International Migration Unions

Nikita Gaponiuk*

September 23, 2020

Abstract

The prevalence of protectionist migration policies lead to the fact that more than 80% of the world population cannot work in any foreign country without a permit. Though a cooperation on labor mobility is an important driver of economic growth, joint economic benefits do not seem to suffice political demands. By endogenizing migration policy in a dynamic gravity setup, we study matching between countries and identify economic and political obstacles of complete liberalization of the labor movement. Based on the analysis of 9 OECD countries, we explain why geographic proximity, trade intensity, similar governmental attitudes and heterogeneity in a dominant type of ownership benefit labor mobility, while difference in technologies and capital intensity does not. Additionally, we argue how redistribution, voting mechanism, and the taste for freedom mutually determine the welfare gains and political viability of migration unions.

Keywords: Migration, union, trade, labor mobility, free movement, cooperation, welfare gains, welfare effects, redistribution, economic growth, voting, Nash equilibrium.

*University of Luxembourg, nikita.gaponiuk@uni.lu. The data on migration flows used in this paper contain noise to prevent a reader from recovering the original dataset until the circumstances under which it can be shared are resolved.

1 Introduction

For more than thirty-five years economic science documents substantial welfare gains from international cooperation on labor mobility[1]. Though free movement agreements became a widespread reality, the right to work in a foreign country still must be earned. This paper attempts to answer the question of why protectionist migration policies, not the cooperation, remain first preference to the majority of countries. Exceptions include European Union, Australia and New Zealand, and a few limited mobility clusters[2]. These are examples of sustainable agreements granting costless visa-free entry and multilateral working permits which we will refer to as migration unions.

The migration literature developed a considerable understanding of general equilibrium welfare effects from two scenarios: complete liberalization of labor movement and border closure. It includes distribution of labor mobility benefits across countries (Docquier et al. (2015)), nationalities and skill types under different trade regimes (Caliendo et al. (2017), Caliendo et al. (2019)), mobile capital (Kennan (2017), Davis & Weinstein (2002), Parro (2013)) and at a precise geographic resolution (Desmet et al. (2018)). Including the latter, some papers analyze partial liberalization (Delogu et al. (2014),Hamilton & Whalley (1984)), however, provide country-specific results given that the liberalization is global. To identify the economic motives of countries to establish a migration union in between two extreme scenarios and isolated from the global exogenous liberalization, we extend a dynamic discrete choice model alike Caliendo et al. (2017) to allow for endogenous migration costs and quantify nationality-, skill- and residence-specific welfare effects from cooperation on labor mobility in the Nash equilibrium. It enables us to analyze multiple counterfactual unions and distinguish between optimal migration policies with and without response to opponents' actions.

Our paper relates to the literature on international unions. Trade unions cooperate on import and export tariffs or quotas to control the movement of goods (Ossa (2014),Wadsworth et al. (2016), Rodrik (2018), Li et al. (2019)), ecological unions (Weale et al. (2002)) may control the emission of negative externalities, economic unions benefit from coordination of public goods (Alesina et al. (2001)) or taxes (Farhi & Werning (2017)), while migration unions cooperate on the labor movement. Regardless the type of the union, its establishment and size condition on economic and political benefits, their trade-off and timing (Picard & Worrall (2016)). Uniformity in preference for public good (Alesina et al. (2005),Harstad

[1] A non-exhaustive list of authors includes Hamilton & Whalley (1984), Moses & Letnes (2004) ,Klein & Ventura (2007), Iranzo & Peri (2009), Clemens (2011), Kennan (2013), Piyapromdee et al. (2014), Di Giovanni et al. (2015), Docquier et al. (2015), Aubry et al. (2016), Caliendo et al. (2017), Kennan (2017), Battisti et al. (2018), Desmet et al. (2018), Lagakos et al. (2020), Zi (2020).

[2] ASEAN, GCC, MERCOSUR and ECOWAS

(2007)), economic size and factor endowment (Gancia et al. (2019)) between countries is argued to positively affect motives for cooperation. On the other hand, excessive centralization, status quo and policy uniformity counteract the emergence of new alliances. This paper examines a vast heterogeneity among countries to provide a comprehensive analysis of novel economic and political factors complicating international cooperation and speculates about the course of actions to improve the political viability of labor mobility liberalization.

Another dimension of cooperation analysis concerns redistribution between and within countries. It is argued that under endogenous migration policies, international transfers are needed for the Nash equilibrium to be Pareto optimal (Casella (2005)). This paper shows that gravity models of migration and trade, employing larger heterogeneity in countries and labor specifications yield Pareto improving outcomes (migration unions) even without international transfers, though a vast majority of results suggests otherwise. Within a destination country, the fiscal policy uniformity is the main obstacle to labor mobility (Guerreiro et al. (2019)). We extend this result to the case of multilateral policy-making.

Our approach stems from trade models alike Artuç et al. (2010), Allen & Arkolakis (2014); Caliendo & Parro (2015) and Dix-Carneiro & Kovak (2017). While in the trade literature, labor mobility is treated as the substitute for trade liberalization, we provide a piece of evidence in favor of complementarity between trade intensity and economic benefits from cooperation on labor mobility.

We assume that households are heterogeneous in their skills, nationalities, and preference of a location. Each period they experience an idiosyncratic shock, incur migration costs to reallocate, supply a unit of labor and consume goods and a value of an option to move. Firms employ both low-skilled and high-skilled labor, rent a country-specific structure resource, and differ in their productivity à la Eaton & Kortum (2002) driven by a dynamic labor density.

Governments provide public goods and maximize the life-time utility of all its nationals over nationality- and skill-specific migration policy costs of entry. Migration costs are expressed in utility units and include costs arising from migration frictions, policy, information asymmetry, cultural and technological gaps. Low migration policy costs make the domestic labor market more vulnerable to economic shocks, while high migration costs put a burden on households, making neither protectionism nor liberalization a universal remedy. Optimal migration policies are determined by the Nash equilibrium.

In the absence of natural experiments, we create a set of counterfactual migration unions to introduce variation in both welfare gains and country-specific characteristics. Each scenario is then solved in the Nash equilibrium using an open-loop procedure and embedded in the sensitivity analysis. In counterfactual equilibrium, the variation in union membership is modeled as an unanticipated exogenous shock. Moreover, we assume that there is full commitment between members of the union. The ability to response can be switched off and on by allowing countries to respond to each others' policy changes. The analysis is extended by the introduction of dynamic political attitudes towards skill and residence that allow the government to individually account for natives' welfare in policy-making.

In the baseline economy of 9 OECD countries and the rest of the world, migration unions are build up based on the Trans-Tasman Travel Arrangement and the Treaty of Amsterdam[3]. Among 29 counterfactual migration unions three are found to be politically viable and nine - economically efficient.

We discover that migration union is beneficial for allying parties similar in technologies, capital intensity, governmental attitudes and opposite in the provision of public goods. Least favorable partners are distant moderately trading countries, having scarce similar types of labor, dominant private ownership, and low capital intensity. Political obstacles to union establishment include bargaining over redistribution, voting mechanism, and the taste for freedom.

Our paper is organized as follows. In Section 2 we expand on the theoretical setup of the model, describe household behavior, the production side, governmental program, and equilibrium conditions in time-differences. Section 3 describes disaggregation and construction of data. Section 4 shows how to estimate parameter and calibrate migration fixed and policy costs. In Section 5 we present results and explain why mismatching between countries largely depends on aforementioned economic and political factors. The rest of the paper consists of a conclusion, bibliography, and appendix.

[3]The paper considers only international agreements valid in 2000. Prior to 2001, the Trans-Tasman Travel Arrangement guaranteed free labor mobility between Australia and New Zealand. The Treaty of Amsterdam entered into force in 1999 and stayed the latest valid treaty between European countries on labor mobility until the Treaty of Nice was enforced in 2003.

2 Migration and trade model

Consider a world with N countries, indexed by i or j, and N nationalities, indexed by n. Each country accommodates S types of workers, literally, high-skilled (h) and low-skilled (l), living for an infinite number of periods denoted by $t = 0, 1, 2....$ Every period begins with a realization of an idiosyncratic taste shock (change of preferences, climate, local economic recession, etc.). Given this information and assumption on the rationality of individuals each government i imposes a set of migration costs $\{\tilde{m}^s_{ji,n,t}\}$ $\forall j, n, t, s$, defined as costs of entering country i from every country j for each period from now on. After the release of migration costs, individuals move across countries, work and consume private and public goods there until a new realization of the shock.

Households optimally decide on the country to work and consume in. Every time they move they face a migration cost and a change in idiosyncratic location preference. The latter allows us to conveniently aggregate household solutions alike Desmet et al. (2018) and others. Along with the paper, we are closely following Caliendo et al. (2017).

In each country, there is a competitive labor market. Firms compete in prices within a country and internationally. They employ CES production function with heterogeneous labor and fixed land resource. As in Eaton and Kortum (2002), production is subject to a country-specific productivity distribution.

In this Section, we setup the dynamic household problem, firms behavior, and explain how wages and prices are chosen in equilibrium. Then we describe how each government chooses optimal migration costs and why they are different from those chosen in the Nash equilibrium.

2.1 Households

Assume that within a period of time t, each household of nationality n and skill type s residing in country i receives a logarithmic utility from the consumption of a private ($C^s_{i,t}$) and a public ($G_{i,t}$) good. A private good is an aggregate of varieties denoted by $\omega \in \Omega_{ji,t}$, where $\Omega_{ji,t}$ is the set of varieties that country i buys from the set of best suppliers[4] denoted by j in period t. Denote $\tilde{c}^s_{j,t}(\omega)$ as the demanded quantity of a variety, then $C^s_{i,t} = \left[\int_{\omega\in\Omega_{ji,t}} \tilde{c}^s_{j,t}(\omega)^{\frac{\rho-1}{\rho}} d\omega\right]^{\frac{\rho}{\rho-1}}$ is the aggregated private good, where ρ is the constant elasticity of substitution between varieties. Household supplies unit of labor inelastically and receive an exogenous amount of a public good from a government[5]. Hence, at each period in time

[4]Each variety ω is imported only from one supplier with the lowest price given transportation costs. Therefore, j does not denote a country, but a set of countries where best suppliers locate.

[5]For details on public good provision see subsection 2.4

a household solves the following problem:

$$\max_{\{\tilde{c}^s_{j,t}(\omega)\}_{\omega\in\Omega_{ji,t}}} \quad \mathrm{U}^s_{i,t} \equiv \alpha_i \log G_{i,t} + (1-\alpha_i)\log C^s_{i,t}, \tag{1}$$

$$\text{s.t.:} \quad \int_{\omega\in\Omega_{ji,t}} \tilde{p}_{j,t}(\omega)\tilde{c}^s_{j,t}(\omega)d\omega \leq \mathrm{w}^s_{i,t}(1-\psi_i). \tag{2}$$

where α_i is the weight of the public good in preferences, $\tilde{p}_{j,t}(\omega)$ is the price of variety ω, $\mathrm{w}^s_{i,t}$ is the nominal wage and ψ_i is the labor income tax.

The Dixit-Stiglitz price index associated with the utility function[6] is given by $P_{i,t} = \left(\int_{\omega\in\Omega_{ji,t}} \tilde{p}_{j,t}(\omega)^{1-\rho}d\omega\right)^{\frac{1}{1-\rho}}$. It can be used to express consumption of the private good via country-specific real wage as follows: $C^s_{i,t} = \frac{\mathrm{w}^s_{i,t}(1-\psi_i)}{P_{i,t}}$.

Let us now assume that the household lives an infinite number of periods and at the beginning of each period it can change the country of residence. Denote $(\tilde{m}^s_{ij,n,t})$ as the migration costs associated with the movement from country i to country j by a household of nationality n and skill type s and $(\epsilon^s_{i,n,t})$ as the idiosyncratic preference of a new location. Both migration costs and shocks to the preferences determine household decision to move. Then, the Bellman equation describing optimal household dynamic behavior will be given by:

$$v^s_{i,n,t} = \mathrm{U}^s_{i,t} + \max_{j=1\ldots N}\left\{\beta E[v^s_{j,n,t+1}] - \tilde{m}^s_{ij,n,t} + \nu\epsilon^s_{j,n,t}\right\}. \tag{3}$$

where $v^s_{i,n,t}$ is the life-time utility of a type s household of nationality n residing in country i at time t, β is a discount and ν is the migration elasticity that scales the variance of the location preference.

To precise the meaning of the equation (3) two fundamental assumptions are needed:

Assumption 1 *Migration costs $\tilde{m}^s_{ij,n,t}$ depend on country of origin (i), country of destination (j), nationality (n), skill type (s) and are additive, measured in terms of utility and endogenously chosen by the country of destination.*

Migration costs $(\tilde{m}^s_{ij,n,t})$ consist of time-varying endogenous policy $(m^{s,pol}_{ij,n,t})$ and fixed exogenous $(m^{s,fix}_{ij,n})$ components such that:

$$\tilde{m}^s_{ij,n,t} = m^{s,pol}_{ij,n,t} + m^{s,fix}_{ij,n}. \tag{4}$$

It is assumed that the fixed part includes costs associated with moving expenses, migration

[6] For detailed derivation see Appendix.

frictions, information asymmetry, cultural and technological gap[7]. All other sources are assumed to be controlled by the government and, therefore, claimed to be endogenous. They will be referred to as migration policy costs. Since an arbitrary change in migration policy of a country would affect all immigrants, including those who are currently residing in there[8], staying in the same location obliges an immigrant to pay policy-related part of migration costs, but does not the native[9]. Moreover, the structure of migration costs also imposes that return migration is free from the policy costs and there are no fixed costs for staying in the same country:

$$m^{s,fix}_{ii,n} = 0, \quad \forall\ i, \tag{5}$$
$$m^{s,pol}_{ji,n,t} = m^{s,pol}_{ki,n,t}, \quad \forall\ j,k, \tag{6}$$
$$\tag{7}$$

In each period a government has to announce a set of migration costs for households to move. The government controls only migration policy costs and can not change the fixed part. Countries can discriminate migrants by nationality and skill type, but not by the origin[10]. In other words, migration policy costs do not depend on migrant's last country of residence, only on her nationality and skill type[11]. Recall, that return migration is free of policy costs[12] and there are no fixed costs for staying in the same country. Migration costs are pure losses and not reinvested into the economy.

More precisely, using (6) and variations of (67) for different combinations of i,j and n policy and fixed migration costs can be derived directly:

$$m^{s,fix}_{ij,n} = (m^{s,fix}_{in,n} - m^{s,fix}_{jn,n}) + \nu \log\left(\frac{\mu^s_{in,n,0}}{\mu^s_{ij,n,0}} \cdot \frac{\mu^s_{jj,n,0}}{\mu^s_{jn,n,0}}\right), \tag{8}$$

[7]Consequently, any quantitative results represented in this paper are valid only within periods short enough to keep the technological gap, moving expenses, and other aforementioned sources of costs, stable among all countries.

[8]Consider, for example, tightening the visa regime or shortening the duration of residence permits.

[9]Payments of migration costs do not diminish in time since migrants are not allowed to acquire multiple citizenship or change it.

[10]Therefore, for example, the model will not be able to account for a fact that an immigrant once resided in Israel, will be forbidden to enter the territory of Lebanon and vice versa.

[11]A non-European citizen entering the EU labor market from homeland or EU-member country will have the same migration policy costs but different fixed migration costs

[12]A househould of nationality i returning from country j to homecountry i pays only fixed part of the migration costs.

$$m^{s,pol}_{ij,n,0} = (1-\beta)m^{s,fix}_{jn,n} + \nu \log\Big(\frac{((G_{j,1})^{\alpha_j}(C^s_{j,1})^{(1-\alpha_j)})^{\beta/\nu}}{((G_{n,1})^{\alpha_n}(C^s_{n,1})^{(1-\alpha_n)})^{\beta/\nu}} \cdot \frac{\mu^s_{jn,n,0}}{\mu^s_{jj,n,0}} \cdot \frac{(\mu^s_{nn,n,1})^{\beta}}{(\mu^s_{jn,n,1})^{\beta}}\Big). \tag{9}$$

It can be seen that fixed migration costs are smaller for those destinations that are more popular among migrants. However, popularity comes with a cost. This cost is accumulated in migration policy costs that become larger if the expected utility of the destination rises. Additionally, equations (68) and (69) capture the equivocal diaspora effects. While the presence of similar migrants $\mu^s_{jj,n,t}$ decrease policy costs, it increases fixed part and reflects the fact that the marginal impact of diaspora decreases with the size of the group.

Assumption 2 *The idiosyncratic preference of a location $\epsilon^s_{i,t}$ is i.i.d. over time and distributed Type-I Extreme Value with zero mean*[13]*.*

This assumption allows us to simply aggregate decisions of heterogeneous agents. Denote $V^s_{i,n,t} \equiv E[v^s_{i,n,t}]$ as the expected life-time utility of a representative household, where the expectation is to be taken over the idiosyncratic preference. Then, it can be shown[14] that:

$$V^s_{i,n,t} = \mathrm{U}^s_{i,t} + \nu \log\left(\sum_{j=1}^{N} e^{\frac{1}{\nu}(\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t})}\right). \tag{10}$$

Due to the properties of Type-I Extreme Value distribution, the stochastic dynamic discrete choice model becomes deterministic. In addition, following the same logic, it is easy to find a closed-form solution for a fraction of people of type s and nationality n moving from country i to country j in period t, denoted by $\mu^s_{ij,n,t}$:

$$\mu^s_{ij,n,t} = \frac{e^{\frac{1}{\nu}(\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t})}}{\sum_{k=1}^{N} e^{\frac{1}{\nu}(\beta V^s_{k,n,t+1} - \tilde{m}^s_{ik,n,t})}}. \tag{11}$$

Therefore, immigration is larger for countries with low entry barriers and high expected lifetime utility. Notice, that the fraction of people moving depends on a country of origin only because of the migration costs. In the absence of those, the equation (11) would become only nationality and destination specific and, consequently, the fraction of people moving to country j would be the same for all countries i.

Denote $L^s_{i,n,t}$ as the current population of nationality n and skill type s residing in country i. Assume that the total world population is fixed. Given equation (11), the allocation of people in the next period can be expressed as follows:

[13] The idiosyncratic preference of a location does not appear in equation (1) and (3) in the current period for the reason that $\int \epsilon f(\epsilon) d\epsilon = 0$. For details on the functional form of probability distribution see Appendix.

[14] For derivations see Appendix 6

$$L^s_{i,n,t} = \Big(\sum_{j=1}^{N} \mu^s_{ji,n,t-1} L^s_{j,n,t-1}\Big), \tag{12}$$

$$L^h_{i,t} = \sum_{n=1}^{N} L^h_{i,n,t}, \qquad L^l_{i,t} = \sum_{n=1}^{N} L^l_{i,n,t}, \qquad L_{i,t} = L^h_{i,t} + L^l_{i,t}. \tag{13}$$

In other words, in period $t+1$ country i will accommodate high-skilled and low-skilled people of all nationalities both who decided to stay and who decided to migrate from all other countries into country i. The vector $L_t \equiv \{L^s_{i,n,t}\}^{N,N}_{i=1,n=1,s\in(l,h)}$ describes the state of the world economy in period t and is determined by a set of decisions taken in period $t-1$. Given the information about the initial allocation of people and migration costs, the model can be solved for equilibrium wages and prices. Let us now describe the firms' behavior and trade dynamics.

2.2 Production

All the firms in the economy are perfectly competitive. They employ low-skilled, high-skilled labor, and fixed endowment of structures to produce goods. Firms' efficiency depends on the country's access to technologies. As in Eaton & Kortum (2002) we assume that $z_{i,t}(\omega)$ is the country i's efficiency in producing variety ω at time t. Then, the production function of each firm producing variety ω is given by:

$$Y_{i,t}(\omega) = z_{i,t}(\omega) \left[\Big(\theta^h_i (L^h_{i,t})^{\frac{\sigma-1}{\sigma}} + \theta^l_i (L^l_{i,t})^{\frac{\sigma-1}{\sigma}} \Big)^{\frac{\sigma}{\sigma-1}} \right]^{(1-\gamma_i)} \Big(H_i\Big)^{\gamma_i}. \tag{14}$$

where $L^h_{i,t}$ is the stock of high-skilled labor, $L^l_{i,t}$ is the stock of low-skilled labor, $\theta^h_i + \theta^l_i = 1$ are shares of high-skilled and low-skilled labor in the production, σ is the elasticity of substitution between labor factors, γ_i is the share of capital payments in value-added, and H_i is the structure resource.

It is assumed that firms pay a competitive rent ($r_{i,t}$) to use a country-specific and time-invariant factor of structures (H_i). Rental payments ($r_{i,t}H_i$) in all countries together constitute a global portfolio $\chi_t = \sum_{i=1}^{N} r_{i,t} H_i$. To match trade imbalances in the data, we assume that only a share ι^i of a global portfolio returns to the country i and then is uniformly distributed among workers of that country. Therefore, the aforementioned labor income tax ψ_i applies to both wages and rents. With the increase in production of a particular variety, the rent rises and shifts the cost function upwards, therefore, limiting the production and trade capability of a country as well as the wage bill of the worker. However, after redistribution of the global portfolio the wage bill rises in all countries but only in some of them sufficiently high to overcome the congestion from limited structures. Losses in wages come from rental payments because they are not being fully reinvested into the home economy.

We assume that a country's productivity $z_{i,t}(\omega)$ follows Frechet distribution such that:

$$F_{z_{i,t}(\omega)}(z) = e^{-T_{i,t}z^{-\theta}}, \quad T_{i,t} = \eta_i\big(D_{i,t}\big)^{\zeta}. \tag{15}$$

The level of technological development ($T_{i,t}$) depends on the constant level of technology (η_i) and is a function of labor density with the static elasticity of productivity to labor density (ζ), where ($D_{i,t}$) is the population density measured as a ratio between labor ($L_{i,t}$) and urban land area. Therefore, firms gain productivity from an increase in the total population.

In general, economic growth highly depends on the stock of the active labor force. A country experiencing a lack of labor may replenish it via two channels: increase the number of natives (by managing fertility and return migration) or increase the number of immigrants. Countries having productivity inelastic to labor density would prefer to replenish the stock of labor force with natives because welfare gains associated with a marginal increase in productivity are to be distributed exclusively to natives. Therefore, those countries would increase migration policy costs and, consequently, crowd out migrants with returned emigrants. Countries having large elasticity would fill in the lack of labor with migrants because even after migrant's consumption and remittances native population still gain from the rise of productivity. And, thus, these countries would prefer to follow pro-migrant policies. Countries in between would partially discriminate some of the migrants' groups. It has to be mentioned, that, even though elasticity of productivity to labor density is an important parameter, it does not solely determine migration policy[15].

CES production function implies that in equilibrium skill premium will be proportional to a skill-labor ratio and the ratio of labor-share parameters:

$$\frac{\mathrm{w}^h_{i,t}}{\mathrm{w}^l_{i,t}} = \frac{\theta^h_i}{\theta^l_i}\Big(\frac{L^l_{i,t}}{L^h_{i,t}}\Big)^{\frac{1}{\sigma}}. \tag{16}$$

The equation (16) characterizes how firms adjust wages for two groups of workers. Given all the rest fixed, a higher supply of one labor type will raise the relative wage of the other type, or an increase in the share of high-skilled labor will raise its importance as well as wages.

The input bundle cost associated with the production function and wage ratio is denoted by $x_{i,t}$. It can be expressed in term of efficiency units or only through wages, rent and share

[15]See details of sensitivity analysis in subsection 5.3

parameters[16]:

$$x_{i,t} = \Omega^i \left[(\theta_i^h)^\sigma (\mathrm{w}_{i,t}^h)^{1-\sigma} + (\theta_i^l)^\sigma (\mathrm{w}_{i,t}^l)^{1-\sigma} \right]^{\frac{1-\gamma_i}{1-\sigma}} r_{i,t}^{\gamma_i}, \quad \text{where} \quad \Omega^i = \frac{1}{1-\gamma_i} \Big(\frac{1-\gamma_i}{\gamma_i} \Big)^{\gamma_i}. \tag{17}$$

The input cost determines the competitive advantage of a firm on the international market. Following Eaton and Kortum (2002), goods will be purchased at the second-best price, therefore, leading to a particular distribution of prices and bilateral trade shares across countries[17]. Given the assumption on the distribution of country-specific productivity and the input bundle costs, it can be shown that the share of income $\pi_{ij,t}$ spent by country i on goods from country j will be given by:

$$\pi_{ij,t} = \frac{X_{ij,t}}{X_{i,t}} = \frac{T_{j,t}/(x_{j,t}k_{ij,t})^\theta}{\sum_{b=1}^N T_{b,t}/(x_{b,t}k_{ib,t})^\theta}, \qquad k_{ij,t} = (1+\tau_{ij,t})d_{ij,t}. \tag{18}$$

where d_{ij} is the geographic barrier or "iceberg" costs between country i and country j, $\tau_{ij,t}$ is the exogenous import tariff imposed by country i on goods imported from country j, $X_{ij,t}$ is the expenditure of country i on goods from country j, $X_{i,t}$ is the total expenditure of country i on goods, including it's own production.

Prices among different firms within one country can be aggregated by a price index, that will depend on the production efficiency, geographic and trade barriers of all countries trading with the country of origin:

$$P_{i,t} = \bar{p} \Big(\sum_{b=1}^N T_{b,t}/(x_{b,t}k_{ib,t})^\theta \Big)^{-\frac{1}{\theta}}, \qquad \bar{p} = \Big[\Gamma(1 + \frac{1-\rho}{\theta}) \Big]^{\frac{1}{1-\rho}}. \tag{19}$$

where ρ is the elasticity of substitution between varieties.

Equations (18) and (19) describe that countries with better technologies have lower input bundle costs and, therefore, trade with the rest of the world more intensively, however, the amount of trade is being constantly discounted by the measure of the distance between countries.

Overall, migration influences firms behavior through labor supply. Low migration costs lead to an inflow if immigrants which in turn decreases wages and the cost of the input bundle. Larger labor density improves technologies and decreases an amount of factors required for a production of one unit of a variety. Eventually, the inflow reduces prices in a country and increases trade.

[16]For details on profit maximization and related derivations see Appendix

[17]For detailed derivations see Appendix.

2.3 Market clearing

First, we assume that to provide public goods the government funds the budget via two channels: import tariffs (τ_{ij}), labor income tax (τ_I^L) such that:

$$P_{i,t}G_{i,t} = \sum_{j=1}^{N} \tau_{ij,t} \frac{\pi_{ij,t}}{(1+\tau_{ij,t})} X_{i,t} + \sum_{s=h,l} \psi_i \mathrm{w}_{i,t}^s L_{i,t}^s, \tag{20}$$

In equilibrium, all goods produced has to be consumed. Since countries can not borrow money in this model and household income comes only from the labor market and rental payments, it is enough to state that for every country i expenditure on goods imported has to equal the value of goods exported. Thus, given the labor allocation in period t, wages and prices can be estimated using the following labor and goods market-clearing conditions.

Goods and structure market-clearing:

$$X_{i,t} = P_{i,t}G_{i,t} + \sum_{s=h,l} (1-\psi_i)\mathrm{w}_{i,t}^s L_{i,t}^s + \iota^i \chi_t, \tag{21}$$

$$\chi_t = \sum_{i=1}^{N} r_{i,t}H_i, \qquad r_{i,t}H_i = \gamma_i \sum_{j=1}^{N} \frac{\pi_{ji,t}}{(1+\tau_{ji,t})} X_{j,t}. \tag{22}$$

Rental payments enter the total expenditure of every country i on all goods, $X_{i,t}$, and then used to calculate equilibrium wages. Labor market-clearing can be expressed as follows:

$$\mathrm{w}_{i,t}^h L_{i,t}^h + \mathrm{w}_{i,t}^l L_{i,t}^l = (1-\gamma_i) \sum_{j=1}^{N} \frac{\pi_{ji,t}}{(1+\tau_{ji,t})} X_{j,t}, \tag{23}$$

$$\frac{\mathrm{w}_{i,t}^h}{\mathrm{w}_{i,t}^l} = \frac{\theta_i^h}{\theta_i^l} \Big(\frac{L_{i,t}^l}{L_{i,t}^h} \Big)^{\frac{1}{\sigma}}. \tag{24}$$

Equation (23) describes the fact that larger rent payments lead to a larger wage bill. However, only a part of rental payments associated with ι^i is, eventually, returned. In this way, rental income per employee is the congestion force. It decreases with labor density and, consequently, reduces private consumption. Equation (24) determines how the wage bill is being distributed between low-skilled and high-skilled workers.

2.4 Migration unions and the role of government

Now, we will introduce the notion of *migration union*. Denote migration union M as a set of countries that have established a multilateral free labor mobility agreement. Then, let M_i be a migration union that country i is a member of. Free mobility agreement implies that there are no migration policy costs of entering member countries for any of the member nationals:

$$m^{s,pol}_{ji,n=k,t} = 0, \quad \forall\ k \in M_i,\ \forall\ j. \tag{25}$$

The assumption on full commitment means that each government is eligible to change migration policy costs only for those nationals that are not members of the migration union. When two parties are set to have a union, migration policy costs between them are being exogenously lowered to zero starting from the first period. Moreover, it is assumed that parties commit to keep an agreement forever ($t = \overline{1,\infty}$). Therefore, a counterfactual is different from the baseline in the set of migration unions and their structure, $M_i\ \forall i$.

Before we proceed to the role of government let us denote $\mathbf{m}_{\neg i,t} \equiv \{m^{s,pol}_{jk,n,t}\}^{N,N,N,\infty}_{j=1,k\neq i,n=1,t,s=l,h}$ as the set of migration policy costs decided by all countries except for country i and considered by country i to be fixed in time regardless of its own migration policy. In this way we build a myopic expectation of each government about other countries' optimal decisions. Then, for simplicity we denote $\mathbf{m}_{i,t} \equiv \{m^{s,pol}_{ji,n,t}\}^{N,N,\infty}_{j=1,n=1,t,s=l,h}$ as the vector of migration policy costs that government i optimizes over. Therefore, $\mathbf{m}_{\neg i,t}$ and $\mathbf{m}_{i,t}$ together constitute the full set of migration policy costs.

We assume, that besides public goods provision, each government decides on the lifetime migration policy ($\mathbf{m}_{i,t}$) to maximize weighted utility of its nationals. Define the welfare function of the government i as:

$$W_{i,0} = \sum_{t=0}^{\infty} \beta^t \Bigg[\sum_{r=1}^{N} \lambda^r_{i,n,t} \Big(\sum_{s=l,h} \lambda^s_{i,t} L^s_{r,n=i,t}\big(\mathbf{m}_{i,t}, \mathbf{m}_{\neg i,t}\big) V^s_{r,n=i,t}\big(\mathbf{m}_{i,t}, \mathbf{m}_{\neg i,t}\big)\Big)\Bigg], \tag{26}$$

where $\lambda^h_{i,q} + \lambda^l_{i,q} = 1$ corresponds to the governmental attitudes towards high-skilled and low-skilled groups in a country, $\sum_{r=1}^{N} \lambda^r_i = 1$ are governmental attitudes towards natives' country of residence (r) and the subscript $n = i$ means that only natives of country i are taken into account.

A recent paper by Alesina & Stantcheva (2020) shows that accounting for governmental attitudes towards unlike groups of people shapes preference for redistribution and, therefore, plays an important role in determining governmental policies. Instead of social marginal welfare weight, our paper assigns weights to every native's welfare and define them as governmental attitudes.

$$\lambda^s_{i,t} = L^s_{i,t}/L_{i,t}, \qquad \lambda^r_{i,n,t} = L_{i,n,t}/L_n. \tag{27}$$

Given equation (27), these weights are dynamic[18], follow the majority, and can only influence the supply of labor. If the number of low-skilled workers in country i increases, the average wage of this group of workers decreases and it motivates the government to decrease low-skilled labor supply by lifting corresponding migration costs. It guarantees that political power will be used to balance the interests of groups. It works similarly with the weights attached to the geographic distribution of natives.

It is reasonable to assume that the government cares about natives differently because they contribute taxes to the homeland budget unequally depending on the current residence and skill. Nevertheless, even though some emigrants do not pay taxes at all, the government can not completely ignore its emigrants' well-being. Almost every government has an Embassy and Consulate in a major emigration country. Apart from citizenship and passport services, they provide legal and medical assistance, child and family matters, voting and political refuge, inform about commercial opportunities, maintain an environment for establishing educational and cultural exchanges. Therefore, even though emigrants' well-being receives less support than natives, it is still an important public concern.

In their Handbook, Rosenblum & Tichenor (2012) highlight the importance of political aspects in immigration studies: *"...even with perfect and complete information about immigrants' roles in the economy, ..., analysts could not calculate an ideal number of immigrants, because the main effects of migration are distributive and thus reflect competing political demands".* We attempt to approximate political demands by assuming that they are included into governmental attitudes.

It can be shown[19] that in equation (26) the value function ($V^s_{r,n=i,t}$) can be expressed as a function of consumption and a share of people willing to stay in the country of residence:

$$V^s_{r,n=i,0} = \sum_{t=0}^{\infty} \beta^t \left[\log \frac{(G_{r,t})^{\alpha_r} (C^s_{r,t})^{(1-\alpha_r)}}{(\mu^s_{rr,n=i,t})^{\nu}} - m^{s,pol}_{rr,n=i,t} \right], \text{ where } m^{s,pol}_{rr,n=i,t} = 0 \ \forall i = r. \tag{28}$$

The equation (28) tells that the value of a worker is influenced by all three forces: utility from total consumption, the value of an option to move, and cost of residing in the same country. Therefore, a higher share of people staying in the same country, $\mu^s_{rr,n=i,q}$, means that the value of an option to move from country r is sufficiently large to ignore some of the destination countries where potential consumption is larger than in r. To solve the equilibrium equation (28) has to converge to a fixed point.

In subsection 2.5 we discuss how to solve numerically system of equations (30) and what

[18]Static weights of 0.5 for skills and $\frac{1}{N}$ for residence would correspond to egalitarian society

[19]See Appendix 6

necessary condition for convergence toward Nash equilibrium is. Now let us describe the market clearing.

2.5 Equilibrium

The state variable that describes the position of the dynamic model in each period of time is the allocation of labor across countries and time $L_t \equiv \{L_{i,t}\}_{i=1}^{N}$. Denote the set of controlled migration policy costs by $m_t \equiv \{m_{ij,n,t}^{pol}\}_{i=1,j=1,n=1}^{N,N,N}$, varying fundamentals by $\Theta \equiv \{\tau_{ij,t}\}_{i=1,j=1}^{N,N}$, and time-invariant fundamentals by $\bar{\Theta} \equiv \{m_{ij,n}^{fix}, d_{ij}, H_i, \psi_i, \eta_i\}_{i=1,j=1}^{N,N}$. The set of time-varying variables contains bilateral trade shares, input bundle costs, prices and wages. Their estimation becomes easier in the model in time differences.

Denote any variable $\dot{y}_{t+1} = \frac{y_{t+1}}{y_t}$ as a variable in time differences. It turns out that any model of this kind has a few convenient properties. First of all, notice that the steady-state value of any variable in time differences equals one by construction. Secondly, it allows us to reduce the dimensionality of transition dynamics by eliminating several time-invariant variables needed to estimate the model. However, to infer levels of endogenous variables, we will still have to execute additional computations that we describe later.

Before going to a solution of the model in time differences, an equilibrium in the baseline model should be defined.

Definition 1. *Given $(L_t, \bar{\Theta}, \Theta)$, a **temporary equilibrium** is a vector of factor prices $\{w_{i,t}^l, w_{i,t}^h, r_{i,t}, P_{i,t}\}_{i=1}^{N}$, that solves the static system of equations (17) to (19) and (21) to (24).*

Wages, rental rate, and price index are then used to calculate aggregate consumption. Then, to solve the model we incorporate information about migration policies and calculate lifetime utilities of households in a conditional sequential equilibrium:

Definition 2. *Given $(L_t, \bar{\Theta}, \Theta)$, levels of variables $\{\pi_{ij,0}, \mu_{ij,n,-1}^s, w_{i,0}^s\}_{i=1,j=1,n=1,s=l,h}^{N,N,N}$ and assuming that a sequence of migration costs $\{m_t\}_{t=0}^{\infty}$ is converging such that $\lim_{t\to\infty} \dot{m}_t = 1$, a **conditional sequential equilibrium** is a sequence of $\{L_t, \mu_{ij,n,t}^s, V_{i,n,t}^s, w_{i,t}^s, P_{i,t}\}_{i=1,j=1,n=1,t=0,s=l,h}^{N,N,N,\infty}$ that satisfies the equilibrium conditions in equations (10) to (12) and temporary equilibrium.*

In the Nash equilibrium, we require all maximizing agents to account for other agents' responses and an infinite horizon of the repeated game. Eventually, each agent will converge to a sequence of controls, by deviating from which it can not obtain a higher payoff.

Denote any arbitrary variable (y^*) as the optimal value. Then, given that all constraints in equations (5), (6) and (25) are satisfied, we can express optimal migration policy costs of entering country i by:

$$\mathbf{m}^*_{i,0} = \text{argmax} \quad W_{i,0}\big(\mathbf{m}_{i,0}, \mathbf{m}_{\neg i,0}\big). \tag{29}$$

However, since countries update their choices given other countries' responses, the Nash equilibrium values of migration policy costs are defined by:

$$\mathbf{m}^*_{i,0} = \text{argmax} \quad W_{i,0}\big(\mathbf{m}_{i,0}, \mathbf{m}^*_{\neg i,0}\big). \tag{30}$$

Definition 3. *A* ***Nash equilibrium*** *is a converging sequence of $\{m_t\}_{t=0}^{\infty}$ that solves simultaneously the welfare maximization problem in (30) for each country in every period of time.*

To ensure that the solution exists and a dynamic system can find a fixed point the welfare function has to be concave in state and control variables. The existence of a solution in Definition 3 allows us to estimate counterfactual by imposing and breaking migration unions as well as by consequently changing its membership.

Numerically, the Nash equilibrium is solved using an open-loop procedure that updates guesses on life-time wages, prices, utilities, and migration policy costs at the same time. The update of migration policy costs follows the gradient calculated using the Broyden-Fletcher-Goldfarb-Shanno algorithm.

In this dynamic model, it is assumed that from some point in time the system should come to a stationary equilibrium.

Definition 4. *A* ***stationary equilibrium*** *is the Nash equilibrium such that for all t there is a constant set of $\{L_t, \mu^s_{ij,n,t}, V^s_{i,n,t}, w^s_{i,t}, P_{i,t}\}_{i=1,j=1,n=1,t=0,s=l,h}^{N,N,N,\infty}$.*

To solve conditional sequential and Nash equilibrium we employ an open-loop procedure. The latter also requires the knowledge of the welfare associated with a choice of optimal controls. From the equation (28) we can observe that the difficulty is to compute the level of consumption since there are prices that are endogenous[20]. To handle this issue notice that (18) together with (19) give:

$$P_{i,0} = T_{i,0}^{-\frac{1}{\theta}} \bar{p} x_{i,0} \pi_{ii,0}^{\frac{1}{\theta}}. \tag{31}$$

Then, real consumption in the initial period can be easily calculated using equation (31). Hence, by consequently multiplying the initial consumption level with the changes in the real wage and public goods we can compute all the values of consumer's value function and

[20]Prices in the model at initial period are normalized to US price index.

corresponding welfare function of the government. However, it is necessary to first estimate a set of parameters that we discuss in Section 4.

An exogenous establishment of a migration union between a pair of countries changes the Nash equilibrium outcomes. From the union member's point of view, the set of controls to optimize over becomes smaller and so more thorough discrimination is possible. It might lead to a situation when a migration union becomes very expensive for the rest of the world. Within the union one country necessarily experience an outflow of one type of labor and another - an inflow. Therefore, it is expected that countries having opposite types of labor mutually scarce would benefit from open borders. The labor in the country with a larger marginal scarcity will gain more in wages that the partner country. Labor movement between allying parties is expected to be influenced by the supply of public goods. People are more probable to move to a country with a high share of public goods. Countries become less exposed to local economic shocks but partially lose independence in migration policy and, thus, are vulnerable to the labor supply shocks in the partner country.

3 Data

This paper uses a unique dataset on bilateral migration flows disaggregated by skill and nationality for years: 2000 and 2005. Additionally, these flows represent economic migration, i.e. migration caused purely by economic reasons (higher wage, better living standards, etc.). Being the most restrictive dataset in the paper, it determines the choice of countries and the period for analysis. The largest set of countries for which all data is available (regardless of a few omitted values) includes Australia, Canada, Israel, Korea (Republic of), New Zealand, Portugal, Sweden, the United Kingdom and the United States. The Rest of the World (RoW) aggregates available data for all other countries. The time begins in 2000 and continues infinitely in increments of 5 years.

It is important to notice that each country in the paper corresponds to a different nationality, which allows for multiple counterfactual scenarios. In their research Caliendo et al. (2017) simulate the model for 17 countries, but migrants can only be disaggregated into 3 nationalities: EU-15, New Member States, and Other. Fortunately, the methodology developed in that paper allows us to enlarge the set of nationalities. Following Caliendo et al. (2017), RoW is treated as a separate aggregated country with its own nationality. Consequently, migrants residing in RoW can only be natives of 9 aforementioned OECD countries. Because RoW accounts for 92% of the world population and aggregates a lot of heterogeneous countries, it's policies assumed to be constant along the whole model's horizon and, likewise, OECD countries keep their policies towards RoW unchanged.

3.1 Migration and Labor Market Data

International migrant stock by nationality and skill

To create the data on the number of migrants disaggregated by skill, nationality, and country of residence one can use the United Nations Database (International migrant stock 2019) and the Database on Immigrants in OECD and non-OECD Countries (DIOC). The UN provides total migrant stock for 232 countries from 1990 to 2019, which makes it suitable for constructing RoW values. The DIOC is a survey conducted across 34 destination countries and more than 200 countries of origin and contains ISCED levels of education. Therefore, it is possible to infer the share of high-skilled (tertiary education) and low-skilled migrants and apply them to a complete UN database. Omitted values occur for Israel, Korea, and New Zealand. They have been substituted with the data from national statistical bureaus. Additionally, in cases where DIOC presented a number of migrants different from UN the largest has been chosen. The number of citizens was calculated by subtracting the stock of migrants from the total population (from UN Population Databases). The mean share of migrants in the total population is 15% for OECD countries and 0.1% for the RoW.

Origin-destination migration flows by nationality and skill

The main source of data on migration flows is the dataset used in Czaika & Parsons (2017) and Czaika & Parsons (2018). In the scope of this paper, it disaggregates bilateral economic migration flows from 180 destinations to 9 OECD countries by skill and nationality, but not by origin. However, the second source of data provided by United Nations[21] does disaggregate migration flows by origin/destination to and from 45 countries. A combination of the latter and stocks of migrants allows to construct a matrix with probability[22] that a migrant of a nationality (n) comes from a particular origin (i) as well as approximate flows to the RoW. Values of return migration and a few migration flows for Israel, Korea, and New Zealand are outsourced from national statistical bureaus.

Wages of low-skilled and high skilled labor

Following Caliendo et al. (2017), for all countries besides Israel, New Zealand and RoW, wages are calculated as a ratio of labor compensation and the number of persons engaged. WIOD Socio-Economic Accounts (Timmer et al. (2015)) provides data on high-skilled and low-skilled labor compensation, the number of hours worked by each group, and the total number of persons engaged in labor. By assuming that both groups work the same amount of time per day, one can easily calculate high-skilled and low-skilled wages.

Israel's statistical bureau provides data on gross income by years of schooling. The high-

[21] International migration flows to and from selected countries: The 2015 revision

[22] For the construction of the probability matrix, see Appendix.

skilled wage was constructed as a population-weighted average gross income of employees having 13 years of schooling or more. By analogy, low-skilled wage contains information on those with less than 13 years of schooling.

World Penn Tables contain data on total labor compensation and the number of persons employed. First, the average wage was calculated as a ratio of both components for New Zealand and as population-weighted ratio for RoW based on 125 countries. Second, using OECD data on skill-premium (OECD (2019)), the population of low-skilled and high-skilled workers, and definition of skill-premium, the total average wage was disaggregated into high-skilled and low-skilled wage. For RoW, the average OECD skill-premium of 151% was used.

Greece - Chassamboulli & Palivos (2013)

3.2 Trade Data

Data on trade volumes are gathered from two sources: IMF Trade Statistics and World Trade Flows (WTF) from UC Davis. IMF Database provides data on imports reported on a cost, insurance, and freight (CIF) for all countries from each corresponding partner and was used for calculation of trade shares.

The methodology for construction of bilateral trade shares can be found in Eaton & Kortum (2002), Bernard et al. (2003) or Waugh (2010). This paper follows the same approach with the only difference that instead of manufactures the data on GDP from OECD National Accounts has been used. One can calculate the share of goods imported by country i from partner j, ($\pi_{ij,t}$), following 2 step procedure:

$$\pi_{ij,t} = \frac{\text{Imports}_{ij,t}}{\text{GDP}_{i,t} + \sum_{j=1}^{N}(\text{Imports}_{ij,t} - \text{Exports}_{ij,t})}, \quad \forall i \neq j, \tag{32}$$

$$\pi_{ii,t} = 1 - \sum_{j \neq i} \pi_{ij,t}. \tag{33}$$

The data on trade shares are also used in the estimation of technology parameters (η_i, α). UC Davis provides data on bilateral trade flows for each year from 1984 to 2016. The number of countries covered by WTF increases around 1992-1993 due to USSR dissolution. Therefore, the time span for calibration covers 1993-2016.

Besides trade shares, the model accounts for the dynamics of trade policies by using a series of trade tariffs. WITS aggregates ad-valorem tariffs from three main sources of: TRAINS, UN Comtrade and WTO. For all countries, except for Israel, the TRAINS data has been used. Tariffs imposed by Israel are supplemented from the WTO database. The series of tariffs include bilateral tariffs for 2000, 2005, 2010 and 2015. They change accordingly with the time in the model and after 2015 rest unchanged.

4 Calibration and Estimation

To solve the model in time differences we require data on the labor allocation by skill and nationality across countries ($L^s_{i,n,t}$), bilateral trade shares ($\pi_{ij,t}$) and bilateral migration shares by skill and nationality ($\mu^s_{ij,n,t}$). Moreover, we need to calibrate the share of high skill-labor payment in total labor payments (θ^h_i), initial level of technology (η_i), structures (H_i) and initial level of both parts of migration costs $\{m^{s,pol}_{ij,n,0}, m^{s,fix}_{ij,n}\}^{N,N,N}_{i=1,j=1,n=1,s=l,h}$. All other global and country-specific parameters we borrow from previous studies or from the data directly. (see Table 4 in Appendix).

In this paper Armington elasticity of demand substitution (ρ) has been chosen to be 3.5 for all simulations. Feenstra et al. (2018) found micro elasticity of demand substitution to be between 3.24 and 4.12, Costinot & Rodríguez-Clare (2014) showed that under assumption on firm-level heterogeneity it varies from 2.9 to 4.2.

Caliendo & Parro (2015) provides an extensive overview of trade elasticity estimates in the literature ranging from 3 to 30. This paper take into account recent works of Simonovska & Waugh (2014) and Waugh (2010) by setting the value for trade elasticity at 4.14.

The fraction of public goods in total consumption (α_i) is the share of collective general government consumption in gross household adjusted disposable income.

Labor income tax (τ_L) is the average personal income tax and social security contribution rates on gross labor income.

Share of labor payments in value-added ($1 - \gamma_i$) can be found as the *Compensation of employees' share of value added* in the *Trade in employment (TiM): Principal indicators* Section of the OECD Input-Output Tables. The value attributed to the rest of the world is constructed out of 55 countries.

4.1 Calibration of the share of high-skilled labor in labor payments

In the calibration of the share of high-skilled labor in labor payments, θ^h_i, this paper follow an approach similar to one in Docquier et al. (2015). Notice that from first-order conditions (24) one can derive the wage premium. Then, θ^h_i is a function of wages, labor, and elasticity of factor substitution:

$$\theta_i^h = \frac{\beta_{i,0}}{\beta_{i,0}+1}, \text{ where } \beta_{i,0} = \frac{\mathrm{w}_{i,0}^h}{\mathrm{w}_{i,0}^l}\Big(\frac{L_{i,0}^h}{L_{i,0}^l}\Big)^{1/\sigma}. \quad (34)$$

4.2 Estimation of the initial state of technology

Many scholars argue that urban sprawl is linked to productivity raise (Ciccone & Hall (1993),Fallah et al. (2011)) and employment density - with the number of patents (Carlino et al. (2007)). In models of trade and migration these features are aggregated in the variety-invariant component of the total factor productivity (Caliendo et al. (2017)) or in the elasticity of innovations to labor density (Desmet et al. (2018)).

Following the latter approach, in this paper the level of technological development ($T_{i,t}$) has a dynamic density-related component ($D_{i,t}^{\zeta}$). The value of elasticity, ζ, is borrowed from Desmet et al. (2018) and equal to 0.06. Nevertheless, there is a need to calibrate country-specific parameter η_i at the initial period of time. One can estimate, first, $T_{i,0}$ using a structural log-linear "gravity" equation alike Eaton & Kortum (2002) and Waugh (2010), and then calculate η_i by combining results with values on labor $L_{i,0}$ and ζ.

However, to account for exogenous variation within a period of time, estimation can be exercised on a panel data. A panel data of trade shares can be constructed using WTF bilateral flows. It contains 24 years of observations and covers 249 countries. First step in estimation of static component of technological development is the fixed effect regression of the following system of trade equations:

$$\log\Big(\frac{X_{ij,t}}{X_{ii,t}}\Big) = S_{j,t} - S_{i,t} - \log(tc_{ij}), \quad (35)$$

where $X_{ij,t}/X_{ii,t}$ is the ratio between share of goods purchased by country i from country j and country i's share of domestic goods in total consumption, $S_{j,t}$ and $S_{i,t}$ are the exporter's and importer's fixed effects, and tc_{ij} is the trade costs.

More precisely, logarithmic trade costs are modeled as a function of distance and borders between countries.

$$\log(tc_{ij}) = d_\kappa + b_{ij} + \epsilon_{ij}, \quad (36)$$

where d_κ is the κ's interval in distance (in miles) between countries i and j ($[0;375]_{\kappa=1}$, $[375;750]_{\kappa=2}$, $[750;1500]_{\kappa=3}$, $[1500;3000]_{\kappa=4}$, $[3000;6000]_{\kappa=5}$, $[6000;+\infty)_{\kappa=6}$), b_{ij} takes value one when countries share a common border and ϵ_{ij} is an exogenous variation, orthogonal to other regressors.

After obtaining an estimate of importer's fixed effect, $\hat{S}_{i,t}$, using data on average total wages from World Penn Tables and estimated importer's fixed effect, $T_{i,t}$ and η_i can be calculated directly from:

$$T_{i,t} = e^{\hat{S}_{i,t}} \mathrm{w}_{i,t}^{\theta}, \tag{37}$$

$$\eta_i = \bar{T}_{i,t} / \left(D_{i,0}\right)^{\zeta}. \tag{38}$$

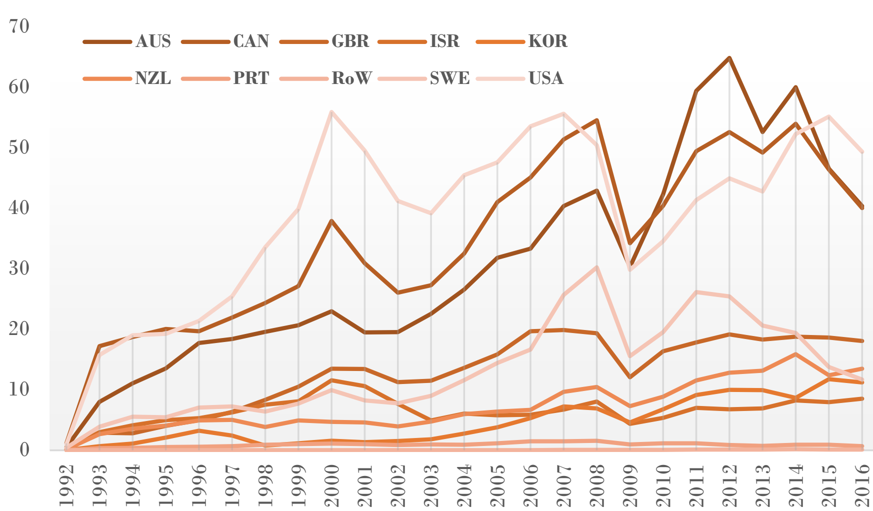
where $\bar{T}_{i,t}$ is the average level of technology across time and D_i is the urban land area.

From equation (38), estimation of the initial technological level requires data on labor density by country. Labor density is a ratio of the number of persons employed and the land area. In the model, it is measured in persons per square meter. To account for the fact that people reallocate to urban areas the data on the urban land area was used as the denominator.

The rest of the world was constructed of 252 countries. Since for 40 countries the data on the urban land area is missing, it has been estimated using the total land area and the median value of the share of urban land in total land area. For the reference, the data on areas are provided by the World Bank and are measured in square kilometers.

The level of technological development represents comparative advantage reflecting into price allocation and trade volumes and, therefore, can be normalized (see Figure 1).

Figure 1: Technological development



The level of technological development is normalized to United States in 1992

4.3 Calibration of the structures

Calculation of the value function at the initial period of time requires knowledge of structures, H_i. To be exact, on the one hand, the logarithmic utility contains structures as a part of housing consumption and, on the other hand, to derive price index, $P_{i,0}$, one needs to calculate marginal costs that also contain structures. "Structures" is the time-invariant and country-specific aggregated variable that corresponds to everything besides labor that allows firms to produce goods and, therefore, can not be directly observed from the data. However, for example, rents and wages can be observed, which makes it possible to calibrate H_i from the system of equations, that can be derived by maximizing the profit function (see Appendix):

$$r_{i,t} = \frac{\gamma_i}{1-\gamma_i}\frac{1}{H_i}\left(\mathrm{w}_{i,t}^{l}L_{i,t}^{l} + \mathrm{w}_{i,t}^{h}L_{i,t}^{h}\right), \quad \forall i, t. \tag{39}$$

Rents ($r_{i,t}$) reflect household expenditure on housing and so can be inferred from the data on the private nominal consumption and the share of housing in private consumption by simply multiplying the both. To construct nominal consumption in USD (C_{USD}^{nom}), one could use the following formula:

$$C_{USD}^{nom} = \frac{XR \times C_{NC}^{real}}{PPP}, \tag{40}$$

where C_{NC}^{real} is the real consumption in the National currency units, XR is the exchange rate and PPP is the purchasing power parities.

The real consumption and the exchange rate can be found in PWT, the PPP's can be found in the OECD Database and are presented in National currency units per US dollar. Since the data in PWT are represented in US dollars in 2011, every value used in the paper was discounted using CPI's from the Bureau of Labor Statistics.

And, finally, the share of housing in private consumption can be found in the OECD Database on household spending. It is measured in the percentage of household disposable income and represents aggregated expenditure on "housing, water, electricity, gas, and other fuels".

4.4 Calibration of the initial level of migration costs

It continues to be challenging in migration literature to estimate international migration costs because they are not directly observed and have a complex structure. Common approximation for migration costs includes travel time, geographic and genetic bilateral distances.

For example, based on internal migration data, costs of moving are about 0.8-1.2% of the mean annual wage, while 84% of them are fixed costs, 9.6% depend on travel time, and 3.5% - on geographic distance (Morten & Oliveira (2016)). If applies, cultural dissimilarity would accumulate for an additional 1-1.5% of annual wage (Falck et al. (2018)). Within a framework of trade models, where amenities and idiosyncratic preferences are implicitly included in utility function, migration costs are estimated to be 177% of relative real income (Tombe & Zhu (2019)).

In the absence of migration policies, international migration is similar to internal migration in a way individuals respond to distances and labor market openness while deciding to move. However, in the presence of policies, people face additional monetary and non-monetary costs (limitation of rights, quotas, etc.), which is difficult to measure and disentangle from the former component. Nevertheless, there are estimates that the incompressible part of migration costs (not related to migration policy) vary from 75% to 90% of total migration costs (Docquier et al. (2015)).

This paper attempts to disaggregate migration costs into policy-related (endogenous) and fixed (exogenous) components by destination, skill, and nationality based on the aforementioned literature and modeling assumptions.

The solution of equations (68) and (69) implies that observations at time $t+1$ are used to estimate migration costs at time t. To satisfy time-consistency, first, assume that the values of utility and migration shares at the initial period and a consecutive one would have been insignificantly different in the absence of changes to migration policies. And, secondly, while modeling, keep migration costs unchanged between periods t and $t+1$ at the calibrated level.

However, together (68) and (69) can not be evaluated since the number of equations is less than the number of variables. Hence, one way to solve the system is to reduce the number of variables by separately fitting a vector of *return* migration costs, $(m_{in,n}^{s,fix})$, to the data. Return migration is a unique movement because, firstly, it does not imply any migration policy-related costs and, secondly, a person comes back into a well-known environment avoiding most of the costs arising from cultural factors and information asymmetry. Moreover, return migration costs are similar to the costs of internal migration when movement is made over a larger distance.

In application to the paper, Australia and New Zealand (The Trans-Tasman Travel Arrangement), as well as countries of the EU: Portugal, Sweden and the United Kingdom (Article 51 of the Amsterdam Treaty), have established a free movement for workers. It implies that for these origin-destination pairs migration policy costs are zero or negligible and, therefore, from equation (69) total of 16 return migration costs can be directly calculated using:

$$m_{jn,n}^{s,fix} = \frac{\nu}{\beta - 1} \log\Big(\frac{((G_{j,1})^{\alpha_j}(C_{j,1}^s)^{(1-\alpha_j)})^{\beta/\nu}}{((G_{n,1})^{\alpha_n}(C_{n,1}^s)^{(1-\alpha_n)})^{\beta/\nu}} \cdot \frac{\mu_{jn,n,0}^s}{\mu_{jj,n,0}^s} \cdot \frac{(\mu_{nn,n,1}^s)^\beta}{(\mu_{jn,n,1}^s)^\beta}\Big), \quad \forall j, n \in M, \tag{41}$$

$$M \equiv \{\text{AUS, NZL}\} \text{ or } \{\text{PRT, SWE, GBR}\}. \tag{42}$$

These observations can be used to get a rough understanding of how geography and amenities influence the costs of return migration. For that purpose, the data on geographical distance, d_{ij}, and quality of air, a_i was collected from CEPII and OECD Environment Databases. Air quality is considered to be a measure of local amenities and considered to be a part of migration costs. Air quality as a measure of local amenities was estimated to be an important factor in migration decisions. According to Bayer et al. (2009), migrants are willing to pay up to 13% of their average monthly wage for a one-unit reduction in average ambient concentrations of particulate matter. The following regression was estimated:

$$m_{jn,n}^{s,fix} = s_{jn} + d_{jn} + d_{jn}^2 + (a_n - a_j) + g_{jn} + \epsilon_{jn}, \tag{43}$$

where s_{jn} is the skill dummy that takes value one if a migrant is high-skilled and zero otherwise, g_{jn} is the genetic distance and ϵ_{jn} is the exogenous variation orthogonal to regressors.

Table 1: Fixed costs of return migration

	OLS (1)	OLS (2)
Skill, (s_{jn})	-5.892	-8.212
	(-1.01)	(-1.97)
Distance, (d_{jn})	0.0547***	0.0643***
	(4.68)	(6.66)
Sq. distance, (d_{jn}^2)	-0.0000145***	-0.0000167***
	(-4.48)	(-6.12)
Amenities, ($a_n - a_j$)	-5.282***	-5.282***
	(-7.10)	(-10.25)
Genetic distance, (g_{jn})		-1351.2***
		(-4.53)
Observations	16	16

Note: The table represents relationship between fixed costs of return migration and most common determinants of migration costs. The sample includes origin-destination pairs {AUS, NZL} and {PRT, SWE, GBR}, where j correspond to origin and n - to destination. Genetic distance data has been taken from Özak (2010). t-statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Although the number of observations is limited, it can be seen that return migration is more costly for low-skilled workers while costs increase with distance and lower air quality at the destination. Results are robust to the addition of genetic distance into regression. Although coefficients represent a crude approximation, they contain important information that can be used in the calibration of other return migration costs.

Summing up, the calibration of migration costs is be executed under the following assumption:

Assumption 3. *Costs of return migration are fitted by constructing an objective cost function*[23] *which takes into account that:*

- *Migration policy costs can take only non-negative values, while fixed migration costs primarily*[24] *take non-negative values;*
- *Fixed migration costs constitute 75-90% of total migration costs;*
- *Distance, amenities and skill influence fixed costs of return migration between all origin-destination pairs similarly to pair of {AUS, NZL} and {PRT, SWE, GBR}.*

Calibrated values of policy and fixed migration costs for low-skilled and high-skilled labor can be found in Figure 2.

For comparison, the average annual log consumption across countries is about 9.88, which makes the cost of migration equal to approximately 2 years wage[25]. Moreover, both policy and fixed migration costs are lower on average for high-skilled labor. It has been argued by (Borjas (1991)) in the 1980s that migration costs, on the contrary, are lower for low-skilled workers. It took about 20 years to promptly address the problem of migrant's self-selection (Chiswick (1999), Hunt & Mueller (2004), Chiquiar & Hanson (2005)) and provide factual and theoretic counterargument in favor of positive self-selection.

5 Results

5.1 New unions and associated economic effects

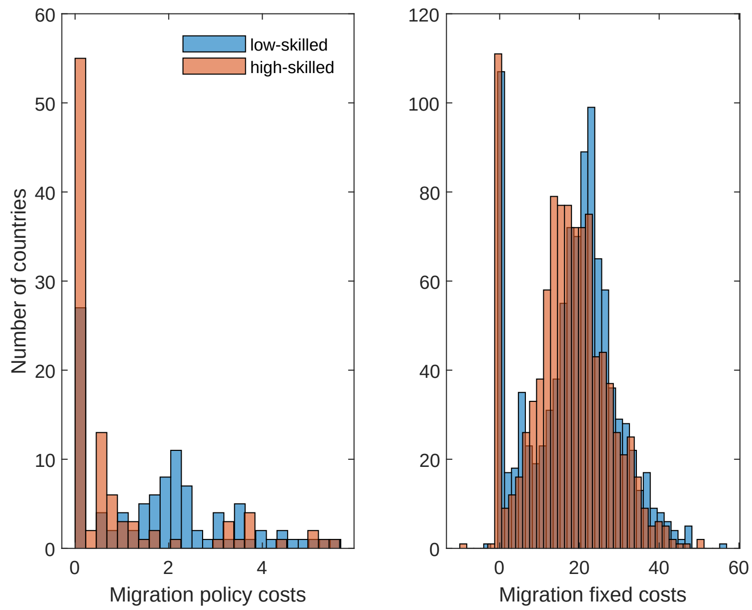
This Section attempts to quantify welfare effects from establishing counterfactual migration unions. For that purpose the total of 30 different setups of the world economy was

[23] The cost function penalizes for deviation of policy and fixed migration costs from non-negative values with a higher weight for policy-related component, for deviation from a 75-90% share of fixed component in total costs and for deviation of coefficients obtained from (43) using whole set of countries from ones in Table (1). Accordingly, in calibration the weights between these functions are distributed as follows: 0.7*(0.2 for fixed and 0.8 for policy component), 0.15, 0.15.

[24] Every year a significant amount of people receive support from national governments to cover moving expenses of return. According to International Association for Migration (IOM), worldwide in 2018 more than 63.000 migrants were assisted to return. These and similar policies explain existence of negative values of fixed migration costs.

[25] For example, Clemens et al. (2019) estimated migration costs of entering US to be at least 9.53 in log consumption equivalent for low-skilled workers in U.S. dollars at PPP

Figure 2: Calibrated values of migration costs in 2000



estimated including the baseline and 4 major restructurings: break of the European Union (consisting of {PRT, SWE, UK}), establishing a union between English speaking countries, total liberalization of the labor movement and total closure of all borders. Then, for each counterfactual, the welfare gains are calculated as a ratio of welfare levels between counterfactual and baseline economy. However, cooperation on labor mobility requires not only economic efficiency but also a political agreement. To reach an agreement the cooperation has to be mutually beneficial. And so, to enhance the analysis, we define *politically viable unions* as unions where both parties are better off in cooperation than without it and the *economically efficient unions* as those where only joint welfare gains are positive.

The Table 2 gathers results on politically viable and economically efficient unions[26]. The represented welfare gains are attributed to nationals only and are aggregated by skill.

Quantitative results show that there exist three viable and efficient ways to restructure the world migration barriers, namely, removing them between Israel and European Union, Brexit and breaking the union between Australia and New Zealand[27]. Consequently, if Israel joins European union, high-skilled and low-skilled natives in aggregate would experience a 0.374% increase in their welfare, while EU's natives would gain 0.883% at almost zero cost

[26]The other results can be found in the Table 3 in Appendix.

[27]The Australian government responded by tightening of migration policy in 2001, restricting requirements for permanent skilled migration from New Zealand (Birrell (2013)).

Table 2: Welfare Gains from establishing a Union

Type	**1st Party**		**2nd Party**		**Aggregate**	**Rest of the world**	**Most vulnerable group**
			Politically Viable				
Join	ISR	0.374	EU	0.883	0.863	∼0	EU (HS)
Exit	AUS	0.159	NZL	3.625	0.678	∼0	AUS (LS)
Exit	UK	0.180	EU	0.127	0.168	∼0	UK & PRT (HS)
			Economically Efficient				
Join	ISR	-2.442	Pacific	1.317	0.770	-0.001	ISR
Join	CAN	-0.005	ISR	5.820	0.658	∼0	CAN (HS)
Exit	Break of the EU (See Figure 3.)				0.156	∼0	PRT
Join	ISR	-0.120	KOR	0.142	0.127	∼0	ISR
Join	CAN	-3.080	US	0.374	0.096	0.001	CAN
Exit	SWE	0.223	EU	0.064	0.080	∼0	PRT (LS)
Join	ISR	-0.071	US	0.070	0.068	0.001	ISR (HS)
Exit	PRT	-0.039	EU	0.079	0.064	∼0	PRT
Exit	Autarky (See Figure 3.)				0.003	-	PRT

Note: The table represents welfare gains in percentage points measured as the ratio of welfare levels between counterfactual and baseline economy. Unions are sorted by the aggregate gains in the descending order. The EU includes {PRT, SWE, UK}, the Pacific union includes {AUS, NZL}, high-skilled and low-skilled labor is denoted by *HS* and *LS* respectively. The sign ~ 0 denotes values having all decimal digits up to the nearest fourth equal zero.

for the rest of the world. Nevertheless, it has to be mentioned that gains are aggregated over skill groups and it is assumed that there is an effective redistribution within every country. However, without redistribution, the largest losses in welfare would be attributed to high-skilled labor having the nationality of any of EU countries. In general, the high-skilled labor happens to be more vulnerable to establishing a new union, while low-skilled labor loses more from exiting the union.

The United Kingdom would increase welfare from Brexit in 2000 by 0.18%. This number aggregates 0.194% gain for low-skilled and 0.023% loss for high-skilled British citizens. Portuguese high-skilled labor would experience a drop of 1.171% in welfare. Meanwhile, the break of the union between pacific countries would only reduce by 0.02% welfare of the high-skilled Australians and increase the welfare of New Zealanders by 3.49% and 4.52% for low-skilled and high-skilled labor respectively.

The second section of the table represents economically efficient unions that create an increase in aggregated welfare large enough to outweigh losses of particular blocking parties. For example, even though the United States can increase the national welfare only by 0.374%, it is enough to compensate Canadians for an aggregate decrease by 3.08% due to the size of the US economy. Portugal appears to be the most vulnerable country to exit among EU members and, in particular, it block Sweden from leaving the EU or both Sweden and the UK, but not the UK solely.

Additionally, establishing new unions influences not only migration policies between parties but also policies towards the rest of the world nationals including OECD countries (Column 5). While most of them have negligible effects, for example, the union between Israel and Pacific countries (Australia, New Zealand) would lead to an average decrease in welfare by 0.001% for every single household outside of the union.

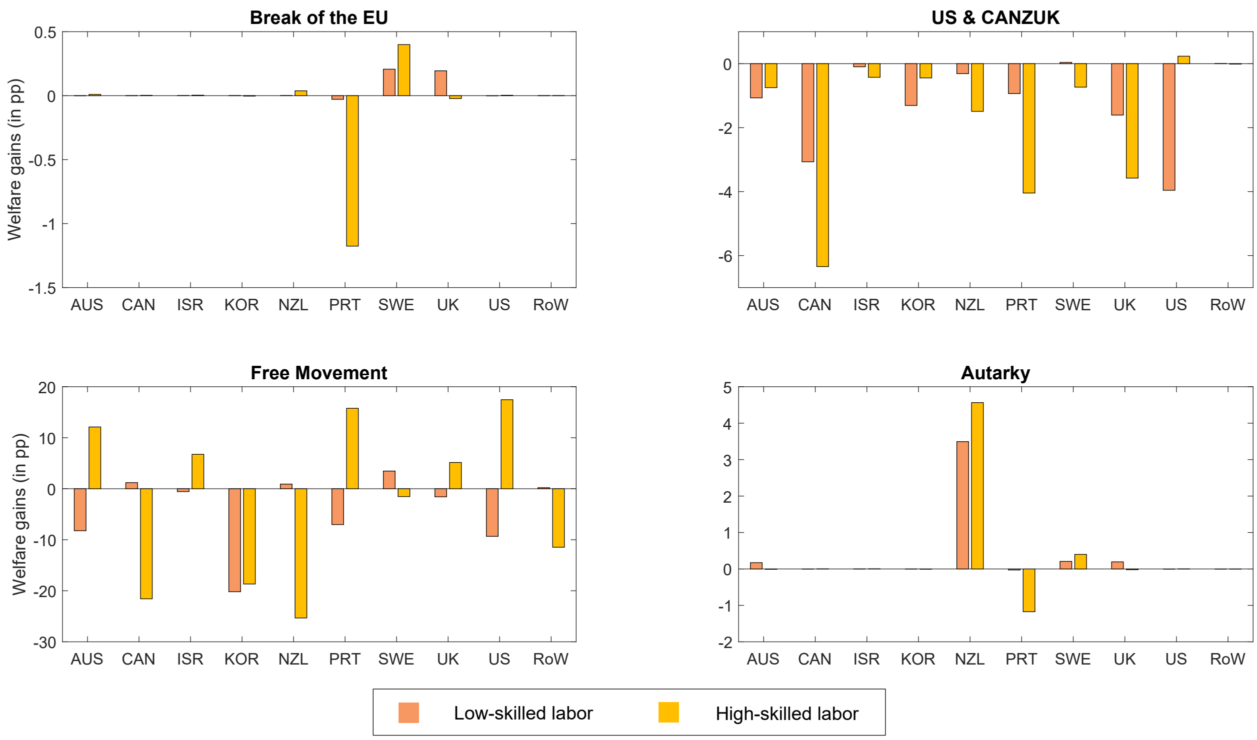
It might seem surprising that the Break of the European Union and the complete closure of borders (Autarky) are economically efficient. In Figure 3 one may find welfare effects disaggregated by skill and country members from a major restructuring. The 1.2% decrease in the welfare of the high-skilled Portuguese population (less than 10% of the total) does not outweigh gains for the Swedish and British populations. Similarly, in Autarky most of the positive effects are attributed to New Zealand, while for 60% of countries there is almost no or negative effects. Therefore, if the coronavirus disease 2019 (COVID-19) would have happened in nineteen years earlier, Sweden and New Zealand could have experienced the closure of international commuting better than the rest of the world.

Liberalizing labor mobility gives us a completely different picture. The first observation is that all countries undergo substantial changes and high-skilled labor exhibits a larger variation in welfare gains. Mainly, it reflects the fact that our world is highly regulated in a way that protects vulnerable groups of people and countries from the unfavorable competition. The 11.5% decrease in the welfare attributed to the high-skilled labor having nationalities of countries aggregated under the rest of the world (RoW) drives most of the losses due to switching to a free movement regime.

The union of Engish speaking countries available in the sample (US, Canada, Australia, New Zealand, UK) yielded no efficient outcomes having both types of labor losing the welfare in most of the countries.

Quantitative results show that simply cooperation on labor mobility may provide OECD countries with substantial welfare gains up to 5.8% for separate countries and up to 0.86% in aggregate. However, due to the small sample size, these numbers can not be treated as the upper bound. Welfare gains largely depend on how countries match to each other. However,

Figure 3: Welfare gains from a major restructuring



Note: The Figure represents welfare gains of each country from one of major changes: (a) break of the {PRT, SWE, UK} union, (b) establishing a union between English speaking countries, (c) total liberalization of labor movement, (d) total closure of all borders.

potential factors contributing to the matching between countries will be discussed in Section 5.4.

Additionally, these quantitative results contribute to the understanding of the role of migration policy concerning redistribution. It has been argued that under endogenous migration policies and international transfers the Nash equilibrium is Pareto optimal (Casella (2005)). This paper shows that gravity models of migration and trade, employing larger heterogeneity in countries and labor specifications yield Pareto improving outcomes (politically viable unions) even without international transfers.

Moreover, the large number of economically efficient but not politically viable unions highlights the importance of coordinated redistribution schemes. They would not be a cornerstone of the political economy if not the questions related to the bargaining on the differentiation of efforts, benefits, commitment to redistribution policy, and strategic delays.

It is argued in the literature that, for example, trade policies or side payments could redistribute gains effectively when uniformity between parties is relatively small and, additionally, in some cases even under the uniform policies (Harstad (2007)). Alesina et al. (2005) opposes political harmonization claiming that it reduces the size of the union, and proposes a notion of the "flexible unions". These solutions are suitable for countries having either intensive trade with the counter-party or a significant share of public goods consumption.

Additionally, Casella (1990) provided a piece of theoretical evidence in favor of currency union as a tool to transfer of seigniorage revenues from small to large economies. A fiscal union is proven to be an important risk-sharing supplement to the currency union (Farhi & Werning (2017)) and in the long run, would increase countries' uniformity and favor labor mobility.

It is important to understand that even if these types of international redistribution are properly working, there is a need to impose such a scheme within each country that would guarantee the reception of benefit by narrow groups of workers, preferably, disaggregated even further beyond skill and nationality.

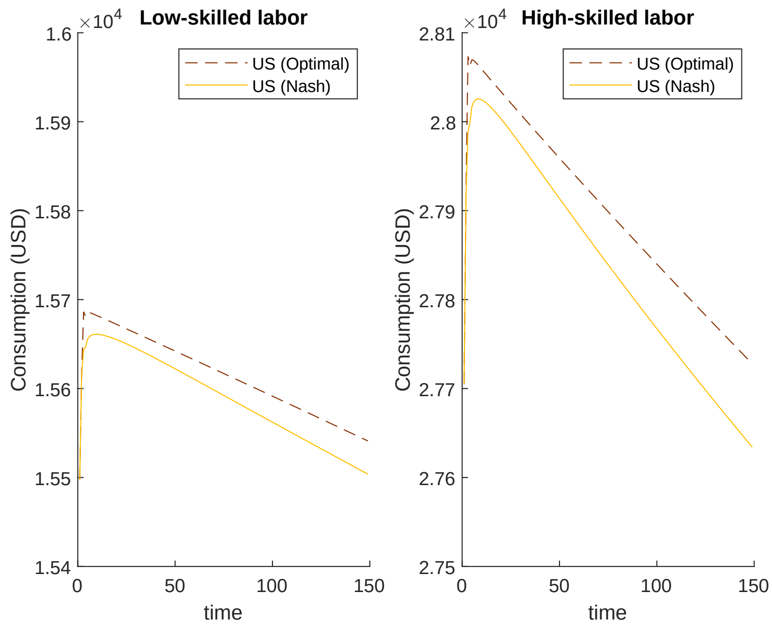
5.2 Migration costs with and without response

The purpose of this Section is to show that optimal migration policy depends on whether countries respond to each other or not. Define response as a process of updating migration policies by other countries given the homeland's actions. Then, in Figure 4 one may see the consumption resulted from the estimation of the baseline economy in the presence and absence of response.

The Figure 4 shows that response reduces potential consumption gains from migration policies. For every low-skilled worker in the United States the drop ranges approximately from 16.9 to 42.3 US dollars per year, while for every high-skilled worker it varies from 30.2 to 93.8 US dollars per year. Retaliation also leads to welfare losses through misallocation of labor and reduced value of an option to move.

The absence of response allows a country to improve the national welfare by more thorough discrimination through migration policy. For example, in the symmetric setup, a unilateral increase in migration policy cost initiated by country **A** towards all nationals **B** will lead to a net outflow of B's due to decreasing rate of immigration in country A, while keeping all the rest stable. Consequently, two motives for nationals A to return back home arise: labor allocation changes such that labor in country B becomes abundant, wages and redistribution of the national wealth per capita drop, while labor in country A becomes scarce and drives wages and social benefits up. Therefore, nationals A can benefit from changed labor market fundamentals and absent policy costs of return migration. The scale of return migration, however, is limited due to side effects. Country B, experiencing relatively more rapid return migration, benefits in technological development due to higher population density and decreased price index. Moreover, if there are nationals A residing in country

Figure 4: Differences in consumption arising from response



Note: The Figure represents losses in consumption arising from response. The dashed line represents US optimal consumption level when other countries do not respond to changes in US migration policies. The solid line is the consumption level in the Nash equilibrium.

B, then a policy tightening will be smoothed based on how much country A cares about its emigrants. Migration costs can not be infinitely high unless a country has substantial technological advances and no emigrants.

At the same time, in an asymmetric setup relatively scarce high-skilled labor and abundant low-skilled labor in country B may be of great interest for country A having an opposite allocation of labor. Without response from B, country A would optimally reduce migration policy costs. Now we question, whether a country B would respond the same way? Seemingly, without an agreement, the incentive to behave opportunistically is extremely large. Therefore, a situation when two countries experience a bilateral decrease of migration policy costs (or establish an economically efficient and politically viable migration union) is hardly feasible[28] without coordinated redistribution or additional latent economic gains arising from being in the union.

[28]Feasible under a sufficiently high level of the intersection of economies in terms of migration, trade, and governmental preferences (see discussion in Section 5).

5.3 Matching by the structure of allied economies

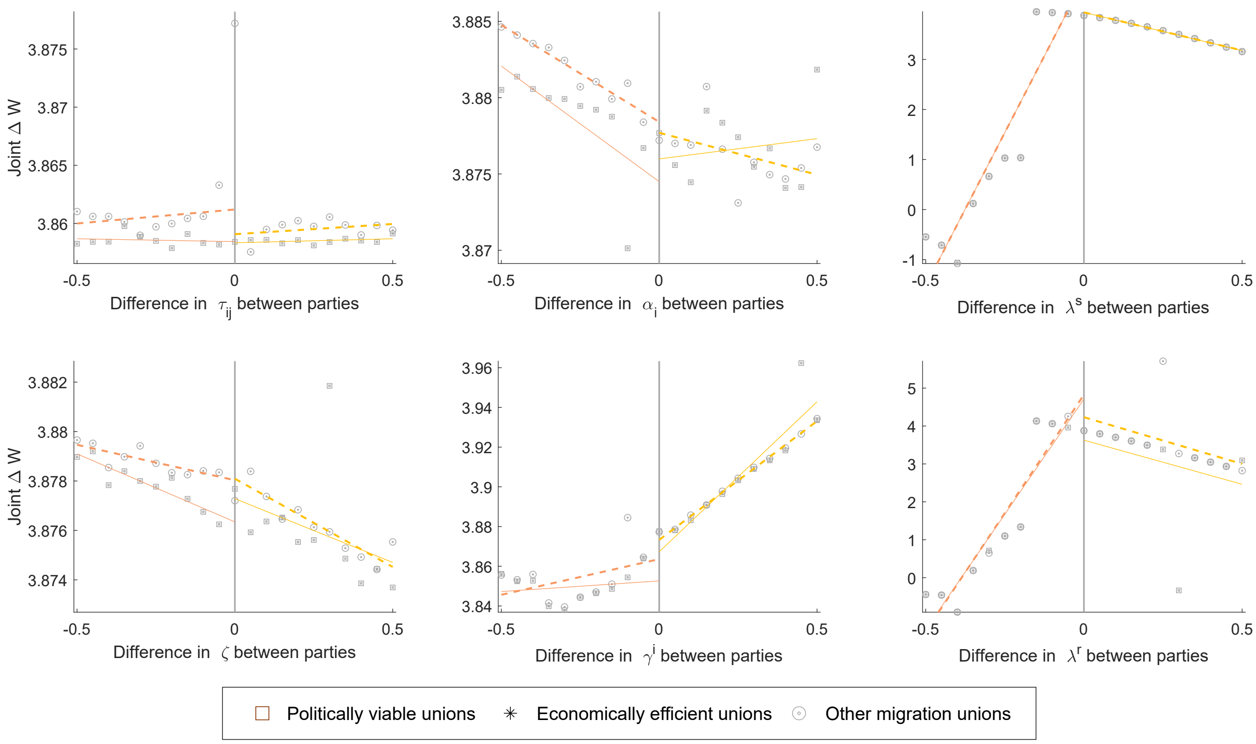
This Section attempts to discover what factors determine joint welfare gains from cooperation on labor mobility and whether heterogeneity between countries is favorable. Each factor will be classified as "opposites attract" or "like attracts like". To disentangle and confirm the influence of economic parameters the sensitivity analysis is represented for symmetric (median) and asymmetric (observed OECD) countries.

Based on the observed data a world of four symmetric median countries was built to discover how changes in underlying economic parameters influence joint welfare gains from establishing a union between two of them. The two other countries were set to play the roles of an active competing country and inactive rest of the world. All countries except one were kept unchanged. The targeted country underwent changes in a single parameter from a 50% decrease to a 50% increase from the base value. Therefore, Figure 5 shows how joint welfare gains change when the distance between allying parties in the parameter value increases by one percent upward and downward. The solid line represents the same sensitivity analysis but when bilateral trade shares between allying parties are 30% higher than the median value.

There is a clear pattern that "like attracts like" in governmental attitudes toward skill and residence of natives based on the inverse "V" shape of the gains. The share of capital payments in value-added (γ_i) also falls in this group even though gains do not have the inverse "V" shape. It is fair if the cooperation of capital intense economies (having large γ_i) is more efficient in general.

While bilateral trade tariffs seem to have no significant effect, the prevalence of socialist policies (α_i) and the ability of dense labor to create innovations (ζ) require a more thorough investigation. It has to be mentioned that in this simulation the union is experiencing an outflow of labor. The extensive production of public goods becomes less efficient as population shrinks and, hence, the right-hand tail is skewed downwards. Therefore, based on the left-hand tail, "opposites attract" and so countries with socialist and capitalist economies would benefit the most from the cooperation on labor mobility. The sensitivity of productivity elasticity to labor density (ζ) also depends on the migration scenario. From a dynamic point of view, a union experiencing an outflow of labor should prefer the low elasticity to slow down the decrease in technological superiority. And so, unless elasticity is high enough to make price drop faster than wages (due to large labor inflow), it will be always better to have a lower elasticity for both parties. This result explains why companies in highly dense areas like Silicon Valley are lobbying liberal migration policies, while the majority finds it inefficient. The important political question arises, because due to the high intersection of

Figure 5: Parameters sensitivity



Note: The Figure represents the sensitivity of welfare gains to isolated changes in economic parameters. Welfare gains are calculated as a ratio between the joint welfare of two allying parties with and without union between them and are measured in percentage points. All countries are symmetric, except for one that undergoes changes in a single parameter. Changes of each parameter vary from 50% decrease to 50% increase from the base value. The base value for governmental attitude towards skill and residence of natives is set to Egalitarian (see Table 4). All data points that are further than three standard deviations from the mean are omitted and imputed with the average of four neighboring points (two points from each side).

labor movement and innovations, the migration policy may determine where the productive regions will eventually locate (Desmet et al. (2018)).

The increase in bilateral trade makes the relationship between gains and all parameters more "V" shaped and so decreases gains from being alike. This effect comes from the argument that both trade and migration policies can be seen as redistribution tools. Then, higher trade being a strategic substitute simply devalues the positive effect of switching to free labor mobility. A similar argument can be applied if migration and trade are considered to be risk-sharing mechanisms (Morten (2019)).

To confirm and enhance the results from the analysis of symmetric countries, a few relationships are built based on the quantified welfare gains between OECD countries. It is seen from Figure 6, that "opposites attract" effect holds for public consumption even when countries are asymmetric and migration scenarios are different. Moreover, it is clear now that the more socialist countries gain more in general from free labor mobility. According to the primary type of ownership in the economic system, the opposite labor income tax policies yield higher gains.

Distance, technological superiority, and aforementioned governmental attitude towards skill fall into "like attracts like" group. The explanation of these results is straightforward. Smaller distance naturally eases bilateral migration. Due to low fixed migration costs, migration policy obtains larger importance in controlling the movement of labor. Therefore, if there are incentives to build barriers, they are relatively high comparing to those countries that are far away from the homeland. As a consequence, marginal welfare effects from liberalizing labor mobility are higher as well. The particularity of technology is that it drives changes in prices and real consumption regardless of the skill type, and, therefore, any difference in the level of technologies between countries results in unilateral migration only, which is rarely beneficial for both parties. Besides, to have larger effects from open borders it is important to have common attitudes toward skill. Otherwise, gains from free movement will be one-sided.

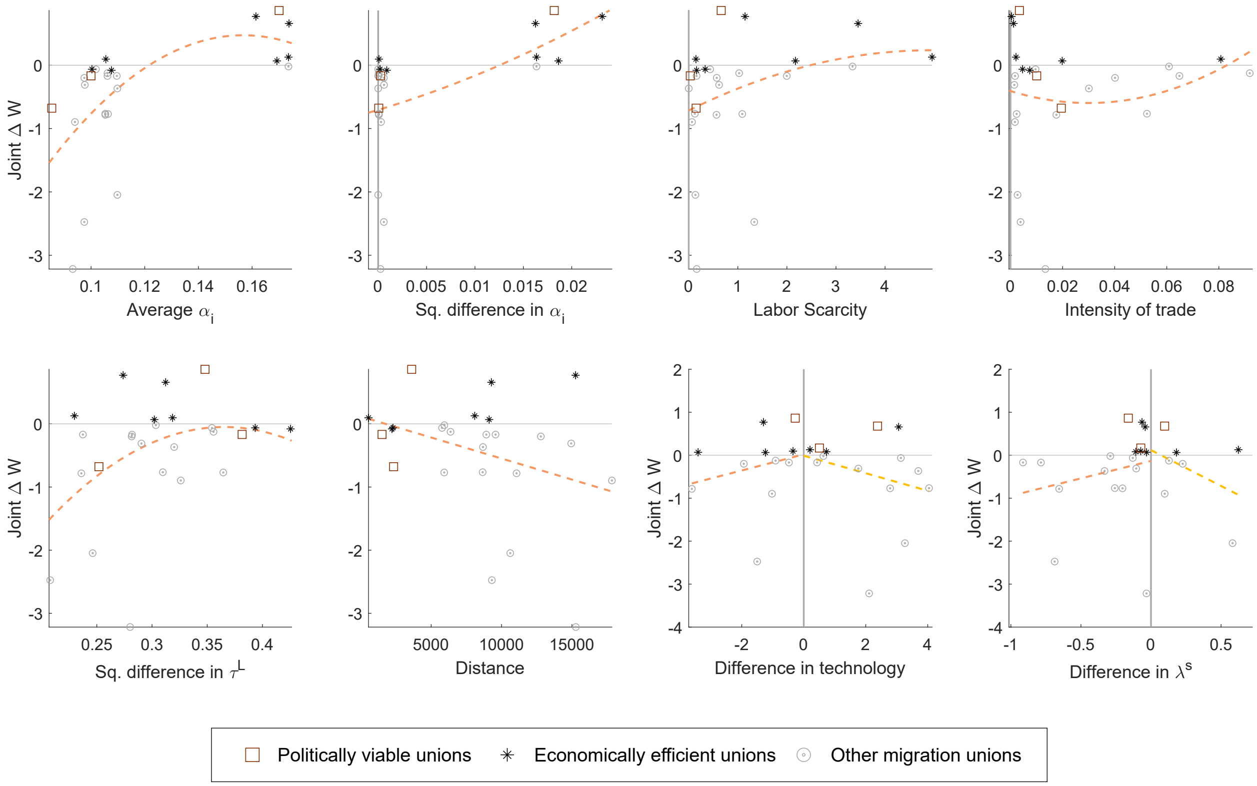
The intensity of trade is computed as an average bilateral trade share between parties engaging in a migration union. Based on the results of the sensitivity analysis with symmetric countries, the intensity of trade may favor the union as well as hinder it. Depending on whether parties marginal welfare effects from an increase in bilateral trade are larger for the "opposites attract" or the "like attracts like" group of factors. Nevertheless, the counterfactual analysis shows that on average parties having intense trade gain more from the free movement of labor.

The labor scarcity (L_{sc}) is constructed using initial migration policy costs and the following formula:

$$L_{sc} = |(\bar{m}_i^l + \bar{m}_j^h) - (\bar{m}_i^h + \bar{m}_j^l)|, \quad \bar{m}_i^s = \sum_{k \setminus M_i}^{N} m_{ik,i,0}^{pol,s} \tag{44}$$

If one type of labor is abundant in a country, it earns a relatively low wage compared to other countries and is constantly seeking opportunities abroad. As a consequence of the high labor supply, it faces large migration policy costs. Therefore, if two countries have scarce mutually opposite types of labor, the labor scarcity would take large values and vice versa if countries are in the exact same situation. Thus, according to the Figure 6, accounting for heterogeneous labor structure is important when deciding on establishing a migration union. Moreover, the more scarce parties are in opposite labor types the larger welfare gains are. The relation between labor scarcity and welfare gains was documented in a recent paper by Price et al. (2020) and confirms our results.

Figure 6: Gravity between asymmetric economies



Note: The Figure represents how heterogeneity between countries can influence joint welfare gains (in percentage points) from establishing a union. For consistency, an "exit" type counterfactual should be accounted as a "join" type with the opposite sign. Whenever a country joins a union of two or more countries, the union's parameter value is taken as an average of members' parameter values. Major restructurings are not included.

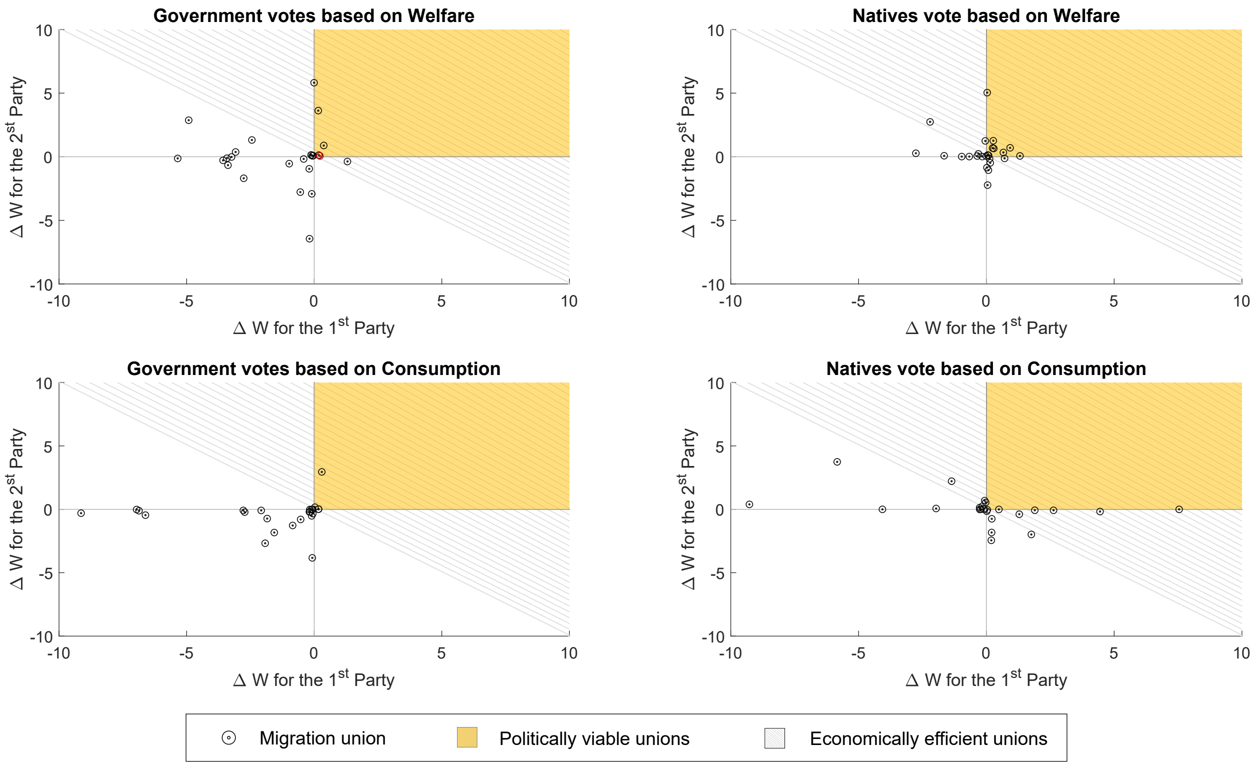
5.4 Non-financial factors in decision to cooperate

This Section attempts to distinguish politically viable and economically efficient unions based on the choice of decision-making authority and on the importance of the freedom to move.

Sections (a) and (c) of Figure 7 represent the situation when a government is authorized to vote for all nationals on establishing a migration union. It means that every native's welfare is being accounted for with an attributed political weight based on governmental attitudes towards skill and residence. Therefore, minorities residing abroad would have less than one vote per person while natives at home more than one. At the Sections (b) and (d) the decision-making process follows the principle of "one man, one vote" regardless of the skill and residence of a native.

Another factor influencing outcomes is the measure of well-being. In this model, welfare is different from consumption in that it also accounts for migration barriers, i.e. how expensive

Figure 7: Voting process and Welfare gains



Note: The Figure represents welfare gains from establishing a union between two parties under different voting processes. Each dot on the graph is a union. When voting is taken by the government, a corresponding welfare and consumption are measured as an average weighted by the political demands. Red dot is politically viable only under the assumption that there exist a fair redistribution within EU countries.Major restructurings are not included.

it is for a household to change a country of work (residence). This division is important to observe the value of free movement as a tool to share risk across locations and how the outcomes would change if risk-sharing would be worthless for households.

It can be seen from Figure 7 that most unions are economically efficient when natives vote for themselves and even more when they value the freedom to move. The most stringent case appears to be when only the government is authorized to vote and no value is assigned to the freedom of movement.

The important question is whether natives have a tool to initiate voting and whether the opposite party can provide its population with a choice to vote for themselves.

Therefore, cooperation on labor mobility ends up being economically inefficient or politically unreasonable because of the absence of the reasonable voting tool or a not yet formed household's preference for freedom of movement.

6 Conclusion

In this paper, we argue that cooperation on labor mobility is an important driver of economic growth. By endogenizing migration policy in a gravity setup, we have been able to study matching between countries and identify economic and political obstacles of complete liberalization of the labor movement.

Having estimated a set of counterfactual economies for 9 OECD countries, we find that differences in factors responsible for absolute advantages between countries like technology and capital intensity decrease welfare gains. Because migration policy costs and transportation costs are strategic substitutes, unions of geographically distant countries appear to have lower policy-related costs as well as lower marginal welfare effects from their removal. According to our results, governmental attitudes should also be included in "like attracts like" group of factors.

On the other hand, countries having scarce mutually opposite types of labor receive larger gains from cooperation on labor mobility. A similar pattern is captured for the dominant types of ownership in the economy implying that socialistic and capitalistic countries would gain most from cooperation, although, on average, more socialistic and at the same time capital intense countries gain more.

We find that, even though trade and migration policy are substitutes in the role of redistribution between countries, the effect of intense trade on welfare gains is positive and stronger than the substitution effect.

Our quantitative results show that at least three new unions are politically viable and nine are economically efficient. From these findings, we conclude that the absence of redistribution between and within economies counteracts the emergence of new alliances.

We also show that the number of economically efficient and politically viable unions strongly depends on whether households consider valuable the freedom of labor movement and whether a political system has a tool to authorize natives to initiate voting for a new union and how votes are weighted based on the location and skill of the voter.

Summing up, the world would agree to open borders when we all will be able to enjoy the same level of technologies, maintain diversity among working teams, specialize in trade and teach each other how to develop high-standard public institutions and remain competitive and innovative, but most importantly, when we learn how to share the benefits.

It has to be mentioned that the set of economic parameters studied in the paper is far from exhaustive and more research is needed to study the impact of other factors like genetics, history, traditions, amenities, investments, common currency, taxation and others on the welfare gains from cooperation on labor mobility.

Tables

List of Tables

Figures

List of Figures

References

Alesina, A., Angeloni, I. & Etro, F. (2001), The political economy of international unions, Technical report, National Bureau of Economic Research.

Alesina, A., Angeloni, I. & Etro, F. (2005), 'International unions', *American Economic Review* **95**(3), 602–615.

Alesina, A. F. & Stantcheva, S. (2020), Diversity, immigration, and redistribution, Technical report, National Bureau of Economic Research.

Allen, T. & Arkolakis, C. (2014), 'Trade and the topography of the spatial economy', *The Quarterly Journal of Economics* **129**(3), 1085–1140.

Artuç, E., Chaudhuri, S. & McLaren, J. (2010), 'Trade shocks and labor adjustment: A structural empirical approach', *American economic review* **100**(3), 1008–45.

Aubry, A., Burzyński, M. & Docquier, F. (2016), 'The welfare impact of global migration in oecd countries', *Journal of International Economics* **101**, 1–21.

Battisti, M., Felbermayr, G., Peri, G. & Poutvaara, P. (2018), 'Immigration, search and redistribution: A quantitative assessment of native welfare', *Journal of the European Economic Association* **16**(4), 1137–1188.

Bayer, P., Keohane, N. & Timmins, C. (2009), 'Migration and hedonic valuation: The case of air quality', *Journal of Environmental Economics and Management* **58**(1), 1–14.

Bernard, A. B., Eaton, J., Jensen, J. B. & Kortum, S. (2003), 'Plants and productivity in international trade', *American economic review* **93**(4), 1268–1290.

Birrell, R. (2013), 'Tougher policy needed for new zealanders wanting to work in australia', *The Conversation. Retrieved from: http://theconversation. com/tougher-policy-needed-for-new-zealanders-wanting-to-work-in-australia-12007* .

Borjas, G. J. (1991), Immigration and self-selection, *in* 'Immigration, trade, and the labor market', University of Chicago Press, pp. 29–76.

Caliendo, L., Dvorkin, M. & Parro, F. (2019), 'Trade and labor market dynamics: General equilibrium analysis of the china trade shock', *Econometrica* **87**(3), 741–835.

Caliendo, L., Opromolla, L. D., Parro, F. & Sforza, A. (2017), Goods and factor market integration: a quantitative assessment of the eu enlargement, Technical report, National Bureau of Economic Research.

Caliendo, L. & Parro, F. (2015), 'Estimates of the trade and welfare effects of nafta', *The Review of Economic Studies* **82**(1), 1–44.

Carlino, G. A., Chatterjee, S. & Hunt, R. M. (2007), 'Urban density and the rate of invention', *Journal of Urban Economics* **61**(3), 389–419.

Casella, A. (1990), Participation in a currency union, Technical report, National Bureau of Economic Research.

Casella, A. (2005), 'Redistribution policy: A european model', *Journal of Public Economics* **89**(7), 1305–1331.

Chassamboulli, A. & Palivos, T. (2013), 'The impact of immigration on the employment and wages of native workers', *Journal of Macroeconomics* **38**, 19–34.

Chiquiar, D. & Hanson, G. H. (2005), 'International migration, self-selection, and the distribution of wages: Evidence from mexico and the united states', *Journal of political Economy* **113**(2), 239–281.

Chiswick, B. (1999), 'Are immigrants favorably self-selected?', *American Economic Review* **89**(2), 181–185.

Ciccone, A. & Hall, R. E. (1993), Productivity and the density of economic activity, Technical report, National Bureau of Economic Research.

Clemens, M. A. (2011), 'Economics and emigration: Trillion-dollar bills on the sidewalk?', *Journal of Economic perspectives* **25**(3), 83–106.

Clemens, M. A., Montenegro, C. E. & Pritchett, L. (2019), 'The place premium: Bounding the price equivalent of migration barriers', *Review of Economics and Statistics* **101**(2), 201–213.

Costinot, A. & Rodríguez-Clare, A. (2014), Trade theory with numbers: Quantifying the consequences of globalization, *in* 'Handbook of international economics', Vol. 4, Elsevier, pp. 197–261.

Czaika, M. & Parsons, C. R. (2017), 'The gravity of high-skilled migration policies', *Demography* **54**(2), 603–630.

Czaika, M. & Parsons, C. R. (2018), 'High-skilled migration in times of global economic crisis', *High-skilled migration: Drivers and policies* pp. 20–47.

Davis, D. R. & Weinstein, D. E. (2002), Technological superiority and the losses from migration, Technical report, National Bureau of Economic Research.

Delogu, M., Docquier, F. & Machado, J. (2014), 'The dynamic implications of liberalizing global migration'.

Desmet, K., Nagy, D. K. & Rossi-Hansberg, E. (2018), 'The geography of development', *Journal of Political Economy* **126**(3), 903–983.

Di Giovanni, J., Levchenko, A. A. & Ortega, F. (2015), 'A global view of cross-border migration', *Journal of the European Economic Association* **13**(1), 168–202.

Dix-Carneiro, R. & Kovak, B. K. (2017), 'Trade liberalization and regional dynamics', *American Economic Review* **107**(10), 2908–46.

Docquier, F., Machado, J. & Sekkat, K. (2015), 'Efficiency gains from liberalizing labor mobility', *The Scandinavian journal of economics* **117**(2), 303–346.

Eaton, J. & Kortum, S. (2002), 'Technology, geography, and trade', *Econometrica* **70**(5), 1741–1779.

Falck, O., Lameli, A. & Ruhose, J. (2018), 'Cultural biases in migration: Estimating non-monetary migration costs', *Papers in Regional Science* **97**(2), 411–438.

Fallah, B. N., Partridge, M. D. & Olfert, M. R. (2011), 'Urban sprawl and productivity: Evidence from us metropolitan areas', *Papers in Regional Science* **90**(3), 451–472.

Farhi, E. & Werning, I. (2017), 'Fiscal unions', *American Economic Review* **107**(12), 3788–3834.

Feenstra, R. C., Luck, P., Obstfeld, M. & Russ, K. N. (2018), 'In search of the armington elasticity', *Review of Economics and Statistics* **100**(1), 135–150.

Gancia, G., Ponzetto, G. A. & Ventura, J. (2019), 'A theory of economic unions', *Journal of Monetary Economics* .

Guerreiro, J., Rebelo, S. & Teles, P. (2019), What is the optimal immigration policy? migration, jobs and welfare, Technical report, National Bureau of Economic Research.

Hamilton, B. & Whalley, J. (1984), 'Efficiency and distributional implications of global restrictions on labour mobility: Calculations and policy implications', *Journal of Development Economics* **14**(1), 61–75.

Harstad, B. (2007), 'Harmonization and side payments in political cooperation', *American Economic Review* **97**(3), 871–889.

Hunt, G. L. & Mueller, R. E. (2004), 'North american migration: returns to skill, border effects, and mobility costs', *Review of Economics and Statistics* **86**(4), 988–1007.

Iranzo, S. & Peri, G. (2009), 'Migration and trade: Theory with an application to the eastern–western european integration', *Journal of International Economics* **79**(1), 1–19.

Kennan, J. (2013), 'Open borders', *Review of Economic Dynamics* **16**(2), L1–L13.

Kennan, J. (2017), Open borders in the european union and beyond: migration flows and labor market implications, Technical report, National Bureau of Economic Research.

Klein, P. & Ventura, G. J. (2007), 'Tfp differences and the aggregate effects of labor mobility in the long run', *The BE Journal of Macroeconomics* **7**(1).

Lagakos, D., Marshall, S., Mobarak, A. M., Vernot, C. & Waugh, M. E. (2020), 'Migration costs and observational returns to migration in the developing world', *Journal of Monetary Economics* .

Li, C., Wang, J. & Whalley, J. (2019), Trade protectionism and us manufacturing employment, Technical report, National Bureau of Economic Research.

Morten, M. (2019), 'Temporary migration and endogenous risk sharing in village india', *Journal of Political Economy* **127**(1), 1–46.

Morten, M. & Oliveira, J. (2016), Paving the way to development: Costly migration and labor market integration, Technical report, National Bureau of Economic Research.

Moses, J. W. & Letnes, B. (2004), 'The economic costs to international labor restrictions: Revisiting the empirical discussion', *World Development* **32**(10), 1609–1626.

OECD (2019), 'Education at a glance: Earnings by educational attainment (edition 2019/1)'. **URL:** *https://www.oecd-ilibrary.org/content/data/d84b9e18-en*

Ossa, R. (2014), 'Trade wars and trade talks with data', *American Economic Review* **104**(12), 4104–46.

Özak, Ö. (2010), 'The voyage of homo-economicus: Some economic measures of distance', *Manuscript in preparation. Department of Economics, Brown University* .

Parro, F. (2013), 'Capital-skill complementarity and the skill premium in a quantitative model of trade', *American Economic Journal: Macroeconomics* **5**(2), 72–117.

Picard, P. M. & Worrall, T. (2016), 'Is a policy of free movement of workers sustainable?', *The Scandinavian Journal of Economics* **118**(4), 718–754.

Piyapromdee, S. et al. (2014), The impact of immigration on wages, internal migration and welfare, Technical report, Working paper.

Price, J., Lehn, C. v. & Wilson, R. (2020), The winners and losers of immigration: Evidence from linked historical data, Technical report, National Bureau of Economic Research.

Rodrik, D. (2018), 'What do trade agreements really do?', *Journal of economic perspectives* **32**(2), 73–90.

Rosenblum, M. R. & Tichenor, D. J. (2012), *The Oxford handbook of the politics of international migration*, Oxford University Press.

Simonovska, I. & Waugh, M. E. (2014), 'The elasticity of trade: Estimates and evidence', *Journal of international Economics* **92**(1), 34–50.

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. & De Vries, G. J. (2015), 'An illustrated user guide to the world input–output database: the case of global automotive production', *Review of International Economics* **23**(3), 575–605.

Tombe, T. & Zhu, X. (2019), 'Trade, migration, and productivity: A quantitative analysis of china', *American Economic Review* **109**(5), 1843–72.

Wadsworth, J., Dhingra, S., Ottaviano, G. & Van Reenen, J. (2016), 'Brexit and the impact of immigration on the uk', *Centre for Economic Performance. LSE* pp. 34–53.

Waugh, M. E. (2010), 'International trade and income differences', *American Economic Review* **100**(5), 2093–2124.

Weale, A., Pridham, G., Cini, M., Konstadakopulos, D., Porter, M., Flynn, B. et al. (2002), 'Environmental governance in europe: an ever closer ecological union?', *OUP Catalogue* .

Zi, Y. (2020), 'Trade liberalization and the great labor reallocation'.

Appendix A. Household problem

A1.1 Household Problem and Price Index

Recall, that the household optimization problem is given by:

$$\max_{\{\tilde{c}^s_{j,t}(\omega)\}_{\omega\in\Omega_{ji,t}}} \quad \alpha_i \log G_{i,t} + (1-\alpha_i)\log C^s_{i,t}, \tag{45}$$

$$\text{s.t.:} \quad \int_{\omega\in\Omega_{ji,t}} \tilde{p}_{j,t}(\omega)\tilde{c}^s_{j,t}(\omega)d\omega \le \mathrm{w}^s_{i,t}(1-\psi_i). \tag{46}$$

Solving Lagrangian yields the following relations:

$$\frac{\partial Ł/\partial \tilde{c}^s_{j,t}(\omega_1)}{\partial Ł/\partial \tilde{c}^s_{j,t}(\omega_2)} = 0 \quad \Longrightarrow \quad \tilde{c}^s_{j,t}(\omega_1) = \tilde{c}^s_{j,t}(\omega_2)\Big(\frac{\tilde{p}_{j,t}(\omega_1)}{\tilde{p}_{j,t}(\omega_2)}\Big)^{-\rho}, \tag{47}$$
$$\frac{\partial Ł}{\partial \lambda} = 0 \quad \Longrightarrow \quad \int_{\omega\in\Omega_{ji,t}} \tilde{p}_{j,t}(\omega)\tilde{c}^s_{j,t}(\omega)d\omega = \mathrm{w}^s_{i,t}(1-\psi_i).$$

By multiplying the former FOC with $\tilde{p}_{j,t}(\omega_1)$, taking integral from both sides and combining with the latter FOC, we get:

$$\tilde{c}^s_{j,t}(\omega_2) = \frac{\mathrm{w}^s_{i,t}(1-\psi_i)\tilde{p}_{j,t}(\omega_2)^{-\rho}}{\int_{\omega\in\Omega_{ji,t}} \tilde{p}_{j,t}(\omega_1)^{1-\rho}d\omega_1}.$$

Then, following a definition of an aggregated price index $P_{i,t}$ denominating nominal income, it has to satisfy $C^s_{i,t} = \mathrm{w}^s_{i,t}(1-\psi_i)/P_{i,t}$ and, therefore, is given by:

$$P_{i,t} = \Big(\int_{\omega\in\Omega_{ji,t}} \tilde{p}_{j,t}(\omega)^{1-\rho}d\omega\Big)^{\frac{1}{1-\rho}}. \tag{49}$$

Following Eaton & Kortum (2002) it will be later used to derive global distribution of priced in general equilibrium.

A1.1 Expected Lifetime Utility and Migration Shares

The lifetime utility of a household of nationality n residing in i is given by the Bellman equation:

$$v^s_{i,n,t} = \log\Big[\big(G_{i,t}\big)^{\alpha_i}\big(C^s_{i,t}\big)^{(1-\alpha_i)}\Big] + \max_{\{j\}_{j=1}^N}\big\{\beta E[v^s_{j,n,t+1}] - \tilde{m}^s_{ij,n,t} + \nu\epsilon^s_{j,n,t}\big\}.$$

Let us denote expected lifetime utility by $V^s_{i,n,t} \equiv E[v^s_{i,n,t}]$. We assume that $\epsilon^s_{j,n,t}$ follows Type-I Extreme Value distribution and i.i.d. over time. Precisely, it has CDF, denoted by $F(\epsilon) = \exp(-\exp(-\epsilon-\bar{\gamma}))$, where $\bar{\gamma} = \int_{-\infty}^{\infty} x\exp(-x-\exp(-x))dx$ is the Euler's constant and PDF denoted by $f(\epsilon)$. We would like to get rid of the stochastic shock by taking expectations over value function. This operation may be considered as an aggregation of heterogeneous households. Due to properties of the EV distribution taking the expectation yields a relatively simple closed-form solution. Let us denote the expected value function by:

$$\Phi^s_{i,n,t} = E\left[\max_{\{j\}_{j=1}^N}\left\{\beta E[v^s_{j,n,t+1}] - \tilde{m}^s_{ij,n,t} + \nu\epsilon^s_{j,n,t}\right\}\right].$$

Then, we can eliminate *max* operator by introducing an indicator function and then opening it with the use of probability distribution:

$$\begin{aligned}
\Phi^s_{i,n,t} &= \sum_{j=1}^N E\left[\left\{\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t} + \nu\epsilon^s_{j,n,t}\right\} \times I\big(\text{max is at } \{j\}\big)\right],\\
&= \sum_{j=1}^N E\left[\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t} + \nu\epsilon^s_{j,n,t}\,\middle|\, \text{max is at } \{j\}\right] \times \text{Prob}(\text{max is at } \{j\}),\\
&= \sum_{j=1}^N \int_{-\infty}^{+\infty} \big(\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t} + \nu\epsilon^s_{j,n,t}\big) f_\epsilon(\epsilon^s_{j,n,t}) d\epsilon^s_{j,n,t} \times \text{Prob}\big(\phi^s_{ij,n,t} \geq \phi^s_{ik,n,t},\ \forall k \neq j\big).
\end{aligned}$$

where $\phi^s_{ij,n,t} = \beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t} + \nu\epsilon^s_{j,n,t}$.

Now let us derive probability at the right hand side of the equation. We will use the following substitution $\bar{\epsilon}^s_{jk,n,t} = \frac{\beta(V^s_{j,n,t+1} - V^s_{k,n,t+1}) - (\tilde{m}^s_{ij,n,t} - \tilde{m}^s_{ik,n,t})}{\nu}$:

$$\begin{aligned}
\text{Prob}\big(\phi^s_{ij,n,t} \geq \phi^s_{ik,n,t},\ \forall k \neq j\big) &= \text{Prob}\big(\epsilon^s_{k,n,t} \leq \bar{\epsilon}^s_{jk,n,t} + \epsilon^s_{j,n,t},\ \forall k \neq j\big),\\
&= \prod_{k \neq j} \text{Prob}\big(\epsilon^s_{k,n,t} \leq \bar{\epsilon}^s_{jk,n,t} + \epsilon^s_{j,n,t}\big),\\
&= \prod_{k \neq j} F\big(\bar{\epsilon}^s_{jk,n,t} + \epsilon^s_{j,n,t}\big).
\end{aligned}$$

Therefore, the expected value function can be expressed as follows:

$$\begin{aligned}
\Phi^s_{i,n,t} &= \sum_{j=1}^N \int_{-\infty}^{\infty} (\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t} + \nu\epsilon^s_{j,n,t}) f(\epsilon^s_{j,n,t}) \prod_{k \neq j} F(\bar{\epsilon}^s_{jk,n,t} + \epsilon^s_{j,n,t}) d\epsilon^s_{j,n,t},\\
&= \sum_{j=1}^N \int_{-\infty}^{\infty} (\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t} + \nu\epsilon^s_{j,n,t}) e^{(-\epsilon^s_{j,n,t} - \bar{\gamma})} e^{\left(-e^{(-\epsilon^s_{j,n,t} - \bar{\gamma})} \sum_{k=1}^N e^{(-\bar{\epsilon}^s_{jk,n,t})}\right)}.
\end{aligned}$$

Let $\lambda^s_{j,n,t} \equiv \log \sum_{j=1}^N \exp(-\bar{\epsilon}^s_{jk,n,t})$ and consider the following change of variables $\zeta^s_{j,n,t} = \epsilon^s_{j,n,t} + \bar{\gamma}$

$$\Phi^s_{i,n,t} = \sum_{j=1}^N \int_{-\infty}^{\infty} (\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t} + \nu(\zeta^s_{j,n,t} - \bar{\gamma})) \exp(-\zeta^s_{j,n,t} - \exp(-(\zeta^s_{j,n,t} - \lambda^s_{j,n,t}))) d\zeta^s_{j,n,t}.$$

Consider an additional change of variables: $\tilde{y}^s_{j,n,t} = \zeta^s_{j,n,t} - \lambda^s_{j,n,t}$, then

$$\Phi^s_{i,n,t} = \sum_{j=1}^N \exp(-\lambda^s_{j,n,t}) \Bigg((\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t} + \nu(\lambda^s_{j,n,t} - \bar{\gamma})) + \nu \int_{-\infty}^{\infty} \tilde{y}^s_{j,n,t} \exp\left(-\tilde{y}^s_{j,n,t} - \exp(-\tilde{y}^s_{j,n,t})\right) d\tilde{y}^s_{j,n,t} \Bigg),$$

$$\Phi^s_{i,n,t} = \sum_{j=1}^N \exp(-\lambda^s_{j,n,t}) \Big(\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t} + \nu \lambda^s_{j,n,t} \Big).$$

By replacing the definition of $\lambda^s_{j,n,t}$, we get:

$$\Phi^s_{i,n,t} = \sum_{j=1}^N \exp\Big(-\log \sum_{k=1}^N \exp(-\bar{\epsilon}^s_{jk,n,t}) \Big) \Big((\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t} + \nu \log \sum_{k=1}^N \exp(-\bar{\epsilon}^s_{jk,n,t}) \Big).$$

Substituting the definition of $\bar{\epsilon}^s_{jk,n,t}$, we get:

$$\Phi^s_{i,n,t} = \nu \Big(\log \sum_{k=1}^N e^{(\beta V^s_{k,n,t+1} - \tilde{m}^s_{ik,n,t})^{\frac{1}{\nu}}} \Big) \sum_{j=1}^N e^{(\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t})^{\frac{1}{\nu}}} \left(\sum_{k=1}^N e^{(\beta V^s_{k,n,t+1} - \tilde{m}^s_{ik,n,t})^{\frac{1}{\nu}}} \right)^{-1}.$$

Which implies:

$$\Phi^s_{i,n,t} = \nu \Big(\log \sum_{k=1}^N e^{(\beta V^s_{k,n,t+1} - \tilde{m}^s_{ik,n,t})^{\frac{1}{\nu}}} \Big).$$

and therefore

$$V^s_{i,n,t} = \mathrm{U}^s_{i,t} + \nu \log \left(\sum_{j=1}^N e^{\frac{1}{\nu}(\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t})} \right).$$

Let us now derive equation for migration shares. Define $\mu^s_{ij,n,t}$ as the fraction of people of nationality n and skill type s moving from country i to j in period t. Then we assume that the share will be equal to a probability that this household will prefer location j out of all other options. In other words, we seek to find a probability that lifetime utility of a household in location j will be the largest.

$$\begin{aligned}
\mu^s_{ij,n,t} &= \text{Prob}\big(\phi^s_{ij,n,t} \geq \phi^s_{ik,n,t},\ \forall k \neq j\big) \\
&= \text{Prob}\Big(\epsilon^s_{k,n,t} \leq \frac{\beta(V^s_{j,n,t+1} - V^s_{k,n,t+1}) - (m^s_{ij,n,t} - m^s_{ik,n,t})}{v} + \epsilon^s_{j,n,t},\ \forall k \neq j\Big) \\
&= \int_{-\infty}^{\infty} \prod_{k \neq j} F\left(\frac{\beta(V^s_{j,n,t+1} - V^s_{k,n,t+1}) - (m^s_{ij,n,t} - m^s_{ik,n,t})}{v} + \epsilon^s_{j,n,t}\right) f(\epsilon^s_{j,n,t}) d\epsilon^s_{j,n,t}
\end{aligned}$$

Then, using the aforementioned substitution, we get:

$$\begin{aligned}
\mu^s_{ij,n,t} &= \int_{-\infty}^{\infty} \exp(-\epsilon^s_{j,n,t} - \bar{\gamma}) e^{(-e^{(\epsilon^s_{j,n,t} - \bar{\gamma})} \sum_{k=1}^{N} e^{(-\bar{\epsilon}^s_{jk,n,t})})} d\epsilon^s_{j,n,t} \\
&= \exp(-\lambda_{j,n,t}) \int_{-\infty}^{\infty} \exp\left(-\tilde{y}^s_{j,n,t} - \exp(-\tilde{y}^s_{j,n,t})\right) d\tilde{y}^s_{j,n,t} \\
&= \exp(-\lambda_{j,n,t})
\end{aligned}$$

By using the definition of $\lambda_{j,n,t}$, we can now derive equation (11):

$$\mu^s_{ij,n,t} = \frac{e^{\frac{1}{\nu}(\beta V^s_{j,n,t+1} - \tilde{m}^s_{ij,n,t})}}{\sum_{k=1}^{N} e^{\frac{1}{\nu}(\beta V^s_{k,n,t+1} - \tilde{m}^s_{ik,n,t})}}$$

A1.2 Value Function in Welfare Estimation

In this Section, we explain how to derive value function used to compute governmental welfare. Let us add and subtract expected lifetime utility inside the value of an option move:

$$V^s_{i,n=i,t} = \text{U}^s_{i,t} + \beta V^s_{i,n=i,t+1} + \nu \log\left(\sum_{j=1}^{N} e^{\frac{1}{\nu}\left(\beta(V^s_{j,n=i,t+1} - V^s_{i,n=i,t+1}) - \tilde{m}^s_{ij,n=i,t}\right)}\right)$$

Notice, that for nationals residing in their homeland:

$$\mu^s_{ii,n=i,t} = \frac{e^{\frac{1}{\nu}(\beta V^s_{j,n,t+1})}}{\sum_{k=1}^{N} e^{\frac{1}{\nu}(\beta V^s_{k,n=i,t+1} - \tilde{m}^s_{ik,n=i,t})}}$$

Inverting previous equation yields the following relation:

$$\nu \log\left(\sum_{j=1}^{N} e^{\frac{1}{\nu}\left(\beta(V^s_{j,n=i,t+1} - V^s_{i,n=i,t+1}) - \tilde{m}^s_{ij,n=i,t}\right)}\right) = -\nu \log \mu^s_{ii,n=i,t}$$

Therefore,

$$V^s_{i,n=i,t} = \text{U}^s_{i,t} + \beta V^s_{i,n=i,t+1} - \nu \log \mu^s_{ii,n=i,t}$$

By iteration this equation forward we obtain:

$$V^s_{i,n=i,t} = \sum_{q=t}^{\infty} \beta^{q-t} \mathrm{U}^s_{i,t} - \nu \sum_{q=t}^{\infty} \beta^{q-t} \log \mu^s_{ii,n=i,t}$$

Finally, we can derive equation (28) for natives residing in their homeland using the definition of our logarithmic utility:

$$V^s_{i,n=i,t} = \sum_{q=t}^{\infty} \beta^{q-t} \log \frac{\left(G_{i,q}\right)^{\alpha_i} \left(C^s_{i,q}\right)^{(1-\alpha_i)}}{(\mu^s_{ii,n=i,q})^{\nu}}$$

Now let us derive value function for emigrants:

$$V^s_{r,n=i,t} = \mathrm{U}^s_{r,n,t} + \beta V^s_{r,n=i,t+1} + \nu \log \left(\sum_{j=1}^{N} e^{\frac{1}{\nu}\left(\beta(V^s_{j,n=i,t+1} - V^s_{r,n=i,t+1}) - \tilde{m}^s_{ij,n=i,t}\right)} \right)$$

By inverting migration share $\mu^s_{rr,n=i,t}$ of emigrants decided to stay in the same foreign country, we obtain slightly different relation:

$$\nu \log \left(\sum_{j=1}^{N} e^{\frac{1}{\nu}\left(\beta(V^s_{j,n=i,t+1} - V^s_{r,n=i,t+1}) - \tilde{m}^s_{ij,n=i,t}\right)} \right) = -\nu \log \mu^s_{ii,n=i,t} - m^{s,pol}_{rr,n=i,t}$$

Then the expected lifetime utility of an emigrant can be expressed as follows:

$$V^s_{r,n=i,t} = \sum_{q=t}^{\infty} \beta^{q-t} \left[\log \frac{\left(G_{r,q}\right)^{\alpha_r} \left(C^s_{r,q}\right)^{(1-\alpha_r)}}{(\mu^s_{rr,n=i,q})^{\nu}} - m^{s,pol}_{rr,n=i,q} \right]$$

$$V^s_{r,n=i,t} = \sum_{q=t}^{\infty} \beta^{q-t} \left[\log \frac{\exp(\mathrm{U}^s_{r,n,q})}{(\mu^s_{rr,n=i,q})^{\nu}} - m^{s,pol}_{rr,n=i,q} \right]$$

Appendix B. Firms problem

Let us recall the production function:

$$Y_{i,t}(\omega) = z_{i,t}(\omega) \left[\left(\theta^h_i (L^h_{i,t})^{\frac{\sigma-1}{\sigma}} + \theta^l_i (L^l_{i,t})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \right]^{(1-\gamma_i)} \Big(H_i \Big)^{\gamma_i} \tag{50}$$

Let us denote $Q_{i,t} = \theta^h_i (L^h_{i,t})^{\frac{\sigma-1}{\sigma}} + \theta^l_i (L^l_{i,t})^{\frac{\sigma-1}{\sigma}}$ for simplicity. Thus, we seek to solve the following profit maximization problem to infer input bundle costs, trade shares and price index:

$$\max_{\{L^l_{i,t}, L^h_{i,t}\}} p_{i,t}(\omega) \left[z_{i,t}(\omega) Q_{i,t}^{\frac{\sigma}{\sigma-1}(1-\gamma_i)} \Big(H_i \Big)^{\gamma_i} \right] - \mathrm{w}^l_{i,t} L^l_{i,t} - \mathrm{w}^h_{i,t} L^h_{i,t} - r_{i,t} H_i$$

Then, using FOCs we get the wage ratio and the output value:

$$\frac{\mathrm{w}^h_{i,t}}{\mathrm{w}^l_{i,t}} = \frac{\theta^h_i}{\theta^l_i}\Big(\frac{L^l_{i,t}}{L^h_{i,t}}\Big)^{\frac{1}{\sigma}} \tag{51}$$

$$p_{i,t}(\omega)Y_{i,t}(\omega) = \frac{Q_{i,t}\mathrm{w}^l_{i,t}(L^l_{i,t})^{\frac{1}{\sigma}}}{\theta^l_i(1-\gamma_i)} \tag{52}$$

From zero profit condition we know that:

$$r_{i,t}H_i + \big(\mathrm{w}^l_{i,t}L^l_{i,t} + \mathrm{w}^h_{i,t}L^h_{i,t}\big) = p_{i,t}(\omega)Y_{i,t}(\omega) \tag{53}$$

Then, by combining equations (51),(52) and (53) we get the value of rental payments:

$$r_{i,t}H_i = \Big(\frac{\gamma}{1-\gamma}\Big)\big(\mathrm{w}^l_{i,t}L^l_{i,t} + \mathrm{w}^h_{i,t}L^h_{i,t}\big) \tag{54}$$

By plugging equation (54) back into equation (52) we get the input bundle costs:

$$p_{i,t}(\omega)z_{i,t}(\omega) = \Omega^i\Bigg[\big(\mathrm{w}^l_{i,t}L^l_{i,t} + \mathrm{w}^h_{i,t}L^h_{i,t}\big)\Big(\theta^h_i(L^h_{i,t})^{\frac{\sigma-1}{\sigma}} + \theta^l_i(L^l_{i,t})^{\frac{\sigma-1}{\sigma}}\Big)^{\frac{\sigma}{1-\sigma}}\Bigg]^{1-\gamma_i} r^{\gamma_i}_{i,t} \tag{55}$$

$$= \Omega^i\Bigg[(\theta^h_i)^{\sigma}(\mathrm{w}^h_{i,t})^{1-\sigma} + (\theta^l_i)^{\sigma}(\mathrm{w}^l_{i,t})^{1-\sigma}\Bigg]^{\frac{1-\gamma_i}{1-\sigma}} r^{\gamma_i}_{i,t} \tag{56}$$

$$\equiv x_{i,t} \tag{57}$$

where $\Omega^i = \frac{1}{1-\gamma_i}\Big(\frac{1-\gamma_i}{\gamma_i}\Big)^{\gamma_i}$.

Following Eaton & Kortum (2002), using distance between trading partners and import tariffs, we express the price of a variety ω produced in country j and shipped to country i as:

$$p_{ji,t}(\omega) = \frac{x_{j,t}}{z_{j,t}}k_{ji} \tag{58}$$

Then, country j presents in country i with a distribution of prices:

$$G_{ji}(p) \equiv Pr\big[p_{ji,t}(\omega) \le p\big] = 1 - e^{-T_{j,t}\big(x_{j,t}k_{ji}\big)^{-\theta}p^{\theta}} \tag{59}$$

However, country i will buy only from those countries that can provide the lowest price. Hence the actual distribution of prices, the country i agree to buy for, will be given by:

$$G_i(p) = 1 - e^{\hat{\Phi}_{i,t}p^{\theta}}, \quad \hat{\Phi}_{i,t} = \sum_{j=1}^{N} T_{j,t}\big(x_{j,t}k_{ji}\big)^{-\theta} \tag{60}$$

Then the probability that country i provides a variety at the lowest price in country j:

$$\pi_{ij,t} = \int_0^{\infty}\prod_{k\neq j}\Big[1 - G_{ki}(p)\Big]dG_{ji}(p) = \frac{T_{j,t}/(x_{j,t}k_{ij,t})^{\theta}}{\sum_{b=1}^{N} T_{b,t}/(x_{b,t}k_{ib,t})^{\theta}} \tag{61}$$

Since the efficiency in producing each variety is drawn from identical distribution, the following is true:

$$P_{i,t} = \Big(\int_{\omega \in \Omega_{ji,t}} \tilde{p}_{j,t}(\omega)^{1-\rho} d\omega\Big)^{\frac{1}{1-\rho}} = \Big(\int_0^{\infty} \tilde{p}_{j,t}^{1-\rho} dG_i(p)\Big)^{\frac{1}{1-\rho}} \tag{62}$$

$$= \bar{\rho}\hat{\Phi}_{i,t}^{-\frac{1}{\theta}}, \quad \text{where } \bar{p} = \Big[\Gamma(1 + \frac{1-\rho}{\theta})\Big]^{\frac{1}{1-\rho}} \tag{63}$$

Appendix C. Additional results

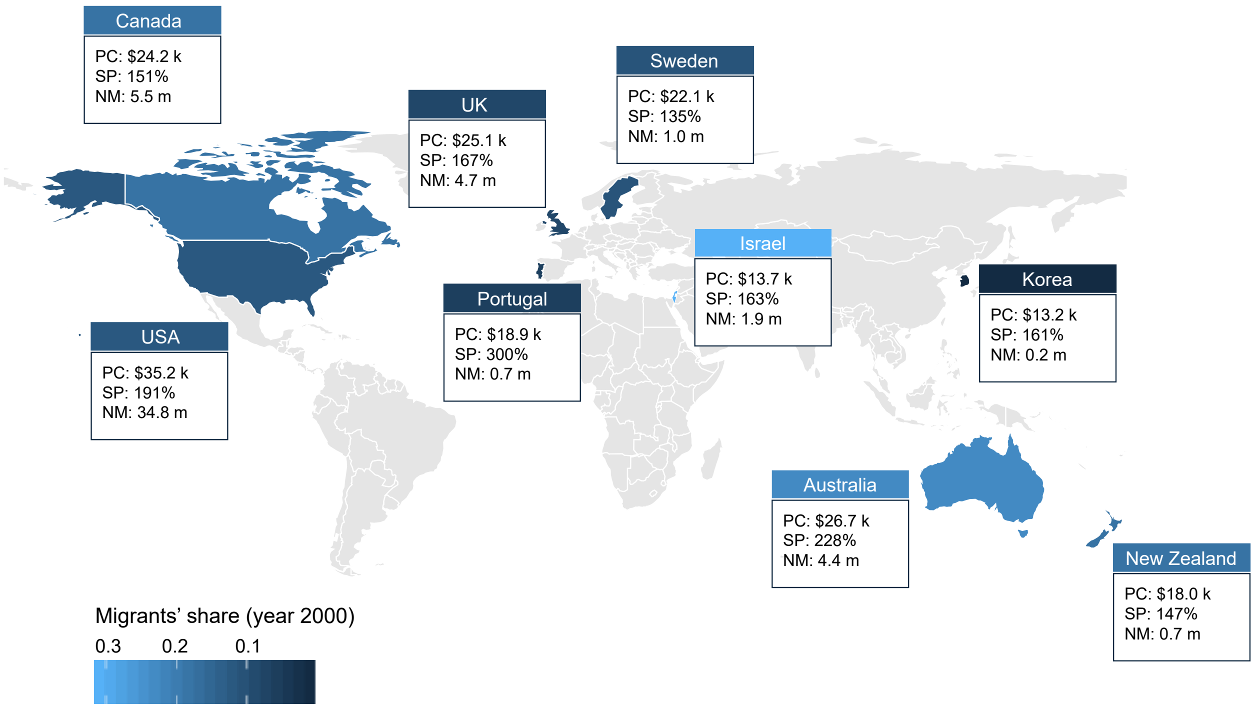
Table 3: Welfare Gains from establishing a Union: Appendix

Type	1st Party		2nd Party		Aggregate	Rest of the world
			Unreasonable			
Join	ISR	-3.239	RoW	-0.019	-0.021	-0.003
Join	US	-0.118	EU	0.129	-0.064	∼0
Join	RoW	-0.095	EU	-2.916	-0.126	0.033
Join	KOR	-5.349	RoW	-0.133	-0.169	0.014
Join	KOR	-4.918	EU	2.872	-0.170	∼0
Join	RoW	-0.184	Pacific	-6.446	-0.200	0.037
Join	Free movement (See Figure 3.)				-0.272	
Join	CAN	-0.412	Pacific	-0.175	-0.310	∼0
Join	CAN	1.303	RoW	-0.373	-0.367	0.006
Join	US	-3.379	RoW	-0.664	-0.766	0.224
Join	CAN	-0.194	EU	-0.950	-0.770	∼0
Join	KOR	-3.572	US	-0.282	-0.783	-0.002
Join	EU	-0.979	Pacific	-0.544	-0.896	-0.002
Join	CAN	-0.546	KOR	-2.779	-2.047	-0.001
Join	KOR	-2.762	Pacific	-1.693	-2.475	-0.021
Join	US & CANZUK (See Figure 3.)				-3.075	-0.003
Join	US	-3.421	Pacific	-0.115	-3.217	-0.014

Note: The table represents welfare gains in percentage points measured as the ratio of welfare levels between counterfactual and baseline economy. Unions are sorted by the aggregate gains in the descending order. The EU includes {PRT, SWE, UK}, the Pacific union includes {AUS, NZL}, high-skilled and low-skilled labor is denoted by *HS* and *LS* respectively.

Appendix D. Map

Figure 8: OECD countries and the rest of the world



Note: PC is the private consumption, SP is the skill premium and NM is the number of migrants.

Appendix E. Migration costs

This paper attempts to disaggregate migration costs into policy-related (endogenous) and fixed (exogenous) components by destination, skill, and nationality based on the aforementioned literature and modeling assumptions. First of all, from equation (11) the ratio between migration share of people moving from country i to j and the share of people staying in the country i is given by:

$$\frac{\mu^s_{ij,n,0}}{\mu^s_{ii,n,0}} = e^{\frac{1}{\nu}\left(\beta V^s_{j,n,1} - \beta V^s_{i,n,1} - (m^s_{ij,n,0} - m^s_{ii,n,0})\right)}. \tag{64}$$

From the equation of the aggregate life-time utility (10):

$$V^s_{j,n,1} - V^s_{i,n,1} = \log\left(\frac{(G_{j,1})^{\alpha_j}(C^s_{j,1})^{(1-\alpha_j)}}{(G_{i,1})^{\alpha_i}(C^s_{i,1})^{(1-\alpha_i)}}\right) + \nu \log\left(\frac{\sum_{k=1}^N e^{\frac{1}{\nu}}(\beta V^s_{k,n,2} - m^s_{jk,n,1})}{\sum_{k=1}^N e^{\frac{1}{\nu}}(\beta V^s_{k,n,2} - m^s_{ik,n,1})}\right), \tag{65}$$

$$= \log\left(\frac{(G_{j,1})^{\alpha_j}(C^s_{j,1})^{(1-\alpha_j)}(\mu^s_{in,n,1})^{\nu}}{(G_{i,1})^{\alpha_i}(C^s_{i,1})^{(1-\alpha_i)}(\mu^s_{jn,n,1})^{\nu}}\right) + \left(m^s_{in,n,1} - m^s_{jn,n,1}\right). \tag{66}$$

By plugging equation (66) back into (64) and using assumption (25) one can derive a system of equations:

$$\nu \log\left(\frac{\mu^s_{ij,n,0}}{\mu^s_{ii,n,0}}\right) = \beta \log\left(\frac{(G_{j,1})^{\alpha_j}(C^s_{j,1})^{(1-\alpha_j)}(\mu^s_{in,n,1})^{\nu}}{(G_{i,1})^{\alpha_i}(C^s_{i,1})^{(1-\alpha_i)}(\mu^s_{jn,n,1})^{\nu}}\right) + \beta(m^{s,fix}_{in,n} - m^{s,fix}_{jn,n}) + (m^s_{ii,n,0} - m^s_{ij,n,0}). \tag{67}$$

More precisely, using (6) and variations of (67) for different combinations of i,j and n policy and fixed migration costs can be derived directly:

$$m^{s,fix}_{ij,n} = (m^{s,fix}_{in,n} - m^{s,fix}_{jn,n}) + \nu \log\left(\frac{\mu^s_{in,n,0}}{\mu^s_{ij,n,0}} \cdot \frac{\mu^s_{jj,n,0}}{\mu^s_{jn,n,0}}\right), \tag{68}$$

$$m^{s,pol}_{ij,n,0} = (1-\beta)m^{s,fix}_{jn,n} + \nu \log\left(\frac{((G_{j,1})^{\alpha_j}(C^s_{j,1})^{(1-\alpha_j)})^{\beta/\nu}}{((G_{n,1})^{\alpha_n}(C^s_{n,1})^{(1-\alpha_n)})^{\beta/\nu}} \cdot \frac{\mu^s_{jn,n,0}}{\mu^s_{jj,n,0}} \cdot \frac{(\mu^s_{nn,n,1})^{\beta}}{(\mu^s_{jn,n,1})^{\beta}}\right). \tag{69}$$

Appendix F. Equilibrium in time differences

Now we describe the equilibrium conditions for the model in time differences. By analogy, in every period the state of economy is given by the distribution of labor across countries $\dot{L}_t = \{\dot{L}^l_{i,t}, \dot{L}^h_{i,t}\}_{i=1}^N$.

Given the labor allocation $\{\dot{L}_t\}_{t=0}^{\infty}$, parameters and initial values of trade shares and labor compensation $\{\pi_{i,0}, \mathrm{w}^l_{i,0}L^l_{i,0}, \mathrm{w}^h_{i,0}L^h_{i,0}\}$, the set of wages and price indexes $\{\dot{w}^l_{i,t}, \dot{w}^h_{i,t}, \dot{r_{i,t}}\dot{P}_{i,t}\}$ solves the **temporary equilibrium** and provides corresponding estimates of co-state variables $\{\dot{x}_{i,t}, \dot{f}^s_{i,t}, \dot{T}_{i,t}, \dot{\pi}_{ij,t}\}$:

$$\dot{x}_{i,t+1} = \Big[\frac{f^h_{i,t}}{f^h_{i,t}+f^l_{i,t}}(\dot{w}^h_{i,t+1})^{1-\sigma} + \frac{f^l_{i,t}}{f^h_{i,t}+f^l_{i,t}}(\dot{w}^l_{i,t+1})^{1-\sigma}\Big]^{\frac{1-\gamma_i}{1-\sigma}}\dot{r}^{\gamma_i}_{i,t+1}, \tag{70}$$

$$\dot{r}_{i,t+1} = \frac{f^h_{i,t}}{f^l_{i,t}+f^h_{i,t}}\dot{f}^h_{i,t+1} + \frac{f^l_{i,t}}{f^l_{i,t}+f^h_{i,t}}\dot{f}^l_{i,t+1}, \qquad f^s_{i,t} = \mathrm{w}^s_{i,t}L^s_{i,t},$$

$$\dot{P}_{i,t+1} = \left(\sum_{j=1}^N \pi_{ij,t}\dot{T}_{j,t+1}/\big(\dot{x}_{j,t+1}\dot{k}_{ij,t+1}\big)^{\theta}\right)^{-\frac{1}{\theta}}, \tag{71}$$

$$\dot{T}_{i,t+1} = \dot{\tau}^T_i\big(\dot{D}_{i,t}\big)^{\zeta} = \left(\frac{L^l_{i,t}}{L_{i,t}}\dot{L}^l_{i,t+1} + \frac{L^h_{i,t}}{L_{i,t}}\dot{L}^h_{i,t+1}\right)^{\zeta}, \quad \text{where} \quad \dot{\tau}^T_i = 1,$$

$$\dot{\pi}_{ij,t+1} = \dot{T}_{j,t+1}\Big(\dot{x}_{j,t+1}\dot{k}_{ij,t+1}/\dot{P}_{i,t+1}\Big)^{-\theta}, \tag{72}$$

$$\Big(1-\sum_{k=1}^N \tau_{jk,t+1}\frac{\pi_{jk,t+1}}{(1+\tau_{jk,t+1})}\Big)X_{j,t+1} = \sum_{s=h,l}\mathrm{w}^s_{j,t}L^s_{j,t}\dot{w}^s_{j,t+1}\dot{L}^s_{j,t+1} + \iota^j\sum_{k=1}^N \dot{r}_{k,t+1}(r_{k,t}H_k), \tag{73}$$

$$\dot{r}_{k,t+1}r_{k,t}H_k = \frac{\gamma_k}{1-\gamma_k}\big(f^l_{k,t}\dot{f}^l_{k,t+1} + f^h_{k,t}\dot{f}^h_{k,t+1}\big),$$

$$\mathrm{w}^h_{i,t}L^h_{i,t}\big(\dot{w}^h_{i,t+1}\dot{L}^h_{i,t+1}\big) + \mathrm{w}^l_{i,t}L^l_{i,t}\big(\dot{w}^l_{i,t+1}\dot{L}^l_{i,t+1}\big) = (1-\gamma_i)\sum_{j=1}^N\frac{\pi_{ji,t+1}}{(1+\tau_{ji,t+1})}X_{j,t+1}, \tag{74}$$

$$\frac{\dot{w}^h_{i,t+1}}{\dot{w}^l_{i,t+1}} = \Big(\frac{\dot{L}^l_{i,t+1}}{\dot{L}^h_{i,t+1}}\Big)^{\frac{1}{\sigma}}.$$

As one can notice, the solution to the system of equations does not require knowledge of geography barriers, labor share parameters or the initial level of technologies. Equilibrium values of nominal wages and prices are then used to compute a set of real wages, denoted by $\{\omega^l_{i,t}, \omega^h_{i,t}\}$ and public goods $\{G_{i,t}\}$ that correspond to consumption $\{C^l_{i,t}, C^h_{i,t}\}$ in our model.

Let us denote the difference in consumer's life-time utility by $\dot{u}^s_{i,t+1} \equiv \exp\left(V^s_{i,t+1} - V^s_{i,t}\right)$

and in migration costs by $\dot{m}^s_{ij,n,t+1} \equiv \dot{\tilde{m}}^s_{ij,n,t+1} = \exp\left(m^{s,pol}_{ij,n,t+1} - m^{s,pol}_{ij,n,t} + m^{s,fix}_{ij,n} - m^{s,fix}_{ij,n}\right)$. Hence, the fixed part of migration costs is eliminated and one can proceed only with policy costs to a conditional sequential equilibrium:

$$\dot{u}^s_{i,t+1} = \left[\left(\dot{G}_{i,t}\right)^{\alpha_i}\left(\frac{\dot{\mathrm{w}}^s_{i,t}}{\dot{P}_{i,t}}\right)^{(1-\alpha_i)}\right]\left(\sum_{j=1}^{N}\mu^s_{ij,n,t}(\dot{u}^s_{j,n,t+2})^{\frac{\beta}{\nu}}/(\dot{m}^s_{ij,n,t+1})^{\frac{1}{\nu}}\right)^{\nu}, \tag{75}$$

$$\mu^s_{ij,n,t+1} = \frac{\mu^s_{ij,n,t}(\dot{u}^s_{j,n,t+2})^{\frac{\beta}{\nu}}/(\dot{m}^s_{ij,n,t+1})^{\frac{1}{\nu}}}{\sum_{k=1}^{N}\mu^s_{ik,n,t}(\dot{u}^s_{k,n,t+2})^{\frac{\beta}{\nu}}/(\dot{m}^s_{ik,n,t+1})^{\frac{1}{\nu}}}, \tag{76}$$

$$L^s_{i,n,t+1} = \left(\sum_{j=1}^{N}\mu^s_{ji,n,t}L^s_{j,n,t}\right). \tag{77}$$

Appendix G. Migration shares and Probability matrix

To construct a matrix of probabilities to migrate and calculate migration shares disaggregated by origin, destination, skill, and nationality, we employ the following procedure:

Step 1. Divide the total flow of migrants into high-skilled and low-skilled flows.

Source: International migration flows to and from selected countries: The 2015 revision.

Let $A^h_{\cdot\to j,n,t}$ be the share of high-skilled workers of nationality n coming from all origins to j. It can be obtained directly from Czaika & Parson database. Also denote $B^h_{j,n,t}$ as the share of high-skilled workers of nationality n residing in j. This measure can be calculated using the dataset on stocks of labor.

Then, assuming that origin i corresponds to nationality n, introduce an averaged share of high-skilled workers moving from origin i to destination j:

$$A^h_{ij,t} = A^h_{nj,t} = \frac{A^h_{\cdot\to j,n,t} + B^h_{j,n,t}}{2}. \tag{78}$$

Thus, flows can be disaggregated by skill in the following way:

$$L^h_{ij,t} = A^h_{ij,t} L_{ij,t} \quad \text{and} \quad L^l_{ij,t} = (1 - A^h_{ij,t}) L_{ij,t}. \tag{79}$$

Step 2. Infer probability that a randomly drawn worker from a country of residence i has nationality n.

Source: Data on stocks of migrants

$$Pr^s_{i,n,t} = L^s_{i,n,t} / L^s_{i,t}. \tag{80}$$

Step 3. Create a table with estimated flows by origin-destination-nationality-skill.

$$L^s_{ij,n,t} = L^s_{ij,t} Pr^s_{i,n,t}. \tag{81}$$

Step 4. Transform flows to a matrix of probabilities that a worker of type s, nationality n moved from origin i to destination j.

$$Pr^s_{ij,n,t} = L^s_{i,j,n,t} / \sum_{i\neq j} L^s_{ij,n,t}. \tag{82}$$

Step 5. Disaggregate Czaika & Parson flows by origin using Probability Matrix (Step 4).

$$\widehat{L^s_{ij,n,t}} = Pr^s_{ij,n,t} \cdot \widehat{\left(\sum_{i\neq j} L^s_{ij,n,t}\right)} \quad \forall i \neq j, \tag{83}$$

where $\widehat{\left(\sum_{i\neq j} L^s_{ij,n,t}\right)}$ are flows from Czaika & Parson.

Step 6. Balance the data by adding return migration flows.

$$\widehat{L^s_{ii,n,t}} = L^s_{i,n,t} - \sum_{i\neq j} \widehat{L^s_{ij,n,t}}. \tag{84}$$

Step 7. Infer the stock of migrants at the beginning of period t ($\underline{t}$).

$$\widehat{L^s_{i,n,\underline{t}}} = \sum_{j=1}^{N} \widehat{L^s_{ij,n,t}}. \tag{85}$$

Step 8. Calculate migration shares.

$$\mu^s_{ij,n,t} = \widehat{L^s_{ij,n,t}} / \widehat{L^s_{i,n,\underline{t}}}. \tag{86}$$

Step 9. Use $\widehat{L^s_{i,n,\underline{t}}}$ as stocks of migrants in the model.

Appendix H. Parameters

Table 4: Global and country-specific parameters

Utility function		
$\nu = 1.89$	Elasticity of migration	Caliendo et al. (2017)
$\rho = 3.5$	Elasticity of demand substitution	Feenstra et al. (2018)
$\beta = .85$	Time-discount factor (5-year equiv. of annual 0.97)	
$\alpha_i = .11^*$	Fraction of public goods in total consumption	OECD National Accounts
$\tau_L^i = .31^*$	Labor income tax	OECD Tax Database
Production function		
$\sigma = 4$	Elasticity of factor substitution	Caliendo et al. (2017)
$\theta^h = .56^*$	Share of high-skilled labor in labor payments	Earnings premium (OECD), $L_{i,0}^s$
$(1 - \gamma_i) = .49^*$	Share of labor payments in value added	OECD Input Output Tables
$\theta = 4.14$	Variation in productivity distribution (*à la* EK(2002))	Simonovska & Waugh (2014)
$\zeta = .06$	Static elasticity of productivity to labor density	Carlino et al. (2007)
$\eta_i = 10.51^*$	Initial state of technology	WTF, CEPII, PWT, ζ
$H_i = .03^*$	Fixed amount of structures	PWT, OECD
d_{ij}	Distance between countries	CEPII
τ_{ij}	Ad-valorem tariffs	WITS (TRAINS, WTO IDB)
Government function		
λ_i^s	Political weight of high-skilled workers	Dynamic: $L_{i,t}^s/L_{i,t}$
λ_i^r	Political weight of natives by country of residence (r)	Dynamic: $L_{i,n,t}/L_n$
$\lambda_i^s = .5$	Political weight of high-skilled workers	for Egalitarian society
$\lambda_i^r = 1/N$	Political weight of natives by country of residence (r)	for Egalitarian society
(*) - median value across countries		

{NOTES}