\partial m_j$ is given by:

$$\frac{\partial \varphi_k(\bar{\mathbf{m}})}{\partial m_j} = -\beta \frac{n_k^h + \theta_k n_k^l}{n_k} u'\big(c_0^l\big) f_0''(l_0) l'_{0,1}\big(n_0^l - \bar{\mathbf{m}}'\, \mathbf{1_J}, 0\big).$$

The summation term in Eq. (11) reflects the fact that the externality is now internalized. The derivative $\partial\varphi_k(\mathbf{m})/\partial m_j$ is the positive effect on the utility in country k when country j welcomes one additional migrant.

Considering only situations in which the participation constraints are binding, i.e. $\lambda_j > 0$ for all j, and rearranging Eq. (11), leads to the characterization of the no-regret allocation as a set of vectors $(\lambda_j, \bar{c}_j^h, \bar{c}_j^l, \bar{m}_j)$ for $j = 1..J$ and a pair $(\bar{l}_0, \bar{c}_0^l)$ satisfying:

$$\frac{1}{\lambda_j} = \frac{\partial \varphi_j(\bar{\mathbf{m}})}{\partial m_j} - \beta u'\big(\bar{c}_0^l\big) f_0''\big(\big(n_0^l - \bar{\mathbf{m}}'\, \mathbf{1_J}, 0\big)\big) \sum_{k \neq j} \frac{\lambda_k}{\lambda_j} \frac{n_k^h + \theta_k n_k^l}{n_k} \tag{12}$$

$$\bar{c}_j^h = f_j'\big(l_j(n_j, \bar{m}_j)\big) l'_{j,1}(n_j, \bar{m}_j) + \frac{(1-\zeta_j)s_j(l_j(n_j, \bar{m}_j)) - (n_j^l + \xi_j \bar{m}_j)\tau_j}{n_j^h} \tag{13}$$

$$\bar{c}_j^l = f_j'\big(l_j(n_j, \bar{m}_j)\big) l'_{j,1}(n_j, \bar{m}_j) + \frac{\zeta_j s_j(l_j(n_j, \bar{m}_j))}{n_j^l} + \tau_j \tag{14}$$

$$0 = \varphi_j(\tilde{\mathbf{m}}) - \varphi_j(\bar{\mathbf{m}}) \tag{15}$$

$$\bar{c}_0^l = f_0'\big(l_0\big(n_0 - \bar{\mathbf{m}}'\, \mathbf{1_J}, 0\big)\big). \tag{16}$$

When there is no altruism, $\partial\varphi_k(\mathbf{m})/\partial m_j$ equals zero in Eq. (11) and the public good nature of South–North migration vanishes. There is no prisoner's dilemma and it is impossible to improve the situation of developing countries without reducing welfare in the rich world. Without altruism, the no-regret allocation is identical to the nationalist allocation. With altruism, Eq. (11) implies that the coordinated value for $\partial\varphi_j(\mathbf{m})/\partial m_j$ becomes negative, i.e. each rich country goes beyond its nationalist equilibrium (in which $\partial\varphi_j(\mathbf{m})/\partial m_j$ equals zero). The set of participation constraints (10) determines the extent to which immigration can be increased. This yields the following proposition:

Proposition 3. *In the absence of altruism, the no-regret allocation coincides with the nationalist allocation, i.e.*

$$\beta = 0 \quad \Rightarrow \quad \bar{\mathbf{m}} = \tilde{\mathbf{m}}.$$

Proof. For country j, Eq. (15) implies $\varphi_j(\bar{\mathbf{m}}) = \varphi_j(\tilde{\mathbf{m}})$. If $\beta = 0$, φ_j depends on m_j, but not on the other elements of the vector $\mathbf{m}$. Since, at $\tilde{m}_j$, φ_j is maximized, $\bar{m}_j = \tilde{m}_j$. □

Let us now consider the case of symmetric countries, for which we define

$$\Phi(m) = \varphi_j\big([m_j..m_j, m_j, m_j.., m_j]'\big) = \varphi(m\mathbf{1_J}),$$

where m is a scalar.

Proposition 4. *Assume all rich countries are identical and $\beta > 0$. Then the symmetric allocations $\tilde{m}$ and $\bar{m}$ satisfy $\bar{m} > \tilde{m}$.*

Proof. See Appendix A.5. □

When countries are different, it is not sure that $\forall j : \bar{m}_j > \tilde{m}_j$. However, in the quantitative exercise below, it is still the case.

2.4. The decentralization scheme

One natural way to implement the no-regret allocation is to allow a supranational agency to manage the immigration policies of all member countries. The rationales are multiple: multilateral agencies are politically more stable than countries; they are better positioned to collect information about recipient countries and monitor their policies (see Rodrik, 1995); delegating policies might also be desirable when public opinion is skeptical about the benefits; and relying on multilateral organizations then reassures voters that their government is making good decisions (Milner, 2006). Another way to decentralize the no-regret allocation is to leave the management of immigration policies to member states, but allow an international agency to design a tax-subsidy scheme. This scheme requires a group of rich countries to agree to put a significant amount of money into a global migration fund and to delegate the responsibility of managing it to an international agency. This agency would have a constitution setting out how money will be paid out according to changes in immigration from developing countries. As in Gersbach and Winkler (2007)'s global refunding system to cope with CO_2 emissions, no further coordination is required, except in administering the system, measuring labor flows, and distributing money. Such a system does not need additional enforcement mechanisms. Rich countries will choose to accept more migrants on their own, and they will not have incentives to go back to the nationalist equilibrium.

More precisely, the set of developed countries can establish a migration agency to organize the tools needed to decentralize the cooperative solution. Each country would pay a lump sum tax to the agency which we assume is levied on high-skilled natives only. The vector of lump-sum taxes is denoted by $\mathbf{q}$. Each country would also enjoy a subsidy per additional migrant welcomed, which is distributed among high-skilled natives. The subsidy rates are country-specific and their vector is $\mathbf{p}$. In other words, each college graduate citizen from country j is allowed to avoid paying the lump-sum tax q_j/n_j^h by welcoming immigrants and getting a subsidy p_j/n_j^h for each of them. Her consumption level becomes:

$$c_j^h = f_j'(l_j)l_{j,1}'(n_j,m_j) + \frac{(1-\zeta_j)s_j(l_j) - (n_j^l + \xi_j m_j)\tau_j + p_j m_j - q_j}{n_j^h}.$$

The global migration fund would only involve migrants originating from developing countries. These immigrants would get a visa only allowing them to reside in host country j, not to move from j to another participating country. Hence, the decentralization problem of the agency is to define the appropriate set of taxes and subsidies such that each country j agrees to host the optimal number of immigrants $\tilde{m}_j$, and such that the agency's global budget constraint is balanced:

$$\bar{\mathbf{p}}'\mathbf{m} - \bar{\mathbf{q}}'\mathbf{1_J} = 0. \tag{17}$$

Introducing the subsidy scheme, the objective for an individual country becomes:

$$\begin{aligned}\hat{\varphi}_j(\mathbf{m}, p_j, q_j) &= \frac{n_j^h}{n_j} u\left(f_j'(l_j)l_{j,1}'(n_j,m_j) + \frac{(1-\zeta_j)s_j(l_j) - (n_j^l + \xi_j m_j)\tau_j + p_j m_j - q_j}{n_j^h}\right) \\ &\quad + \frac{n_j^l}{n_j}\theta_j\, u\left(f_j'(l_j)l_{j,1}'(n_j,m_j) + \frac{\zeta_j(f_j(l_j) - f_j'(l_j)l_j)}{n_j^l} + \tau_j\right) \\ &\quad + \frac{n_j^h + \theta_j n_j^l}{n_j}\left(\beta u\big(f_0'\big(n_0^l - \mathbf{m}'\,\mathbf{1_J}\big)l_{0,1}'\big(n_0^l - \mathbf{m}', 0\big)\big) - \varepsilon_j\left(\frac{m_j}{n_j}\right)^2\right),\end{aligned} \tag{18}$$

with $\hat{\varphi}_j(\mathbf{m}, 0, 0) = \varphi_j(\mathbf{m})$. This decentralization problem can be written as follows:

Proposition 5. *The no-regret allocation $\bar{\mathbf{m}}$ can be decentralized with subsidy rates $\bar{\mathbf{p}}$ and lump-sum taxes $\bar{\mathbf{q}}$ satisfying $\bar{p}_j\bar{m}_j = \bar{q}_j\ \forall j$, and*

$$-\frac{p_j n_j^h}{n_j^2}u'\big(\bar{c}_j^h\big) - \beta u'\big(\bar{c}_0^l\big) f_0''(\dot{l}_0) \sum_{k\neq j} \frac{\lambda_k}{\lambda_j}\frac{n_k^h + \theta_k n_k^l}{n_k} = \frac{1}{\lambda_j} \quad \text{for } j = 1..J, \tag{19}$$

where the multipliers λ's are the solution to Eqs. (11).

Proof. The first-order condition for an interior maximum is:

$$\begin{aligned}\frac{\partial\hat{\varphi}_j(\mathbf{m}, p_j, q_j)}{\partial m_j} &= \frac{n_j^h}{n_j}u'\big(c_j^h\big)\bigg(f_j''(l_j)l_{j,2}'(n_j,m_j)l_{j,1}'(n_j,m_j) + f_j'(l_j)l_{j,12}''(n_j,m_j) \\ &\quad - \frac{(1-\zeta_j)f_j''(l_j)l_{j,2}'(n_j,m_j)l_j + \xi_j\tau_j - p_j}{n_j^h}\bigg) + \frac{n_j^l}{n_j}\theta_j \Psi_j^l(.) + \frac{n_j^h + \theta_j n_j^l}{n_j}\Psi_j^\beta(.) \\ &= 0.\end{aligned} \tag{20}$$

Comparing (20) with (9), we observe that:

$$\frac{\partial\hat{\varphi}_j(\mathbf{m}, p_j, q_j)}{\partial m_j} = \frac{\partial\varphi_j(\mathbf{m})}{\partial m_j} + \frac{p_j n_j^h}{n_j^2} u'\left(\bar{c}_j^h\right) = 0.$$

Replacing l_j by $\bar{l}_j$ (and the corresponding $\bar{l}_0$, $\bar{c}_j$ and $\bar{c}_0$ from Eqs. (14)–(16)) and using Eq. (12), we obtain Eq. (19). This condition determines the subsidy rate such that each country will choose the no-regret level of migration. In addition, we also need to satisfy the participation constraint (15). Doing so requires $\bar{q}_j = \bar{m}_j/p_j$, as $\hat{\varphi}_j(\bar{\mathbf{m}}, p_j, \bar{m}_j/p_j) = \varphi_j(\bar{\mathbf{m}}) = \varphi_j(\tilde{\mathbf{m}})$. As we have $\bar{p}_j\bar{m}_j = \bar{q}_j\ \forall j$, the global budget constraint (17) is automatically satisfied. □

Our problem is not the same as the usual one in the environmental literature in which a planner targets a global level of pollution and decentralizes the optimal policy, either through a tax on pollutants, or through tradable quotas. In our case, the fact that we take the participation constraints (15) explicitly into account prevents the decentralization of the policy through a uniform subsidy scheme. Indeed, Proposition 5 shows that, in the case of asymmetric host countries, it is not possible to decentralize the no-regret allocation with a uniform subsidy p (or alternatively with a system of tradable immigration duties in which each country would sell/buy migrants at a market price p). The constraints (15) require the use of additional instruments, such as the differentiated subsidy rates p_j.

The same participation constraint explains why there are no monetary transfers between countries at the no-regret solution. Compared with the nationalist allocation, Eq. (15) implies that citizens enjoy the same level of utility by bearing more domestic immigration costs and benefiting from a lower level of extreme poverty. In a command optimum, a benevolent planner could reach this solution by imposing a larger number of migrants on each host country, with no need to transfer money between them. Hence, in a market context, we distort immigration decisions by subsidizing immigrants, but we need to avoid additional income effects. This can be done by levying a lump-sum tax on destination countries and allowing them to retrieve the amount paid to the global fund by hosting more migrants.

3. Quantitative assessment

In this section, we simulate the effect of the no-regret treaty on the number of South–North migrants. We consider that the initial nationalist allocation of labor is the allocation observed in the year 2000. The outcome of our treaty depends on the number of countries, country-specific parameters, and the global parameter of aversion to extreme poverty. Country-specific parameters are calibrated so as to match data on income per capita, inequality and redistribution, the size and skill composition of the labor force, and the number of immigrants originating from developing countries. The key parameter of altruism, β, is fixed in line with US data on donations to charities protecting people in a precarious situation. In the next two sub-sections, we describe our data sources and explain how country-specific and global parameters are identified. Then, we simulate the no-regret treaty and its decentralization; the quantitative analysis is first conducted in the theoretical framework of Section 2, without remittances and development aid. These extensions will be introduced in Section 4.

3.1. Data

The aggregate developing country ($j = 0$) includes all developing countries as defined in the World Bank classification.[11] For the J developed countries, we considered two different possibilities: the G7 countries (Canada, France, Germany, Italy, Japan, the United Kingdom and the United States), and 24 high-income members of the OECD (Organization for Economic Co-operation and Development).[12] For these $J + 1$ countries, Appendix B presents the data used to calibrate the model. The data were obtained from the following sources.

The data on the size of the native population, $N_j = n_j + m_j$, were obtained from the United Nations Population Division database; we only consider individuals aged 25 and over as a proxy for the labor force. As shown in Column 1, we identify a total of 630 million adults in the set of OECD, high-income countries, and 2491 million adults in developing countries.

The data on the proportion of high-skilled workers in the labor force, n_j^h/N_j, were obtained from Barro and Lee (2010) and are reported in Column 2. We use the proportion of individuals aged 25 and over who completed college. The average proportion amounts to 18.4 percent in rich countries and ranges from 3.9 percent in Portugal to 30.6 percent in the United States. It amounts to 3.3 percent in developing countries.

The data on m_j (the number of working-age immigrants from developing countries) were obtained from Parsons et al. (2007). They provide a 226 × 226 matrix of migration stocks by origin and destination. The data are generated by disaggregating the information on migrant stock in each destination country or economy as given in its census. The reference period is the 2000 round of population censuses.[13] This allows us to identify immigrants originating from developing countries.

[11] Available from http://data.worldbank.org/about/country-classifications/.

[12] We have excluded four OECD members that joined the OECD after 2000 (Chile, Estonia, Israel and Slovenia) and three members that were not considered as high-income countries in 2000 (Czech and Slovak Republics and Poland).

[13] Four versions of the database are available, at increasing levels of completeness, but decreasing levels of accuracy as the missing data are added via assumptions and interpolation with each successive version. We use the most comprehensive bilateral matrix (version 4).

Since the database does not provide information on the number who are of working age, we assume that the proportion of working-age migrants is the same as that of citizens in the origin country. In some receiving countries, the national census accounts for undocumented immigrants (e.g. the US and, to a lesser extent, other settlement countries). It is not the case in Europe. The "clandestino" study reveals that illegals represent about ten percent of the total immigration stock and an uncertain proportion of them originate from developing countries. For this reason, we ignore illegals that are not captured in the national census. As shown in Column 3, the largest immigrant-to-population ratios were observed in Canada and Switzerland (0.101). On the contrary, the ratio was particularly low in Korea (0.002) and Japan (0.005).

In addition, the migration data set does not distinguish low-skilled and high-skilled migrants. We assumed that South–North migrants are low skilled, either because they are indeed, or moving from one country to the other implies losing one's human capital endowment. As far as receiving countries are concerned, the data show that about one third of OECD immigrants from low-income countries have a college education (see Docquier et al., 2009). However many of them tend to find jobs in occupations which are typically staffed by low-skilled natives, indicating that their actual "human capital" does not correspond to their education level (see Dustmann et al., 2008 for the UK, and Mattoo et al., 2008 for the US). Mattoo et al. (2008) computed the probability of an immigrant from country 0 with a home-country college education obtaining a skilled job in the US labor market. The proportion is always below 50 percent for developing countries, and below 30 percent for the majority of them. Hence, the actual share of high-skilled migrants from poor countries is around 10 percent only. As far as sending countries are concerned, low-skilled workers represent more than 95 percent of the population. It is true that a brain drain can affect income (see e.g. de la Croix and Docquier, 2012), but the effect is ambiguous: recent theoretical and empirical literatures show that high-skilled emigration need not deplete a country's human capital stock and can generate positive network externalities (Docquier and Rapoport, 2012). In addition, if a brain drain is at work, it will not be affected by our no-regret treaty which exclusively targets low-skilled migrants. In sum, m_j denotes all South–North migrants assuming that they are employed as low-skilled workers in rich countries, and n_j is proxied by the remaining population (including North-North migrants).

To identify the parameters of the production function, we use data on the average return per year of schooling ρ_j, obtained from Hendricks (2004) and reported in Column 4. The data are not available for 3 host countries in our sample: we use the Dutch rate for Belgium and Luxembourg, and the Swedish rate for Iceland. To compute the wage ratio between high-skilled and less educated workers, ω_j^h, we combine Hendricks's data with data from Barro and Lee on the average years of schooling by educational level. Barro and Lee's data give us the difference in average years of schooling between the high-skilled and low-skilled g_j (see Column 5).

The data on GDP in purchasing power parity (PPP) in 2000 were obtained from the World Development Indicators (World Bank, 2008) and are reported in Column 6. For GDP per worker, we obtain a weighted average level of USD 36,840 in rich countries, against USD 7070 in developing countries.

Immigrants and natives have different average productivity levels for two reasons. First, for a given level of education, they might have a lower productivity because of lower quality of education. Second, they are employed in different occupations and have comparative advantages and disadvantages with respect to natives on the labor market in the host country. We will explain how we address the second source of heterogeneity in the next sub-section. Parameter δ_j was included in the model to capture the first source of heterogeneity. To identify δ_j, we use the estimates from Coulombe and Tremblay (2009). This study compares the skill intensity and schooling level of Canadian immigrants and natives who were both subjected to standardized tests in literacy, math, and problem solving. These tests provide measures of proficiency that are comparable across countries and over time. They estimate a 'skill-schooling gap' for each origin country, defined as the difference between the mean years of schooling of the immigrant subgroup and the typical native with the same proficiency level. For example, a skill-schooling gap of x years means that Canadian nationals are as productive as immigrants with x more years of schooling. Assuming that one year of schooling generates a productivity gain of 8 percent, we estimate the relative productivity of native workers in each developed country j (compared to a Canadian worker) and of immigrants from the developing world (a weighted average of immigrants from developing source countries to country j). In the production function, we multiply the number of South–North immigrants, m_j, by δ_j, the ratio ability between immigrants and nationals. The ratio ranges from 0.60 in Luxembourg to 0.88 in Korea, and has a mean value of 0.74 (see Column 7).

We use data on the ratio of Gini indexes before and after taxes/transfers (denoted by R_j) to calibrate the size of redistributive transfers, τ_j. The data are taken from Immervoll and Richardson (2011) and are available for 18 countries (see Column 8). A simple linear regression shows that the ratio is a decreasing function of the share of social expenditures in GDP. For the remaining rich countries, we predict the ratio of Gini indexes using the estimated coefficients. The average ratio amounts to 0.83 in developing countries and 0.74 in rich countries, ranging from 0.613 in Belgium to 0.871 in Switzerland.

Data on the income share of natives and immigrants are also used to calibrate the preference parameter for immigrants in the technology. We use the wage shares of low-skilled natives, shl_j, and immigrants, shm_j, provided in Docquier et al. (2014): Table A4 of their online appendix). Combining these income shares with the numbers of immigrants and low-skilled natives, we obtain the wage ratio between natives and immigrants, $\omega_j^m = w_j^l/w_j^m$. Indeed, the ratio of wage shares is equal to $shm_j/shl_j = \omega_j^m n_j/m_j$, so that $\omega_j^m = (shm_j m_j)/(shl_j n_j)$. As shown in Column 9, the average wage ratio amounts to 1.05 in rich countries, ranging from 0.86 in Canada to 1.35 in Iceland.

Data on ξ_j, the relative amount of welfare benefits received by immigrants as the percentage of low-skilled natives, are used to calibrate the fiscal impact of immigration. The data are available for 13 European countries in

Barbone et al. (2009). For the other countries, we predict ξ_j using an OLS regression on the share of redistributive public spending in GDP. As a proxy for redistributive spending, we aggregate social security contributions by employees, employers, and self-employed individuals (Source: IMF), public education spending (Source: UNESCO), and public health spending (Source: WHO). On average, the amount of welfare benefits received by immigrants is 52 percent lower than the amount received by low-skilled natives (see Column 10).

Our income indicators are expressed in PPP units. Hence, when monetary transfers occur between country j and country 0, such as remittances and development aid, we need to account for real exchange rates (γ_j), capturing cross-country heterogeneity in prices. We first compute the exchange rate of each host country vis-a-vis the United States by dividing its GDP per capita in PPP value by its level in current dollar. We obtain an average of $\gamma_{US} = 2.32$. We then compare this value to the exchange rate computed for other rich countries vis-a-vis the United States. For example, one dollar earned in Denmark (resp. Portugal) has a purchasing value of USD 1.4 (resp. 0.85) in the United States, and a value of USD 3.32 (resp. 1.98) in the developing country. Real exchange rates γ_j are reported in Column 11.

3.2. Calibration

We now select specific functional forms and assign values to the parameters in order to match a series of targets. The calibrated country-specific parameters are reported in Appendix C.

As far as technology is concerned, we use a Cobb–Douglas production function with constant return to scale:

$$y_j = \overline{\mu}_j k_j^{\alpha_j^k} h_j^{\alpha_j^h} l_j^{\alpha_j^l}, \qquad \alpha_j^k + \alpha_j^h + \alpha_j^l = 1 \tag{21}$$

where $\overline{\mu}_j$ denotes total factor productivity, α_j^k, α_j^h and α_j^l are the elasticities of output to physical capital, human capital and raw labor, respectively.

The variable l_j stands for the quantity of raw labor used in the economy. It combines national labor and immigrants through a nested CES function. Immigrants are included with a discount δ_j to account for possible lower productivity because, for example, of language barriers:

$$l_j = \left(\nu_j (n_j)^{\frac{\sigma-1}{\sigma}} + (1-\nu_j)(\delta_j m_j)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{22}$$

where σ is the elasticity of substitution between citizens' and immigrants' labor, and ν_j is a share parameter of the production function. We assume there are no immigrants in country 0. Hence, labor input in the poor country is made up of the citizens minus those who emigrate to each of the developed countries:

$$l_0 = \nu_0^{\frac{\sigma}{\sigma-1}} \left(n_0^l - \mathbf{m}' \mathbf{1_J}\right).$$

Given this production function, the before-tax wage ratios between low-skilled natives and immigrants (ω_j^m), and between college graduates and low-skilled natives (ω_j^h) are equal to

$$\omega_j^m = \frac{\nu_j}{1-\nu_j} \frac{1}{\delta_j} \left(\frac{n_j}{\delta_j m_j}\right)^{-\frac{1}{\sigma}} \tag{23}$$

$$\omega_j^h = 1 + \frac{\alpha_j^h}{\alpha_j^l} \frac{l_j}{h_j} \frac{1}{\nu_j} \left(\frac{l_j}{n_j}\right)^{-\frac{1}{\sigma}}. \tag{24}$$

Two structural parameters are taken from the empirical literature, σ and α_j^k. The elasticity of substitution between natives and immigrants, σ, has been the focus of several recent papers and has generated a certain level of debate. Borjas et al. (2012), Peri (2011), and Ottaviano and Peri (2012) use US data and Manacorda et al. (2012) use UK data in their estimations. The first study finds a value of infinity; the second and third papers estimate an elasticity between 10 and 20, and the paper on UK data finds a value around 6. We use the intermediate value, $\sigma = 20$. We assume a constant elasticity of output to physical capital, α_j^k, equal to 0.365 in all countries.

Second, we use country-specific data on wage shares by education level and country of origin to compute the wage ratio between low-skilled natives and immigrants (ω_j^m). Using data on n_j, m_j and δ_j, we identify ν_j as a residual from (23). Once ν_j is identified (see Column 2), we can compute l_j using (22).

Third, we compute the wage ratio between college graduates and less educated natives as $\omega_j^h = (1+\rho_j)^{g_j}$, where ρ_j is the Mincerian return of one additional year of schooling and g_j is the schooling gap between the high-skilled and low-skilled. As shown in Column 1, the average wage ratio amounts to 2.0 in OECD, high-income countries (a mean return to schooling of 8 percent) and to 3.5 in developing countries (12 percent per year of schooling). Subject to the constraint $\alpha_j^l + \alpha_j^h = 1 - \alpha_j^k$, α_j^l and α_j^h can be identified from (24). We obtain an average value of 0.56 for α_j^l, and a value around 0.5 for the most advanced countries (see Column 3).

Fourth, we use (21) and calibrate $\mu_j = \overline{\mu}_j k_j^{\alpha_j^k}$ to match cross-country differences in GDP in the developed world. Hence, μ_j captures differences in total factor productivity and the other potential fixed factors (such as physical capital) each country is endowed with. The calibrated values for μ_j for each country are given in Column 4. Technological parameters are thus perfectly compatible with country-specific data on GDP per capita, labor and income shares in host and sending countries.

Fifth, we calibrate the share ζ_j of the surplus $s_j(l_j)$ received by the low-skilled. Remember that the surplus remunerates two fixed factors, human and physical capital. We assume that the return to human capital goes entirely to high-skilled natives whereas the return to physical capital is shared proportionally to wages. This implies:

$$\zeta_j s_j(l_j) = \alpha_j^k y_j \frac{n_j^l}{n_j^l + \omega_j^h n_j^h}. \tag{25}$$

The rest of the surplus, $(1-\zeta_j)s_j(l_j)$, goes to the high-skilled. In Column 5, we report $\hat{\zeta}_j = \zeta_j/(1-\zeta_j) \times (n_j^h/n_j^l)$, which measures the ratio of surplus per capita between low-skilled and high-skilled natives. We obtain an average ratio of 0.3. The ratio varies between 0.118 in Portugal and 0.435 in Norway; unsurprisingly, it is highly correlated with the skill premium (correlation rate of −0.91).

As far as redistributive transfers are concerned, we use the data on the composition of the population (n_j^l, n_j^h) and compute the Gini index of each economy. We then calibrate the level of the lump-sum net benefit τ_j paid by college graduates to each low-skilled worker so as to match the data on the ratio of Gini indexes before and after taxes/transfers. The results for τ_j are given in Column 6.

The calibration of the common altruism factor β is based on US data. According to the US Census Bureau, the percentage of US citizens living in extreme poverty (i.e. with less than \$2 per day or \$730 per year) was equal to 1.5 percent in 2012. One sixth of them were homeless and one third used an emergency shelter or a housing program. We can reasonably estimate that extremely poor people live with at most \$700 per year (i.e. 0.8 percent of the average income per American worker). In its annual report on Philanthropy for 2013, Giving USA (2013) estimated that the average American gives 2 percent of her income to charities. Two thirds of giving go to Religion, Education, Arts, Environment, Foundations or International Affairs; one third goes to organizations protecting individuals in a precarious situation (through health programs, human services, public society benefits, and giving to individuals). Hence, we estimate that Americans give 0.66 percent of their income to fight extreme poverty. With the logarithmic utility function and assuming that there are about 33 active Americans per person living in extreme poverty (the labor force represents 50 percent of the US population and the population living in extreme poverty represents 1.5 percent), a t percent gift to charities should maximize $\ln(y - ty) + \beta \ln(0.008y + 33ty)$. This is the case if $\beta \simeq 0.007$. In a previous version of this paper (de la Croix and Docquier, 2010), we calibrated β using insights from the experimental literature on social preferences. We equalized the marginal rate of substitution between the payoff of the rich and that of the poor found in the experimental literature (Charness and Rabin, 2002 and 2005) and the marginal rate of substitution from our model $(\beta u'(c_0^l)/u'(c_{US}^h) = \beta c_{US}^h/c_0^l)$. The β obtained was also equal to 0.013. In our simulations, we use $\beta \simeq 0.007$ as the benchmark value and produce results for β varying between 0 and 0.013.

As for the extrinsic ratio of political power between low-skilled and college-educated nationals (θ_j), we assume it is inversely proportional to the wage ratio (i.e. $\theta_j = 1/\omega_j^h$) (see Column 7). Finally, we use the nationalist optimal migration condition defined in Definition 2 (see Eq. (9)) and calibrate the ϵ_j to match the observed immigration stocks in developed countries. The results for ϵ_j are shown in Column 8.

3.3. Comparative statics and fit of the model

Our model accounts for the economic effects of immigration, aversion to extreme poverty, and non-economic factors affecting citizens' vote on immigration. In this sub-section, we run comparative static experiments to identify the sensitivity of the immigration policy to each economic determinant, and the proportion of variability in observed immigration stocks explained by the full set of economic factors.

Table 1 gives some comparative static results for the G7 countries. The nationalist allocation is reported in Column 1. Column 2 shows that the role of altruism is small when considering rich countries in isolation. It affects the stock of immigrants by less than one percent in all countries, except in the United States (2.3 percent). Columns 3 and 4 show that the redistributive policy plays an important part. Without transfers (i.e. $\tau_j = 0$), college graduates would welcome more migrants and the number of immigrants would be 5.4 percent greater in Italy (where τ_j amounts to USD 237) and 150 percent greater in France (where τ_j equals USD 1237). Similar effects are obtained when $\xi_j = 0$; as the percentage of deviation from the nationalist allocation, the correlation rate between the effects of τ_j and ξ_j equals 0.979. Column 5 gives the number of immigrants that is preferred by high-skilled citizens: if low-skilled citizens did not vote (i.e. $\theta_j = 0$), the stock of immigrants would be three times as large as in the nationalist allocation (except in Japan where ε_j is high).

Finally, Columns 6 and 7 illustrate the global effect of economic factors. We simulated the number of immigrants if ε_j only (Column 6) or both ε_j and θ_j (Column 7) were equal to the mean of the sample. The only sources of variation between countries are the differences in economic factors. The correlation rate between the changes in immigration obtained in these

Table 1
Comparative statics.

G7 countries	Nationalist allocation $\tilde{m}$	No altruism $\beta = 0$	No redistribution		Power to high-skilled $\theta_j = 0$	Removing non-economic factors	
			$\tau_j = 0$	$\xi_j = 0$		$\varepsilon_j = \text{med}[\varepsilon_i]$	$\varepsilon_j = \text{med}[\varepsilon_i]$ $\theta_j = \text{med}[\theta_j]$
Canada	2108.7	2107.3	2675.2	3024.9	6282.3	1715.2	2050.7
France	2967.4	2952.4	7303.1	8182.4	9834.6	799.2	693.7
Germany	4753.8	4718.1	11354.0	12687.7	15393.2	1177.7	1256.0
Italy	917.9	916.6	967.8	1014.5	3493.2	1589.1	1603.3
Japan	502.5	500.8	571.9	663.6	1048.7	830.6	805.4
UK	1838.1	1833.7	2365.2	3301.2	7056.4	1732.7	1116.8
USA	16009.7	15649.2	34322.0	37864.4	48301.3	3361.5	2870.5

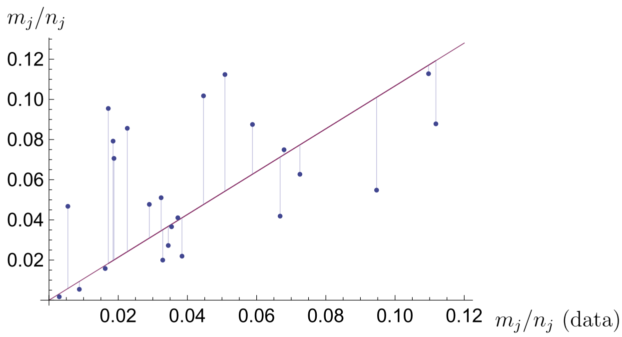


Fig. 2. Fit of the model when non-economic differences across countries are removed.

two counterfactual simulations is high (0.974); this reveals that ε_j is the main non-economic determinant of immigration. At the mean of non-economic factors, Japan's immigration would be 65 percent greater, while the US and French immigration stocks would be about 80 percent lower.

Fig. 2 illustrates this result for the full sample of OECD, high-income countries. The observed immigration rates are represented on the horizontal axis; the counterfactual immigration rates obtained when θ_j and ε_j equal the sample mean are represented on the vertical axis. Removing the variability in non-economic factors induces significant changes in immigration in some countries. However, the correlation between the counterfactual and observed immigration rate equals 50 percent. This means that economic determinants account for half of the variability in immigration across countries.

3.4. No-regret allocation

Before looking at the detailed implications of our treaty, we first look at the global effect of the policy, as a function of the altruism factor β, assuming that the set of participating countries is the OECD, high income group (hosting about 35.2 million low-skilled immigrants from the South in 2000). We let β vary between 0 and its value obtained from experimental games (0.013). For each value of β, we recalibrate the ε_j in such a way that the nationalist allocation matches the observed data. We then compute the no-regret allocation.

Fig. 3 reports the results. It draws the change in total migration compared to the nationalist allocation, $\bar{\mathbf{m}}' \mathbf{1_J} - \tilde{\mathbf{m}}' \mathbf{1_J}$.

The relation between the gain in migration and the altruism factor is almost linear (for this range of β's). The total increase in migration varies between 0 (for $\beta = 0$) and 5.9 million (for $\beta = 0.013$); in the latter case, the no-regret treaty increases the stock of low-skilled migrants by 17 percent in 2000. It is worth noticing that, while β has a minor effect in the nationalist allocation,[14] we do not need implausibly large values of β for our no-regret treaty to have a significant effect.

Assuming that β takes its benchmark value (0.007), we now discuss the no-regret allocation in detail. Appendix D gives the allocations obtained for each host country. We consider two alternative sets of participating countries, as mentioned above. Not surprisingly, an absolute increase in immigration is related to the population size. In relative terms, changes in immigration are negatively correlated with the level of redistributive transfers, productivity and nationalist immigration rates. The largest responses are observed in Korea and Japan (+36 and +21 percent, respectively). The immigrant stock increases by around 10 percent in six countries (Belgium, Denmark, Finland, Iceland, Ireland and Sweden). The smallest changes are observed in countries with high nationalist immigration rates (+4.0 percent in Canada, +4.5 in Australia, +4.3 in Switzerland). The US immigrant stock increases by 6.7 percent.

[14] Remember that at the nationalist allocation, altruism (i.e. $\beta = 0.007$) affects the level of immigration by less than one percent in most destination countries (see Table 1).

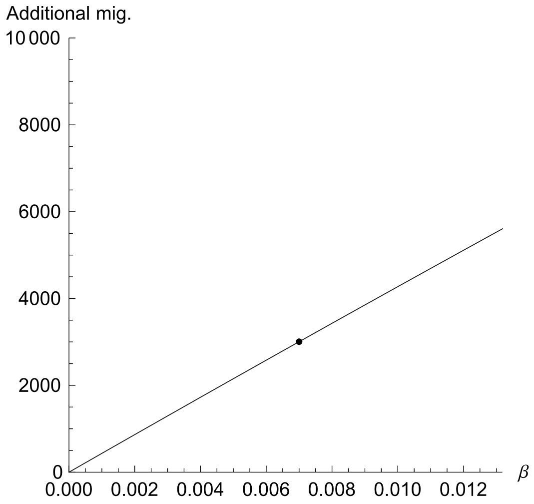


Fig. 3. How many more migrants (in thousands) does the no-regret treaty bring?.

Table 2
Global effect under the various scenarios.

	Nationalist	No-regret	
		G7	OECD
Total additional working-age migration (×1000)	0	1660	2603
Ratio of US GDP to developing world GDP	7.07	7.05	6.95

Note: Total additional working-age migration (×1000) is given by $\bar{\mathbf{m}}'\mathbf{1_J} - \tilde{\mathbf{m}}'\mathbf{1_J}$. Ratio of US GDP to developing world GDP is computed as $\frac{f_{US}(n^l_{US}+\delta_{US}m_{US})}{n_{US}+m_{US}} / \frac{f(n^l_0-\mathbf{m}'\mathbf{1_J})}{n_0-\mathbf{m}'\mathbf{1_J}}$.

Table 2 summarizes our global findings. The additional migration made possible by the no-regret treaty ranges from 1.7 to 2.6 million working-age migrants depending on the set of countries considered. These numbers are important compared to the 35.2 million migrants originating from developing countries and living in high-income countries today (+5.7 percent with the G7 sample, +7.4 percent with the OECD, high-income sample). Larger responses will be obtained when remittances are factored in. This shows that, with a reasonable level of aversion to extreme poverty, the public good externality of migration policy is very significant.

We also computed the effect of such large migration flows on the ratio between US GDP and the average GDP in developing countries in the most favorable case which is OECD, high-income countries participating in the scheme. Indeed, there is a difference in the effectiveness of the treaty if all OECD, high-income countries are involved compared with the situation in which only the G7 countries participate in the treaty. By broadening the pool of participating countries, we make externalities stronger, which translates into a bigger difference between the no-regret and the nationalist allocation.[15] The GDP ratio is reduced from 7.07 to 6.95 (i.e. −1.7 percent). The effect is relatively small because the total population ratio between developing and high-income countries is large.

3.5. Decentralization

We then computed the vectors of subsidies $\bar{\mathbf{p}}$ and lump-sum taxes $\bar{\mathbf{q}}$ that would make the no-regret allocation incentive compatible (Proposition 5). Appendix D gives the result. The subsidies p_j are computed for 1,000,000 additional migrants and expressed in USD per citizen.[16]

We first observe that households in small countries should receive more per immigrant than households in large countries. Indeed, welcoming 1000 additional migrants is much more significant in Luxembourg than in the United States. To a lower extent, the subsidy per migrant increases with the level of productivity (correlation rate of 0.58): citizens from richer countries require higher subsidies to welcome additional migrants. Finally, we also observe that subsidies are higher

[15] In a previous version of this paper, we showed that broadening the pool of rich countries further (by including non-OECD, high income countries) does not bring many additional gains, mainly because these additional countries are relatively small.

[16] The subsidy per citizen should be considered as an annual amount because we used data on annual GDP. If migrants stay on average 30 years in the destination countries and the real interest rate is 2.5 percent, a subsidy of 4.44 (Australia) implies a total discounted subsidy of about 100 dollars for welcoming one million additional immigrants.

when the set of countries participating in the arrangement is larger. This occurs because the size of the positive externality increases with the size of the set of participating countries.

4. Extensions

In this section, we study whether adopting a utilitarian point of view rather than our no-regret allocation would modify the international allocation of migrants, how the inclusion of migrants' remittances and levels of development aid affects our results, and to what extent development aid and migration are substitutes.

4.1. Utilitarianism

In the no-regret allocation, the efficiency gains from coordinating immigration policies are totally used to improve the welfare of the worst-off in the world (i.e. low-skilled workers living in developing countries). Another possible strategy for the international agency could be to maximize a social welfare function aggregating the utility levels of rich countries' citizens. In this section, we study the utilitarian optimum maximizing the aggregate welfare of rich countries. Remember the outcome of the vote in country j is equal to the weighted sum of voters' utilities, W_j, defined in (7). We compute the utilitarian optimum using two variants of the social welfare function and assuming that political participation rates are identical to those in national votes. First, in the Benthamite (or total utilitarianism) tradition, we maximize the weighted sum of the utilities of all voters,

$$\max_{\mathbf{m}} \sum_j n_j W_j.$$

This Benthamite social welfare function characterizes the outcome of a supranational probabilistic vote in which citizens from all developed countries decide on the number of additional immigrants from all possible destinations. Under this criterion, political power increases with country size. Second, in the Millian (or average utilitarianism) tradition, we maximize the unweighted sum of country-specific outcomes,

$$\max_{\mathbf{m}} \sum_j W_j.$$

This Millian social welfare function assigns one vote per country whatever its size. The numerical results are presented in Appendix E for the whole set of high-income countries. On the left panel, we compare the levels of immigration observed in the nationalist, no-regret, Benthamite and Millian utilitarian solutions. The right panel provides the country-specific levels of social welfare and the comparison with the no-regret (or Nationalist) solution.

Under the Benthamite criterion, the total number of immigrants increases by 1.9 million (rather than 2.6 million in the no-regret). If the efficiency gains from coordination are used to maximize the welfare of rich countries' citizens, the number of additional immigrants falls in all countries (except Sweden). Compared with the non-regret treaty, the decrease in migration is more severe in countries where immigrants receive generous welfare benefits (correlation with ξ_j equal to -0.90). Note that a loss of welfare is observed in four OECD, high-income states (France, Germany, Sweden and the United States). Clearly, these countries would have no incentive to participate in such a Benthamite–Utilitarian treaty.

Under the Millian criterion, the total number of immigrants increases drastically by 9.8 million (3.8 times more than in the no-regret). Given their increased political power, smaller countries impose many more immigrants to large countries. A loss of social welfare is observed in the United States and Germany which would host together about 6 million more immigrants than in the no-regret solution. Compared with the no-regret, an increase in immigration would also be observed in France. These countries would have no (or fewer) incentives to join the Millian-Utilitarian treaty. These results comfort our no-regret strategy based on a whole set of implementation constraints (10).

4.2. Remittances

The number of additional migrants generated by the no-regret allocation depends on the value of the welfare externality, i.e. on the rate of aversion to extreme poverty and the effect of migration on income in developing countries. In the previous sections, the latter is only achieved by increasing the marginal productivity of labor in developing countries. Here, we evaluate the robustness of our results to the inclusion of remittances for different values of β. Remittances are quantitatively important at the global level and may well have a stronger impact on developing countries than the labor market responses to emigration. They may strongly modify the value of the externality. To model remittances, we assume the each migrant remits a fraction η of her income to low-skilled stayers in country 0. The consumption level of those left behind becomes:

$$c_0^l = f_0'(l_0) + \eta \sum_j \frac{m_j c_j^m \gamma_j}{l_0}$$

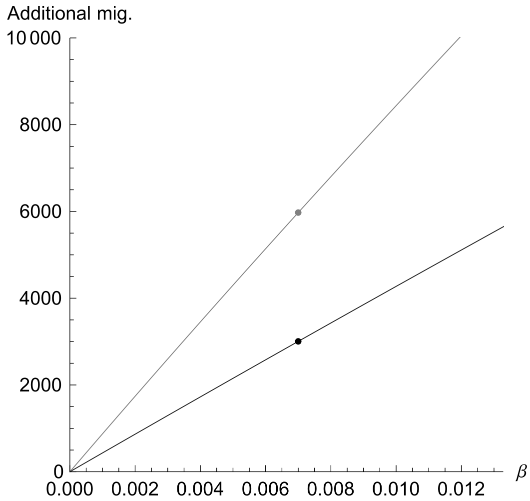


Fig. 4. How many more migrants (in thousands) does the no-regret treaty bring?

where η is the fraction of income remitted by each migrant to those left behind in country 0, and γ_j is a destination-specific real exchange rate capturing the fact that one dollar sent by immigrants to country j has a different purchasing value in country 0.

To calibrate η, we rely on Bollard et al. (2011) who put together a micro-level database on the remitting behavior of 33,000 immigrants from developing countries based on 14 surveys in 11 OECD destination countries. They distinguish between immigrants with a university degree and those who are less educated. Pooling the data and accounting for the intensive and extensive margins, they show that less educated immigrants remit 3 percent of their income on average (i.e. $\eta = 0.03$). Given η and the set of γ_js reported in Appendix B, remittances affect the term $\Psi_j^\beta(.)$ in Eq. (9). For each value of β, we recalibrated ε_j in such a way that the nationalist allocation with remittances matches the observed data, and then computed the no-regret allocation. It is worth noticing that the new ε_js are virtually equal to those in the benchmark (correlation rate of 0.999).

Fig. 4 reports the result. Including remittances strongly increases the number of additional migrants generated by the no-regret treaty. Migration increases by 6 million (rather than +2.6 million in the benchmark) with $\beta = 0.007$ (i.e. +17 percent). When β is calibrated using the results of the experimental literature (0.013), the number of additional migrants exceeds 10 million (+33 percent). With remittances, even a small rate of altruism allows important increases in global labor movements without any utility cost for the rich. Moreover, the actual effect can exceed that of Fig. 4 since official records underestimate the full scale of remittances. Conservatively adding 50 percent to official flows (see World Bank, 2006) would further increase η and the effectiveness of our no-regret treaty.

4.3. Development aid

In a similar vein, we can assess whether accounting for country-specific levels of development aid can modify our results. By reducing income in developed countries and increasing income in developing countries, development aid can reduce extreme poverty and incentives to host immigrants.

We include development aid in the model, assuming that it is financed by taxes on high-skilled workers in the donor country and fully benefits low-skilled workers in country 0. Denoting by a_j the amount of development aid paid by each high-skilled citizen from country j and applying the real exchange rate γ_j, the consumption of high-skilled citizens in the North and low-skilled workers in country 0 becomes:

$$c_j^h = f_j'(l_j) l_{j,1}'(n_j, m_j) + \frac{(1-\zeta_j)s_j(l_j) - (n_j^l + \xi_j m_j)\tau_j}{n_j^h} - a_j$$

$$c_0^l = f_0'\big(n_0^l - \mathbf{m}'\,\mathbf{1_J}\big) l_{0,1}'\big(n_0^l - \mathbf{m}'\,\mathbf{1_J}, 0\big) + \eta \sum_j \frac{m_j c_j^m \gamma_j}{l_0} + \sum_j \frac{n_j^h a_j \gamma_j}{l_0}.$$

To recalibrate the model with aid, we use data on development assistance from the World Development Indicators (World Bank, 2012) and divide them by the number of high-skilled citizens (we subtract aid from the surplus going to the highly skilled). We obtain an average amount of USD 491 per high-skilled citizen in the OECD, high-income countries, with a maximal value of USD 4525 in Luxembourg and a minimal value of USD 189 in the United States. Aid modifies the marginal utility of high-skilled natives in the North and low-skilled workers in the South. We then adjust the ε_js to match observed migration data in the nationalist equilibrium.

Our numerical experiments reveal that accounting for development aid does not affect the effectiveness of our no-regret treaty. We first investigated whether development aid and immigration are substitutable in our framework. When a_j is included in (8), we can compute the derivative of (8) with respect to a_j and numerically assess the willingness of voters to increase or decrease the amount of development aid. At the observed level of development aid, we obtain:

$$\left.\frac{\partial \varphi_j(\tilde{\mathbf{m}})}{\partial a_j}\right|_{a_j\text{obs}} < 0 \quad \forall j, \qquad \left.\frac{\sum_i n_i^h \partial \varphi_i(\tilde{\mathbf{m}})}{\partial a_j}\right|_{a_j\text{obs}} < 0 \quad \forall j$$

If citizen voters had to decide whether to adjust a_j, they would prefer to decrease it. It could be argued that this result holds because citizen voters do not take into account the externality that development aid given by one country increases the utility of all developed countries. However, taking the externality into account leads to a similar conclusion: in the Benthamite utilitarian case, a supranational vote on aid would also lead to a decrease in development aid.

We also evaluated the same derivatives at $a_j = 0$:

$$\left.\frac{\sum_i \partial \varphi_i(\tilde{\mathbf{m}})}{\partial a_j}\right|_{a_j=0} < 0, \qquad \left.\frac{\sum_i \partial \varphi_i(\mathbf{0})}{\partial a_j}\right|_{a_j=0} < 0, \qquad \left.\frac{\sum_i n_i^h \partial \varphi_i(\mathbf{0})}{\partial a_j}\right|_{a_j=0} < 0 \quad \forall j$$

At the nationalist level of immigration (or even in a no-migration configuration in which the worst-off in the world are poorer), citizen voters would vote for a negative amount of aid. This also holds true in the Benthamite utilitarian case. These results are due to a dilution effect: each dollar paid by citizens of developed countries is divided by the demographic ratio. Even the United States (183.5 million adult citizens) only represents 7.3 percent of the developing world adult population. And the whole set of high-income countries represents about 25.2 percent of the developing world population. The small rate of aversion to extreme poverty cannot be used to explain positive amounts of development aid.[17]

It is worth noticing that the same dilution effect explains the low impact of altruism on the nationalist allocation. However, the main difference with development aid is that low-skilled immigration also induces an economic surplus and improves the welfare of citizens in the host country. Contrary to aid, a positive level of immigration is observed even if $\beta = 0$. Starting from that interior solution, the coordination of immigration policies induces significant efficiency gains.

5. Conclusion

Most of the literature on the political economy of migration is written from the point of view of the host country. For example, Benhabib (1996) and Crettez (2011) consider the immigration policy which citizens voted for. These studies give a richer description of what we called a nationalist policy above, with different instruments and heterogeneous citizens, but they do not consider ways to improve upon such a policy to increase global welfare. A few papers adopt the point of view of a benevolent planner maximizing world welfare as Benhabib and Jovanovic (2012) for example. They show how the optimal policy depends on technology (constant or decreasing social return, the relative backwardness of poor countries, and externalities), and on the weight the planner attaches to rich countries. However, they do not care about rich countries' incentives to implement such a policy.

Our study relies on the fact that economic incentives and political mechanisms are essential to understand or influence immigration policies in rich countries. Compared with these two strands of the literature, we adopt a different point of view: we minimize extreme poverty in the world subject to implementation constraints, i.e. accounting for (limited) incentives to host immigrants in rich countries. We accordingly propose a political mechanism that improves the international allocation of labor by considering solutions that are politically acceptable in rich countries.

Increasing migration and reducing extreme poverty subject to the participation constraint is possible provided that households are averse to inequality. Nevertheless, in the absence of any coordinating device, citizens in any given rich country lack incentives to welcome additional migrants as this policy is a global public good which benefits all rich countries. We showed that we can construct a no-regret allocation taking the public good nature of immigration policy better into account, and that this allocation can be supported by a tax-subsidy scheme.

In order to quantitatively assess the increase in migration that such a no-regret allocation can bring, we calibrated the model using statistics on immigration, working-age population, output, skill premia and inequality. We simulated the proposed scheme on different sets of rich countries. We showed that even a small rate of altruism allows important increases in global labor movements without any utility cost for the rich. Applying the treaty to the set of OECD, high-income countries increases migration by 17 percent when remittances are accounted for. And if altruism is as strong as predicted by experimental games, our policy would produce a greater increase in global migration.

[17] Explaining aid requires a different model of altruism (accounting for bilateral cultural/colonial ties between country pairs), or a model of self-interest (improving the situation of the worst-off benefits the donor, as in Andreoni, 2006). For example, the development objectives of the 1970s and 1980s obviously included shoring up support from poor countries in the geopolitics of the Cold War. After the fall of the Berlin Wall, other targets appeared such as the War on Terror.

Hence, when all rich countries are averse to global inequality, albeit to a very small extent, setting up a global migration fund designed to coordinate national immigration policies is likely to generate welfare gains for millions of additional migrants and for billions of individuals left in the poor world.

Appendix A. Proofs

A.1. Proof of Lemma 1

A sufficient condition for an interior solution is $\partial\tilde{\Phi}_j(m_j)/\partial m_j = \Psi_j > 0$ at $m_j = 0$. For $\beta = 0$ and $l''_{j,12}(n_j, m_j) = 0$, the condition is $\Xi > 0$ with:

$$\Xi = \left(u'\left(c_j^h\right)n_j^h + n_j^l\theta_j u'\left(c_j^l\right)\right) f_j''(l_j) l'_{j,1}(n_j, 0) - u'\left(c_j^h\right)(1-\zeta_j) f_j''(l_j) l_j - \theta_j u'\left(c_j^l\right)\zeta_j f_j''(l_j) l_j - \frac{\xi_j \tau_j}{l'_{j,2}(n_j, 0)},$$

with $l_j = l(n_j, 0)$. The value of ξ such that $\Xi = 0$ is:

$$\bar{\xi}_j = \frac{f_j''(l_j) l'_{j,2}(n_j, 0)}{\tau_j}\left(\left(u'\left(c_j^h\right)n_j^h + n_j^l\theta_j u'\left(c_j^l\right)\right) l'_{j,1}(n_j, 0) - u'\left(c_j^h\right)(1-\zeta_j) l_j - \theta_j u'\left(c_j^l\right)\zeta_j l_j\right).$$

$\bar{\xi}_j$ is positive if

$$\left(u'\left(c_j^h\right)n_j^h + n_j^l\theta_j u'\left(c_j^l\right)\right) l'_{j,1}(n_j, 0) - u'\left(c_j^h\right)(1-\zeta_j) l_j - \theta_j u'\left(c_j^l\right)\zeta_j l_j < 0.$$

As $l_j(\cdot)$ is homogeneous of degree one, one has:

$$l_j\left(n_j^h + n_j^l, 0\right) = l'_{j,1}(n_j, 0)\left(n_j^h + n_j^l\right) + l'_{j,2}(n_j, 0)0 = l'_{j,1}(n_j, 0)\left(n_j^h + n_j^l\right).$$

Replacing l_j by this value and dividing by $l'_{j,1}(n_j, 0)$ we obtain:

$$\left(u'\left(c_j^h\right)n_j^h + n_j^l\theta_j u'\left(c_j^l\right)\right) - u'\left(c_j^h\right)(1-\zeta_j)\left(n_j^h + n_j^l\right) - \theta_j u'\left(c_j^l\right)\zeta_j\left(n_j^h + n_j^l\right) < 0.$$

which simplifies into:

$$u'\left(c_j^h\right)\left(\zeta_j n_j^h - (1-\zeta_j)n_j^l\right) - \theta_j u'\left(c_j^l\right)\left(\zeta_j n_j^h - (1-\zeta_j)n_j^l\right) < 0.$$

The term $\zeta_j n_j^h - (1-\zeta_j)n_j^l$ is negative because we have assumed that $(1-\zeta_j)/\zeta_j > n_j^h/n_j^l$. Hence the inequality becomes:

$$u'\left(c_j^h\right) - \theta_j u'\left(c_j^l\right) > 0,$$

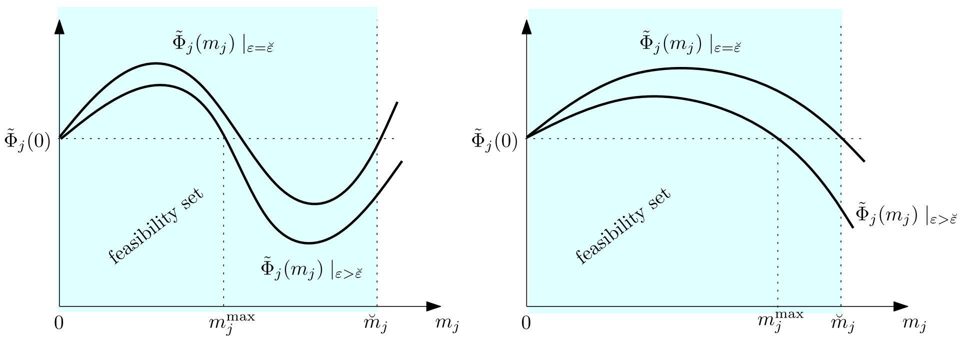
which is satisfied when $\theta_j < \bar{\theta}_j$. Hence, $\theta_j < \bar{\theta}_j$ implies $\bar{\xi} > 0$.

Once $\bar{\xi} > 0$, imposing $\xi < \bar{\xi}$ suffices to guarantee $\Xi > 0$.

If either β or $l''_{j,12}(n_j, m_j)$ are positive (more altruism or less substitution between natives and migrants), Ψ_j at $m_j = 0$ is even larger.

A.2. Proof of Lemma 2

Under the conditions of Lemma 1, we know that $\tilde{\Phi}_j(m_j)$ is increasing at $m_j = 0$ for any value of ε. From Eq. (9), we have $\tilde{\Phi}_j(\check{m}_j)/\partial\varepsilon < 0$.



If the function $\tilde{\Phi}_j(m_j)$ intersects the value $\tilde{\Phi}_j(0)$ only once, the two possible cases are depicted above. If $\partial\tilde{\Phi}_j(\breve{m}_j)/\partial\breve{m}_j > 0$ (left picture), then, by continuity of φ and by Lemma 1, there exists $m_j^{\max} < \breve{m}_j$ such that $\tilde{\Phi}_j(m_j^{\max}) = \tilde{\Phi}_j(0)$ even for $\varepsilon = \breve{\varepsilon}$ and for any ε greater than that level. If $\partial\tilde{\Phi}_j(\breve{m}_j)/\partial\breve{m}_j < 0$ (right picture), the result follows from the implicit function theorem.

If the function $\tilde{\Phi}_j(m_j)$ intersect several times the value $\tilde{\Phi}_j(0)$, there are several solutions to the equation $\tilde{\Phi}_j(m_j) = \tilde{\Phi}_j(0)$ within the interval $[0, \breve{m}]$. For the purposes of the lemma, we however only need to have at least one solution $\tilde{\Phi}_j(m_j) = \tilde{\Phi}_j(0)$ within the interval $[0, \breve{m}]$; we accordingly define $m_j^{\max}$ as the largest solution to equation $\tilde{\Phi}_j(m_j) = \tilde{\Phi}_j(0)$ within the interval $[0, \breve{m}]$.

A.3. Proof of Proposition 1

The first-order condition (9) can be rewritten as:

$$\frac{n_j^h}{n_j}\big(\Psi_j^h(.) + \Psi_j^\beta(.)\big) + \frac{n_j^l}{n_j}\theta_j\big(\Psi_j^l(.) + \Psi_j^\beta(.)\big) = 0.$$

It implies that either $\Psi_j^h(.) + \Psi_j^\beta(.) > 0$ and $\Psi_j^l(.) + \Psi_j^\beta(.) < 0$ (the case of the proposition), or the opposite, abstracting from the knife-edge case in which both would be nil. If $\Psi_j^h(.) > \Psi_j^l(.)$, i.e. if the economic gains of migration are greater for the high-skilled, we have that, at the nationalist allocation, $\Psi_j^h(.) + \Psi_j^\beta(.) > 0$ and $\Psi_j^l(.) + \Psi_j^\beta(.) < 0$, i.e. each high-skilled national prefers to increase the level of immigration, whereas each low-skilled national prefers to decrease it. The condition $\Psi_j^h(.) > \Psi_j^l(.)$ can be rewritten as

$$\frac{u'(c_j^l)}{u'(c_j^h)} > \frac{f_j''(l_j)l_{j,2}'(n_j,m_j)l_{j,1}'(n_j,m_j) + f_j'(l_j)l_{j,12}''(n_j,m_j) - \frac{(1-\zeta_j)f_j''(l_j)l_{j,2}'(n_j,m_j)l_j + \xi_j\tau_j}{n_j^h}}{f_j''(l_j)l_{j,2}'(n_j,m_j)l_{j,1}'(n_j,m_j) + f_j'(l_j)l_{j,12}''(n_j,m_j) - \frac{\zeta_j f_j''(l_j)l_{j,2}'(n_j,m_j)l_j}{n_j^l}}.$$

Under (A), this reduces to:

$$\frac{u'(c_j^l)}{u'(c_j^h)} > \frac{f_j''(l_j)\delta_j - \frac{(1-\zeta_j)f_j''(l_j)\delta_j l_j + \xi_j\tau_j}{n_j^h}}{f_j''(l_j)\delta_j - \frac{\zeta_j f_j''(l_j)\delta_j l_j}{n_j^l}}.$$

As $u'(c_j^l) > u'(c_j^h)$, we have either

$$f_j''(l_j)\delta_j - \frac{(1-\zeta_j)f_j''(l_j)\delta_j l_j + \xi_j\tau_j}{n_j^h} > f_j''(l_j)\delta_j - \frac{\zeta_j f_j''(l_j)\delta_j l_j}{n_j^l} > 0, \tag{26}$$

or

$$f_j''(l_j)\delta_j - \frac{(1-\zeta_j)f_j''(l_j)\delta_j l_j + \xi_j\tau_j}{n_j^h} < f_j''(l_j)\delta_j - \frac{\zeta_j f_j''(l_j)\delta_j l_j}{n_j^l} < 0. \tag{27}$$

In case (26), the inequality can be rewritten as:

$$-f_j''(l_j)\delta_j l_j\left(1 - \frac{\zeta_j n_j^h}{(1-\zeta_j)n_j^l}\right) > \xi_j\tau_j. \tag{28}$$

Given that we have assumed $f'''x + f'' > 0$, a sufficient condition on the parameters for (28) to hold is:

$$-f_j''\big(n_j + m_j^{\max}\big)\delta_j\big(n_j + m_j^{\max}\big)\left(1 - \frac{\zeta_j n_j^h}{(1-\zeta_j)n_j^l}\right) > \xi_j\tau_j,$$

i.e. $\xi_j < \hat{\xi}_j$. In case (27), the left part of the inequality is always verified because $(1-\zeta_j)/\zeta_j > n_j^h/n_j^l$. The second part holds if:

$$f_j''(l_j)\delta_j\left(1 - \frac{\zeta_j l_j}{n_j^l}\right) < 0 \quad \Leftrightarrow \quad \zeta_j < \frac{n_j^l}{n_j^l + \tilde{m}_j}.$$

A sufficient condition for this to hold is:

$$\zeta_j < \frac{n_j^l}{n_j^l + m_j^{\max}}.$$

A.4. Proof of Proposition 2

Applying the implicit function theorem, we obtain:

$$\frac{d\tilde{m}_j}{d\theta_j} = -\frac{\partial\Psi_j(.)}{\partial\theta_j} / \frac{\partial\Psi_j(.)}{\partial m_j} < 0,$$
$$\frac{d\tilde{m}_j}{d\xi_j} = -\frac{\partial\Psi_j(.)}{\partial\xi_j} / \frac{\partial\Psi_j(.)}{\partial m_j} \lesseqgtr 0,$$

with

$$\frac{\partial\Psi_j(.)}{\partial m_j} = \frac{\partial^2\varphi_j(\mathbf{m})}{\partial m_j^2} < 0,$$
$$\frac{\partial\Psi_j(.)}{\partial\theta_j} = \frac{n_j^l}{n_j}\left(\Psi_j^l(.) + \Psi_j^\beta(.)\right) < 0,$$
$$\frac{\partial\Psi_j(.)}{\partial\xi_j} = \frac{n_j^h}{n_j}\left(-\frac{u''(c_j^h)m_j\tau_j}{n_j^h}\left(f_j''(l_j)\delta_- \frac{(1-\zeta_j)f_j''(l_j)\delta_j l_j + \xi_j\tau_j}{n_j^h}\right) - \frac{u'(c_j^h)\tau_j}{n_j}\right). \tag{29}$$

The first term of (29) is negative in case (26) of Proposition 1. Case (26) arises when

$$f_j''(l_j)\delta_j - \frac{\zeta_j f_j''(l_j)\delta_j l_j}{n_j^l} > 0$$

that is when

$$n_j^l < \zeta_j l_j.$$

The first term of (29) is equal to zero if $u''(c_j^h) = 0$. The second term is negative. The sum of the two terms is negative in case (26), or if $u''(c_j^h) = 0$, and indeterminate otherwise.

A.5. Proof of Proposition 4

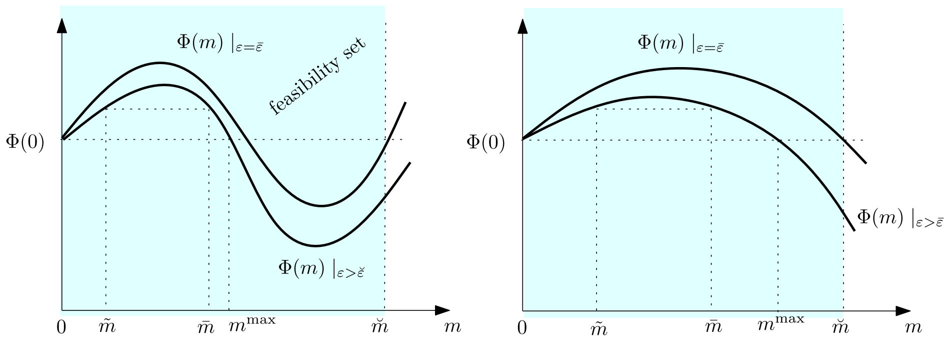
With symmetric countries and allocations, the system of J equalities $\varphi_j(\tilde{\mathbf{m}}) = \varphi_j(\bar{\mathbf{m}})$ reduces to $\Phi(\tilde{m}) = \Phi(\bar{m})$, where $\tilde{m}$ and $\bar{m}$ are scalars. Using Eq. (8) we have

$$\begin{aligned}\Phi(m) = {} & \frac{n^h}{n}\,u\left(f'(l)l'_{j,1}(n,m) + \frac{(1-\zeta)(f(l) - f'(l)l) - (n^l + \xi m)\tau}{n^h}\right) \\ & + \frac{n^l}{n}\theta\,u\left(f'(l)l'_{j,1}(n,m) + \frac{\zeta(f(l) - f'(l)l)}{n^l} + \tau\right) \\ & + \frac{n^h + \theta n^l}{n}\left(\beta u\left(f_0'\left(n_0^l - m - (J-1)m\right)l'_{0,1}\left(n_0^l - m - (J-1)m, 0\right)\right) - \varepsilon\left(\frac{m}{n}\right)^2\right).\end{aligned}$$

The nationalist allocation satisfies $\Phi_j'(\tilde{m})\,|_{(J-1)m} = 0$ where terms $(J-1)m$ are constant when taking the derivative. Hence, when we compute the derivative with respect to all m's, we have $\Phi_j'(\tilde{m}) > 0$ as long as $\beta > 0$, as:

$$\Phi_j'(\tilde{m}) = \Phi_j'(\tilde{m})\,|_{(J-1)m} - \beta(J-1)\,\frac{n^h + \theta n^l}{n}u'\left(c_0^l\right)\left(f_0''(l_0)l'_{0,1}(l_0,0) + f_0'(l_0)l''_{0,11}(l_0,0)\right) > 0.$$

Moreover, following the same argument as the one developed for Lemma 2, there exists an value $\bar{\varepsilon}$ such that $\Phi(\breve{m}) = \Phi(0)$, where $\breve{m}$ is the maximum feasible migration (leaving migrants indifferent between staying and leaving). For any $\varepsilon > \bar{\varepsilon}$, there exists an $m^{\max}$ such that $\Phi(m^{\max}) = \Phi(0)$, where $m^{\max} < \breve{m}$, and $\Phi(m)$ is decreasing at $m^{\max}$. Hence there exists $\bar{m} < m^{\max}$ such that $\bar{m} > \tilde{m}$ and $\Phi_j(\bar{m}) = \Phi_j(\tilde{m})$.



Appendix B. Cross-country data

	N_j	$\frac{n_j^h}{N_j}$	$\frac{m_j}{N_j}$	ρ_j	g_j	$\frac{y_j}{N_j}$	δ_j	R	ω_j^m	ξ_j	γ_j
Developing	2490590.0	0.033	0.000	0.120	11.000	7.067	1.000	0.829	1.000	0.000	1.000
Australia	12520.8	0.186	0.081	0.070	6.300	35.933	0.767	0.712	0.862	0.353	2.247
Austria	5801.9	0.077	0.092	0.090	8.600	34.823	0.708	0.703	1.020	0.579	2.558
Belgium	7233.2	0.179	0.049	0.070	7.900	33.657	0.719	0.613	0.971	0.584	2.597
Canada	20804.7	0.196	0.101	0.070	7.100	37.356	0.741	0.781	0.862	0.436	2.326
Denmark	3747.8	0.112	0.039	0.065	7.900	38.391	0.678	0.733	1.351	0.489	3.322
Finland	3579.6	0.071	0.016	0.105	10.400	32.473	0.716	0.791	1.000	0.513	2.817
France	40417.9	0.088	0.073	0.077	9.400	32.413	0.723	0.690	1.220	0.620	2.667
Germany	60268.7	0.132	0.079	0.084	7.500	32.512	0.709	0.692	1.235	0.520	2.500
Greece	7749.6	0.141	0.080	0.085	8.800	21.648	0.873	0.751	0.952	0.433	2.062
Iceland	173.6	0.116	0.027	0.070	8.700	43.714	0.718	0.746	1.351	0.450	3.650
Ireland	2309.2	0.168	0.020	0.115	7.000	43.694	0.745	0.625	0.893	0.398	2.915
Italy	42626.9	0.057	0.021	0.065	9.100	31.089	0.794	0.850	0.952	0.508	2.500
Japan	92337.4	0.190	0.005	0.094	7.700	33.202	0.645	0.776	1.010	0.359	2.725
Korea Rep	42288.8	0.148	0.002	0.107	8.300	21.630	0.879	0.802	1.010	0.280	1.789
Luxembourg	302.9	0.110	0.059	0.070	8.200	66.830	0.601	0.703	0.971	0.477	2.747
Netherlands	11109.3	0.149	0.070	0.070	7.300	36.700	0.717	0.705	0.971	0.518	2.584
New Zealand	2399.7	0.208	0.052	0.065	7.600	27.449	0.823	0.761	0.862	0.403	2.463
Norway	3050.6	0.122	0.040	0.055	7.400	41.032	0.679	0.709	1.000	0.533	3.205
Portugal	6888.7	0.039	0.045	0.110	12.800	24.119	0.875	0.725	0.962	0.510	1.976
Spain	28838.9	0.121	0.035	0.100	9.000	26.204	0.830	0.809	0.962	0.457	2.208
Sweden	6219.2	0.155	0.066	0.070	7.100	31.604	0.718	0.629	1.190	0.569	2.915
Switzerland	5200.4	0.128	0.101	0.104	8.000	37.012	0.636	0.871	1.020	0.448	3.247
UK	40352.9	0.098	0.046	0.091	9.400	32.462	0.782	0.757	1.163	0.458	2.688
USA	183564.0	0.306	0.087	0.110	6.300	49.988	0.714	0.829	1.220	0.457	2.320

Note: $N_j = n_j + m_j$. R_j = gini net/gini gross. All population figures include individuals aged 25+.

Appendix C. Identified parameters – benchmark

	ω_j^h	ν_j	α_j^l	$\hat{\mu}_j$	$\hat{\zeta}_j$	τ_j	θ_j	ε_j
Developing	3.479	1.000	0.587	7.067				
Australia	1.532	0.431	0.579	59.380	0.420	1.354	0.653	0.535
Austria	2.098	0.452	0.585	56.729	0.260	1.291	0.477	0.311
Belgium	1.707	0.452	0.564	53.823	0.347	1.800	0.586	0.300
Canada	1.617	0.420	0.568	61.942	0.388	1.314	0.619	0.439
Denmark	1.645	0.523	0.592	57.794	0.366	0.967	0.608	0.407
Finland	2.825	0.473	0.562	50.578	0.168	0.832	0.354	1.683
France	2.008	0.504	0.583	49.604	0.273	1.267	0.498	0.185
Germany	1.831	0.502	0.571	49.562	0.313	1.535	0.546	0.175
Greece	2.050	0.486	0.553	32.797	0.267	0.948	0.488	0.429
Iceland	1.802	0.541	0.581	63.819	0.315	1.153	0.555	0.594
Ireland	2.143	0.451	0.533	68.191	0.244	2.817	0.467	1.284
Italy	1.774	0.481	0.608	49.611	0.324	0.237	0.564	1.573
Japan	1.997	0.464	0.534	51.130	0.268	1.212	0.501	4.162
Korea Rep	2.325	0.551	0.531	30.163	0.216	0.707	0.430	15.716
Luxembourg	1.742	0.407	0.587	116.081	0.338	2.079	0.574	0.531
Netherlands	1.639	0.446	0.580	59.858	0.374	1.336	0.610	0.414
New Zealand	1.614	0.453	0.564	43.617	0.381	0.917	0.620	0.712
Norway	1.486	0.448	0.599	68.040	0.435	0.901	0.673	0.691
Portugal	3.803	0.497	0.573	36.674	0.118	0.792	0.263	0.763
Spain	2.358	0.487	0.545	39.523	0.215	0.805	0.424	0.843

	ω_j^h	ν_j	α_j^l	$\hat{\mu}_j$	$\hat{\zeta}_j$	τ_j	θ_j	ε_j
Sweden	1.617	0.498	0.579	48.553	0.380	1.528	0.619	0.146
Switzerland	2.207	0.426	0.550	60.575	0.242	0.947	0.453	0.363
UK	2.268	0.517	0.564	48.141	0.228	1.103	0.441	0.481
USA	1.930	0.499	0.493	72.116	0.289	2.634	0.518	0.164

Note: In the table, we report $\hat{\mu}_j = \mu_j/(n_j + m_j)^{1-\alpha_j^l}$ which does not depend on the size of the country. We also report $\hat{\zeta}_j = \zeta_j/(1-\zeta_j) \times (n_j^h/n_j^l)$ which is the individual share of surplus of a low-skilled worker compared to that of a high-skilled one.

Appendix D. Migration stocks and subsidies

Country	Migration stocks m_j			Subsidy p_j/n_j	
	Nationalist	No-regret		G7	OECD
		G7	OECD		
Australia	1010.8		1056.6		4.44
Austria	535.9		564.4		6.22
Belgium	351.2		391.1		4.49
Canada	2108.7	2175.4	2192.5	2.03	2.55
Denmark	147.8		162.9		14.11
Finland	55.6		62.5		25.80
France	2966.9	3176.8	3233.3	0.45	0.57
Germany	4752.7	5065.0	5151.9	0.29	0.37
Greece	623.6		655.1		3.80
Iceland	4.6		5.2		430.34
Ireland	45.2		50.3		44.34
Italy	914.9	982.1	999.4	1.65	2.08
Japan	498.4	584.4	606.7	1.33	1.67
Korea Rep	71.8		97.7		4.67
Luxembourg	17.9		19.0		329.33
Netherlands	774.3		821.5		4.28
New Zealand	124.7		132.3		21.01
Norway	122.6		132.2		24.11
Portugal	307.7		327.7		5.90
Spain	1017.9		1097.5		1.81
Sweden	410.2		466.1		2.99
Switzerland	525.4		547.9		9.85
UK	1837.9	1953.9	1984.2	0.91	1.14
USA	16000.2	16802.7	17071.9	0.13	0.17

Note: Migrants are measured in thousands. Subsidy is per citizen for 1,000,000 additional migrants (in USD).

Appendix E. Results with utilitarianism

Country j	Migration levels				Objective W_j				
	$\tilde{m}_j$	$\bar{m}_j$	Benth.	Millian	No regret	Benthamite		Millian	
Australia	1010.8	1056.6	1031.5	1020.2	3.63247	3.63247	Gain	3.63248	Gain
Austria	535.9	564.4	553.6	539.5	3.62568	3.62568	Gain	3.62569	Gain
Belgium	351.2	391.1	377.7	358.4	3.56717	3.56718	Gain	3.56719	Gain
Canada	2108.7	2192.5	2151.7	2140.9	3.68630	3.68630	Gain	3.68631	Gain
Denmark	147.8	162.9	155.1	148.8	3.69413	3.69413	Gain	3.69414	Gain
Finland	55.6	62.5	57.1	55.8	3.55818	3.55819	Gain	3.55820	Gain
France	2966.9	3233.3	3192.7	3316.9	3.55118	3.55118	Loss	3.55119	Gain
Germany	4752.7	5151.9	5103.7	5584.0	3.55777	3.55777	Loss	3.55777	Loss
Greece	623.6	655.1	639.5	628.1	3.15933	3.15933	Gain	3.15934	Gain
Iceland	4.6	5.2	4.8	4.6	3.82485	3.82485	Gain	3.82486	Gain
Ireland	45.2	50.3	46.5	45.3	3.84248	3.84249	Gain	3.84250	Gain
Italy	914.9	999.4	933.7	947.6	3.46664	3.46664	Gain	3.46665	Gain
Japan	498.4	606.7	511.2	552.6	3.56033	3.56034	Gain	3.56034	Gain
Korea Rep	71.8	97.7	73.5	74.8	3.14666	3.14667	Gain	3.14668	Gain
Luxembourg	17.9	19.0	18.4	17.9	4.25290	4.25290	Gain	4.25291	Gain
Netherlands	774.3	821.5	799.0	784.4	3.65752	3.65752	Gain	3.65753	Gain
New Zealand	124.7	132.3	127.6	125.0	3.36106	3.36106	Gain	3.36107	Gain
Norway	122.6	132.2	126.2	123.0	3.74914	3.74914	Gain	3.74915	Gain
Portugal	307.7	327.7	314.6	309.5	3.28320	3.28320	Gain	3.28321	Gain
Spain	1017.9	1097.5	1043.5	1046.9	3.34812	3.34812	Gain	3.34813	Gain
Sweden	410.2	466.1	472.3	424.4	3.51102	3.51102	Loss	3.51103	Gain
Switzerland	525.4	547.9	537.8	527.6	3.72060	3.72060	Gain	3.72061	Gain
UK	1837.9	1984.2	1900.8	1938.4	3.55623	3.55623	Gain	3.55624	Gain
USA	16000.2	17071.9	16955.1	24335.6	4.02424	4.02424	Loss	4.02413	Loss
All	35227.0	37830.8	37128.6	45050.3					

References

Agrawal, Ajay, Kapur, Devesh, McHale, John, Oettl, Alexander, 2011. Brain drain or brain bank? The impact of skilled emigration on poor-country innovation. Journal of Urban Economics 69 (1), 43–55.

Andreoni, James, 2006. Philanthropy. In: Kolm, S.C., Mercier-Ythier, J. (Eds.), Handbook of the Economics of Giving, Altruism and Reciprocity, vol. 2. Elsevier North-Holland, pp. 1201–1269.

Azam, Jean-Paul, Laffont, Jean-Jacques, 2003. Contracting for aid. Journal of Development Economics 70, 25–58.

Barbone, Luca, Bontch-Osmolovsky, Misha, Zaidi, Salman, 2009. The foreign-born population in the European Union and its contribution to national tax and benefit systems. World Bank Policy Research Working Paper no. 4899. The World Bank, Washington, DC.

Barro, Robert J., Lee, Jong-Wha, 2010. A New Data Set of Educational Attainment in the World, 1950–2010. Nber Working Papers 15902. National Bureau of Economic Research, Inc.

Ben-Gad, Michael, 2008. Capital-skill complementarity and the immigration surplus. Review of Economic Dynamics 11 (2), 335–365.

Benhabib, Jess, 1996. On the political economy of immigration. European Economic Review 40 (9), 1737–1743.

Benhabib, Jess, Jovanovic, Boyan, 2012. Optimal migration: a world perspective. International Economic Review 53 (2), 321–348.

Bollard, Albert, McKenzie, David, Morten, Melanie, Rapoport, Hillel, 2011. Remittances and the brain drain revisited: the microdata show that more educated migrants remit more. World Bank Economic Review 25 (1), 132–156.

Borjas, George J., 1995. The economic benefits from immigration. Journal of Economic Literature 9 (2), 3–22.

Borjas, George J., 2009. The analytics of the wage effect of immigration. NBER Working Papers 14796. National Bureau of Economic Research, Inc.

Borjas, George J., Friedman, Richard B., Katz, Lawrence F., 1997. How much do immigration and trade affect labor market outcomes? Brookings Papers on Economic Activity 28 (1), 1–90.

Borjas, George J., Grogger, Jeffrey, Hanson, Gordon H., 2012. Comment: on estimating elasticities of substitution. Journal of the European Economic Association 10 (1), 198–210.

Bubb, Ryan, Kremer, Michael, Levine, David I., 2011. The economics of international refugee law. The Journal of Legal Studies 40 (2), 367–404.

Charness, Gary, Rabin, Matthew, 2002. Understanding social preferences with simple tests. The Quarterly Journal of Economics 117 (3), 817–869.

Charness, Gary, Rabin, Matthew, 2005. Expressed preferences and behavior in experimental games. Games and Economic Behavior 53, 151–169.

Clemens, Michael A., Montenegro, Claudio E., Pritchett, Lant, 2008. The place premium: wage differences for identical workers across the U.S. Border. Working Paper Number 148. Center for Global Development.

Clemens, Michael A., 2011. Economics and emigration: trillion-dollar bills on the sidewalk? The Journal of Economic Perspectives 25 (3), 83–106.

Cortès, Patricia, Tessada, José, 2011. Low-skilled immigration and the labor supply of highly skilled women. American Economic Journal: Applied Economics 3 (3), 88–123.

Coughlin, Peter, Nitzan, Shmuel, 1981. Electoral outcomes with probabilistic voting and Nash social welfare maxima. Journal of Public Economics 15 (1), 113–121.

Coulombe, Serge, Tremblay, Jean-Francois, 2009. Migration and skills disparities across the Canadian provinces. Regional Studies 43 (1), 5–18.

Crettez, Bertrand, 2011. Is selling immigration rights politically sustainable? Recherches économiques de Louvain 77 (4), 33–55.

de la Croix, David, Docquier, Frederic, 2010. An incentive mechanism to break the low-skill immigration deadlock. CReAM Discussion Paper Series 1008. Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College, London.

de la Croix, David, Docquier, Frédéric, 2012. Do brain drain and poverty result from coordination failures? Journal of Economic Growth 17 (1), 1–26.

di Giovanni, Julian, Levchenko, Andrei, Ortega, Francesc, forthcoming. A global view of cross-border migration. Journal of the European Economic Association. http://dx.doi.org/10.1111/jeea.12110.

Docquier, Frédéric, Rapoport, Hillel, 2012. Globalization, brain drain, and development. Journal of Economic Literature 50 (3), 681–730.

Docquier, Frédéric, Lindsay Lowell, B., Marfouk, Abdeslam, 2009. A gendered assessment of highly skilled emigration. Population and Development Review 35 (2), 297–321.

Docquier, Frédéric, Özden, Çaglar, Peri, Giovanni, 2014. The labour market effects of immigration and emigration in OECD countries. The Economic Journal 124 (579), 1106–1145.

Dustmann, Christian, Frattini, Tommaso, Preston, Ian, 2008. The effect of immigration along the distribution of wages. CReAM Discussion Paper Series 0803. Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London, April.

Facchini, Giovanni, Mayda, Anna Maria, 2008. From individual attitudes towards migrants to migration policy outcomes: theory and evidence. Economic Policy 23, 651–713.

Facchini, Giovanni, Mayda, Anna Maria, 2009. Does the welfare state affect individual attitudes towards immigrants? Evidence across countries. Review of Economics and Statistics 91 (2), 295–314.

Fehr, Ernst, Schmidt, Klaus M., 2006. The economics of fairness, reciprocity and altruism: experimental evidence and new theories. In: Kolm, S.C., Mercier-Ythier, J. (Eds.), Handbook of the Economics of Giving, Altruism and Reciprocity, vol. 1. Elsevier/North-Holland, pp. 615–691.

Gersbach, Hans, Winkler, Ralph, 2007. On the design of global refunding and climate change. Discussion Paper 6379. CEPR.

Giving USA, 2013. The Annual Report on Philanthropy for the Year 2012. Indiana University, IUPI.

Hatton, Timothy J., 2004. Seeking asylum in Europe. Economic Policy 19, 5–62.

Hatton, Timothy J., Williamson, Jeffrey G., 2006. Refugees, asylum seekers, and policy in Europe. In: Foders, Federico, Langhammer, Rolf J. (Eds.), Labour Mobility and the World Economy. Springer, Kiel, pp. 250–284.

Hendricks, Lutz, 2004. A database of Mincerian earnings regressions. http://www.lhendricks.org/Mincer.htm.

Hirschman, Charles, 2001. The educational enrollment of immigrant youth: a test of the segmented-assimilation hypothesis. Demography 38 (3), 317–336.

Immervoll, Herwig, Richardson, Linda, 2011. Redistribution policy and inequality reduction in OECD countries: what has changed in two decades? OECD Social, Employment and Migration Working Papers 122. OECD Publishing, October.

Javorcik, Beata S., Özden, Çaglar, Spatareanu, Mariana, Neagu, Cristina, 2011. Migrant networks and foreign direct investment. Journal of Development Economics 94 (2), 231–241.

Kerr, William R., 2008. Ethnic scientific communities and international technology diffusion. Review of Economics and Statistics 90 (3), 518–537.

Klein, Paul, Ventura, Gustavo, 2007. TFP differences and the aggregate effects of labor mobility in the long-run. The B.E. Journals in Macroeconomics 7 (1), 1–38.

Kremer, Michael, Watt, Stanley, 2009. The globalization of household production. Mimeo.

Kugler, Maurice, Rapoport, Hillel, 2007. International labor and capital flows: complements or substitutes? Economics Letters 94 (2), 155–162.

Manacorda, Marco, Manning, Alan, Wadsworth, Jonathan, 2012. The impact of immigration on the structure of wages: theory and evidence from Britain. Journal of the European Economic Association 10 (1), 120–151.

Mattoo, Aaditya, Neagu, Ileana Cristina, Özden, Çaglar, 2008. Brain waste? Educated immigrants in the US labor market. Journal of Development Economics 87 (2), 255–269.

Milner, Helen V., 2006. Why multilateralism? Foreign aid and domestic principal–agent problems. In: Hawkins, Darren G. (Ed.), Delegation and Agency in International Organizations. Cambridge University Press, New York, pp. 107–139.

Ortega, Francesc, 2010. Immigration, citizenship, and the size of government. The B.E. Journal in Economic Analysis & Policy 10 (1).
Ottaviano, Gianmarco I.P., Peri, Giovanni, 2012. Rethinking the effect of immigration on wages. Journal of the European Economic Association 10 (1), 152–197.
Parsons, Christopher R., Skeldon, Ronald, Walmsley, Terrie L., Winters, Alan, L., 2007. Quantifying international migration: a database of bilateral migrant stocks. Policy Research Working Paper Series 4165. The World Bank.
Peri, Giovanni, 2011. Rethinking the area approach: immigrants and the labor market in California. Journal of International Economics 84 (1), 1–14.
Pritchett, Lant, 2006. Let their people come: breaking the gridlock on global labor mobility. Center for Global Development. University Presses Marketing, distributor.
Rodrik, Dani, 1995. Why is there multilateral lending? NBER Working Papers 5160. National Bureau of Economic Research, Inc.
Skaksen, Jan, Rose, Malchow-Møller, Nikolaj, Jensen Aastrup, Claus, 2007. Does coordination of immigration policies among destination countries increase immigration? Working Paper 2007-15. Copenhagen Business School, Department of Economics.
Walmsley, Terri Louise, Winters, Alan L., 2003. Relaxing the restrictions on the temporary movements of natural persons: a simulation analysis. CEPR Discussion Papers 3719. C.E.P.R. Discussion Papers.
Winters, L. Alan, 2001. Assessing the efficiency gain from further liberalization: a comment. In: Suave, P., Subramanian, A. (Eds.), Efficiency, Equity and Legitimacy: The Multilateral Trading System and the Millenium. Chicago University Press, Chicago, pp. 106–113.
World Bank, 2006. Global Economic Prospects 2006: Economic Implications of Remittances and Migration. The World Bank, Washington, DC.
World Bank, 2012. World Development Indicators 2012. The World Bank, Washington, DC.
Yaari, Menahem, Bar-Hillel, Maya, 1984. On dividing justly. Social Choice and Welfare 1 (1), 1–24.