LISER

LUXEMBOURG INSTITUTE OF SOCIO-ECONOMIC RESEARCH

n° 2019-10 July 2019

WORKING PAPERS

Mexican Migration to the United States: Selection, Assignment, and Welfare

Michał **BURZYŃSKI** [1]
Paweł **GOLA**[2]

[1] Luxembourg Institute of Socio-Economic Research (LISER), Luxembourg
[2] Norwegian Business School (BI), Norway

www.liser.lu

Mexican Migration to the United States: Selection, Assignment, and Welfare*

Michał Burzyński

Luxembourg Institute of Socio-Economic Research (LISER)

Paweł Gola

Norwegian Business School (BI)

June 2019

Abstract

This paper quantifies the effects of Mexican migration to the United States on individual welfare along the continuous distribution of skills in both countries. We develop a model that focuses on the sorting of workers within and across national labor markets. Mexican workers self-select into migration, and then, within each country, all workers match with productivity-differentiated firms. Firms operate in monopolistically competitive international markets, which they can freely enter or exit. These features of the model ensure that workers with similar skills are substitutes and dissimilar workers are complements. Thus, migration redistributes welfare in the source and host country. In particular, the observed Mexican immigration to the United States depresses the wages of below-median local workers. However, the welfare losses in the United States are modest in scope: A \$1.70 per day lump-sum tax on Mexican immigrants is sufficient to finance a compensating transfer for all U.S. citizens.

Keywords: Migration; Matching; Selection; Welfare.
JEL classification: C68, C78, F22, J24.

*Contact: Michał Burzyński: michal.burzynski@liser.lu; Paweł Gola: pawel.gola@bi.no.

1 Introduction

Immigrants do not make up a random sample of the population from the sending country; neither are workers randomly assigned to firms within countries. These two facts are closely connected: the distribution of skills among migrants affects which jobs are performed by the native-born workers, whereas the kind of jobs offered to immigrants determines who decides to migrate. While the literature has long recognized the importance of selection into migration (Borjas, 1987; Chiquiar and Hanson, 2005; Moraga, 2011; Kaestner and Malamud, 2014; Borjas et al., 2018), it has been much slower to investigate the impact that migration has on the within-country assignment of workers to jobs (Peri and Sparber, 2009; Foged and Peri, 2016), and has been silent on the interacting effects.

In this paper, we analyze the impact that migration has on the labor markets in sending and destination countries, by proposing a theoretical model that accounts for how workers are sorted both across and within countries. We embed this labor market into a tractable general equilibrium framework in which agents consume domestically produced and internationally traded goods. We then calibrate the model and quantify the impact that more restrictive and more liberal U.S. immigration policies targeted towards Mexico have on the distributions of real wages for U.S. citizens, Mexican migrants, and Mexican stayers.[1]

Formally, we build a two-country assignment model that endogenizes the within-country supply of skills and firms by embedding the framework of Gola (2018) into a general equilibrium framework with international trade. Workers are endowed with continuously distributed vectors of skills. Within each country, firms agree on how the workers are ranked, but these rankings vary across countries, exactly as in Roy (1951), Heckman and Sedlacek (1985), and Borjas (1987). Productivity is continuously distributed among firms and differs both within- and across countries, and high-productivity firms are complements to the workers whom they rank highly.[2] Consequently, workers match with firms positively and assortatively, as in Sattinger (1979), Dupuy (2012), Dupuy (2015), and Mak and Siow (2017). The goods market is monopolistically competitive, and the supply of firms is endogenized in the same way as in Melitz (2003). All individuals exhibit love of variety over a continuous set of imperfectly substitutable consumption goods, following

[1] Throughout the paper, we refer to the non-Mexican workers that reside in the United States as "U.S. citizens". Of course, this designation is a simplification, because this group consists not only of U.S. citizens but also migrants from other countries, and because some U.S. citizens live outside of the United States.

[2] The marginal product of worker's rank increases with the productivity of the firm she is assigned to.

Dixit and Stiglitz (1977). All active firms serve domestic and foreign markets, as in Krugman (1980).

The model is calibrated to represent the U.S. and Mexican economies in 2015. We conduct three migration policy experiments. First, in order to quantify the economic impact of Mexican migration on both economies, we set the costs of migration from Mexico to the United States to infinity. The results for the partial labor market equilibrium—where the demand for each variety of the consumption good is held constant—paint a fairly grim picture of the distributional impact of migration: 68 percent of U.S. citizens (located in the left tail of the wage distribution) lose from Mexican migration, while only 23 percent are better off. In general equilibrium, however, there is an additional gain from migration, as the greater number of workers translates into a larger number of varieties produced in that country. Additionally, Mexican migration has also a (slightly) negative effect on the fiscal position of the U.S. economy, which we quantify in an out-of-equilibrium exercise.[3] Once these market size and fiscal effects are accounted for, the share of U.S. citizens that gain from Mexican immigration rises to 44 percent. The fact that the share of U.S. population that gains from Mexican migration is just short of majority arguably rationalizes the polarization of U.S. voters on this issue. In Mexico, 75 percent of native stayers (the low- and medium-skilled) gain from emigration in the partial labor market equilibrium, whereas 20 percent of them (predominantly high earners) are worse-off. Once the market size and fiscal effects are factored in, the share of better-off Mexicans is reduced to 43 percent.

Second, we consider the effect of small changes in the monetary cost of migration: this could be caused, for example, by an increase in the cost of a visa. In the partial labor market equilibrium, an increase in the per-day migration cost of 3 USD reduces the population of Mexican immigrants in the United States by 30 percent and decreases the share of U.S. citizens that gain from immigration to approximately 20 percent. Once all the economic effects are taken into account, such an increase improves the real wages of virtually all U.S. citizens. In our third experiment, we assume that the increase in migration costs is caused by a tax on migration, the proceeds of which are then redistributed among U.S. citizens. We find that a tax of just 1.7 USD per-day (4 USD when compensating for labor market effects only) is sufficient to make all U.S. workers better off from Mexican immigration. This result illustrates that even though our model shows that a large share of the U.S. population is worse off due to Mexican immigration, their losses

[3]This is modeled as amount of extra transfer payments that each U.S. citizen needs to pay to balance the budget once Mexican migrants pay their taxes and receive their benefits.

are small in magnitude, and are easily compensated.

We will now explain how our theory relates to other work, especially the constant elasticity of substitution (CES) model, which is the workhorse model in the migration literature (Borjas, 2003; Ottaviano and Peri, 2012; Docquier et al., 2014). The key difference between the standard CES model and the assignment model employed in this paper is that in the CES model workers are *pre-assigned* to their jobs. This means that there is no natural ranking of workers: low-skilled workers, if in scarce supply, can earn more than high-skilled workers.[4] In the assignment model, every worker can perform any job, and high-skilled workers enjoy an absolute advantage over low-skilled workers. This difference proves to be fundamental for two reasons. First, it gives rise to very different patterns of complementarity between low- and high-skilled workers. In the CES model low-skilled workers are not similar to medium- or high-skilled workers. Consequently, as shown by Dustmann et al. (2013), if the supply of capital is perfectly elastic, then an influx of low-skilled migrants only lowers the wages of low-skilled natives but raises the wages of all other worker types by the same proportion.[5] In the assignment model, the impact of low-skilled migration on the wage of a given native-born worker depends on how similar her skills are to the migrants' skills. This feature introduces distance-dependent elasticities of substitution (DIDES) (Teulings, 2005). In particular, if firms' entry is endogenous, an influx of low-skilled migrants will lower the wages of native-born workers with the same or similar skills, and increase the wages of natives with very different skills (Teulings, 1995, 2005; Costrell and Loury, 2004; Gola, 2018).[6]

To see why the elasticity of substitution depends on the distance in skill endowments, it is instructive to decompose the labor market response to an influx of low-skilled immigrants. Holding the supply of firms constant, low-skilled native workers are replaced by the migrants, and the native-born end up with worse jobs. Furthermore, all natives earn lower wages, as all firms have now better outside op-

[4]The problem with the lack of a natural ranking of skills in the CES model has been acknowledged by Dustmann et al. (2013). They overcome this difficulty by formulating an empirical test of whether the ranking of skills is preserved after an inflow of immigrants ("rank insensitivity" test). Interestingly, an alternative solution proposed by Dustmann et al. (2013) allows for high-skilled workers to perform low-skilled jobs, which is precisely what the assignment model does.

[5]Note that the reasoning above applies directly only to the standard CES model, which has a single elasticity of substitution. A nested CES model imposes some ranking on worker types but cannot be straightforwardly extended to settings with continuous skill types; further, the ranking is not completely exogenous, implying that standard Roy model techniques cannot be employed even with nested CES.

[6]Figure 2 in Teulings (2005) is particularly illuminating as far as the difference in the patterns of complementarity between an assignment and a CES model is concerned.

tions than before. After allowing for adjustments in the number of firms, however, the decrease in wages raises firms' profits, thus prompting the entry of new firms. This increase in the supply of firms has a positive effect on the matches of *all* workers (compared to the initial effect) and, overall, low-skilled natives end up in worse jobs and high-skilled natives find better jobs. Thus, in equilibrium, workers with sufficiently high skill levels will receive a higher wage than they received prior to the influx of immigrants.[7]

Furthermore, the existence of an exogenous ranking of workers allows us to apply standard Roy model techniques (the separation function) to model self-selection into migration. In fact, our paper is first to consider selection into migration (à la Roy) in a setting where low- and high-skilled workers are q-complements.[8] In existing models of selection, workers are either perfect substitutes, or—if they are complements—their skills are perfectly correlated across countries.[9] Note that both selection and worker complementarity are necessary to investigate the redistributive effects of migration in the sending and destination countries: Without complementarity all local workers in the host country earn lower wages due to immigration, whereas without selection the model is silent on how small changes in migration policies affect the skill distribution in each country.

The labor market response to migration constitutes the fundamental question in the contemporary discussion regarding the welfare implications of migration (Card, 2001, 2009; Borjas, 2003; Llull, 2018). Nevertheless, as noted by Aubry et al. (2016), the overall implications of migration emerge as an interplay across several economic channels. Accordingly, our model generates three additional general equilibrium effects. First, assuming that all individuals have al love of variety and consume horizontally differentiated baskets of goods, the change in the mass of firms in the market impacts real wages through the ideal price index, as in Krugman (1980). Papers by Iranzo and Peri (2009), di Giovanni et al. (2015),

[7]Not all workers with better jobs earn higher wages, however: Wages depend both on job quality and the firms' outside option, and the latter improves due to the influx of low-skilled migrants. Indeed, some native workers will end up employed by a more productive firm and yet earning a lower wage.

[8]Two worker types are q-complements, when an increase in the supply of one type leads to an increase in the marginal product of the other type and *vice versa*.

[9]The fact that workers of different skill levels are complements has been widely documented, both in the migration literature (Ottaviano and Peri, 2012; Borjas and Monras, 2017) and the broader labor literature (Katz and Murphy, 1992). Moreover, perfect correlation is a knife-edge assumption and seems particularly questionable given that the returns to various skill sets are likely to differ across countries (e.g., one could earn high wage in Mexico without proficiency in English, which is unlikely to happen in the United States). The empirical evidence on this correlation is scant, though Hanushek et al. (2015) show in Table A3 that there are significant differences in returns to numeracy, literacy, and problem-solving skills across countries.

Aubry et al. (2016), and Biavaschi et al. (2019) predict that the *market size effect* might act as a significant component of the total welfare effect of migration.[10]

Second, macroeconomic shocks induced by migration policies are propagated through *trade* linkages. Many papers in the economic geography literature investigate this concept in a regional context (Allen and Arkolakis, 2014; Redding and Rossi-Hansberg, 2017; Burstein et al., 2017). The theoretical literature on the international context is far more scarce (di Giovanni et al., 2015; Burzyński, 2018; Heiland and Kohler, 2018).[11] In our model, as in the pioneering work by Melitz (2003), trade acts as a substitute for migration, since we abstract from various externalities indicated in the empirical literature. Finally, we experiment with the *fiscal consequences* of immigration, emphasized by Auerbach and Oreopoulos (1999), Storesletten (2000), Rowthorn (2008), and Dustmann and Frattini (2014). Our approach essentially involves calibrating an instantaneous net government surplus per capita using the exact tax schedules and lump-sum and proportional benefits for the United States and Mexico in 2015. In our model, this out-of-equilibrium fiscal module shows that Mexican immigrants have an almost neutral net fiscal position on the United States economy, while in Mexico immigration induces additional fiscal costs.

The rest of the paper is organized as follows. Sections 2 and 3 discuss the theoretical model and its numerical calibration. In Section 4, we analyze the economic consequences of experimenting with the costs of Mexican migration to the United States. Section 5 concludes. Proofs of all statements are available in Appendix A. Appendix C provides details of the calibration procedure, Appendix D reports the results of several robustness checks and Appendix E compares numerically how wages change in response to a supply shock in the assignment and CES models.

2 The Model

The model consists of fours building blocks: (1) Mexican and (2) U.S. citizens (*workers*), and (3) Mexican- and (4) U.S.-based *firms*.[12] Mexican citizens are

[10]The market size effect estimated by these models is, by construction, homogeneous across all types of workers. The effect gives further insights into global efficiency gains from migration without altering redistribution across heterogeneous individuals. The same is true in our approach. The magnitude of the effect depends on the elasticity of substitution between the varieties of consumption goods.

[11]For a review of the empirical literature on the relationship between migration and trade, refer to Parsons et al. (2014).

[12]As indicated in footnote 1, the designation "U.S. citizens" includes also immigrants from other countries than Mexico—we treat migration from such countries as exogenous.

mobile and decide in which country to work by maximizing their real wages net of the migration costs. The U.S. citizens can only work in the United States. Firms first choose whether to enter the market, and later set the prices of the goods variety they produce and decide which worker to employ. The goods produced by each company are traded with the other country and the rest of the world (ROW).[13]

2.1 Workers and Firms

Workers There is a unit measure of Mexican citizens, each endowed with a vector of skills $(x_U, x_M) \in [0,1] \times [0,1]$. The skill x_U determines the worker's productivity in the U.S. labor market and the skill x_M determines her productivity in Mexico.[14] The joint distribution of X_U, X_M—conditional on the workers being Mexican citizens, denoted by C—has full support on $[0,1]^2$, and is twice continuously differentiable. Without loss of generality, we assume that the marginal distributions of X_U and X_M in the population of Mexican citizens are standard uniform.[15] This means that C is a copula (Sklar, 1959).

There is also a measure $R_U^W > 0$ of U.S. citizens. For simplicity, we assume that these individuals cannot move to Mexico, and thus are fully described by their U.S. skill $x_U \in [0,1]$. The distribution of X_U, conditional on the workers residence in the United States, is labeled as $F(\cdot)$. F is twice continuously differentiable and strictly increasing.[16]

[13]Only the Mexican and the U.S. economy are modeled explicitly. The prices of the goods traded by the rest of the world are given exogenously and their production is not modeled: ROW is only included in the model to allow for a trade imbalance between Mexico and the United States.

[14]It is best to think of the country specific skills x_U, x_M as indexes of basic skill sets (cognitive, manual, social, language). As the industrial structure of each country differs, firms require the basic skills in different proportions, giving rise to two sector-specific indexes x_U, x_M. In this sense x_U, x_M are akin to the tasks in Heckman and Sedlacek (1985). Section 2 in Gola (2018) provided the formal assumptions that are sufficient for such an aggregation of skills into two indexes to be without loss of generality.

[15]As shown in Section 2 of Gola (2018) the assumption of uniform marginal distributions is simply a normalization: We define the type of a Mexican citizen in terms of the quantile she occupies in the distribution of each skill. Consequently, any model written in terms of general joint distributions has an equivalent representation in terms of uniformly distributed skills. Here, we restrict attention to this *canonical* formulation of the model.

[16]Note that if F first-order stochastically dominates the standard uniform distribution, then we can meaningfully say that the population of U.S. citizens is more proficient than the population of Mexican citizens in the skill set used in the United States. To see why, suppose that $F(0.5) = 0.25$. Then half of the Mexican citizens and only a quarter of U.S. citizens are less skilled than a worker with skill $x_U = 0.5$.

Firms There exists an unlimited supply of *ex ante* identical companies. In order to enter the market in country $i \in \{U, M\}$, a firm needs to pay a fixed cost of $c_i^e > 0$ units of the composite consumption good (defined in Section 2.2). The measure of all firms that enter the market in country i is denoted by R_i^F. Once their type is known, firms decide whether to remain active and produce, or to exit the market. The set of active firms in country i will be denoted by $\mathscr{H}_i \subset [0,1]$. Active firms incur a fixed pruduction cost of $c_i^f > 0$ units of the consumption good and employ a single worker whom they pay the competitive wage for the skill she provides. If a country i-based firm of type h_i hires a worker with country i-specific skill x_i, they produce $f_i(x_i, h_i)$ units of a firm-specific variety of the consumption good. We assume that $\partial f_i/\partial x_i, \partial f_i/\partial h_i, \partial^2 f_i/\partial x_i \partial h_i > 0$, which ensures that the production function is strictly increasing and supermodular in workers' and firms' type.[17] We further assume that $f_i(0, h_i) < 0$, which means that the least-skilled workers will never be hired.

In line with Melitz (2003), each firm produces a unique variety of the consumption good, which implies that the measure of all varieties produced in a country is equal to the measure of all active firms in that market.[18] Therefore, the goods market is monopolistically competitive, with each firm freely setting the price of its variety to maximize their profits.

2.2 Goods Market

Welfare and Demand for Varieties People have homothetic preferences over the set of all available varieties (domestic and imported). Let ε be the elasticity of substitution between any two varieties. We can thus define a composite consumption good Q by taking the individual varieties as inputs, with

$$Q \equiv \left[R_U^F \int_{\mathscr{H}_U} q_U(h_U)^{\frac{\varepsilon-1}{\varepsilon}} \, \mathrm{d}h_U + R_M^F \int_{\mathscr{H}_M} q_M(h_M)^{\frac{\varepsilon-1}{\varepsilon}} \, \mathrm{d}h_M + q_W^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}, \qquad (1)$$

where $q_i(h_i)$ denotes the consumption of a variety produced in country $i \in \{U, M\}$ by firm h_i and q_W denotes the consumption of goods produced in the rest of the

[17] Supermodularity is a standard assumption in the matching literature. It implies that equilibrium matching between firms and workers will be positive and assortative in each country, which ensures that in equilibrium wage functions in each country will depend on migration decisions in a tractable way.

[18] Unlike in Melitz (2003), however, there is no fixed cost of export, so that all active firms export part of their production.

world.[19] The utility of employed workers depends positively on the consumption of Q, and negatively on migration costs. Specifically, the utility of a worker with skill x_i employed in country $i \in \{U, M\}$ and born in country $j \in \{U, M\}$ is

$$U_{ij}(x_i) \equiv \ln(Q_i(x_i) - \delta_{ij}) - \Delta_{ij}. \tag{2}$$

In Equation (2) Δ_{ij} represents the personal (utility) migration cost of moving from country j to country i and is measured in utils, whereas δ_{ij} represents the monetary cost of legal migration barriers and is measured in units of the numeriare (Q). Of course, $\delta_{ii} = \Delta_{ii} = 0$, so that remaining in one's country of birth is costless. Unemployed workers do not earn any money and hence cannot afford to buy Q.[20] They do, however, receive reservation utility from leisure/home production, with an unemployed country j citizen's utility equal to $\bar{U}_j$.

A worker supplying skill x_i in country i earns a wage $w_i(x_i)$ and maximizes her consumption of Q subject to the budget constraint

$$\sum_{k\in\{U,M\}} R_k^F \int_{\mathscr{H}_k} \tau_{ik} p_k(h_k) q_k(h_k)\, \mathrm{d}h_k + p_W \tau_{iW} q_W = w_i(x_i), \tag{3}$$

where τ_{ik} denotes the ice-berg trade cost of shipping a good from country k to country i, whereas $p_k(h_k)$ denotes the price of the variety produced by firm h_k in country k. The price of the ROW variety p_W is treated as exogenously given.[21] The standard solution of the individual utility maximization problem reveals that a worker with skill x_i who is employed in country i demands

$$q_i^j(x_i, h_j) = (\tau_{ij} p_j(h_j))^{-\varepsilon} \cdot P_i^{\varepsilon-1} \cdot w_i(x_i) \tag{4}$$

units of a variety produced by firm h_j in country j, where:

$$P_i = \left[\sum_{k\in\{U,M\}} R_k^F \int_{\mathscr{H}_k} (\tau_{ik} p_k(h_k))^{1-\varepsilon}\, \mathrm{d}h_k + (\tau_{iW} p_W)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}. \tag{5}$$

[19] Regarding the ROW economy, we assume that the total production is given exogenously, and demand depends on prices in the same way as in Mexico and the United States. A simple microfoundation would have ROW consisting of a single representative consumer, who produces a constant quantity q_W of a single variety, and has the same preferences as the Mexican and U.S. consumers.

[20] Our model cannot distinguish between unemployed and inactive individuals. In what follows, we treat both groups of people alike and differentiate them from the working (active) population.

[21] This is, of course, just a normalization, as only relative prices matter.

Finally, recall that in order to enter the market and produce, firms need to acquire a fixed amount of the composite good. Their cost-minimization problem is dual to the workers' utility maximization problem. Therefore, in order to purchase the amount of Q needed to pay the entry costs, every firm in country i will demand $(\tau_{ij} p_j(h_j))^{-\varepsilon} P_i^{\varepsilon} c_i^e$ of the variety produced by firm h_j in country j. To cover the production cost, active firms will also demand $(\tau_{ij} p_j(h_j))^{-\varepsilon} P_i^{\varepsilon} c_i^f$ of said variety.

Firms' Pricing Decisions Aggregate demand for variety h_i is given by

$$q_i^A(h_i) = p_i(h_i)^{-\varepsilon} \cdot \sum_{k \in \{U,M,W\}} Y_k \left(\tau_{ki}/P_k\right)^{1-\varepsilon}, \tag{6}$$

where Y_k is the total expenditure in country k, and is equal to the sum of all of consumers' and all of firms' spending on the composite good (see Equation (24) in Section 2.3).[22] This allows us to write the inverse demand function:

$$p_i(h_i) = q_i^A(h_i)^{-1/\varepsilon} \left(\sum_{k \in \{U,M,W\}} Y_k \left(\tau_{ki}/P_k\right)^{1-\varepsilon} \right)^{1/\varepsilon}. \tag{7}$$

In equilibrium, the demand for variety h_i, $q_i^A(h_i)$, must be equal to its supply, $f_i(x_i, h_i)$, implying that the revenue produced by a worker-firm match (x_i, h_i) is equal to:

$$\begin{aligned} r_i(x_i, h_i) \equiv & p_i(h_i) f_i(x_i, h_i) - P_i c_i^f \\ = & f_i(x_i, h_i)^{\frac{\varepsilon-1}{\varepsilon}} \left(\sum_{k \in \{U,M,W\}} Y_k \left(\tau_{ki}/P_k\right)^{1-\varepsilon} \right)^{1/\varepsilon} - P_i c_i^f. \end{aligned} \tag{8}$$

The price levels set by producers are equal to constant markups over marginal cost, as in Melitz (2003).[23]

[22] Note that the demand for each good is multiplied by the iceberg trade cost, because in order to sell q units of good to a consumer in country i, the firm needs to produce $\tau_{ki} q$ units of the good.

[23] It follows from Equation (17) in Section 2.3 and Equation (8) that: $\frac{\partial}{\partial x_i} w(\mu_i(h_x)) = \frac{\partial}{\partial x_i} \pi_i(\mu_i(h_i), h_i) = \frac{\partial}{\partial x_i} f_i(x_i(h_i), h_i) p(h_i) \frac{\varepsilon-1}{\varepsilon}$, so: $p(h_i) = \frac{\varepsilon-1}{\varepsilon} \frac{w'(x_i(h_x))}{\frac{\partial}{\partial x_i} f_i(x_i(h_i), h_i)} = \frac{\varepsilon-1}{\varepsilon} \cdot MC(h_i)$.

2.3 Labor Market

Supply of Skills All workers decide whether to work or remain unemployed; additionally, Mexican citizens choose their country of residence. In reaching their labor supply decisions, workers maximize their utility and take the nominal wage functions $w_i : [0,1] \to \mathbb{R}$ and the price indexes P_i as given. It will be convenient to define the real wage in terms of the units of Q, $\bar{w}_i(x_i) \equiv w_i(x_i)/P_i$, as follows from Equations (1), (4) and (5). Therefore, a Mexican citizen with skill vector (x_U, x_M) migrates to the United States if and only if

$$e^{-\Delta_{UM}}(\bar{w}_U(x_U) - \delta_{UM}) \geq \max\{\bar{w}_M(x_M), \bar{w}_M^c\},$$

where $\bar{w}_i^c = \exp(\bar{U}_i)$. This inequality defines the no-arbitrage condition of Mexican citizens' sorting into the two labor markets. If it is not satisfied, the worker becomes employed in Mexico if $\bar{w}_M(x_M) \geq \bar{w}_M^c$ and remains unemployed otherwise.

The cumulative supply of skill x_i provided by country j citizens in country i—$S_{ij}(x_i)$—is defined as the measure of country j citizens employed in country i whose country-i specific skill is higher than x_i. It follows from the considerations above that:

$$S_{UM}(x_U) \equiv \Pr\left[X_U \geq x_U, e^{-\Delta_{UM}}(\bar{w}_U(X_U) - \delta_{UM}) \geq \max(\bar{w}_M(X_M), \bar{w}_M^c)\right], \quad (9)$$

$$S_{MM}(x_M) \equiv \Pr\left[X_M \geq x_M, \bar{w}_M(X_M) \geq \max\{e^{-\Delta_{UM}}(\bar{w}_U(X_U) - \delta_{UM}), \bar{w}_M^c\}\right]. \quad (10)$$

Similarly U.S. citizens choose to be employed only if their real wage is greater than their reservation wage, implying that:

$$S_{UU}(x_U) = R_U^W \Pr\left[X_U \geq x_U,\ w_U(X_U) \geq w_U^c\right]. \quad (11)$$

Since we assume that the total cost of moving from the United States to Mexico is prohibitive, only Mexican citizens are present in Mexico, and thus $S_{MU}(x_M) = 0$ for all x_M. Note that $S_{ij}(0)$ gives the measure of all country j citizens employed in country i.

Finally, *the cumulative supply of skill x in country i*—$S_i(x_i)$—is defined as the measure of workers of either origin living in country i with a skill level greater than x_i, so that

$$S_i(x_i) = S_{iU}(x_i) + S_{iM}(x_i). \quad (12)$$

Demand for Skills Fixing firms' entry (i.e., for a given R_i^F), the demand for skills in each country is determined by the firms' hiring decisions, which in turn are driven by profit maximization, with firms taking the wage function as given. Denote the operating profit of firm h_i by $\pi_i(h_i)$ and the skill of the worker it hires by $\mu_i(h_i)$. The operating profit is equal to the revenue net of the wage paid to the worker, with

$$\pi_i(h_i) = \max_{x_i \in [0,1]} \; r_i(x_i, h_i) - w_i(x_i), \tag{13}$$

$$\mu_i(h_i) \in \arg\max_{x_i \in [0,1]} \; r_i(x_i, h_i) - w_i(x_i). \tag{14}$$

The *demand for skill* of level x_i in country i, $D_i(x_i)$, is equal to the measure of country i firms that hire workers with country i-specific skill of at least x_i, for a given wage function w_i and firm measure R_i^F:

$$D_i(x_i) \equiv R_i^F \Pr\left[\mu_i(H_i) \geq x_i, \pi_i(h_i) \geq 0\right].^{24} \tag{15}$$

Firms' Entry The expected operating profit in country i is

$$\pi_i^E = \int_0^1 \max\{\pi_i(h_i), 0\} \, \mathrm{d}h_i.$$

Firms enter the country that maximizes their expected profits (i.e. expected operating profit net of entry cost). The monetary cost of entry is equal to $c_i^e P_i$: Therefore, if entry is positive in country i ($R_i^F > 0$), then the expected profit must be equal to the monetary cost of entry: $\pi_i^E = P_i c_i^e$.

Partial Labor Market Equilibrium Taking the revenue functions and price indexes as given, we can now define the partial labor market equilibrium.

Definition 1. For a given pair of revenue functions r_U, r_M and price indexes P_U, P_M the *partial labor market equilibrium* is characterized by

1. the supply of skills $S_i : [0,1] \to [0,1]$ in each country, which is determined by workers' sorting decisions and is given by Equations (9)—(12);

2. the demand for skills $D_i : [0,1] \to [0,1]$ in each country, which is determined by firms' profit maximization and is given by Equation (15);

[24] Because revenue increases strictly in firm's type, so will profit—thus firms with $\pi_i(h_i) = 0$ are of measure zero.

3. firms' measures R_i^F, consistent with the zero-expected-profits-condition, such that $\pi_i^E = P_i c_i^e$ if $R_i^F > 0$ and $\pi_i^E \leq P_i c_i^e$ otherwise;

4. wages $w_i : [0,1] \to \mathbb{R}$ in each country, which are set to clear the markets: $S_i(x_i) = D_i(x_i)$ for $i \in \{U, M\}$ and all $x_i \in [0,1]$.

We will now show the existence and uniqueness of the labor market equilibrium. Define the set $\mathbb{E}$ of *labor equilibrium allocations* as the set of *allocations*—that is, quintuples of supply functions and firms' measures $(S_{UU}, S_{UM}, S_{MM}, R_U^F, R_M^F)$—for which there exists a pair of wage functions w_U, w_M that—given R_U^F, R_M^F—induces both the supply and the demand functions to be equal to S_U, S_M and satisfies the zero-expected-profit-condition. We will first characterize a superset of $\mathbb{E}$, and then show that among all allocations in that superset, any partial equilibrium allocation maximizes uniquely the sum of revenues (when appropriately weighted). This immediately proves that $\mathbb{E}$ contains at most one element (uniqueness). That it is non-empty (existence) will follow from the fact that the total net income function must attain a maximum.

We will start by deriving the wage functions that equate supply and demand within each country, for a given allocation. Let the *critical skill of country j citizens in country i* be the country-specific skill of the least skilled country j citizen employed in country i:

$$x_{ij}^c = \sup\{x_i \in [0,1] : S_{ij}(x_i) = S_{ij}(0)\}. \tag{16}$$

The *critical skill in country i* is then defined as the skill level of the least-skilled worker employed in country i, with $x_i^c = \min\{x_{iU}^c, x_{iM}^c\}$. The inverse of the hiring function μ_i will be called the *matching function* and denoted by m_i: A worker with skill x_i matches with the firm $m_i(x_i)$, and they jointly generate $r_i(x_i, m_i(x_i))$. It is well-known that with supermodular revenue functions, matching must be positive and assortative (PAM); that is, the matching function must be strictly increasing.[25] This condition and market clearing immediately gives

$$m_i^*(x_i) = 1 - S_i(x_i)/R_i^F \text{ for } x_i \geq x_i^c.\text{[26]}$$

[25] This was first shown by Sattinger (1979). For a proof that uses the same vocabulary as employed here, see proof of Proposition 1 in Gola (2018).

[26] Because of market clearing and the fact that the revenue function—and hence profits—strictly increase in firm type, it follows that under equilibrium wages $h_i < (>) m_i(x_i^c)$ implies that $\pi_i(h_i) < (>) 0$. As μ_i is strictly increasing, market clearing allows us to write $S_i(x_i) = D_i(x_i) = R_i^F(1 - h_i^*(x_i))$.

The first-order condition of the firm's hiring decision then implies that

$$\partial w_i(\mu_i(h_i))/\partial x_i = \partial r_i\Big(\mu_i(h_i), 1 - S_i(\mu_i(h_i))/R_i^F\Big)/\partial x_i. \tag{17}$$

The difference in the wages paid to workers of marginally different skill is equal to the difference in the revenue they produce. Integrating from x_i^c to x_i gives

$$w_i(x_i) = \int_{x_i^c}^{x_i} \partial r_i(r, 1 - S_i(r)/R_i^F)/\partial x_i \, \mathrm{d}r + w_i(x_i^c) \qquad \text{for } x_i \geq x_i^c.^{27} \tag{18}$$

The wage paid to the worker with critical skill level x_i^c depends on the reservation wages of the workers employed in that country. In particular, if $x_i^c < 1$ then $w_M^c(x_M^c) = w_M^c$, whereas $w_U(x_U^c) = \min\{w_U^c, e^{\Delta_{UM}} P_U\,(w_M^c/P_M + \delta_{UM})\}$, where $w_i^c \equiv P_i^c \bar{w}_i^c$.[28] Finally, because $x_{MM}^c = x_M^c$ we have that if $x_{MM}^c < 1$, then $w_M(x_{MM}^c) = w_M^c$. This further implies from Equation (9) that if $x_{UM}^c < 1$, then $w_U(x_{UM}^c) = P_U\left(e^{\Delta_{UM}} w_M^c/P_M + \delta_{UM}\right)$.

Now we will derive a set of conditions that equilibrium supply functions must satisfy. First, note that for $x_i \geq x_i^c$ wages are strictly increasing in each country. Therefore, for any Mexican citizen with Mexican-specific skill $x_M > x_{MM}^c$, there will exist a cut-off value $\psi(x_M)$ of the U.S.-specific skill such that she will strictly prefer to remain in Mexico if $x_U < \psi(x_M)$ and strictly prefer to migrate to the United States if $x_U > \psi(x_M)$. Therefore, the migration decisions of Mexican citizens can easily be expressed by the means of the critical skills x_{MM}^c, x_{UM}^c and the *separation function* $\psi : [x_{MM}^c, 1] \to [x_{UM}^c, 1]$, which takes the Mexican-specific skill as an argument and returns the corresponding cut-off value of the U.S.-specific skill. For $x_M \geq x_{MM}^c$ the separation function depends on the sectoral wage functions as follows:

$$\psi(x_M) = \max\{x_U \in [x_{UM}^c, 1] : e^{-\Delta_{UM}}\,(w_U(x_U)/P_U - \delta_{UM}) \leq w_M(x_M)/P_M\}. \tag{19}$$

Note that for $x_M \leq x_M^s$, where $x_M^s = \inf\{x_M \geq x_M^c : \psi(x_M) = 1\}$, this implies

[27] For $x_i < x_i^c$ market clearing implies: $w_i(x_i) \geq w_i(x_i^c) + r_i\Big(x_i, 1 - S_i(x_i^c)/R_i^F\Big) - r_i\Big(x_i^c, 1 - S_i(x_i^c)/R_i^F\Big)$.

[28] Suppose that $i = M$ (the argument for $i = U$ is analogous). It follows immediately from Equations (10) and (16) that if $x_i^c < 1$ then $w_M^c(x_M^c) \geq P_M w_M^c$. As $f_M(0, h_M) < 0$, it follows that in equilibrium workers with skill x_M close to 0 cannot be employed in Mexico; thus $x_M^c > 0$. It follows that $w_M^c(x_M^c) \leq P_M w_M^c$—the continuity of the revenue function implies that otherwise workers with skill slightly lower than x_M^c would strictly prefer to be employed in Mexico than remain unemployed, which contradicts the definition of the critical skill.

that

$$e^{-\Delta_{UM}}\left(w_U(\psi(x_M))/P_U - \delta_{UM}\right) = w_M(x_M)/P_M. \tag{20}$$

As the critical skills and the separation function are sufficient to characterize the migration decisions of all Mexican citizens, they are also sufficient to derive the supply of skill from Mexican citizens in each country.

Lemma 1. Given the critical skills of Mexican citizens x^c_{UM}, x^c_{MM} and the separation function ψ, the supply of skills in both countries is respectively

$$\begin{aligned} S_{UM}(x_U) &= \begin{cases} \int_{x_U}^1 \frac{\partial}{\partial x_U} C(r, \phi(r))\, \mathrm{d}r, & x_U \geq x^c_{UM} \\ S_{UM}(x^c_{UM}) & x_U < x^c_{UM}, \end{cases} \\ S_{MM}(x_M) &= \begin{cases} \int_{x_M}^1 \frac{\partial}{\partial x_M} C(\psi(r), r)\, \mathrm{d}r, & x_M \geq x^c_{MM} \\ S_{MM}(x^c_{MM}) & x_M < x^c_{MM}, \end{cases} \end{aligned} \tag{21}$$

where $\phi : [x^c_{UM}, 1] \to [x^c_{MM}, 1]$ depends on ψ as follows:

$$\phi(x_U) = \sup\{x_M \in [x^c_{MM}, 1] : \psi(x_M) < x_U\}.$$[29]

It follows that for any $x_M \geq x^c_{MM}$:

$$S_{UM}(\psi(x_M)) + S_{MM}(x_M) = 1 - C(\psi(x_M), x_M). \tag{22}$$

Finally, the same reasoning used above yields that if $w_U(1) \geq w^c_U$, then $x^c_{UU} = w_U^{-1}(w^c_U)$. It follows from the fact that the wage functions are increasing that:

$$S_{UU}(x_U) = \begin{cases} R^W_U(1 - F(x_U)), & x_U \geq x^c_{UU} \\ R^W_U(1 - F(x^c_{UU})) & x_U < x^c_{UU}. \end{cases} \tag{23}$$

Equations (21)—(23), together with the requirement that labor markets must clear, put strong restrictions on the partial equilibrium allocations. In particular, the restrictions on allocations imply that for any $A = (S_{UU}, S_{UM}, S_{MM}, R^F_U, R^F_M) \in \mathbb{E}$, it must be the case that (1) $S_{ij} : [0,1] \to [0,1]$ is non-decreasing, absolutely continuous, and semi-differentiable on the interior, with $S_{ij}(1) = 0$; (2) the function $\psi : [x^c_{MM}, 1] \to [x^c_{UM}, 1]$ that satisfies $\partial C(\psi(x_M), x_M)/\partial x_M = -\partial_+ S_{MM}(x_M)/\partial x_M$ is well-defined, with a continuous, strictly positive derivative at $x_M \in (x^c_M, x^s_M)$;

[29]Note that for $x_M \in [x^c_M, x^s_M]$ we have that $\phi(\psi(x_M)) = x_M$, that is ϕ is the inverse of ψ on the image of $[x^c_M, x^s_M]$.

(3) S_{MU}, S_{MM} satisfy Equation (22) and $-\partial_+ S_{Mi}(x_i)/\partial x_M \in (0,1)$; (4) $S_{UU}(x_U)$ satisfies (23), and (4) $S_i(0) \leq R_i^F$, where $S_i(x_i) = S_{ii}(x_i) + S_{ij}(x_i)$ and $S_{MU}(x_M) = 0$.[30] The allocations meeting conditions (1)–(5) will be called *feasible* and the set of all feasible allocations will be denoted by $\mathbb{A}$. Clearly, $\mathbb{E} \subset \mathbb{A}$.

The total revenue produced in country i (expressed in real terms) under a feasible allocation $A \in \mathbb{A}$ is equal to the sum of real revenues produced by all workers in that country:

$$T_i(A) \equiv P_i^{-1} \int_1^0 r_i\Big(x_i, 1 - S_i(x_i)/R_i^F\Big)\, \mathrm{d}S_i(x_i) \qquad \text{if } R_i^F > 0.^{31}$$

In contrast to Gola (2018), the partial labor market equilibrium does not maximize the sum of the total revenues produced in the two countries (net of entry costs). This is due to the costs of migration: As is well-known since at least Costrell and Loury (2004), the wage paid in country i to worker of skill x_i is equal to the marginal revenue produced in country i due to the addition of worker i. Thus, in order to maximize the sum of revenues, Mexican workers should migrate if $\bar{w}_U(x_U) \geq \bar{w}_M(x_M)$, whereas the actual equilibrium condition subtracts the migration costs from the United States wage.[32] This leads to the educated guess that the partial labor market equilibrium might maximize the sum of (net) revenues if they are appropriately weighted:

$$\begin{aligned} V(A) \equiv e^{-\Delta_{UM}} \left[T_U(A) + \bar{w}_U^c F(x_{UU}^c) R_U^W - R_U^F c_U^e\right] \\ + T_M(A) + \bar{w}_M^c C(x_{UM}^c, x_{MM}^c) - R_M^F c_M^e - \delta_{UM} S_{UM}(0). \end{aligned}$$

Proposition 1. A worker and firm allocation $A^* \in \mathbb{A}$ can hold in the partial labor

[30]Condition (1) follows from Equations (20), (21) , (23), the and market clearing. Condition (2) follows from differentiating Equation (20) and noting that wages must be non-decreasing in equilibrium; (3) follows from (21) and (22); and (4), obviously, from Equation (23).

[31]If $R_i^F = 0$ then, trivially, $T_i(A) = 0$.

[32]For slightly more complicated reason, the equilibrium does not maximize the total revenue net of both entry and migration costs either. In short, assigning a marginal Mexican worker to the United States creates additional revenue ($\bar{w}_U(x_U)$), an additional pecuniary migration cost (δ_{UM}), and the additional utility cost (worth $(1 - e^{-\Delta_{UM}})(\bar{w}_U(x_U) - \delta_{UM})$ units of the consumption good). These are the only changes that are reflected in the no-arbitrage condition (Equation (20)); on top of that, however, the addition of a marginal Mexican worker changes the wages of all other Mexican workers in the United States, and thus also the monetary equivalent of the utility migration cost for all of them. This additional change in costs is not included in the no-arbitrage condition, creating a wedge between maximizing revenue net of all the costs and the equilibrium.

market equilibrium if and only if it maximizes the weighted sum of (net) revenues:

$$A^* \in \mathbb{E} \Leftrightarrow V(A^*) - V(A) > 0 \text{ for all } A \in \mathbb{A} \setminus \{A^*\}.$$

The "if" condition is proved in a manner similar to the standard proof of the First Welfare Theorem (see, e.g. Mas-Colell et al., 1995). To prove the "only if" we use the calculus of variations to derive first-order conditions, and show that they imply that all the equilibrium conditions are met. It is worth stressing that, in contrast to Proposition 10 in Gola (2018), our Proposition 1 does not imply that the partial labor market equilibrium maximizes total welfare (for given prices), for the reasons outlined in footnote 32. The (non-transformed) total welfare function would weight the welfare of Mexican stayers and United States citizens in the same way, which is not the case in Equation (24).[33]

Theorem 1. The equilibrium defined in Definition 1 exists and is unique.

Given Proposition 1, Theorem 1 follows by a straightforward application of Weirestrass' Theorem.

Finally, note that in a partial equilibrium of the labor market the total expenditure on the composite good Q by firms in country i is equal to:

$$P_i(S_i(0)c_i^f + R_i^F c_i^e) = \int_1^{x_i^c} P_i c_i^f + \pi_i(1 - S_i(x_i)/R_i^F)\, \mathrm{d}S_i(x_i).$$

By the definition of operating profit it follows that the total expenditure on Q in country i is equal to

$$\begin{aligned} Y_i &= P_i(S_i(0)c_i^f + R_i^F c_i^e) + \int_1^{x_i^c} w_i(x_i)\, \mathrm{d}S_i(x_i) \\ &= \int_1^{x_i^c} r(x_i, 1 - S_i(x_i)/R_i^F) + P_i c_i^f\, \mathrm{d}S_i(x_i). \end{aligned} \tag{24}$$

The expenditure in ROW is equal to its income, and exogenously set to Y_W.

2.4 General Equilibrium

The economy is in general equilibrium if the goods market is in equilibrium given the total expenditures resulting from the labor market, and the labor market is

[33] In broad strokes, this is reminiscent of the results in Dupuy et al. (2017), where imperfect transferability of utility also rules out the efficiency of the equilibrium, but the equilibrium nevertheless corresponds to the optimum of an assignment problem in which the welfare of all agents is appropriately weighted.

in equilibrium given the revenue functions and price indexes resulting from the goods market.

The following condition, which must hold in equilibrium for any $i \in \{U, M, W\}$ by Equations (5), (6), (8), and (24), provides the link between the goods and labor markets:

$$P_i = \left[\frac{(\tau_{iU})^{1-\varepsilon} Y_U}{\sum_k Y_k \tau_{kU}^{1-\varepsilon} P_k^{\varepsilon-1}} + \frac{(\tau_{iM})^{1-\varepsilon} Y_M}{\sum_k Y_k \tau_{kM}^{1-\varepsilon} P_k^{\varepsilon-1}} + (\tau_{iW} p_W)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}. \tag{25}$$

Definition 2. The economy is in general equilibrium if the revenue functions (r_U, r_M), price indexes (P_U, P_M, P_W), and total expenditures (Y_U, Y_M) are such that:

(i) the labor market is in partial equilibrium given the revenue functions and price indexes (Definition 1);

(ii) price indexes are consistent with individual preferences, consumers' utility maximization problem and goods market clearing, given the total expenditure (Equation (25));

(iii) the revenue functions are consistent with market clearing conditions in the goods market, given the total expenditure (Equation (8)).[34]

The equilibrium exists, but is not necessarily unique.

Theorem 2. The general equilibrium exists. It is unique if trade is costless ($\tau_{ij} = 1$ for all $i, j \in \{U, M, W\}$).

If trade is costless, then $P_U = P_M = P_W$ irrespective of Y_U, Y_M and uniqueness follows directly from the uniqueness of the labor market equilibrium (Theorem 1).[35] If trade costs are greater than 1, the equilibrium might not be unique. As pointed out by Krugman (1980), if trade is costly, then *ceteris paribus* real wages are higher in the larger country than in the smaller country, as the larger country has cheaper access to a greater range of varieties. This creates a force for

[34] In addition to thee requirements, Equation (6) must hold for the ROW as well; that is, the market for the foreign variety must clear. However, Walras's Law ensures that if all other markets clear, the market for the foreign variety does too, and thus the above definition of equilibrium is sufficient.

[35] To see this, note that we can divide revenue in each country by $(Y_U + Y_M + Y_W)/P_W$ without loss of generality.

any general equilibrium with very high emigration from Mexico to become self-enforcing. In particular, as trade costs approach infinity, a complete out-migration of all employed Mexican citizens must constitute an equilibrium.[36]

It is worth pointing out that, while unfortunate, the multiplicity of equilibria does not pose a big problem for what we set out to accomplish in this paper. There are two reasons for this. Firstly, the equilibrium of our calibrated model turns out to be unique, which is likely caused by the large volume of trade between Mexico and the United States, as well as beetwen Mexico and the ROW. Secondly, even if there were multiple equilibria, our calibration procedure would select the one most closely resembling the data.

3 Calibration

This section contains a detailed discussion of the numerical calibration of the model. After specifying and motivating the chosen forms of key functions (Section 3.1), we provide a description of the datasets used to calibrate the model (Section 3.2) and comment on the results of the benchmark calibration (Section 3.3).

3.1 Functional Forms

Copula The general version of the model presented in Section 2 allows for any relationship between the marginal distributions of skills among Mexicans. In our quantitative exercise, we impose positive dependence between the skills used in each country.[37] However, the strength of this relationship might not be identical across quantiles. For our baseline calibration, we select the Clayton copula, which is a member of Archimedean copula family.[38] Several reasons motivate our choice.

[36] With infinite trade costs, every country can consume only the varieties created at home. But if ever working Mexican is employed in the United States, then any hypothetical match in Mexico would create only a single variety, which is of measure zero. Therefore, the deviating worker's consumption of Q would be zero, implying that either remaining unemployed in Mexico or migrating to the United States must be the better option. Therefore, no Mexican citizen would like to deviate from this equilibrium. Put differently, the combination of trade costs and love of variety introduce an externality, as the real revenue produced by any match depends not only on that match, but also on all the other matches in the entire economy.

[37] Assuming a negative correlation between skill levels would imply that Mexicans who are highly ranked in the Mexican skill distribution would be characterized by a relatively low level in U.S. skill rankings, and vice versa. We consider this case as being less probable than the positive relationship, since the process of accumulation of skills in each country requires a similar set of individual traits (ability to learn, adapt, and develop manual capabilities or educational proficiency in a particular environment).

[38] In Appendix B we graphically compare the Clayton, Gaussian, and Gumbel copulas.

First, this copula is described by a simple formula:

$$C(x_U, x_M) = \left(x_U^{-\theta} + x_M^{-\theta} - 1\right)^{-1/\theta}, \quad \theta > 0, \tag{26}$$

which imposes no computational difficulties in calculating its cumulative distribution function (CDF) and marginal probability density functions (PDFs). Second, this copula allows for the co-existence of a very strong dependence in the bottom-left corner of the two-dimensional distribution (Mexicans ranked low in their domestic skill will almost surely be ranked low in the foreign skill level) with a weaker dependence in the top-right corner (Mexicans ranked high in their domestic skills have a great chance of being ranked high in the other country's skill distribution, but this pattern is not a necessity).[39] Finally, the Clayton copula is characterized by a single parameter, θ, that defines the strength of the relationship between the two marginal distributions. In particular, this parameter can be easily mapped onto a well-known rank correlation measure, Kendall's τ. One can show that $\tau = \theta/(\theta + 2)$.

Production Functions The model from Section 2 introduces country-specific production functions that map the rankings of workers and firms into the units of goods produced by a specific match. In Section 2 these production functions are only required to be strictly increasing in both arguments and weakly supermodular. In the numerical exercises we will limit attention to multiplicatively separable functions of the following form:

$$f_i(x_i, h_i) = \exp\left(\Phi_i^{-1}(x_i; k_i, s_i)\right) \cdot \Psi_i^{-1}(h_i; \gamma_i), \tag{27}$$

where: Φ_i denotes the CDF of normal distribution with location k_i and dispersion s_i, and $\Psi_i^{-1}(h_i; \gamma_i) = (1 - h_i)^{-1/\gamma_i}$ denotes the inverse of the CDF of Pareto distri-

[39] Think of a Mexican medical doctor who is highly skilled and works in Mexico. She is probably ranked very high in the Mexican native population. Had she chosen to migrate to the United States, she might have encountered significant difficulties in having her diploma recognized. Thus, after migrating, she might have taken a job that does not fully capitalize on her high skill level and could have worked as a guardian or a nurse, which significantly reduces her in the ranking of Mexican immigrants in the United States. Conversely, imagine a construction worker in Mexico who is ranked below the median. If he is located in the left tail of the skill distribution (is thus paid very low wages in Mexico), he has presumably not acquired enough skills to achieve a better ranking after migrating to the United States (and consequently accepts a low wage). If he is closer to the median in Mexico (which would reveal his additional, specific skills, i.e., he might be a crane operator), it is more probable that he would overtake a significant number of his peers in the U.S. skill rankings because the specific skill that he possesses is easily transferable across the border. In other words, he does not need U.S.-specific skills to make a full use of his abilities, unlike the medical doctor in our previous example.

bution with scale parameter 1 and shape parameter γ_i.[40] Under these assumptions our model is equivalent to a model in which skills are log-normally distributed, firm productivity is Pareto distributed, and the number of units produced by any match is equal to the product of the worker's skill and the firm's productivity.

In the case of skills, it is well known that wage distributions generally resemble a log-normal distribution, but are more skewed. Given that the skewness of the wage distribution will be produced in our model by selection and positive assortative matching, using a log-normal appears to be the most natural choice for the distribution of skill.[41] In the case of productivity, our choice is motivated by the extensive body of literature which advocates that various characteristics of firms follow a Pareto distribution (see di Giovanni et al., 2011; di Giovanni and Levchenko, 2013).[42]

3.2 Data Description

The calibration represents a static state of the Mexican and U.S. economies in 2015. We differentiate between two types of empirical moments that we use to calibrate the model. First, we collect a set of observable variables that exogenously determine some model parameters. We call this set of observables group (A). Second, we construct empirical moments that directly correspond to model variables and enable us to identify the majority of model parameters. These values are gathered into group (B).

Below we briefly describe our data sources. Table 1 summarizes the values of all moments collected from several sources of data. Based on Labor Force

[40] Both of these distributions are assumed to be truncated at three-sigmas, in accordance with the data truncation, which is discussed in the next section. Formally, our existence proofs require Lipschitz continuity, which means that there needs to be a limit on the support of the distributions.

[41] As shown first by Heckman (1979), the empirical distribution of wages deviates from a log-normal distribution largely due to workers' selection into the labor market. A micro-founded derivation of a skew-normal distribution that naturally follows from workers' sorting decisions, as in Roy (1951) and Borjas (1987), can be found in Azzalini and Valle (1996). Positive and assortative matching boosts wage inequality and significantly increases the third moment of the wage distribution in relation to the marginal distribution of skills (see, e.g. Sattinger, 1975). Finally, the log-normal distribution of marginal skills can be theoretically justified, as the product of many independently distributed random variables is distributed log-normally. Thus, if one believes that the aggregate workers' skill is the product of many independent characteristics, it follows that it is log-normally distributed.

[42] Note that setting parameter $\gamma_i = 0$ imposes both no matching and a degenerated distribution of firms' productivity, which brings our framework back to the general selection model by Roy (1951) with truncated, exogenously given log-normal distributions of wages. Shutting down matching precludes the analysis of firms' endogenous profit distribution and their entry and exit decisions.

Surveys (LFS), the Database in Immigrants in OECD and non-OECD Countries (DIOC) reports that there are 51.5 million working-age Mexicans (employed and unemployed natives and migrants), 42.4 million of whom are employed in Mexico, and slightly above 7 million are (legally) employed in the United States (see the second section of Table 1). In the United States, 149 million people are working-age residents (natives and non-Mexican immigrants), 136 million of whom are actually employed. By normalizing Mexico's population to unity, we set the size of the total U.S. population to $R_U^W = 2.892$, while the working-age populations in both countries are equal to: $S_M(0) = 0.832$ and $S_U(0) = 2.643 + 0.137 = 2.78$, respectively.

We divide the gross domestic product (GDP) into three broad groups: wage shares, corporate profit shares, and the consumption of fixed capital (capital investment shares). The first element must match the model's share of total revenue transferred back to all workers in the form of remunerations. The second element pins down the share of total revenue that accrues to firms, whereas the latter identifies the investment costs that cover the depreciation of fixed capital (which is not explicitly modeled in our approach). Capital investments represent non-labor costs incurred by firms in the production process. Using the FRED classification of capital expenditures into "equipment," "intellectual property," and "structures," we categorize non-labor costs into their variable and fixed part. The first two groups can be though of as being proportional to the size and output of a firm, thus we include them in the variable part. Investments in "structures" are assumed to relate to the fixed part of production costs. Thus, 65 percent of capital expenditures are variable in the United States, while in Mexico this share amounts to 35 percent.

With the above data on the structure of GDP in Mexico and the United States, we can compute the fixed cost of production, c_i^f, which serves as an important component of our identification strategy. Furthermore, country-specific measures of potential firms are set equal to the amount of employed and unemployed individuals plus the number of active job vacancies (available for the United States from the Bureau of Labor Statistics, for Mexico, we generate a proportional number of vacancies). This procedure yields $R_U^F = 3.1$ and $R_M^F = 0.993$.

We collect some moments describing the labor markets in both countries. The unemployment rate in the Mexican labor market equals $u_M = 4.0$ percent, whereas in the United States it is equal to $u_U = 8.62$ percent, following the DIOC. Then, we compute the distributions of wages for the three groups of individuals being analyzed: residents in the United States, natives in Mexico, and Mexican migrants

Table 1: Empirical moments for model calibration

Object Name	Symbol	Value US	Value MEX	Source	Group
Demographics					
Total Population	R_i^W	2.892	1.000	DIOC	(A)
Working Population	$S_{ii}(0)$	2.643	0.832	"	(A)
Structure of GDP					
Wage Share	w_i^{share}	0.56	0.52	OECD & FRED	(B)
Profit Share	π_i^{share}	0.27	0.305	"	(A)
Capital Investment Share:	ci_i^{share}	0.17	0.175	"	(A)
Equipment	$ci_i^{share,E}$	0.39	0.24	FRED & imp.	(A)
Structures	$ci_i^{share,S}$	0.35	0.65	"	(A)
Intellectual Property	$ci_i^{share,IP}$	0.26	0.11	"	(A)
Firms					
Fixed Production Costs	c_i^f	$4,846$	$1,513$	imp.	(B)
Potential Firms	R_i^F	3.100	0.993	DIOC & BoLS	(A)
Labor Market					
Unemployment Rate	u_i	8.62%	4.00%	DIOC & LFS	(A)
Minimal Wage	w_i^c	\$4,133	\$605	IPUMS	(B)
Maximal Wage	w_i^m	\$183,380	\$45,834	"	(A, B)

Object Name	Symbol	Value	Source	Group
Migration, Trade, and the Rest of World				
Migration from MEX to US	$S_{UM}(0)$	0.137	DIOC	(B)
Goods' Elasticity of Substitution	ε	7	literature	(A)
Price Index, ROW	P_W	$0.28 \cdot P_U$	WITS & WDI	(A)
Gross Domestic Product, ROW	Y_W	$4 \cdot Y_U$	WITS	(A)
Bilateral Trade Matrix	Y_{ij}	-	TiVA	(A)
Wage Distributions				
Wage Distribution U.S. Residents	$w_{UU}(\cdot)$	-	IPUMS	-
Wage Distribution MEX Immigrants	$w_{UM}(\cdot)$	-	"	-
Wage Distribution MEX Residents	$w_{MM}(\cdot)$	-	IPUMSint	-
Copula				
Conditional Probability of Migration	$P(\cdot)$	-	MMP	(B)

Notes: imp.= imputation; DIOC = Database on Immigrants in OECD and non-OECD Countries by the OECD; BoLS = Bureau of Labor Statistics by U.S. Dep. of Labor; U.S. census = U.S. census Bureau; FRED = The Federal Reserve Bank of St. Louis; MES = Mexico Enterprise Survey by the World Bank; LFS = Labor Force Survey by Eurostat; IPUMS and IPUMSint by Institute for Social Research and Data Innovation; WITS = World Integrated Trade Solutions by the World Bank; WDI = World Development Indicators by the World Bank; TiVA = Trade in Value Added by the OECD; MMP = Mexican Migration Project by Princeton University.

to the United States. We calibrate the model on publicly available intercensal data from Mexico and the United States. We use the 2015 1 percent American Community Survey (ACS) data provided by IPUMS, Ruggles et al. (2017), and we compute yearly wage data for 1.23 million U.S. resident workers (excluding managers) and 52,000 Mexican immigrants living in the United States.[43] For Mexico,

[43]In the US census, wages are presented as "Wages or salary income last year" and are quoted in USD. However, there is a significant heterogeneity of individuals with respect to actual hours

we use the Mexican 2015 intercensal survey provided by IPUMS International, from which we extract a single variable: earnings per month in Mexican pesos for 5.78 million Mexican natives (excluding managers). To make the results comparable with those obtained for the United States, we multiply it by 12 and divide this outcome by the average 2015 exchange rate of 1 USD to 17.81 Mexican pesos, according to the OECD.[44] The constructed wage distributions give us the minimal and maximal values of annual wages: in the United States the minimal (maximal) wage is equal to $\$4,133$ USD ($\$183,380$ USD), whereas in Mexico it is equal to $\$45,834$ USD.

Furthermore, we feed the model with international trade data. The actual trade flows across the United States, Mexico and the ROW are taken from the OECD's Trade in Value Added database.[45] Mexico's price index is normalized to unity, while for the ROW it is determined by the trade-weighted purchasing power parity (PPP) differentials with the United States, resulting in $P_W = 0.28 \cdot P_U$. The U.S. price index, P_U, is computed to match the domestic production in the ROW economy, yielding a value of $P_U = 4.97$ and $P_W = 1.38$. We also take $Y_W = 4 \cdot Y_U$, as a rough approximation based on the World Development Indicators and the World Integrated Trade Solution databases maintained by the World Bank. The full trade matrix for the United States, Mexico and the ROW enables us to fit bilateral trade costs, τ_{ij}, assuming that $\forall i \; \tau_{ii} = 1$.

We need additional empirical moments to pin down the two-dimensional distribution of Mexican skill levels represented by the Clayton copula. To this end, we compute conditional probabilities of migration for Mexican natives (the probability of migrating conditional on being classified in a particular quantile of the Mexican wage distribution). We exploit the data provided by the Mexican Migra-

worked per week. In our model there is no intensive margin decision on labor supply, thus we have to control for wages per equivalent unit of time. To solve this, we store information about individual-specific typical hours worked and the number of weeks worked in the year being analyzed. From these data, we calculate a wage rate that is earned in a standard, 40-hour-per-week shift, assuming a person is employed for all 52 weeks: $w = w_{monthly} \cdot 40/hours_{PERweek} \cdot 52/weeks_{PERyear}$.

[44] For all wage distributions we remove 1 percent of the lowest and 2 percent of the highest values to skip outliers. Then, we smooth the data by interpolating the missing values locally (in case of sudden jumps in the empirical distribution). Finally, with these upgraded data, we compute kernel densities that allow us to generate $K = 100,000$ density points for $100,000$ quantiles of each distribution. Since we must discretize the whole model, we choose a grid of $K = 100,000$ points on which all the functions and distributions are going to be computed. This approach enables us to retain considerable accuracy in our computations while allowing our calibration and simulation algorithms to solve relatively fast.

[45] The added value exported from the United States to Mexico amounts to 1 percent of U.S. GDP (7.5 percent to ROW), while Mexican added value in exports to the United States constitutes almost 10 percent of Mexican GDP (11 percent to ROW).

tion Project (MMP), which indicates the wages Mexican immigrants to the United States earned before and after moving. This approach allows us to locate these particular individuals in our two-dimensional skill space (x_M, x_U) and, knowing that they migrated, compute the frequency from which quantile of Mexican wages they originated. The collection of these probabilities indicates the exact pattern of migrants' self-selection. We obtain a result that is very much in line with the findings of Moraga (2011) and Kaestner and Malamud (2014), both of whom provide evidence that the probability of a Mexican emigrating to the United States is inversely related to the quantile of the Mexican wage distribution in which she is located.

For the out-of-equilibrium fiscal part of our model, we collect the actual income and corporate tax rates and thresholds for the United States and Mexico from the OECD.[46]

3.3 Identification of the Reference Calibration

The identification strategy relies on the functional forms imposed in Section 3.1 and the datasets described in Section 3.2.[47] The functional structure includes parametric expressions for the copula function (the joint distribution of Mexican and U.S. labor market skills in the Mexican population), production functions, and individual utility functions (assuming a constant elasticity of substitution between varieties and the monetary and non-monetary costs of migration). This subsection provides a general discussion of the identification strategy. Appendix C offers a detailed description of the calibration algorithm and identification, as well as a graphical analysis of the chosen parameter vector.

The calibration algorithm has a simple conceptual design. First, we assign the values of the empirical moments from group (A) to the respective model parameters; these values are displayed in Table 1. Second, we use all degrees of freedom from the wage distribution of U.S. citizens to compute the skill distribution among U.S. citizens, $F(\cdot)$.[48] Third, in order to calibrate the vector of the

[46]The fiscal module enables us to compute indicative, first-round effects that migration policies would have on the net fiscal position of individuals in Mexico and the United States. Taxes and transfers are not internalized by individuals; migration decisions are reached by comparing gross real wages. For details, refer to Appendix C.

[47]Our model is only parametrically identified, similarly to self-selection models analyzed by Heckman (1979), Heckman and Sedlacek (1985), Heckman and Honore (1990), and Borjas (1987). It impossible to identify non-parametrically a selection model using a single cross-section (Heckman and Honore, 1990).

[48]Thus, the values of $F(\cdot)$ for all skill levels $x_U \in [x_U^c, 1]$ are residuals from Equation (17), taking all other model parameters and the actual wage distribution of U.S. residents as given.

remaining model parameters $\Xi = \{k_U, s_U, \gamma_U, k_M, s_M, \gamma_M, \theta, \Delta_{UM}, \delta_{UM}\}$, we conduct a numerical search through the nine-dimensional parameter space and seek for values of Ξ that minimize a loss function (C.4).[49] The loss function summarizes the distance between empirical moments obtained from the data (denoted by respective variables with a hat) and the values generated by the model. Our model is over-identified (we fit two country-specific fixed costs and two wage shares in GDP, followed by 100,000 quantiles for each wage distribution).[50] In particular, if the value of the loss function is zero, then the vector Ξ would (among others) need to solve the system of nine equations (C.I1)-(C.I9).

In our Monte Carlo calibration procedure, we find a close relationship between respective moments from the data and the parameters in Ξ, as depicted in Figure C.2. Table C.4 summarizes the fit of replicated moments with the calibrated parameters. Some parameters are precisely identified by the respective model equations and data moments, while others emerge as a solution to a subsystem of simultaneous equations. The first group of exactly identified parameters includes γ_U and γ_M, pinpointed by w_U^{share} and w_M^{share}, respectively, and θ, which is predominantly determined by the conditional probabilities of emigrating from Mexico, $P(\cdot)$.[51] Migration costs, Δ_{UM} and δ_{UM} are also exactly identified by the minimal and maximal wages received by Mexicans in Mexico and in the United States. Despite this precise identification, we decide to keep the latter two degrees of freedom relatively slack (not fitted exactly), as they would be extremely sensitive to the cuts that we impose on the observable wage distributions. We trade this less precise fit to gain a better fit for the remaining quantiles of wage distributions. Finally, two parameters of production functions per country remain to be identified: k_i, s_i. Since they affect not only country-specific moments (fixed costs and the dispersion of wage distributions) but also influence the location of the separation function (related to the skewness of wage distributions), they cannot be individ-

[49] The search is done using a Monte Carlo algorithm with a Simulated Annealing Optimization method (a variant of the Metropolis algorithm).

[50] Heckman and Honore (1990) prove that a two-country, log-normal Roy (1951) model is exactly identified with three country-specific parameters that determine the location, dispersion, and skewness of wage distributions and the number of migrants. It stands to reason that Roy's model with log-normal marginals and the (one-parametric) Clayton copula is also identified exactly by these moments. Our model extends Roy (1951) through the inclusion of worker-firm matching, which enlarges the set of parameters by γ_i's to account for firms' non-degenerated profit distributions. Actually, the parameter γ_i is exactly the one that determines the "distance" between the classic Roy (1951) model and our model with matching. In this sense, we need at least nine moments in the data to identify the model. We provide more, as we fit numerous quantiles of wage distributions, which despite being characterized by strong cross-sectional correlations, generate at least three moments per country: location, dispersion, and skewness.

[51] We also compute: x_U^c from $F(x_U^c) = u_U$; x_M^c from $C(x_U^c, x_M^c) = u_M$; and h_i^c from $1 - S_i/R_i^F$.

ually determined. In fact, they must fulfill the zero profit conditions for cutoff firms (Equations (C.I3) and (C.I4)), allow for the full range of wages in Mexico and the United States (Equation (C.I5)) and ensure that the actual measure of Mexican migrants fits the observable one (Equation (C.I6)). Even though there is a close relationship between these four parameters and particular moments (see Figure C.2), one has to bear in mind that they are a solution to a system of four simultaneous equation rather than an explicit one-to-one identification.

The chosen vector of parameters allows us to compute wage distributions that fit the actual data well. As shown in Figure 1, in which the empirical distributions are depicted in gray, while the model distributions are in black (long-dashed for Mexicans working in Mexico, double-dashed for Mexicans working in the United States, and dotted for U.S. citizens), in almost all instances the quantiles are well matched, apart from the left tail of the Mexican native wage distribution. The model visibly overestimates the annual minimal wage in Mexico and slightly underestimates the minimal remuneration Mexican immigrants receive in the United States. Since we solve the model from the right-hand side (starting from $x_U = 1$), the maximal wages in the United States are taken as given, while the highest Mexican wage is well matched in the calibration.

Figure 1: Evaluation of Model Fit: Distributions of Annual Wages

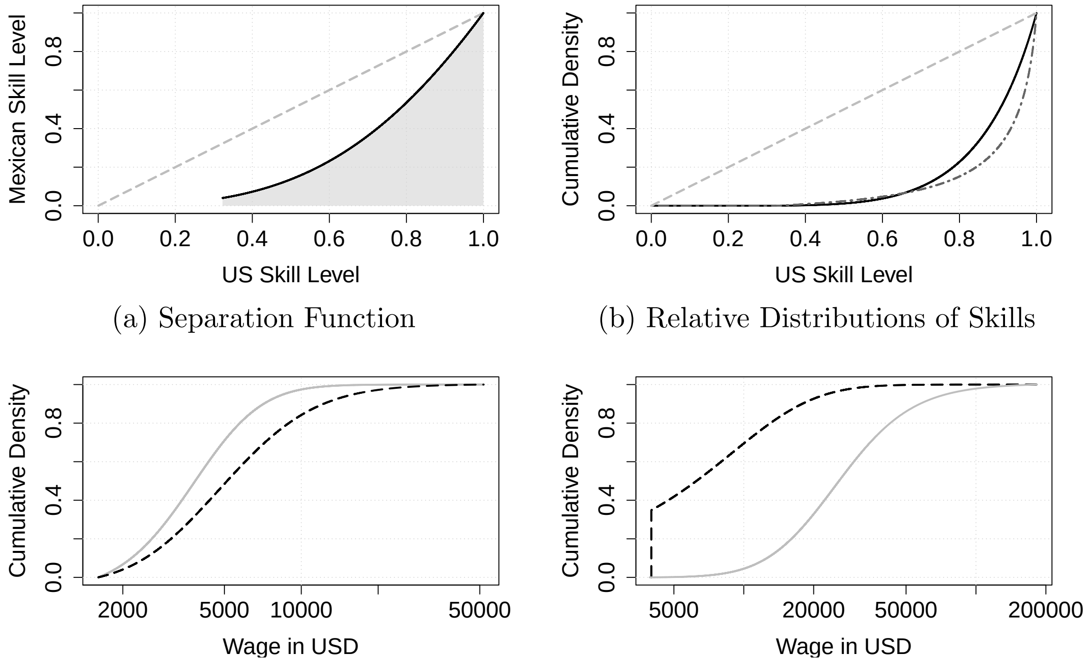
Note: Figure 1 depicts the fit of our model to the data on wage distributions. The black long-dashed line represents the model wage distribution in Mexico, the black double-dashed line represents the Mexican immigrants' wage distribution in the United States, and the black dotted line represents U.S. citizens' wage distribution. The gray lines depict respective distributions from the data. Horizontal axis is in log units of USD.

The selection pattern of Mexican immigrants to the United States generated by our calibrated model is depicted in Figure 2. Figure 2a plots the inverse separation function ϕ as a solid black line, in the (x_U, x_M) space, and shows who decides to migrate (gray surface to the right from the solid black line) and who decides to stay

in Mexico (left from the solid black line). We find that Mexican emigrants to the United States are strongly and positively selected among their peers with respect to the U.S. skill x_U. This unsurprising result is illustrated in Figure 2b, in which we compare the U.S. skill distribution across three populations of interest: Mexican natives (long-dashed gray), Mexican migrants (solid black), and U.S. citizens (two-dashed dark gray). Although Mexican emigrants possess a significantly higher level of U.S. skills than Mexicans in total (immigrants are *positively selected with regards to American skills*), they underperform in comparison with U.S. citizens. One can conclude that Mexican emigrants to the United States provide a generally lower intensity of U.S. skills than U.S. workers (apart from the very low skill levels, where U.S. workers are more numerous). This feature is preserved mainly because Mexicans as a whole are *strongly downgraded* in terms of U.S.-relevant skills compared to U.S. natives. In particular, this result means that the overall measure of Mexicans who are proficient in U.S.-specific tasks is very limited. At the same time, Mexican emigrants are significantly and *negatively selected in terms of Mexican skills*, a phenomenon stressed by recent empirical studies (Moraga, 2011; Kaestner and Malamud, 2014). Had the migrants returned to Mexico, they would have earned lower wages than stayers. The latter finding is depicted in Figure 2c, in which the long-dashed black line represents the observed distribution of wages for Mexican stayers, while the solid gray curve plots the counterfactual distribution of wages for Mexican emigrants to the United States had all of them returned to Mexico. A confirmation of a substantial positive selection of Mexican immigrants with respect to their U.S. skill level is shown in Figure 2d. There, we compare the wage distributions of current Mexican migrants in the United States (the solid gray curve) and Mexican natives in a counterfactual state of the world in which they all move to the United States (long-dashed black line). More than 35 percent of these individuals would be characterized by a U.S.-specific skill level that falls below the threshold required to join the U.S. labor market.

4 Main Results

This section provides a quantitative evaluation of the impact of Mexican immigration to the United States on both economies, with a particular focus on welfare. We construct a counterfactual scenario which removes Mexican immigrants from the U.S. labor market and returns them to Mexico. We compare this counterfactual with the *status quo*. In Section 4.2, we decompose the aggregate effects of the presence of Mexican workers in the United States into these three chan-



(a) Separation Function

(b) Relative Distributions of Skills

(c) Selection in Mexican Skills

(d) Selection in U.S. Skills

Figure 2: Selection of Mexican Emigrants to the United States

Note: Figure 2 depicts the patterns of Mexican migrants' selection. Figure 2a plots the separation function (in solid black) compared to the 45-degree line (dashed gray) in the space of Mexicans' skill levels (x_U, x_M). Figure 2b presents the distributions of U.S. skill levels among U.S. citizens (two-dashed dark gray), Mexican citizens (dashed gray), and Mexican immigrants (solid black). Figure 2c illustrates the distribution of Mexican citizens' wages (dashed black) and the counterfactual distribution of Mexican immigrants' wages after their return to Mexico (solid gray). Figure 2d compares the wage distribution of Mexican immigrants' wages in the United States (solid gray) with the counterfactual distribution of Mexican citizens' wages had all of them moved to the United States (dashed black).

nels: the labor market effect, the love of variety effect, and the fiscal effect. In Section 4.3, we simulate the U.S. economy under a range of different immigration policies and report their implications for U.S. citizens. Finally, we propose a feasible redistributive policy under which all U.S. citizens benefit from Mexican migration.[52]

4.1 The Economic Effects of Mexican Migration to the United States

We commence with a discussion of the economic effects generated by Mexican immigrants working and consuming in the United States. The results presented

[52]Additional results, in which we compare our model to CES labor markets in the spirit of Acemoglu and Autor (2011) and Dustmann et al. (2013), are available in Appendix E.

in the remainder of this subsection document the difference between the current situation (represented by our calibrated economy) and the counterfactual case, in which there is no Mexican migration to the United States (this is modeled by setting migration costs to infinity). Thus, positive values represent gains from migration, and negative values represent losses.

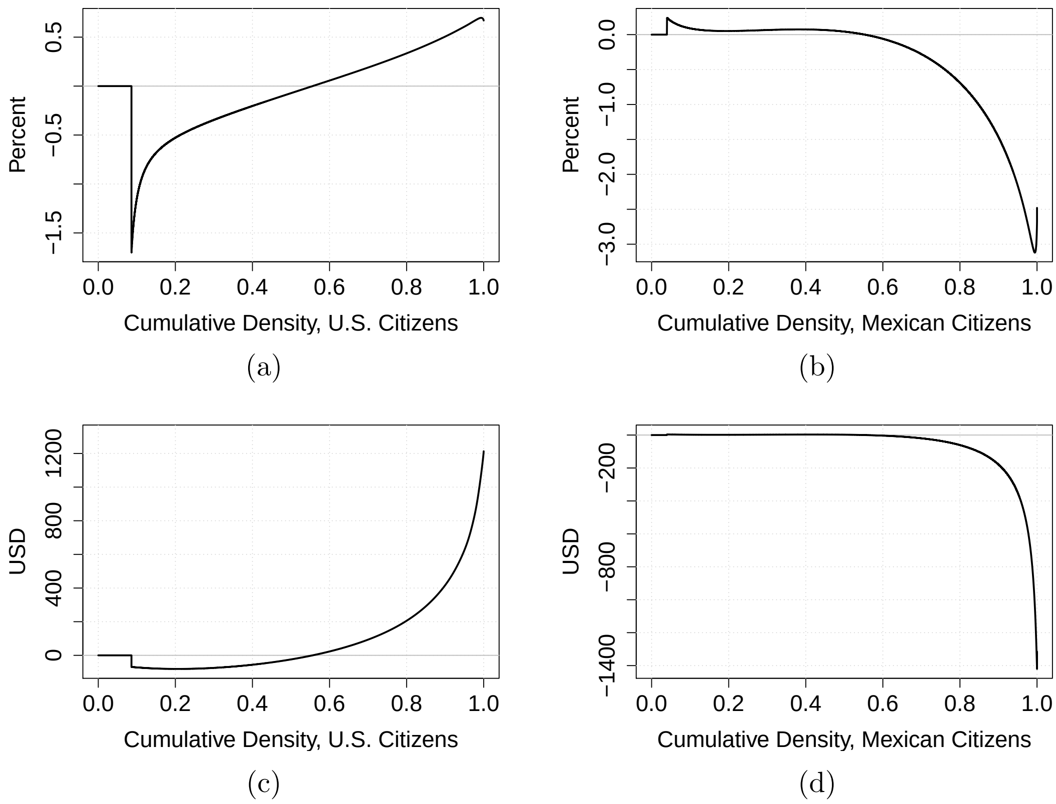


Figure 3: Welfare Effects of Mexican Immigration to the United States

Note: Figure 3 presents the overall economic effects of Mexican migration to the United States (difference between the current situation and the no-migration counterfactual). Figure 3a (3b) contains relative changes in U.S. citizens' and Mexican stayers' real wages along the distribution, while Figure 3c (3d) provides absolute variations in U.S. citizens' and Mexican stayers' real wages (in USD). Horizontal axes represent quantiles of respective wage distributions.

The overall economic effects (including the out-of-equilibrium fiscal effect) of Mexican migration to the United States, expressed as changes in real wages denominated in U.S. dollars, are depicted in Figures 3a and 3c.[53] The group of losers comprises U.S. citizens who earned low wages, while the group of winners consists of high-earners.[54] Given what we know about the migrants' self-selection

[53]Its decomposition into the labor market effect, the love of variety effect and the fiscal effect is discussed in Section 4.2.

[54]Most of the losers simply earn lower wages than before, while a small fraction is forced into unemployment.

in the calibrated economy (see Figure 2 in in Section 3.3), the general pattern of redistribution is as expected. The distribution of migrants' U.S.-relevant skills trends lower compared to the locals, which means that low-earning U.S. citizens are close substitutes with most of the immigrants. The high-earning natives, in contrast, gain not just due to the love-of-variety effect, but also due to the fact that they possess skill levels that are high enough, so they act as complements to the Mexican migrants.

The real wage effects of Mexican migration in the United States range from –1.6 percent (roughly –80 USD of annual remuneration for the ninth percentile of U.S. wage distribution) to 0.7 percent (approximately 1,200 USD for the 99th percentile). Our model suggests that the employment effects from Mexican immigration are extremely small; the U.S. unemployment rate rises by only 0.006 percentage points. Even though the level of wage inequality among U.S. citizens increases, the average wage earned by U.S. citizens moderately grows by 0.27 percent (approximately 100 USD).

The emigration of Mexican workers to the United States also triggers significant effects for the Mexican stayers; these effects are depicted in Figures 3b and 3d. In this case, one observes results that qualitatively opposite to those in the United States: low-earning stayers gain from emigration, while high-earning stayers lose. This effect is, again, driven mostly by the patterns of selection: Mexican migrants are negatively selected with respect to the skills used in the Mexican economy (see Figure 2c).[55] The gains of low earning stayers are, however, very low in magnitude: just 0.25 percent (4 USD in the fifth percentile of the Mexican wage distribution). The most skilled Mexicans lose up to –3.1 percent of their real wage (amounting to around –1,400 USD). Unemployment decreases modestly (by 0.06 percentage points), and the overall amount of firms decreases by 13 percent, which is roughly the size of the migration wave.

The greatest gains from migration accrue to the Mexican migrants themselves. Figure 4a illustrates the magnitudes of the net economic gains collected by Mexicans, who migrate to the United States. Note that the extent to which an immigrant becomes better off from moving abroad is only partly dependent on her absolute skill endowment.[56] People characterized by a skill set that is exactly on

[55]It is worth noting that, using the nomenclature from Borjas (1987), this does not imply a negative selection pattern. In contrast, our calibrated models suggests refugee sorting: the Mexican migrants are selected positively with respect to the skills needed in the U.S. economy; however the supply of such skills is so much better among U.S. citizens than among Mexicans, that despite the positive selection, the migrants' skill distribution is still worse than those of U.S. citizens.

[56]This statement and the following interpretation is true only if the separation function is

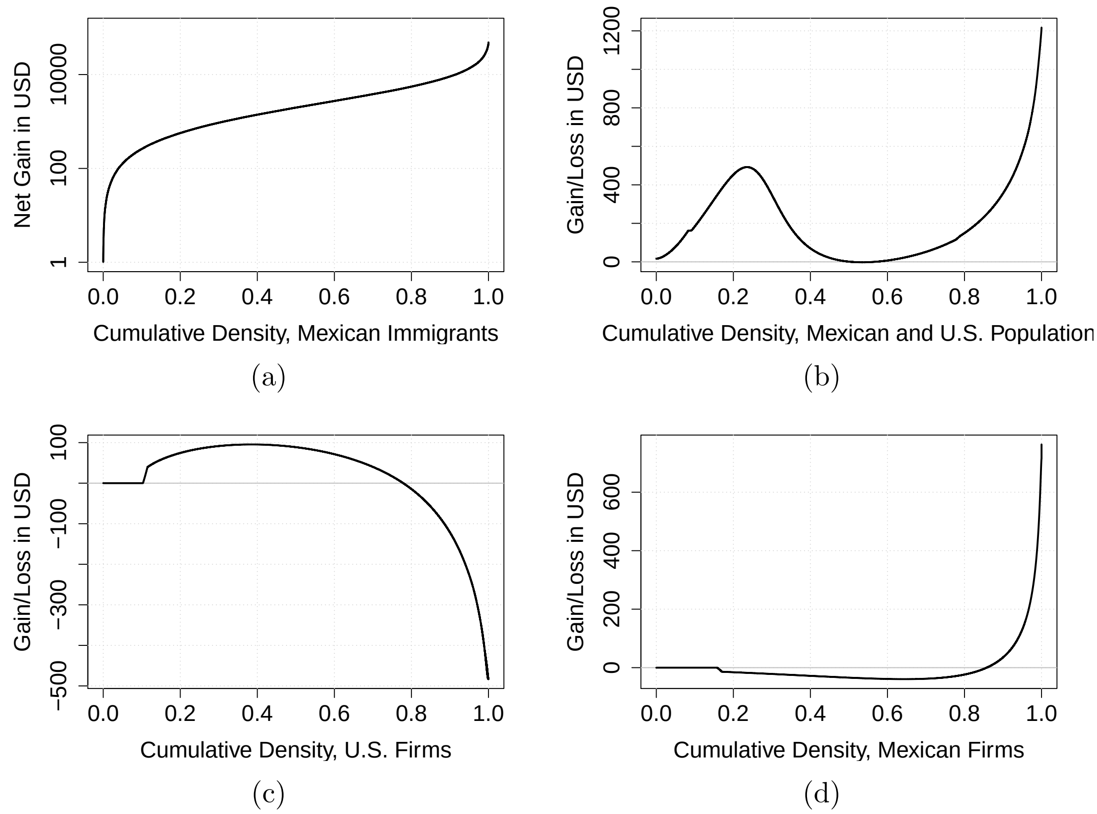


Figure 4: Further Economic Effects of Mexican Immigration to the United States

Note: Figure 4 presents the distributions of net gains/losses from Mexican migration to the United States for: Mexican immigrants (4a), the pooled population of U.S. citizens and Mexicans (4b), U.S. firms (4c), and Mexican firms (4c). Horizontal axes represent quantiles of respective distributions.

the separation line are indifferent between staying and moving, as depicted in Figure 2a. This means that their net gain from moving to the United States equals zero, the difference between U.S. and Mexican wage rate corrected for migration costs (see Equation (20)). Individual gains from migration increase substantially with a worker's distance from the separation function. Overall, the net benefits from migration attain sizable magnitudes, with the maximum being approximately 48,000 USD and the median benefit close to 1,975 USD.

A question of great social and political importance relates to the change in earnings inequality: within each country as well as when the combined population living in Mexico and the United States is taken into account. The results reported so far imply that Mexican migration to the United States increases wage

constant in both states of the world. The latter is necessarily true if the number of migrating workers is zero. Otherwise, the individual accounting of costs and benefits from migration is subject to labor market equilibrium and general equilibrium effects.

inequality in the United States and reduces it in Mexico. The aggregate picture is analyzed in Figure 4b, which reports the change in the wages that correspond to each quantile of the combined wage distributions in both countries. Strikingly, Mexican migration to the United States improves the global distribution of wages in first-order stochastic dominance sense. This finding is perfectly consistent with the results reported above: Note that a worker can occupy different quantiles in the combined wage distribution prior to and after migration. For example, with Mexican immigrants, low-skilled U.S. workers will occupy lower quantiles than previously; and despite the fact that they earn less than previously, their wages are still higher than those of the previous occupants of these quantiles. The increases in wages are particularly pronounced in the left and the right tail of the aggregate wage distribution, but there is almost no change around the median. This result is reflected in the change in the Gini coefficient of wage inequality, which decreases from 0.473 in the no-migration counterfactual to 0.471 in the calibrated economy. The current Mexican-U.S. migration pattern reduces overall wage inequalities, since workers sort from a relatively high-inequality country to the relatively low-inequality country.

Firms also respond to changes in the supply of skills when Mexicans are permitted to move to the United States. Opening the border increases the supply of low- and medium-skilled workers in the United States and invites low- and medium-productivity firms to reduce the offered wages. These companies are allowed to form matches with cheaper, yet more productive employees, which boosts their profits between the 10th and the 80th percentile, as presented in Figure 4c. The effect that dominates from the 80th percentile onwards relates to the equilibrium number of new firms entering the market. Recall that a higher number of workers in the United States increases the number of firms in the market, which symmetrically extends the supply of firms along the whole productivity distribution. Accordingly, there is a larger number of highly productive firms and the new ones that enter the market face higher competition for the high-skilled workers. This forces firms to bid up high-skilled wages, which erodes firms' profits. A reverse situation occurs in Mexico (Figure 4d), where an outflow of negatively selected emigrants discourages firms from staying in the market, benefits high-productivity firms, and depresses the profits of the least productive employers. Comparing the results in Figures 4c and 4d to those presented in Figure 3, one observes a clear pattern of redistribution between capital (shareholders of firms) and labor (workers) after accounting for Mexican-U.S. migration. In the United States, the gains are transferred from the least-skilled workers to the least-productive firms and

from more productive firms to high-skilled workers. In Mexico, the outcomes are exactly opposite.

Mexican migration to the United States turns out to act as a substitute for trade in our calibrated economy. As reported in Table 2, bilateral trade flows between the United States and Mexico decrease by more than 6 percent due to Mexican migration.[57] Pointedly, this reduces the U.S. trade deficit with Mexico by 0.26 percentage points. In addition to boosting domestic absorption by almost 5 percent, the U.S. exports to the ROW grow by 2.4 percent, while imports from the ROW increase by 2.5 percent.

Table 2: Percent Changes in Trade, Bilateral Matrix

From:\To:	ROW	MEX	US
ROW	–0.06%	–8.99%	2.52%
MEX	–9.36%	–19.17%	–6.58%
US	2.44%	–6.32%	4.91%

4.2 Decomposition of the Total Effect

In this subsection, we investigate how much each of the three main channels featured in our model—the labor market effect, the market size effect, and the fiscal effect—contributes to the overall results presented in the previous section. These three effects are depicted in Figure 5.

We start with the *labor market effect*, meaning the impact of migration on wages when the revenue produced by each match (given by Equation (8)) and the price indexes are held constant.[58] The labor market effect can be further decomposed into the change in wages under a fixed supply of firms, along with the firm entry and exit effect. The red double-dashed lines in Figure 5 depict the change in *wages if firm supply remains constant*: Mexican migration puts wages in the United States under severe pressure, and increases the wages of Mexican stayers. The intuition for this is standard: an inflow of migrants worsens the match quality for all workers, while at the same time improving the outside options of all firms. This necessarily results in a fall in wages for all U.S. citizens. The

[57] Fewer Mexican exports to the United States is a consequence of U.S. consumers' substitution of Mexican imported goods with more differentiated and cheaper U.S. domestic varieties. Exports to Mexico drop because of the demand effect (a smaller population of Mexican stayers), even though the prices of U.S. varieties go down.

[58] This can be fruitfully thought of as the change that would happen if the demand function each firm faces does not change (Equation (6)).

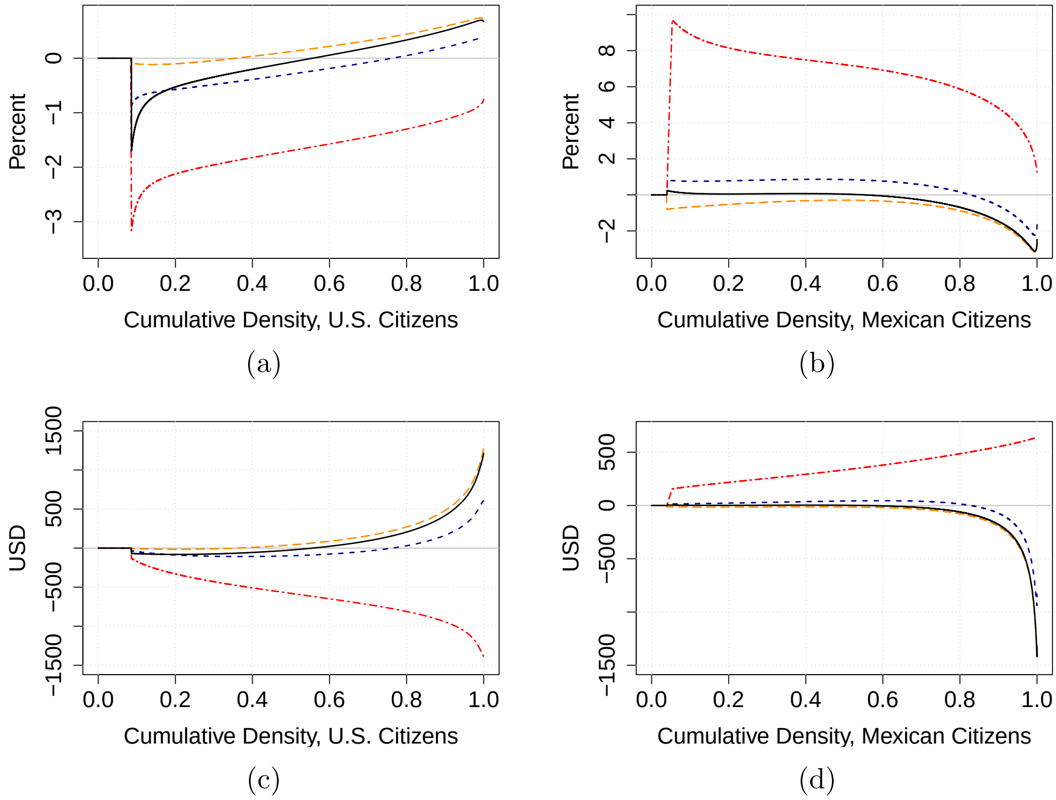


Figure 5: Decomposing the Total Effects of Mexican migration to the United States

Note: Figure 5 depicts the economic effects of Mexican migration to the United States decomposed into three economic effects. The red double-dashed lines depict labor market effect with fixed supply of firms. The blue dashed lines present the overall labor market effect. The orange long-dashed lines add the market size effects (this includes the effects of trade) to labor market effects. The black solid lines summarize the total effect of all four sub-effects (including the fiscal impact). Figure 5a (5b) depicts relative changes in U.S. citizens (Mexican stayers') wages along the distribution, while Figure 5c (5d) presents absolute variations in U.S. citizens' (Mexican stayers') wages in USD. Horizontal axes represent quantiles of respective wage distributions.

heterogeneity in the effect is caused by the fact that the skill distribution among migrants differs from the skill distribution among the U.S. citizens, and workers of different skill levels act as imperfect substitutes when the supply of firms is fixed. The nominal wage effects range from over –3.1 percent (–130 dollars) for the U.S. citizen that become unemployed, to slightly less than –1 percent (–1,400 USD) for the most skilled U.S. worker. For the same reasons, Mexican emigration generates strong wage gains for Mexican stayers. The absolute magnitudes reach the level of 640 USD for the most skilled Mexican stayer.

Firms' entry and exit, however, mitigates these severe wage effects. As wages fall in the United States, so do labor costs for U.S. firms. This effect increases

the profits generated by incumbent U.S. firms, which prompts more entry into the U.S. market. In Mexico, contrarily, labor costs increase, profits fall, and the supply of firms falls.[59] In Figure 5, the overall labor market effect is depicted by the blue dashed lines (the difference between the blue and the red lines represents the exact magnitude of the firms' entry and exit effect alone). As the panels in Figure 5 show, the role of firms' entry and exit is crucial not only in terms of magnitudes. More importantly, it changes the signs of the economic effects experienced by a significant proportion of workers in both countries. One can now observe that both populations clearly polarize into a well-defined group of winners (high-skilled workers in the United States, low- and medium-skilled workers in Mexico) and losers (low- and medium-skilled workers in the United States, high-skilled workers in Mexico) as a result of Mexican migration. In the United States, approximately two out of three citizens remain strictly worse off, while in Mexico, less than 20 percent lose out. On the whole, the equilibrium adjustment in visibly reshapes the overall outcomes in both labor markets. Nevertheless, even with the firm entry and exit effect taken into account, the overall labor market effect of Mexican migration makes the majority of U.S. citizens worse off.

The third element of our decomposition adds the *market size effect* to the previous analysis (illustrated by the orange long-dashed lines in Figure 5). The latter works through the ideal price indexes induced by a change in the available varieties of consumption goods, similar to Melitz (2003), di Giovanni et al. (2015), and Aubry et al. (2016). In our model, the addition of the market size effect results in a strikingly optimistic evaluation of Mexican migration on the welfare of U.S. citizens, but also makes all Mexican stayers worse off. A higher supply of workers in the United States, followed by a firm entry effect, generates a positive consumption externality as the set of domestically produced varieties becomes richer. People tend to consume fewer units of more differentiated goods, which diminishes the ideal price index, and boosts real wages. In quantitative terms, the expanded variety of goods in the United States induces a 0.35 percent decrease in prices (and a respective increase in real wages). This price effect reduces the share of U.S. losers from Mexican migration from about 68 percent to approximately 27 percent. In Mexico, the market size effect significantly depresses the wages of the Mexican workers who remain in the country and renders the overall result unambiguously negative for everyone's welfare, as there is a 1.03 percent jump in prices.

[59]This firm entry and exit effect is best thought of as in Melitz (2003): over time there is attrition of firms, so the decrease in entry causes a fall in overall supply.

Finally, we add the *net fiscal effect*, which completes the total economic effect that is depicted by a black solid line in Figure 5 (identical to the outcome in Figure 3). Specifically, we calculate how much the budget deficit in country i would change in response to Mexican immigration, and then redistribute this difference across all workers in country i. This is an out-of-equilibrium exercise, since in our model Mexican workers only take gross wages into account when making their migration decision, but it nevertheless provides an indication about the order of magnitudes. Despite being quantitatively small, the change in net benefits received by incumbent citizens sets the final share of winners and losers to approximately 50:50 in both countries. Due to Mexican immigration, U.S. citizens are forced to pay 65.5 USD of the budget-balancing lump-sum tax, while Mexican stayers benefit from a 16.5 USD lump-sum transfer. Note that this is calculated given the existing U.S. tax and transfers structure. However, the presence of Mexican immigrants generates unambiguous increases in the average wages in both economies, even though immigrants are less skilled than incumbent populations. Thus, there must necessarily exist a redistributive policy between winners and losers that makes all U.S. citizens benefit from Mexican immigration. We will consider this questions after discussing the effects of alternative U.S. migration policies toward Mexicans in the following subsection.

4.3 Changes to Migration Policy

In this section we evaluate the impact of U.S. migration policies that moderately change the monetary cost of migration (δ_{UM}).[60] First, we assume that the change in migration cost is "burned", that is, that it constiutes a change in cost not just for each migrant, but also the economy as a whole. Later, we consider an alternative scenario, in which the increase in migration cost is caused by a tax on migration, and the revenues from that tax are redistributed among U.S. citizens as a lump sum transfer.[61]

Figure 6 summarizes the results: panel (a) includes only the overall labor market effects of migration, whereas panel (b) considers all channels of the economic impact of migration (total labor market effects as well as market size and fiscal effects). The horizontal axes represent the difference in the monetary cost of migration between the counterfactual and the calibrated economy, while the vertical

[60] Very large increases in δ_{UM} imply the same outcomes as discussed in Section 4.1.

[61] Equivalently, the increase in the cost of migration could be thought of as an increase in the visa cost above the average cost of running the visa program, with the additional revenue redistributed among U.S. citizens.

axes compare the counterfactual outcomes with the no-migration scenario. The solid black line depicts the effect when the additional cost is "burned", whereas the gray dashed line depicts the case when the additional revenue collected from migrants is redistributed among U.S. citizens.

Focusing exclusively on the labor market effects (Figure 6a), we find that higher monetary migration costs significantly reduce the share of Mexican immigrants in the United States. Simultaneously, the average wages received by U.S. citizens increase slightly. In contrast, lower immigration costs invite more Mexican immigrants to the United States and decrease average wages. Changes in δ_{UM} also affect the migrants' skill distribution. Recall that in Roy's model with normally distributed rewards, a rise in the additive cost differentiates more the average migrant from the average Mexican. This remains true in our counterfactual exercise: because in our calibrated economy the average Mexican migrant is more skilled in the U.S. skill set than the average Mexican is (see Figure 2), the increase in migration costs raise the average U.S. skill set possessed by Mexican migrants. However, since the Mexican population as a whole is less well-equipped in the U.S. skill set than the U.S. citizen population is, this leads—for increases of about 2,000 USD annually—to a skill distribution among Mexican migrants that is more concentrated around the median skill level than is the corresponding skill distribution among U.S. citizens. For this reason, only the medium-skilled U.S. citizens lose from migration, whereas both the least- and most-skilled U.S. citizens gain. This leads to an increase in the overall share of people gaining from migration, as depicted in the lowest panel of Figure 6a.

Adding market size and fiscal effects from Mexican immigration to the United States (Figure 6b) alters significantly the general picture. Changes in average wages are now dominated by the market size effect and remain positive (in relation to the no-migration scenario) even for very high rates of immigrant inflows. Surprisingly, an increase in migration costs of 1,000 USD maximizes the average gain from Mexican migration, improving the welfare of every U.S. citizen, but corresponds to removing around one-third of current Mexican migrants in the United States. Decreases in migration costs steadily erode the group of U.S. winners from Mexican migration to the United States.

Now consider the alternative scenario, in which the increase in migration cost is caused by a tax on migration.[62] This means that the overall cost is not "burned", but is redistributed among the U.S. citizens. We consider the simplest and easiest

[62]To avoid any stigma connected to being a migrant, this tax could be charged up-front, or framed as an increase in the cost of visa.

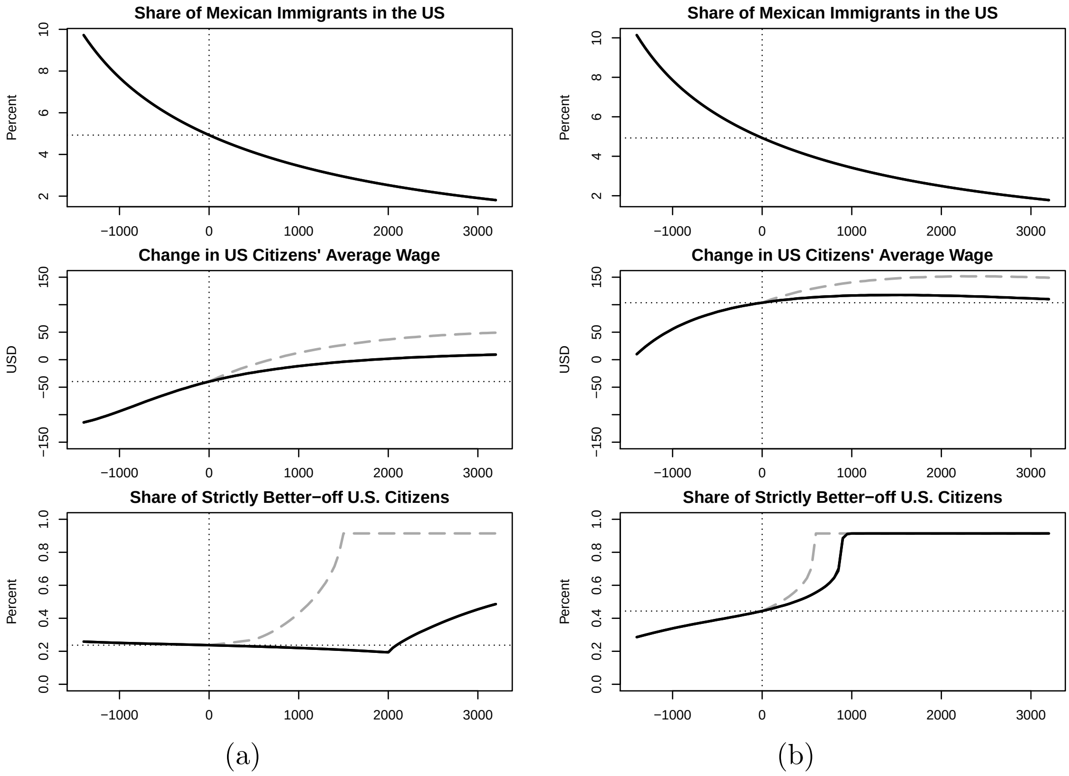


Figure 6: The Aggregated Effects of Alternative Migration Policies in the United States

Note: Figure 6 presents aggregated measures of U.S. citizens' welfare after implementing alternative costs for Mexican immigrants, relative to the no-migration scenario. The first row depicts changes in the share of Mexican immigrants in the United States in percent; the second row contains the changes in average wages of U.S. citizens in USD; and the third row depicts the fraction of U.S. citizens who are strictly better off in percent. Panel (a) summarizes the results with the labor market effect only, while panel (b) considers all economic effects including market size and fiscal externalities. Black solid lines consider the case of "burned" migration costs, while the gray dashed lines denote the case of redistributing additional visa costs as transfers for U.S. citizens. Horizontal axes represent deviations in monetary costs of migration, δ_{UM}, relative to the *status quo*.

way to implement such a redistributive scheme, in which all U.S. citizens receive the same transfer. Such a scheme, while undoubtedly very flawed from the standpoint of fairness, has the advantage of being easily implementable. Specifically, it does not require identifying the actual U.S. winners and losers from Mexican immigration because all U.S. citizens are treated alike.

As can be seen in Figure 6a, an annual tax of 1,480 USD levied on Mexican immigrants is sufficient to ensure that all U.S. citizens gain from migration even if only the labor market effects are taken into account. Once the market size and fiscal effect are also factored in, the required tax burden reduces to 624 USD (1.7 USD per day). In the case of labor market effects, this "migration benefits

everyone" state is achieved when the current size of the Mexican immigrants population in the United States is reduced by 41 percent; in the case of all economic effects, only a 21 percent decrease in Mexican migration is required. Finally, notice that the above-mentioned estimates are very conservative. Multiple features of our model motivate this assertion—for instance, we assume a perfect substitutability between migrants and U.S. citizens within a skill group of measure zero. Including a slight degree of skill complementarity, as indicated by Ottaviano and Peri (2012) or Manacorda et al. (2012), would cause low- and medium-skilled U.S. citizens to lose much less from Mexican immigration. Observe also that the full model assumes a very conservative market size effect and a rather pessimistic net fiscal position from the presence of Mexican immigrants in the United States and overlooks some important effects induced by immigration (innovation effects, technology spillovers, and network externalities or gains from diversity). In this sense, a migration cost of 624 USD per year (1.7 USD per day) can also be seen as an upper bound on the tax necessary imposed on Mexicans to make all U.S. citizens better off.

5 Conclusions

International migration has become a key subject of contemporary economic, social, and political discussions. It has recently gained unprecedented societal recognition and media coverage, and affected many electoral results over the last few years. Nonetheless, how to evaluate the welfare impact of international migration has remained an intensively debated problem. In this paper, we propose a new model that provides a fresh look into how the size and quality of the migrant population affects people living in the origin and destination countries in terms of the respective wage distributions. Our approach extends the model of Gola (2018) by combining the selection model by Roy (1951), the matching model in Sattinger (1979), and the trade theory in Krugman (1980) and Melitz (2003). By calibrating the model using Mexican and U.S. data for 2015, we verify that the self-selection of Mexican immigrants to the United States is the key factor that shapes the overall welfare implications of migration in both countries.

We find that Mexican immigration has a significantly negative effect on wages in the United States and corresponding positive effect on wages in Mexico if the supply of firms is kept constant, which is then attenuated by a beneficial firm supply adjustment in the United States and a detrimental adjustment in Mexico. The firm entry and exit effect induces complementarities between low- and high-skilled

individuals and generates welfare gains and losses along the wage distributions. Moreover, immigration (emigration) induces positive (negative) consumption externalities in the form of the market size effect and affects fiscal budgets in both countries. Overall, Mexican immigration to the United States benefits 44 percent of the U.S. population – predominantly high-skilled individuals – while approximately 47 percent of U.S. citizens are harmed.

Our evaluation of alternative immigration policies between Mexico and the United States reveals that any liberalization of the cost that a Mexican citizen pays for a U.S. visa will idecrease the share of U.S. citizens who gain from migration. Conversely, increasing the pecuniary costs of migration (thus reducing the number of Mexican immigrants and improving their average quality) increases the share of winners from immigration in the United States. In this sense, conducting a democratic vote on various U.S. immigration policies towards Mexico is expected to support further restrictions. However, our estimations suggest that the welfare losses for low- and medium-skilled U.S. citizens generated by Mexican immigration can be easily compensated for through a small lump-sum transfer payment. A redistributive tax of 1.7 USD per day imposed on Mexican immigrants is sufficient to make all U.S. citizens better off.

References

Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Allen, T. and Arkolakis, C. (2014). Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics*, 129(3):1085–1140.

Aubry, A., Burzyński, M., and Docquier, F. (2016). The Welfare Impact of Global Migration in OECD Countries. *Journal of International Economics*, 101:1–21.

Auerbach, A. J. and Oreopoulos, P. (1999). Analyzing the Fiscal Impact of US Immigration. *The American Economic Review*, 89(2):176–180.

Azzalini, A. and Valle, A. D. (1996). The Multivariate Skew-Normal Distribution. *Biometrika*, 83(4):715–726.

Biavaschi, C., Burzyński, M., Elsner, B., and Machado, J. (2019). Taking the Skill Bias out of Global Migration. *Journal of Development Economics, forthcoming.*

Borjas, G. J. (1987). Self-Selection and the Earnings of Immigrants. *The American Economic Review*, 77(4):531–553.

Borjas, G. J. (2003). The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *The Quarterly Journal of Economics*, 118(4):1335–1374.

Borjas, G. J., Kauppinen, I., and Poutvaara, P. (2018). Self-Selection of Emigrants:

Theory and Evidence on Stochastic Dominance in Observable and Unobservable Characteristics. *The Economic Journal*, 129(617):143–171.

Borjas, G. J. and Monras, J. (2017). The Labour Market Consequences of Refugee Supply Shocks. *Economic Policy*, 32(91):361–413.

Burstein, A., Hanson, G., Tian, L., and Vogel, J. (2017). Tradability and the Labor-Market Impact of Immigration: Theory and Evidence from the US.

Burzyński, M. (2018). Trading Goods or Human Capital: The Gains and Losses from Economic Integration. *The Scandinavian Journal of Economics*, 120(2):503–536.

Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, 19(1):22–64.

Card, D. (2009). Immigration and Inequality. *American Economic Review*, 99(2):1–21.

Chiquiar, D. and 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.

Costrell, R. M. and Loury, G. C. (2004). Distribution of Ability and Earnings in a Hierarchical Job Assignment Model. *Journal of Political Economy*, 112(6):1322–1363.

di Giovanni, J. and Levchenko, A. A. (2013). Firm Entry, Trade, and Welfare in Zipf's World. *Journal of International Economics*, 89(2):283–296.

di Giovanni, J., Levchenko, A. A., and Ortega, F. (2015). A Global View of Cross-Border Migration. *Journal of the European Economic Association*, 13(1):168–202.

di Giovanni, J., Levchenko, A. A., and Ranciere, R. (2011). Power Laws in Firm Size and Openness to Trade: Measurement and Implications. *Journal of International Economics*, 85(1):42–52.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic Competition and Optimum Product Diversity. *The American Economic Review*, 67(3):297–308.

Docquier, F., Ozden, Ç., and Peri, G. (2014). The Labour Market Effects of Immigration and Emigration in OECD Countries. *The Economic Journal*, 124(579):1106–1145.

Dupuy, A. (2012). A Microfoundation for Production Functions: Assignment of Heterogeneous Workers to Heterogeneous Jobs. *Economica*, 79(315):534–556.

Dupuy, A. (2015). The Assignment of Workers to Tasks with Endogenous Supply of Skills. *Economica*, 82(325):24–45.

Dupuy, A., Galichon, A., Jaffe, S., and Kominers, S. D. (2017). Taxation in Matching Markets. *SSRN Electronic Journal*.

Dustmann, C. and Frattini, T. (2014). The Fiscal Effects of Immigration to the UK. *The Economic Journal*, 124(580).

Dustmann, C., Frattini, T., and Preston, I. P. (2013). The Effect of Immigration Along the Distribution of Wages. *Review of Economic Studies*, 80(1):145–173.

Foged, M. and Peri, G. (2016). Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data. *American Economic Journal: Applied Economics*, 8(2):1–34.

Gola, P. (2018). Supply and Demand in a Two-Sector Matching Model.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., and Woessmann, L. (2015). Returns to Skills Around the World: Evidence from PIAAC. *European Economic Review*, 73:103–130.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153.

Heckman, J. J. and Honore, B. E. (1990). The Empirical Content of the Roy Model. *Econometrica*, 58(5):1121.

Heckman, J. J. and Sedlacek, G. (1985). Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market. *Journal of Political Economy*, 93(6):1077–1125.

Heiland, I. and Kohler, W. (2018). Global Economy.

Iranzo, S. and Peri, G. (2009). Migration and Trade: Theory with an Application to the Eastern–Western European Integration. *Journal of International Economics*, 79(1):1–19.

Kaestner, R. and Malamud, O. (2014). Self-Selection and International Migration: New Evidence from Mexico. *Review of Economics and Statistics*, 96(1):78–91.

Katz, L. F. and Murphy, K. M. (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The Quarterly Journal of Economics*, 107(1):35–78.

Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. *The American Economic Review*, pages 950–959.

Llull, J. (2018). Immigration, Wages, and Education: A Labour Market Equilibrium Structural Model. *The Review of Economic Studies*, 85(3):1852–1896.

Mak, E. and Siow, A. (2017). Occupational Choice and Matching in the Labor Market. Technical report, IZA Discussion Papers.

Manacorda, M., Manning, A., and Wadsworth, J. (2012). The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain. *Journal of the European Economic Association*, 10(1):120–151.

Mas-Colell, A., Whinston, M. D., Green, J. R., et al. (1995). *Microeconomic Theory*, volume 1. Oxford University Press New York.

Melitz, M. (2003). The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations. *Econometrica*, 71(6):1695–1725.

Moraga, J. F.-H. (2011). New Evidence on Emigrant Selection. *The Review of Economics and Statistics*, 93(1):72–96.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the Effect of Immigration on Wages. *Journal of the European Economic Association*, 10(1):152–197.

Parsons, C. R., Winters, L. A., et al. (2014). International Migration, Trade and Aid: A Survey. *International Handbook on Migration and Economic Development*, pages 65–112.

Peri, G. and Sparber, C. (2009). Task Specialization, Immigration, and Wages. *American Economic Journal: Applied Economics*, 1(3):135–169.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative Spatial Economics. *Annual Review of Economics*, 9:21–58.

Rowthorn, R. (2008). The Fiscal Impact of Immigration on the Advanced

Economies. *Oxford Review of Economic Policy*, 24(3):560–580.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3(2):135–146.

Ruggles, S., Genadek, K., Goeken, R., Grover, J., and Sobek, M. (2017). Integrated Public Use Microdata Series: Version 7.0 [dataset]. *University of Minnesota, Minneapolis.*

Sattinger, M. (1975). Comparative Advantage and the Distributions of Earnings and Abilities. *Econometrica*, 43(3):455.

Sattinger, M. (1979). Differential Rents and the Distribution of Earnings. *Oxford Economic Papers*, 31(1):60–71.

Sklar, A. (1959). *Fonctions de Repartition a n-Dimensions et Leurs Marges.*

Storesletten, K. (2000). Sustaining Fiscal Policy Through Immigration. *Journal of Political Economy*, 108(2):300–323.

Teulings, C. N. (1995). The Wage Distribution in a Model of the Assignment of Skills to Jobs. *Journal of Political Economy*, 103(2):280–315.

Teulings, C. N. (2005). Comparative Advantage, Relative Wages, and the Accumulation of Human Capital. *Journal of Political Economy*, 113(2):425–461.

Appendices to: Mexican Migration to the United States: Selection, Assignment, and Welfare*

Michał Burzyński

Luxembourg Institute of Socio-Economic Research (LISER)

Paweł Gola

Norwegian Business School (BI)

June 2019

Abstract

Appendix A provides proofs of all formal statements. Appendix B provides a graphical comparison of the Clayton, Gaussian, and Gumbel copulas. Appendix C provides details of the calibration procedure, Appendix D reports the results of several robustness checks and Appendix E compares numerically how wages change in response to a supply shock in the assignment and CES models.

*Contact: Michał Burzyński: michal.burzynski@liser.lu; Paweł Gola: pawel.gola@bi.no.

Appendix A Proofs

Proof of Lemma 1

Follows immediately by the exact same reasoning as the proof of Lemma 2 in Gola (2018).

Proof of Proposition 1

The proof will consist of two steps. First, we will prove that

$$A^* \in \mathbb{E} \Rightarrow V(A^*) - V(A') \geq 0 \text{ for all } A' \in \mathbb{A}. \tag{A.1}$$

and, further, that if $V(A^*) = V(A')$ then $A' \notin \mathbb{E}$. Second, we will prove

$$A^* \in \mathbb{E} \Leftarrow V(A^*) - V(A') \geq 0 \text{ for all } A' \in \mathbb{A} \setminus A^*, \tag{A.2}$$

which will complete the proof.

"If"

Assume that $\mathbb{E}$ is non-empty and consider some A^*, A' such that $A^* \in \mathbb{E}, A' \in \mathbb{A}$ and $A^* \neq A'$.[1] The tuple $w = (w_M, w_S)$ that clears markets for A^* is denoted as w^*.[2]

We can write the total real wage bill of country j citizens who work in country i, under wage function w_i and supply function S_{ij} as:

$$\bar{w}_{ij}^A(w_i, S_{ij}) = \int_1^0 \frac{w_i(t)}{P_i} \, \mathrm{d}S_{ij}(t).$$

Define the weighted average of (net) real wages of all workers as:

$$\begin{aligned}\bar{w}^A(w, A) = \; & e^{-\Delta_{UM}} \left[\bar{w}_{UM}^A(w_U, S_{UM}) + \bar{w}_{UU}^A(w_U, S_{UU}) + \bar{w}_U^c F(x_U^c(S_{UU}))R_U^W\right] \\ & + \bar{w}_{MM}^A(w_M, S_{MM}) + \bar{w}_M^c C(x_{UM}^c(A), x_M^c(A)) - \delta_{UM} S_{UM}(0).\end{aligned}$$

As $S_{UM}^*, S_M^*, S_{UU}^*$ and w^* are the equilibrium supply and wage functions, respec-

[1] Here, and in the remainder of this proof, by $\neq$ we mean the negation of "equal almost everywhere".

[2] Or some selection from the set of such functions, for cases when $S_i(0) = 0$ for some $i \in \{U, M\}$.

tively, it follows from the first equilibrium condition (Definition 1) that

$$\bar{w}^A(w^*, A^*) \geq \bar{w}^A(w^*, A'). \tag{A.3}$$

Profit maximization implies that, if $R_i^{F'} > 0$, then

$$\begin{aligned} \frac{\pi_i^{E*}}{P_i} - c_i^e &= \frac{1}{P_i} \int_0^1 \max\{r_i(\mu_i(h), h) - w_i^*(\mu_i(h)), 0\} \, \mathrm{d}h - c_i^e \\ &\geq \frac{T_i(S_i', R_i^{F'}) - w_{ij}^A(w_i^*, S_{ij}') - w_{ii}^A(w_i^*, S_{ii}')}{R_i^{F'}} - c_i^e, \end{aligned} \tag{A.4}$$

where μ_i is the optimal hiring function defined in Section 2.3. Suppose that $R_U^{F*}, R_M^{F*} > 0$, the other cases are considered in footnote 3. Note that if $R_U^{F*}, R_M^{F*} > 0$, then

$$\mu_i(h) = (S_i^*)^{-1}((1-h)R_i^{F*}) \qquad \text{for } h \in [1 - S_i^*(0)/R_i^{F*}, 1], \tag{A.5}$$

whereas for $h \in [0, 1 - S_i^*(0)/R_i^{F*}]$ we have $r_i(v, h) - w_i^*(v) \leq 0$ for all $v \in [0, 1]$. This gives:

$$\frac{\pi_i^{E*}}{P_i} - c_i = \left(T_i(S_i^*, R_i^{F*}) - \bar{w}_{ij}^A(w_i^*, S_{ij}^*) - \bar{w}_{ii}^A(w_i^*, S_{ii}^*)\right) / R_i^{F*} - c_i^e. \tag{A.6}$$

Note also that $e^{-\Delta_{UM}} R_U^{F'}(\frac{\pi_U^{E*}}{P_U} - c_U^e) + R_M^{F'}(\frac{\pi_M^{E*}}{P_M} - c_M^e) \geq V(A') - \bar{w}^A(w^*, A')$. If $R_M^{F'}, R_S' > 0$ this follows directly from Equation (A.4). If $R_i^{F'} = 0$, then it follows as $T_i(S_i', R_i^{F'}) - R_i^{F'} c_i^e - w_{ij}^A(w_i^*, S_{ij}') - w_{ii}^A(w_i^*, S_{ii}') \leq 0 = R_i^{F'}(\pi_i^{E*} - c_i^e)$. Using the fact that $\pi_i^{E*} - c_i = 0$ by the definition of equilibrium, we can write

$$\begin{aligned} V(A^*) - \bar{w}^A(w^*, A^*) &= e^{-\Delta_{UM}} R_U^{F*}(\frac{\pi_U^{E*}}{P_U} - c_U^e) + R_M^{F*}(\frac{\pi_M^{E*}}{P_M} - c_M^e) \\ &= e^{-\Delta_{UM}} R_U^{F'}(\frac{\pi_U^{E*}}{P_U} - c_U^e) + R_M^{F'}(\frac{\pi_M^{E*}}{P_M} - c_M^e) \\ &\geq V(A') - \bar{w}^A(w^*, A'). \end{aligned} \tag{A.7}$$

This proves implication (A.1) by Equation (A.3).[3]

Finally, suppose that $A' \in \mathbb{E}$ and that $V(A^*) = V(A')$. If $S_i^* \neq S'$ for any i, then Equation (A.3) must hold strictly, and thus $V(A^*) > V(A')$. Hence, $S_i^* = S_i'$

[3]For $R_i^{F*} = 0$ we have by the definition of equilibrium that $\pi_i^{E*} - c_i^e \leq 0$. If $R_i^{F'} > 0$ we have that

$$\begin{aligned} 0 &= T_i(S_i^*, R_i^{F*}) - \bar{w}_{ij}^A(w_i^*, S_{ij}^*) - \bar{w}_{ii}^A(w_i^*, S_{ii}^*) - R_i^{F*} c_i^e \\ &\geq R_i^{F'}(\pi_i^{E*} - c_i^e) \geq T_i(S_i', R_i^{F'}) - w_{ij}^A(w_i^*, S_{ij}') - w_{ii}^A(w_i^*, S_{ii}') - R_i^{F'} c_i^e. \end{aligned}$$

for all i and $R_i^{F*} \neq R_i^{F'}$ for some $i \in \{U, M\}$. However, as the profit holding under allocation A is

$$\pi_i^E(A) = \int_{1-\frac{S_i(0)}{R_i^F}}^{1} \int_{1-\frac{S_i(0)}{R_i^F}}^{h} \frac{\partial}{\partial h} r_i(S_i^{-1}((1-p)R_i^F), p) \mathrm{d}p \, \mathrm{d}h + r_i(S_i^{-1}(R_i^F), 0) - w_i^c \tag{A.8}$$

and surplus increases strictly with firm type, it follows that if $R_i^{F*} \neq R_i^{F'}$ then $\pi_i^E(A^*) \neq \pi_i^E(A') = P_i c_i^e$, implying that $A^* \notin \mathbb{E}$; contradiction!

"Only If"

This part of the proof will proceed in two steps. First, we will decompose the optimization problem into inner and outer problems, derive the first-order conditions for the inner problem, and show that any maximizer of the inner problem must satisfy conditions (1), (2) and (4) of the competitive equilibrium. Second, we show that any maximizer of the outer problem needs to additionally meet condition (3), thus completing the proof.

"Inner" Problem Denote the set of all functions that meet conditions (1) and (2) of the set of feasible allocations (page 15) by $\mathbb{S}_{MM}$, and the set of all functions that meet condition (1) only by $\mathbb{S}$. Further, denote by $\mathbb{S}_{UM}(S_{MM})$ the set of functions $S_{UM} \in \mathbb{S}$ that satisfy condition (3) of set $\mathbb{A}$ for a given $S_{MM} \in \mathbb{S}_{MM}$. Note that if $x_{MM}^c < 1$, then the set $\mathbb{S}_{UM}(S_{MM})$ is a singelton, which will be denoted by $S_{UM}(S_{MM})$.

For given R_M^F, R_U^F we can then define the set $\mathbb{A}(R_U^F, R_M^F)$ of all such $S_{MM}, S_{UU} \in \mathbb{S}$ that there exists some $S_{UM} \in \mathbb{S}_{UM}(S_{MM})$ such that $(S_{UU}, S_{UM}, S_{MM}, R_U^F, R_M^F) \in \mathbb{A}$. Then the optimization problem $\max_{A \in \mathbb{A}} V(A)$ is equivalent to the optimization problem:

$$\underbrace{\max_{(R_M^F, R_U^F) \in \mathbf{R}^2_{\geq 0}}}_{\text{outer problem}} \underbrace{\max_{(S_{UU}, S_{MM}) \in \mathbb{A}(R_U^F, R_M^F)} V(S_{UU}, S_{MM}, R_U^F, R_M^F)}_{\text{inner problem}},$$

Also, trivially, if $R_i^{F'} = 0$, then

$$\begin{aligned} 0 &= T_i(S_i^*, R_i^{F*}) - \bar{w}_{ij}^A(w_i^*, S_{ij}^*) - \bar{w}_{ii}^A(w_i^*, S_{ii}^*) - R_i^{F*} c_i^e \\ &= T_i(S_i', R_i^{F'}) - w_{ij}^A(w_i^*, S_{ij}') - w_{ii}^A(w_i^*, S_{ii}') - R_i^{F'} c_i^e. \end{aligned}$$

Thus, it follows that $V(A^*) - \bar{w}^A(w^*, A^*) \geq V(A') - \bar{w}^A(w^*, A')$.

where

$$V(S_{UU}, S_{MM}, R_U^F, R_M^F) \equiv \max_{S_{UM} \in \mathbb{S}_{UM}(S_{MM})} V(S_{UU}, S_{UM}, S_{MM}, R_U^F, R_M^F)$$
$$\text{s.t. } S_U(0) \leq R_U^F.$$

Definition 1. The interior $\text{int}(\mathbb{A}(R_U^F, R_M^F))$ of set $\mathbb{A}(R_U^F, R_M^F)$ consists of all such $S_{MM}, S_{UU} \in \mathbb{A}(R_U^F, R_M^F)$ that $x_{UU}^c, x_{UM}^c, x_{MM}^c < 1$ and $S_i(0) < R_i^F$.

We will show in detail that all interior solutions of the inner problem satisfy conditions (1), (2) and (4) of the competitive equilibrium. The proof for corner (i.e., not interior) solutions is conceptually identical but requires small tweaks for each of the possible cases.

Fix $(R_M^F, R_U^F) \in \mathbf{R}^{\mathbf{2}}_{>\mathbf{0}}$ and consider a maximizer $(S_{UU}^*, S_{MM}^*) \in \text{int}(\mathbb{A}(R_U^F, R_M^F))$ of the inner problem.[4] Consider a one-parametric family of functions $S_{MM}(\cdot; t_M)$ such that (a) for each $t_M \in [0,1]$, $(S_{UU}^*, S_{MM}(t_M)) \in \text{int}(\mathbb{A}(R_U^F, R_M^F))$, and (b) there exists some t_M^* that corresponds to S_{MM}^*. It follows that

$$t_M^* \in \arg\max_{t_M} V(S_{UU}^*, S_{UM}(S_{MM}(t_M)), S_{MM}(t_M), R_U^F, R_M^F),$$

and any maximizer of the original problem has to satisfy the first-order conditions of this single-variable problem. A family $S_{MM}(\cdot; t_M)$ that satifies the conditions above can be constructed for any interior (S_{UU}^*, S_{MM}^*).[5] Further, the very same exercise can be also conducted for a family of US citizens' supply functions, $S_{UU}(\cdot; t_U)$.

Define the function

$$V(t_M; S_{UU}^*, R_U^F, R_M^F) = V(S_{UU}^*, S_{UM}(S_{MM}(t_M)), S_{MM}(t_M), R_U^F, R_M^F),$$

and analogously function $V(t_U; S_{MM}^*, R_U^F, R_M^F)$. In the remaining analysis of the inner problem we will supress (S_{UU}^*, R_U^F, R_M^F) from notation. The optimal matching function that holds under $(S_{UU}(t_U), S_{UM}(S_{MM}(t_M)), S_{MM}(t_M))$ will be denoted by $m_i(x_i; t_M) = \mu_i^{-1}(x_i; t_M)$ (see Equation (A.5)). Note that as t_M changes, the implied separation function $\psi(\cdot; t_M)$ changes as well. With this in mind, it can be

[4] Note that interior solutions exist only if $(R_M^F, R_U^F) \in \mathbf{R}^{\mathbf{2}}_{>\mathbf{0}}$.

[5] Consider a family of separation functions, such that $\psi(x_M; t_M) = \psi^*(x_M) + (x_M - x_M^{c*})^2 (x_M - x_M^{s*})^2 \left((t_M - 1)\underline{\epsilon} + t_M \bar{\epsilon}\right)$ for $x_M \in (x_M^{c*}, x_M^{s*})$, and $\psi(x_M; t_M) = \psi^*(x_M)$ otherwise. Of course, each $\psi(\cdot; t_M)$ gives raise to a supply function $S_{MM}(\cdot; t_M)$. It follows by the definition of x_M^{s*} that if $(S_{UU}^*, S_{MM}^*) \in \text{int}(\mathbb{A}(R_U^F, R_M^F))$ then there must exist small enough $\underline{\epsilon}, \bar{\epsilon} > 0$ that $(S_{UU}^*, S_{MM}(t_M)) \in \text{int}(\mathbb{A}(R_U^F, R_M^F))$ for all $t_M \in [0,1]$.

shown easily that

$$\frac{\partial}{\partial t_M} m_U(x_U) = \frac{\frac{\partial}{\partial t_M} S_{MM}(\phi(x_U))}{R_U^F}.$$

Further, note that by integrating $T_i(A)$ by substitution, and denoting $\frac{r_i(x_i,h_i)}{P_i}$ by $\bar{r}_i(x_i,h_i)$ we get that

$$T_i(A) = R_i^F \int_{m_i(x_i^c)}^1 \bar{r}_i(\mu_i(h),h)\mathrm{d}h.$$

Differentiating wrt t_M yields

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t_M} T_i(A) &= -R_i^F \bar{r}_i(x_i^c, m_i(x_i^c)) \frac{\mathrm{d}}{\mathrm{d}t_M} m_i(x_i^c) \\
&\quad + R_i^F \int_{m_i(x_i^c)}^1 m_i'(\mu_i(h)) \frac{\mathrm{d}}{\mathrm{d}t_M} m_i(\mu_i(h)) \frac{\partial}{\partial x_i} \bar{r}_i(\mu_i(h),h)\mathrm{d}h \\
\text{[by substitution]} &= \bar{r}_i(x_i^c, m_i(x_i^c)) \frac{\partial}{\partial t_M} S_i(0) + R_i^F \int_{x_i^c}^1 \frac{\mathrm{d}}{\mathrm{d}t_M} m_i(x_i) \frac{\partial}{\partial x_i} \bar{r}_i(x_i, m_i(x_i))\mathrm{d}x_i.
\end{aligned}$$

Thus it can be shown that:

$$\begin{aligned}
\frac{\partial}{\partial t_M} V &= \frac{\mathrm{d}}{\mathrm{d}t_M} \left[e^{-\Delta_{UM}} T_U(A) + T_M(A) + \bar{w}_M^c C(x_{UM}^c, x_{MM}^c) - \delta_{UM} S_{UM}(0) \right] \\
&= e^{-\Delta_{UM}} \int_{x_{MM}^c}^1 \frac{\partial}{\partial t_M} S_{MM}(x_M) \psi'(x_M) \frac{\partial}{\partial x_U} \bar{r}_U(\psi(x_M), m_U(\psi(x_M)))\mathrm{d}x_M \\
&\quad - \int_{x_{MM}^c}^1 \frac{\partial}{\partial t_M} S_{MM}(x_M) \frac{\partial}{\partial x_M} \bar{r}_M(x_M, m_M(x_M))\mathrm{d}x_M \\
&\quad + \frac{\partial}{\partial t_M} S_{UM}(0) e^{-\Delta_{UM}} \left(\int_{x_U^c}^{x_{UM}^c} \frac{\partial}{\partial x_U} \bar{r}_U(x_U, m_U(x_U))\mathrm{d}x_M + \bar{r}_U(x_U^c, m_U(x_U^c)) \right) \\
&\quad - \frac{\partial}{\partial t_M} S_{UM}(0) \delta_{UM} \\
&\quad + \bar{r}_M(x_M^c, m_M(x_M^c)) \frac{\partial}{\partial t_M} S_{MM}(0) + \bar{w}_M^c \frac{\mathrm{d}}{\mathrm{d}t_M} C(x_{UM}^c, x_{MM}^c), \qquad \text{(A.9)} \\
\frac{\partial}{\partial t_U} V &= \frac{\mathrm{d}}{\mathrm{d}t_U} e^{-\Delta_{UM}} \left[T_U(A) + w_U^c F(x_{UU}^c) R_U^W \right] \\
&= e^{-\Delta_{UM}} \frac{\partial}{\partial t_M} S_{UU}(0) \bar{r}_U(x_U^c, m_U(x_U^c)) \\
&\quad + e^{-\Delta_{UM}} \frac{\partial}{\partial t_M} S_{UU}(0) \int_{x_U^c}^{x_{UU}^c} \frac{\partial}{\partial x_U} \bar{r}_U(x_U, m_U(x_U))\mathrm{d}x_M - \bar{w}_U^c \qquad \text{(A.10)}
\end{aligned}$$

Because $V(t_i)$ is a simple, single-variable function, it follows that $\frac{\partial}{\partial t_i} V(t_i) \leq 0$ if

$t_i^* \in [0,1)$ and $\frac{\partial}{\partial t_i}V(t_i) \geq 0$ if $t_i^* \in (0,1]$. Crucially, these conditions must hold for *all* families $S_{ii}(\cdot;t_i)$ that meet conditions (a) and (b) above.

Lemma 1. For any interior maximizer (S^*_{UU}, S^*_{MM}) of the inner problem, it is the case that if $x_M \in (x^{c*}_{MM}, x^{s*}_{MM})$, then

$$o(x_M) \equiv e^{-\Delta_{UM}}\psi^*_{x_M}(x_M)\frac{\partial}{\partial x_U}\bar{r}_U(\psi^*(x_M), m^*_U(\psi^*(x_M))) - \frac{\partial}{\partial x_M}\bar{r}_M(x_M, m^*_M(x_M)) = 0. \tag{A.11}$$

Proof. Consider such family $S_{MM}(\cdot;t_M)$ that $x^c_{MM}(t_M) = x^{c*}_{MM}$, $S_{MM}(0;t_M) = S^*_{MM}(0)$, $\frac{\partial}{\partial x_M}S_{MM}(x^{c*}_{MM};t_M) = \frac{\partial}{\partial x_M}S^*_{MM}(x^{c*}_{MM})$, and $S_{MM}(x_M;t_M) = S^*_{MM}(x_M)$ for all $x_M \geq x^{s*}_{MM}$.[6] This implies that x^c_{MM}, x^c_{UM} and $S_{MM}(0)$ do not change with t_M, and thus neither does $S_{UM}(0)$. It follows that Equation (A.9) reduces to

$$\frac{\partial}{\partial t_M}V(t_M) = e^{-\Delta_{UM}}\int_{x^c_{MM}}^{x^s_{MM}}\frac{\partial}{\partial t_M}S_{MM}(x_M)\psi'(x_M)\frac{\partial}{\partial x_U}\bar{r}_U(\psi(x_M), m_U(\psi(x_M)))\mathrm{d}x_M.$$
$$-\int_{x^c_{MM}}^{x^s_{MM}}\frac{\partial}{\partial t_M}S_{MM}(x_M)\frac{\partial}{\partial x_M}\bar{r}_M(x_M, m_M(x_M))\mathrm{d}x_M \tag{A.12}$$

Suppose that there exists some $x_M \in (x^{c*}_{MM}, x^{s*}_{MM})$ such that $o(x_M) \neq 0$. Note that because $\psi(\cdot)$ is continuously differentiable, $o(\cdot)$ is continuous. This implies that there exists some $\delta > 0$ and some $\bar{x}_M$ such that $o^*(x_M) \neq 0$ for all $x_M \in [\bar{x}_M - \delta, \bar{x}_M + \delta]$. We can always construct a feasible family $S_{MM}(\cdot;t_M)$ such that for all $t''_M > t'_M$, $\mathrm{sgn}(S_{MM}(x_M;t'_M) - S_{MM}(x_M;t''_M)) = \mathrm{sgn}(o(x_M))$ and $t^*_M \in (0,1)$.[7] For such a family (a) $\frac{\partial}{\partial t_M}V(t^*_M) = 0$, and (b) $\frac{\partial}{\partial t_M}V(t^*_M) \neq 0$ by Equation (A.12); contradiction! $\square$

We are now ready to show that for any interior (S^*_{UU}, S^*_{MM}) there exists a pair of wage functions (w_U, w_M) which together with (S^*_{UU}, S^*_{MM}) satisfy conditions (1), (2) and (4) of the equilibrium. It follows from the discussion in Section 2

[6] $S_{MM}(\cdot;t_M) = S^*_{MM}(\cdot)$ meets these conditions trivially.

[7] That is, $S_{MM}(x_M;t''_M) - S_{MM}(x_M;t'_M)) = 0$ only if $o(x_M) = 0$, and if $o(x_M) \neq 0$ then $S_{MM}(x_M;t''_M) - S_{MM}(x_M;t'_M))$ has the same sign. To construct such a family, consider any interval $[x'_M, x''_M] \subset [x^{c*}_{MM}, x^{s*}_{MM}]$ such that $o(x'_M) = o(x''_M) = 0$ and $o(x_M) \neq 0$ and is of the same sign for all $x_M \in [x'_M, x''_M]$. Then let $\psi(x_M;t_M)$ solve

$$\frac{\partial}{\partial x_M}C(\psi(x_M,t_M), x_M) = \mathrm{sgn}(o(x_M))(x_M - \frac{x'_M + x''_M}{2})^3(x_M - \bar{x}'_M)^2(x_M - \bar{x}''_M)^2\left((t_M - 1)\underline{\epsilon} + t_M\bar{\epsilon}\right)$$
$$-\frac{\partial_+}{\partial x_M}S^*_{MM}(x_M)$$

for some positive but very small $\underline{\epsilon}, \bar{\epsilon}$. If x_M belongs to an interval on which $o(x_M) = 0$, then set $\psi(x_M;t_M) = \psi^*(x(M))$.

that the wage functions w_U, w_M for which conditions (2) and (4) of equilibrium are satisfied, are given by Equation (18), where $w_M^c(x_M^c) = w_M^c$ and $w_U(x_U^c) = \min\{w_U^c, e^{\Delta_{UM}} P_U (w_M^c/P_M + \delta_{UM})\}$. For condition (1) to be satisfied, it must be the case that these w_U, w_M satisfy (i) Equation (19) as well as (ii) $w_U(x_{UU}^{c*}) = w_U^c$.

First, consider $\frac{\partial}{\partial t_U} V(t_U)$. It follows immediatelly from Equation (A.10) that

$$r_U(x_U^{c*}, m_U(x_U^{c*})) + \int_{x_U^{c*}}^{x_{UU}^{c*}} \frac{\partial}{\partial x_U} r_U(x_U, m_U(x_U)) \mathrm{d}x_U = w_U^c. \tag{A.13}$$

Let us turn attention to $\frac{\partial}{\partial t_M} V(t_M)$ and consider such family $S_{MM}(\cdot; t_M)$ that $S_{MM}(x_{MM}^{c*}; t_M) = S_{MM}^*(x_{MM}^{c*})$, $x_{UM}^c(t_M) = x_{UM}^{c*}$ and $S_{MM}(x_M; t_M) = S_{MM}^*(x_M)$ for all $x_M \geq x_{MM}^{s*}$. This implies that (a) $t_M^* \in (0,1)$ (b) that $\frac{\mathrm{d}}{\mathrm{d}t_M} S_{UM}(0; t_M^*) = 0$ and (c) that $\frac{\mathrm{d}}{\mathrm{d}t_M} S_{MM}(0) = \frac{\partial}{\partial t_M} x_{MM}^c(t_M) \frac{\partial}{\partial x_M} C(x_{MM}^c(t_M), x_{MU}^{c^*})$. Substituting this and Equation (A.11) into Equation (A.9) yields

$$r_M(x_M^{c*}, m_M(x_M^{c*})) = w_M^c. \tag{A.14}$$

Consider such family $S_{MM}(\cdot; t_M)$ that $S_{UM}(0; t_M) \neq S_{UM}^*(0)$ and $S_{MM}(x_M; t_M) = S_{MM}^*(x_M)$ for all $x_M \geq x_{MM}^{s*}$. Then subsituting Equations (A.11) and (A.14) into Equation (A.9) yields

$$\begin{aligned} \frac{\partial}{\partial t_M^*} V(t_M^*) = e^{-\Delta_{UM}} \frac{\partial}{\partial t_M^*} S_{UM}(0) \left[\int_{x_U^{c*}}^{x_{UM}^{c*}} \frac{\partial}{\partial x_U} \bar{r}_U(x_U, m_U(x_U)) \mathrm{d}x_M + \bar{r}_U(x_U^{c*}, m_U(x_U^{c*})) \right] \\ - \frac{\partial}{\partial t_M^*} S_{UM}(0) \left(\bar{w}_M^c + \delta_{UM} \right) = 0 \end{aligned} \tag{A.15}$$

Substituting Equation (A.13) into (A.15) yields

$$\frac{P_U e^{\Delta_{UM}}}{P_M} \left(w_M^c + P_M \delta_{UM} \right) - w_U^c = w(x_{UM}^{c*}) - w(x_{UU}^{c*}),$$

which implies that $x_{UM}^{c*} \geq x_{UU}^{c*}$ if and only if $P_U e^{\Delta_{UM}} \left(w_M^c / P_M + \delta_{UM} \right) \geq w_U^c$.

Suppose that $x_{UM}^{c*} \geq x_{UU}^{c*}$. Then $w_U(x_U^{c*}) = P_U w_U^c$ and condition (ii) follows immediately. Further, Equation (A.13) reduces to $r_U(x_U^{c*}, m_U(x_U^{c*})) = w_U^c$. Substituting this into Equation (A.15) ensures that

$$w_U(x_{UM}^{c*}) = P_U e^{\Delta_{UM}} \left(w_M(x_{MM}^{c*}) / P_M + \delta_{UM} \right).$$

As Equation (A.11) is the same as the first derivative of Equation (20), it follows that condition (ii) must be satisfied as well.

Suppose that $x^{c*}_{UM} < x^{c*}_{UU}$. Then $w_U(x^{c*}_{UM}) = P_U e^{\Delta_{UM}}(w_M(x^{c*}_{MM})/P_M + \delta_{UM})$, which reduces Equation (A.15) to $r_U(x^{c*}_U, m_U(x^{c*}_U)) = w^c_U$. Substituting this into Equation (A.15) ensures that $w_U(x^{c*}_{UU}) = w^c_U$. Again, Equation (A.11) is the same as the first derivative of Equation (20), which together with $w_U(x^{c*}_{UM}) = P_U e^{\Delta_{UM}}(w_M(x^{c*}_{MM})/P_M + \delta_{UM})$ ensures that condition (ii) must be satisfied as well.

"Outer" Problem The proof that the maximizers of the outer problem satisfy condition (3) of the equilibrium, follows the logic of the proof of Lemma 20 in Gola (2018). Consider some maximizer (R^{F*}_U, R^{F*}_M) of the outer problem and some $R^{F'}_i$. Define the function $R^F_i(t_R) = t_R R^{F*}_i + (1 - t_R) R^{F'}_i$. Note that

$$\bar{T}_i(A) \equiv \int_1^0 \bar{r}_i\Big(x_i, \max\{1 - S_i(x_i)/R^F_i, 0\}\Big)\,\mathrm{d}S_i(x_i) = T_i(A) \qquad \text{if } R^F_i > 0$$

which allows us to drop condition (4) from the definition of the set of feasible allocations as $\mathbb{A}$. Denote this modified set of feasible allocations by $\bar{\mathbb{A}}$, and by $\bar{V}$ the the total weighted net revenue function, in which T_i has been replaced by $\bar{T}_i$. Then define

$$V^I(S_{UU}, S_{MM}, t_R) = \max_{S_{UM} \in \mathbb{S}_{UM}(S_{MM})} \bar{V}(S_{UU}, S_{UM}, S_{MM}, R^F_U(t_R), R^F_M(t_R)).$$

It is easy to show that $V^I(S_{UU}, S_{MM}, t_R)$ is differentiable for all t_R but at most 4 ($t_R \in \{0, 1\}$ and $R^F_i(t_R) = S_i(0)$), and also that whenever $V^I_t(S_U, S_M, t)$ does exist we have that

$$\begin{aligned} V^I_t(S_M, S_S, t) = (R^{F*}_U - R^{F'}_U)(\frac{1}{P_U}\pi^E_U(S_U, R_U(t)) - c^e_U) \\ +(R^{F*}_M - R^{F'}_M)(\frac{1}{P_M}\pi^E_M(S_S, R_S(t)) - c^e_M), \end{aligned}$$

where

$$\pi^E_i(S_i, R^F_i) = \begin{cases} \int_0^1 \int_0^h \frac{\partial}{\partial h} r_i(S_i^{-1}((1-p)R^F_i), p)\mathrm{d}p + r_i(S_i^{-1}(R^F_i), 0)\,\mathrm{d}h & \text{for } R_i \in (0, S_i(0)), \\ \int_{1-\frac{S_i(0)}{R^F_i}}^1 \int_{1-\frac{S_i(0)}{R^F_i}}^h \frac{\partial}{\partial h} r_i(S_i^{-1}((1-p)R^F_i), p)\mathrm{d}p\,\mathrm{d}h & \text{for } R^F_i > S_i(0). \end{cases} \tag{A.16}$$

Thus $V(S_{UU}, S_{MM}, \cdot)$ is absolutely continuous for any $(S_{UU}, S_{MM}) \in \bar{\mathbb{A}}$ and any choice of $R_i^{F'}$. Clearly, $\frac{1}{P_i}\pi_i^E(S_i, R_i^F(t)) - c_i^e \in [-c_i, \bar{r}_i(1,1) - c_i^e]$, implying

$$|V_t(S_{UU}, S_{MM}, t)| \leq (R_U^{F*} - R_U^{F'}) \max\{c_U^e, \bar{r}_U(1,1)\} + (R_M^{F*} - R_M^{F'}) \max\{c_M^e, \bar{r}_M(1,1)\}$$

which proves

$$V(t) \equiv \max_{(S_{UU}, S_{MM}) \in \mathbb{A}} V^I((S_{UU}, S_{MM}, t_R)$$

is absolutely continuous by Theorem 2 in Milgrom and Segal (2002).

Define $S_{UU}(t), S_{MM}(t) \in \arg\max_{(S_{UU}, S_{MM}) \in \mathbb{A}} V^I((S_{UU}, S_{MM}, t_R), T(t_R) \equiv V(t) - c_U^e R_U^F(t_R) - c_M^e R_M^F(t_R)$ and pick any $t \in (0,1)$ for which $V(t)$ is differentiable. Consider two $\tilde{c}_U^e, \tilde{c}_M^e \in \mathbf{R}_{\geq 0}$ such that $\tilde{c}_i^e = \pi_i^E(R_U^F(t), R_M^F(t))$. For entry costs $\tilde{c}_M^e, \tilde{c}_S^e$, the allocation $A(t) = (S_{UU}(t), S_{UM}(S_{MM}(t)), S_{MM}(t), R_U^F(t), R_M^F(t))$ is a partial labor market equilibrium, implying that it maximizes the function $\tilde{V}(t_R) = T(t_R) - \tilde{c}_U^e R_U^F(t) - \tilde{c}_M^e R_M^F(t)$. Clearly, both $\tilde{V}(\cdot)$ and $T(\cdot)$ are differentiable at t_R as well. It follows from first-order conditions that $\tilde{V}_{t_R}(t_R) = 0$ implying that

$$\begin{aligned} T_{t_R}(t_R) &= (R_U^{F*} - R_U^{F'})\tilde{c}_U^e + (R_M^{F*} - R_M^{F'})\tilde{c}_M^e \\ &= (R_U^{F*} - R_U^{F'})\pi_U^E(R_U^F(t), R_M^F(t)) + (R_M^{F*} - R_M^{F'})\pi_M^E(R_U^F(t), R_M^F(t)). \end{aligned}$$

This proves that

$$V_t(t) = (R_U^{F*} - R_U^{F'})(\pi_U^E(R_U^F(t), R_M^F(t)) - c_U^e) + (R_M^{F*} - R_M^{F'})(\pi_M^E(R_U^F(t), R_M^F(t)) - c_M^e).$$

Note, by the way, that because we can induce an equilibrium for any values of (R_U^F, R_M^F) by an appropriate choice of $(\tilde{c}_U^e, \tilde{c}_M^e)$, it follows from the "if" part of this proof, that the set $\arg\max_{(S_{UU}, S_{MM}) \in \mathbb{A}} V^I(S_{UU}, S_{MM}, t_R)$ is a singleton.

Now, let us show that if $R_U^{F*} > 0$ then $\pi_M^E \geq c_M^e P_M$. First, pick some $R_M^{F'} < R_M^{F*}$ and define $V(t)$ for (R_U^{F*}, R_M^{F*}) and $(R_U^{F*}, R_M^{F'})$. From the definition of maximum follows that there exists some $t'_R \in (0,1)$ such that for any $t_R > t'_R$ we have $\pi_M^E(R_U^F(t_R), R_M^F(t_R)) \geq c_M^e P_M$. Recall that for any allocation $A(t)$ the average profit of firms in country i is given by Equation (A.8) it follows from continuity of $(S_{UU}(t), S_{MM}(t))$ that $\pi_M^E(R_U^{F*}, R_M^{F*}) \geq c_M^e P_M$.[8] It remains to show that if $R_U^{F*} \geq 0$ then $\pi_M^E \leq c_M^e P_M$, but the proof is completely analogous, because $\pi_i(t)$ is continuous even for $R_M^F = 0$, in the sense that the limit of the average

[8]It follows from Berge's (1963) maximum theorem that the correspondence $\mathbf{S}(t_R) \equiv \arg\max_{(S_{UU}, S_{MM}) \in \mathbb{A}} V^I((S_{UU}, S_{MM}, t_R)$ is upper-hemicontinuous. However, as this correspondence is singleton valued, this implies that it is continuous.

profit that holds for $R_M^F > 0$ as $R_M^F \to 0$ is an equilibrium for $R_M^F = 0$.[9] The proof for U.S. is analogous.

Proof of Theorem 1

Consider the set $\bar{\mathbb{A}}(R_U^F, R_M^F)$ of all functions S_{UU}, S_{UM}, S_{MM} that meet conditions (1)–(4) on page 15 given (R_U^F, R_M^F). As all functions in $\bar{\mathbb{A}}(R_U^F, R_M^F)$ are absolutely continuous, differentiable almost everywhere and their derivative lies in $[-1, 0]$, it follows that they are Lipschitz continuous with the same Lipschitz constant. Hence, by the Arzela-Ascoli theorem $\bar{\mathbb{A}}(R_U, R_M)$ is compact. Therefore, it follows from the Extreme Value theorem that the set

$$V(R_U^F, R_M^F) \equiv \underset{(S_{UM}, S_{UM}, S_{MM}) \in \bar{\mathbb{A}}(R_U^F, R_M^F)}{\arg\max} V(S_{UU}, S_{UM}, S_{MM}, R_U^F, R_M^F)$$

is non-empty. We have shown on page 10 that $V(R_U^F, R_M^F)$ is a singleton, and in footnote 8 that it is continuous in R_U^F, R_M^F. Thus, employing the same logic as in the proof of Theorem 2 in Gola (2018) it can be easily shown that there exists a compact set $\bar{R} \in \mathbf{R^2}_{\geq 0}$ such that:

$$\max_{\bar{R}} V(R_U, R_M) = \max_{\mathbf{R^2}_{\geq 0}} V(R_U, R_M).$$

It follows from the Extreme Value theorem that $\arg\max_{\mathbf{A}} V(A)$ is non-empty. It follows trivially from Proposition 1 that the equilibrium exists and is unique.

Proof of Theorem 2

Define a map $\mathscr{F}(\mathbf{P}) \equiv Y_W - p_W^{1-\varepsilon} \sum_k Y_k P_k^{\varepsilon-1} \tau_{kW}^{1-\varepsilon}$. Because $p_w q_W^A = Y_W$ Equation (6) for $i = W$ can be rewritten as

$$\mathscr{F}(\mathbf{P}) = 0.$$

[9] This is because the equilibrium wage function that holds in the non-degenerate country (U.S.) is trivially continuous in R_M^F, and the services wage function determines the lowest wage function in Mexico that prevents any worker from remaining in that country. A similar reasoning holds even if both countries are degenerate.

Substituting this into Equation (25) results in

$$P_i = \left[\frac{(\tau_{iU})^{1-\varepsilon} Y_U}{a\mathscr{F}(\mathbf{P}) + \sum_k Y_k \tau_{kU}^{1-\varepsilon} P_k^{\varepsilon-1}} + \frac{(\tau_{iM})^{1-\varepsilon} Y_M}{a\mathscr{F}(\mathbf{P}) + \sum_k Y_k \tau_{kM}^{1-\varepsilon} P_k^{\varepsilon-1}} + (\tau_{iW} p_W)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}, \tag{A.17}$$

where $a \in (0, \min_{i \in \{U,M,W\}, j \in \{U,M\}} \{(\frac{\tau_{ji}}{\tau_{iW}})^{1-\varepsilon}\})$. It is easy to show that any vector $\mathbf{P} = (P_U, P_M, P_W)$ that solves the system of three Equations given by (A.17) must also satisfy Equation (6).[10] Therefore, it follows trivially that any such $\mathbf{P}$ solves also the system given by (25).

Lemma 2. Consider the set $\mathbb{P}$ of all $\mathbf{P} \in \mathbf{R^3}_{>0}$ that solve Equation (A.17) for all $i \in \{U, M, W\}$. The set $\mathbb{P}$ is non-empty.

Proof. Consider the interval $I_i = [\left[\frac{(\tau_{iU})^{1-\varepsilon} Y_U}{a\mathscr{F}(\mathbf{P})} + \frac{(\tau_{iM})^{1-\varepsilon} Y_M}{a\mathscr{F}(\mathbf{P})} + (\tau_{iW} p_W)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}, \tau_{iW} p_W]$ and define the map $\mathscr{T} : I_U \times I_M \times I_W \to I_U \times I_M \times I_W$ such that

$$\mathscr{T}_i(P) \equiv \left[\frac{(\tau_{iU})^{1-\varepsilon} Y_U}{a\mathscr{F}(\mathbf{P}) + \sum_k Y_k \tau_{kU}^{1-\varepsilon} P_k^{\varepsilon-1}} + \frac{(\tau_{iM})^{1-\varepsilon} Y_M}{a\mathscr{F}(\mathbf{P}) + \sum_k Y_k \tau_{kM}^{1-\varepsilon} P_k^{\varepsilon-1}} + (\tau_{iW} p_W)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}. \tag{A.18}$$[11]

Clearly, $\mathscr{T}$ is increasing. Thus, the Lemma follows from Tarski's (1955) fixed point theorem.

$\square$

The largest vector of price indexes solving Equation (A.17) for $Y_U, Y_M, Y_W \in \mathbb{R}^3_{\geq 0}$ is denoted by $\bar{\mathbf{P}}(Y_U, Y_M, Y_W)$, and continuous in all arguments. Define the map $\mathscr{B}_i : \mathbb{R}^3_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\mathscr{B}_i(\mathbf{Y}) = \sum_{k \in \{U,M,W\}} \frac{Y_k \tau_{ki}^{1-\varepsilon}}{\bar{P}_k(\mathbf{Y})^{1-\varepsilon}}$. Note that $\mathscr{B}_i$ is homogenous of degree 1 (because $\bar{P}(\cdot)$ is homogeneous of degree zero) and increasing. Therefore, $\max_{\mathbf{Y} \leq \lambda \mathbf{1}} \mathscr{B}_i(\mathbf{Y}) = \lambda \mathscr{B}_i(\mathbf{1})$.

Denote by $A(\mathbf{Y})$ the allocation that holds in the equilibrium of the partial labor equilibrium under price index vector $\bar{\mathbf{P}}(\mathbf{Y})$ and expenditure vector $\mathbf{Y}$. Then

[10] Multiplying voth sides of (A.17) by $P_i^{1-\varepsilon} Y_i$, summing by i and rearranging results in

$$\mathscr{F}(\mathbf{P}) \left[1 + \frac{Y_U}{a\mathscr{F}(\mathbf{P}) + \sum_k Y_k \tau_{kU}^{1-\varepsilon} P_k^{\varepsilon-1}} + \frac{Y_M}{a\mathscr{F}(\mathbf{P}) + \sum_k Y_k \tau_{kM}^{1-\varepsilon} P_k^{\varepsilon-1}} \right] = 0,$$

which implies that $\mathscr{F}(\mathbf{P}) = 0$.

[11] To see that $\mathscr{T}_i$ always maps into I_i, first note that it is increasing in P for all $P \geq 0$, and that $\mathscr{T}(0,0,0)$ is equal to the lower bound of I_i. Secondly, $Y_i \left(a\mathscr{F}(\mathbf{P}) + \sum_{k \in \{U,M,W\}} \frac{Y_k \tau_{ik}^{1-\varepsilon}}{P_k^{1-\varepsilon}} \right)^{-1}$ is always positive, implying that $\mathscr{T}_i(\mathbf{P}) \leq \tau^{Wi} p_W$.

we can define the map $\mathscr{K} : \mathbb{R}^3_{\geq 0} \to \mathbb{R}^3_{\geq 0}$ such that

$$\mathscr{K}_i \equiv \begin{cases} \mathscr{B}_i(\mathbf{Y})^{\frac{1}{\varepsilon}} \int_1^0 f_i(x, 1 - S_i(x;\mathbf{Y})/R_i^F(\mathbf{Y}))^{\frac{\varepsilon-1}{\varepsilon}} \, \mathrm{d}S_i(x;\mathbf{Y}) & \text{if } i \in \{U, M\}, \\ p_W q_W & \text{if } i = W. \end{cases}$$

Any fixed point of this map characterizes a general equilibrium of this model.

For $i \in \{U, M\}$ denote $\int_1^0 f_i(x, 1 - S_i(x)/R_i^F)^{\frac{\varepsilon-1}{\varepsilon}} \, \mathrm{d}S_i(x)$ by $Q_i(S_i)$. Then $\bar{Q}_i = \max_{S_i \in \mathbf{S}_i} Q_i(S_i)$, where $\mathbf{S}_i$ is the set of all feasible supply functions in country i. Set

$$\lambda = \max\{\max_{i \in \{U,M\}} [\bar{Q}_i^{\varepsilon} \mathscr{B}_i(\mathbf{1})]^{\frac{1}{\varepsilon-1}}, p_W q_W\}.$$

Thus if $\mathbf{Y} \leq \lambda$ then $\mathscr{K}_i(\mathbf{Y}) \leq \lambda$.[12] It follows that we can define a restriction $\mathscr{K}^R : [0, \lambda]^3 \to [0, \lambda]^3$ of map $\mathscr{K}$. $\mathscr{K}^R$ must have a fixed point by Brouwer's fixed-point theorem, and – therefore – so does $\mathscr{K}$.[13] This concludes the existence proof.

It can be easily shown that the equilibrium must be unique if $\tau_{ij} = 1$ for all $i, j \in \{U, M, W\}$. First note that then Equation (25) is solved uniquely by $P_U = P_M = P_W$. This further implies that $\mathscr{B}_U(\mathbf{Y}) = \mathscr{B}_M(\mathbf{Y}) = \mathscr{B}_W(\mathbf{Y})$. Clearly, we can re-normalize the revenue functions in each country by dividing them all by $\mathscr{B}_U$ without any impact on the supply functions holding in the labor market equilibrium. Thus the supply and demand functions holding in equilibrium are unique by Theorem 1. From this follows trivially that Y_i/Y_j is given uniquely for any $i, j \in \{U, M, W\}$. However, as $Y_W = p_W q_W$, this determines Y_M, Y_U uniquely.

[12] Trivially for $i = W$. For $i \in \{U, M\}$ we have $\mathscr{K}_i(\mathbf{Y}) \leq [\bar{Q}_i \mathscr{B}_i(\mathbf{1})]^{\frac{1}{\varepsilon}} \lambda^{\frac{1}{\varepsilon}} \leq \lambda$.

[13] $\mathscr{K}$ is absolutely continuous by the same reasnoning as that in the proof of Theorem 1.

Appendix B Copula Functions

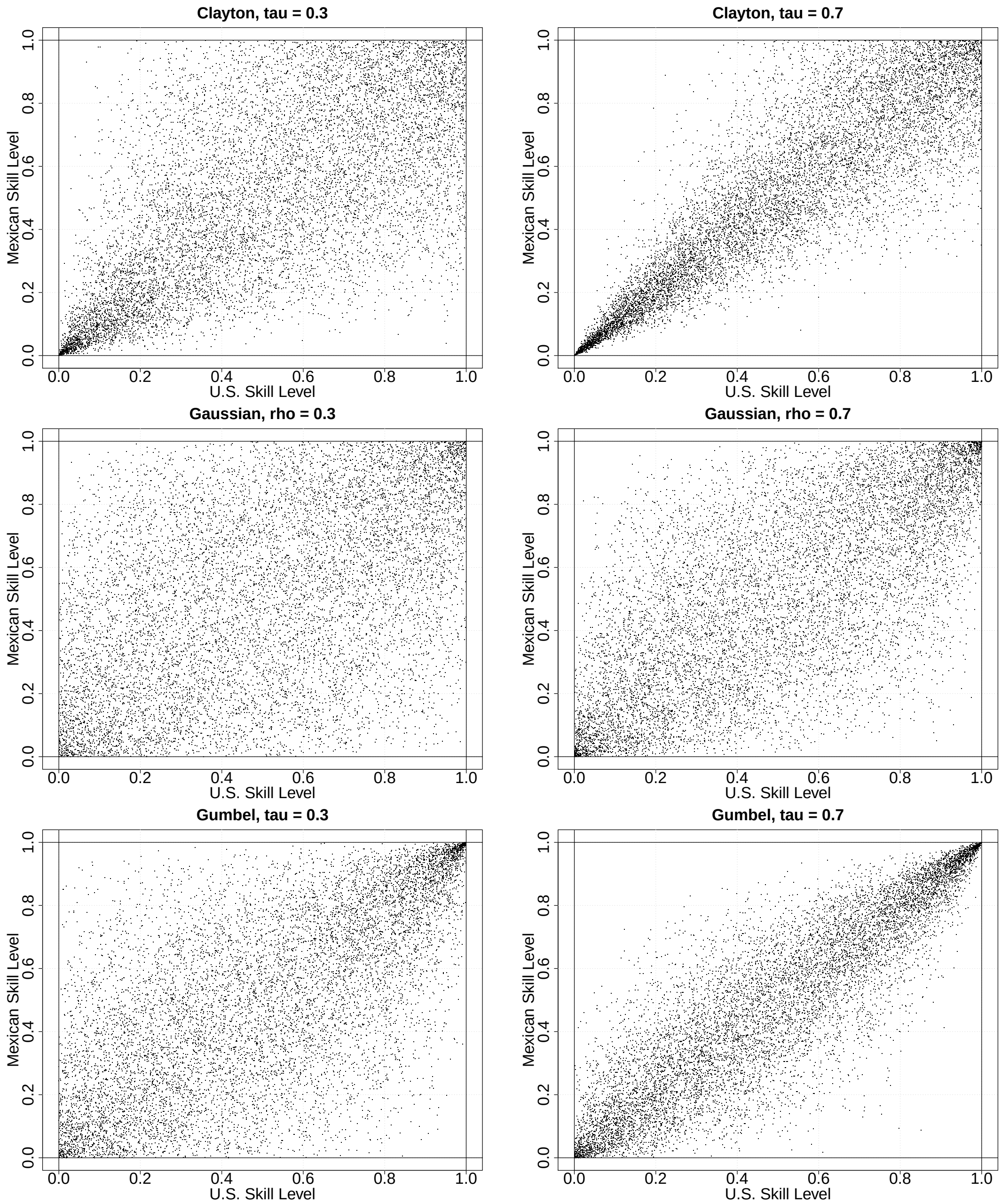


Figure B.1: Two-dimensional Distributions for Clayton, Gaussian and Gumbel Copulas

Note: Figure B.1 presents the distributions of skills assuming different copula functions (row 1: Clayton, row 2: Gaussian, row 3: Gumbel), and low (column 1) and high (column 2) correlations between skills.

Appendix C Calibration Details

Fiscal module For the fiscal part, we collect the data on income and corporate tax rates and thresholds for the United States and Mexico from the OECD, see Table C.1. We match the actual income structure in the United States quite well by calculating that corporate taxes constitute approximately 20.5 percent of tax revenues, while the summary by the National Priorities Project (provided using data from the Office of Management and Budget) reveal this figure to be close to 18.8 percent.[14] The remainder of tax revenues is financed by personal income taxes (in particular, our calibration reveals that 3 percent of total U.S. tax revenues originate from Mexican immigrants). Thirty percent of all U.S. government spending is related to Social Security, while 16 percent is related to military and security. Thus, we broadly assume that 50 percent of all transfers are lump sum payments, identical for all workers in the United States.[15] The rest, which is approximately 50 percent of governmental expenditures (including Medicare, education, and science), are treated as proportional to incomes. We further assume that the margin of adjustment is incorporated in the lump-sum transfers, taking the tax schedule and wage-proportional transfers as constants in the counterfactual scenario. In Mexico, we assume a similar structure of benefits.

Solution of the model For a given vector of parameters Ξ, the solution algorithm starts with exploiting the distribution of U.S. citizens' wages – the only one that is not affected by the selection mechanism.[16] Using Equation (17), we arrive at the following differential equation:

$$\frac{\partial}{\partial x_U} w_U(x_U) = \frac{\partial}{\partial x_U} \hat{w}_U(F(x_U)) \leftrightarrow \frac{\partial}{\partial x_U} r_U(x_U, h_U(x_U)) = \hat{w}'_U(F(x_U))F'(x_U), \tag{C.1}$$

where the left hand side function is the derivative of the surplus with respect to its first argument (skill ranking x_U), while the right hand side function is the observed inverse distribution of wages $\hat{w}_U(\cdot)$ being a function of the distribution

[14] For more information on the structure of inflows and outflows to/from the U.S. governmental budget, please consult *https://www.nationalpriorities.org/budget-basics/federal-budget-101/spending/*. Note that we are concerned only with the two types of taxes, individual income and corporate, that together constitute 89 percent of all U.S. taxes collected and 49 percent of total governmental revenues.

[15] This set of assumptions ignores the fact that immigrants might have a different net fiscal impact through a lower probability of collecting welfare benefits. In this sense, our estimates of the fiscal effects from immigration would be the lower bound of the real estimates.

[16] Recall that $\Xi = \{k_U, s_U, \gamma_U, k_M, s_M, \gamma_M, \theta, \delta_{UM}, \Delta_{UM}\}$, and for $i \in \{U, M\}$: $k_i, s_i, \gamma_i, \theta, \Delta_{UM} > 0$; $\delta_{UM} \in \mathbb{R}$.

Table C.1: Income tax schedules in the United States and Mexico

US Rate	Threshold	MEX Rate	Threshold
10%	9,225	1.92%	334.1
15%	37,450	6.4%	2,835
25%	90,750	10.88 %	4,983
28%	189,300	16%	5,792
33%	411,500	17.92%	6,935
35%	413,200	21.36%	13,987
39.6%	-	23.55%	22,046
		30%	42,090
		32%	56,121
		34%	168,363
		35%	-
Corporate Rate (flat):			
35%		30%	

Note: Thresholds are in 2015 USD. source: OECD.

of American skills $F(\cdot)$, multiplied by the density of skills supplied by Americans: $F'(x_U)$. Equation (C.1) is the first equation in the system of two differential equations, and is solved with an initial condition: $\hat{w}_U(1) = w_U(1)$. The solution is discretized on the assumed grid, and computed using the Euler method.[17]

The second step is to reveal the underlying selection mechanism induced by a tuple: $\{\Xi, F(\cdot)\}$. We therefore proceed with exhausting the migration condition (20), and taking its first derivative:

$$\frac{\partial}{\partial x_U}\bar{w}_M(\phi(x_U)) = e^{-\Delta_{UM}}\frac{\partial}{\partial x_U}\bar{w}_U(x_U) \leftrightarrow$$

$$\frac{\partial}{\partial x_U} r_M(\phi(x_U), h_M(\phi(x_U)))\phi(x_U)' = \frac{P_M}{P_U} e^{-\Delta_{UM}} \frac{\partial}{\partial x_U} r_U(x_U, h_U(x_U)). \quad \text{(C.2)}$$

The latter serves as the second equation in the two-dimensional system, solved simultaneously with Equation (C.1), using the Euler method on the assumed grid, and taking the initial condition: $\phi(1) = 1$.[18] For the given solution for selection

[17] Euler method is the simplest numerical way to solve an ordinary differential equation (ODE) with a given initial condition. For a given ODE: $y'(x) = f(x)$, $y(1) = f(1)$, and a given series of grid points: $\{x(1), ..., x(K)\}$, one computes the values of y by setting: $y(x(t)) = y(x(t-1)) + (x(t) - x(t-1))f(x(t-1))$.

[18] Our model approximates the model with unbounded, log-normally distributed skills, in which $\phi(1) = 1$ (this means that $\forall x_U \leq 1\, \exists x_M : (x_U, x_M)$ stays in Mexico). Thus, setting $\phi(1) = 1$ amounts to imposing a condition that, in this dimension at least, we consider only specifications

pattern, determined by the separation function $\phi(\cdot)$, the mass of Mexican immigrants in the U.S. can be computed by using Equation (21), discretized in the following way:

$$S_{UM}(x_U - dx_U) = S_{UM}(x_U) + dx_U \partial C(x_U, \phi(x_U))/\partial x_U, \tag{C.3}$$

for all rankings x_U ranging from 1 down to x^c_{UM}, with step $dx_U = 1/K$. The starting point requires that: $S_{UM}(1) = 0$.

At this stage, we can use the Euler discretization of country-specific Equations (17) to determine the wage distributions of Mexican workers in the U.S. and in Mexico. The final result of the calibration for a given vector of parameters Ξ is a set of three wage distributions: U.S. citizens, $w_U \equiv (w_U(x_U), F(x_U))$, Mexican immigrants in the U.S., $w_{UM} \equiv (w_U(x_U), F_{UM}(x_U))$, and Mexican stayers, $w_M \equiv (w_M(x_M), F_M(x_M))$.[19]

Calibration algorithm Our goal in the calibration procedure is to find such a vector of parameters Ξ that gives the best possible fit of w_U, w_{UM} and w_M to the observed distributions $\hat{w}_U$, $\hat{w}_{UM}$ and $\hat{w}_M$, along with perfectly matching identifying moments in the data (see subsection *Identification* for the correspondence between model parameters and empirical moments that establish identification of the model). The solution of the model requires finding functions and distributions, therefore in performing the calibration we cannot escape solving the model for each proposed vector of parameters Ξ.

The calibration procedure assumes a search through a $\dim \Xi = 9$ dimensional space of parameters, and each vector requires a full solution of the model on the defined grid. To maximize the performance of such a computationally-intensive search, we propose a version of a basing-hopping algorithm, enriched with a Monte Carlo search procedure, with a given goal function.[20] Our implementation of

that retain this important feature of the model with (untruncated) log-normally distributed skills.

[19]The proposed notation includes the relative densities of the analyzed groups of workers. $F_{UM}(x_U) = (S_{UM}(x^c_{UM}) - S_{UM}(x_U))/S_{UM}(x^c_{UM})$, while: $F_M(x_M) = (S_M(x^c_M) - S_M(x_M))/S_M(x^c_M)$.

[20]Standard, one-dimensional selection models can be calibrated using a Maximum Likelihood Estimation (MLE). In the case of our model this is not feasible because the selection patterns cannot be solved for analytically. This means that we are unable to obtain closed form solutions for the distributions of wages, which makes it impossible to use a standard MLE algorithm. Instead, we set the model parameters to match the full distributions of the three groups of workers that we observe. This method is computationally less demanding, but arrives at a similar outcome: a MLE of Ξ would aim at equalizing the model distribution of wages to the observed ones, so that the probability of selecting an individual from a given wage distribution

the random search through the parameter space is in principle a variant of the Simulated Annealing Optimization method (being a member of the Metropolis algorithms family), used extensively in engineering, physics and chemistry.

Each vector Ξ is evaluated using a subjective goal function:[21]

$$\begin{aligned}\zeta(\Xi) = & \, p_1|c_U^f - \hat{c}_U^f| + p_2|c_M^f - \hat{c}_M^f| + p_3|w_M^{min} - \hat{w}_M^{min}| + p_4|w_M^{max} - \hat{w}_M^{max}| \\ & + p_5|w_{UM}^{min} - \hat{w}_{UM}^{min}| + p_6|w_U^{share} - \hat{w}_U^{share}| + p_7|w_M^{share} - \hat{w}_M^{share}| + p_8 e(P - \hat{P}) \\ & + p_9|S_{UM}(0) - \hat{S}_{UM}(0)| + p_{10}e(w_U) + p_{11}e(w_{UM}) + p_{12}e(w_M),\end{aligned} \tag{C.4}$$

where $e(\cdot)$ is an error function that computes the squared difference between an object from the model and its empirical counterparty in the data, and p's are subjective weights.[22] The $P(\cdot)$ function computes the conditional probabilities of emigration from Mexico (for which we construct empirical counterpart using the data from the Mexican Migration Project), while functions $w_i(\cdot)$ represent the group-specific distributions of wages. The goal function aims at minimizing: (i) the distance between eight model parameters and corresponding moments in the data (multiplied by weights: $p_1, ..., p_8$); (ii) the absolute difference between the number of Mexican migrants in the U.S. from the model and from the data (weighted by p_9) and the distances between model and data wage distributions in three populations (weighted by $p_{10}, ..., p_{12}$). Absolute priority is given to full identification of the model (components $p_1, ..., p_8$), while the actual number of migrants and the evaluation of the fit to distributions serves as a choice rule between (otherwise equally good) vectors of parameters Ξ.

The proposed Monte Carlo search method assumes the following procedure:

1. Select a randomly draw vector of parameters Ξ_0.

2. If $\zeta(\Xi_0) < \mathit{threshold}$ continue; else go to step 1.

3. Search for a new vector of parameters in a small neighborhood of the current vector of parameters: $\Xi_1 : e(\Xi_0, \Xi_1) < \epsilon(\zeta(\Xi_0))$, where the imposed distance is a function of the current "goodness of fit" of the model.

4. If $\zeta(\Xi_1) < \zeta(\Xi_0)$ then $\Xi_0 \leftarrow \Xi_1$ and go to step 3.

(that is an ordered pair of wage rate and ranking) is maximized.

[21] $p_1 = p_2 = 100, p_3 = p_4 = p_5 = 1, p_6 = p_7 = 5 \cdot 10^4, p_8 = 10^4, p_9 = 4 \cdot 10^5, p_{10} = 500, p_{11} = 3, p_{12} = 2$.

[22] For $P(\cdot)$ the function $e(\cdot)$ returns the Euclidean distance between model vector of probabilities and data. For distributions, for every grid point we compute Euclidean distances between quantiles of data and model distributions.

5. If no better vector Ξ_1 found after a given number of replications, return the best fitting vector Ξ_0 and go to step 1.

The algorithm settled on the vector of parameters indicated in Table C.2. For a graphical analysis of the loss function minimum achieved by the best parameter vector, consult Figure C.1, where we disturb the best vector of parameters (deviation of which is normalized to zero in the figures) with small positive and negative deviations. Location (k_i) and spread (s_i) of the skill-component in the U.S.-based surplus function take higher values than their counterparts in Mexico. The former is driven by a significant first-order stochastic dominance of the wage distribution of Mexican emmigrants relative to Mexican stayers, while the latter indicates a slightly larger dispersion in skills pricing on the American market comparing to Mexico. Then, firms' component in wages appears to be only slightly more important in the U.S. than in Mexico, as the wage rate in Mexico is lower than in the U.S., while wage range is greater in the U.S. Interestingly enough, our best calibration returns a rather low value of the copula parameter θ. Its value close to 1.2 indicates that American and Mexican skills are weakly related with an average rank correlation of 0.375. Migration costs take values in expected ranges: the multiplicative one equals $1 - \Delta_{UM} = 72\%$ of migrant's wage in the U.S., while the additive one is $\delta_{UM} = 499$. Trade costs, reported in Table C.3, take values ranging between 0.27 and 10.47, they are solely determined by the bilateral trade matrix for a given combination of price indexes, aggregated productions and the elasticity of substitution between product varieties. Table C.1 gathers the fiscal thresholds used to calibrate the tax schedules in both countries.

Identification A close inspection of Equation (C.4) reveals that a hypothetical parameter vector that achieves no loss, would need to satisfy a system of 200,019 equations. While in our calibration procedure all of these equations matter, the first 19 (spelled out fully in (C.I1)–(C.I9)), would, on their own, identify the model's parameters. And indeed, in our Monte-Carlo calibration procedure, we find very close relations between the moments from the data featured in these 19 equations and parameters in Ξ, as depicted in Figure C.2 and summarized in Table C.4. Some parameters are precisely identified by respective model equations and data moments, other emerge as a solution to a subsystem of simultaneous equations.

Equations (C.I1) and (C.I2) are jointly solved by Δ_{UM} and δ_{UM} for given values of minimal and maximal wages achieved by Mexicans in Mexico and the U.S. There exists a close relation between the multiplicative (additive) migration

cost and the maximal (minimal) calibrated wage attainable in Mexico (the U.S.), as summarized in Table C.4 and depicted in Figure C.2, graph 8 and 9. These two parameters determine the relative positioning of distributions of wages for Mexican stayers and emigrants. They are identified by the extremes of these distributions, as the no-arbitrage migration equation has to be fulfilled for the least and the most skilled Mexican worker (and we know the values of the separation function in these extreme points).

Table C.2: Calibrated values of parameters

US Market	MEX Market	Migration Parameters
$k_U = 13,091.53$	$k_M = 6,166.57$	$\theta = 1.1913$
$s_U = 0.8004$	$s_M = 0.4596$	$\delta_{UM} = 499.61$
$\gamma_U = 0.3526$	$\gamma_M = 0.3481$	$\Delta_{UM} = 0.2804$

Table C.3: Calibrated trade costs

From:\To:	ROW	MEX	US
ROW	1.00	0.99	5.12
MEX	3.97	1.00	10.47
US	0.49	0.27	1.00

$$e^{-\Delta_{UM}}\left(\hat{w}_U^c/P_U - \delta_{UM}\right) = \hat{w}_M^c/P_M, \tag{C.I1}$$

$$e^{-\Delta_{UM}}\left(\hat{w}_U^{max}/P_U - \delta_{UM}\right) = \hat{w}_M^{max}/P_M, \tag{C.I2}$$

$$r_U\left(x_U^c, h_U^c; k_U, s_U, \gamma_U\right) = \hat{w}_U^c + P_U\hat{c}_U^f, \tag{C.I3}$$

$$r_M\left(x_M^c, h_M^c; k_M, s_M, \gamma_M\right) = \hat{w}_M^c + P_M\hat{c}_M^f, \tag{C.I4}$$

$$\hat{w}_M^{max} - \hat{w}_M^c = \int_{x_M^c}^1 \partial/\partial x_M r_M\left(r, m_M(r); k_M, s_M, \gamma_M\right) \mathrm{d}r, \tag{C.I5}$$

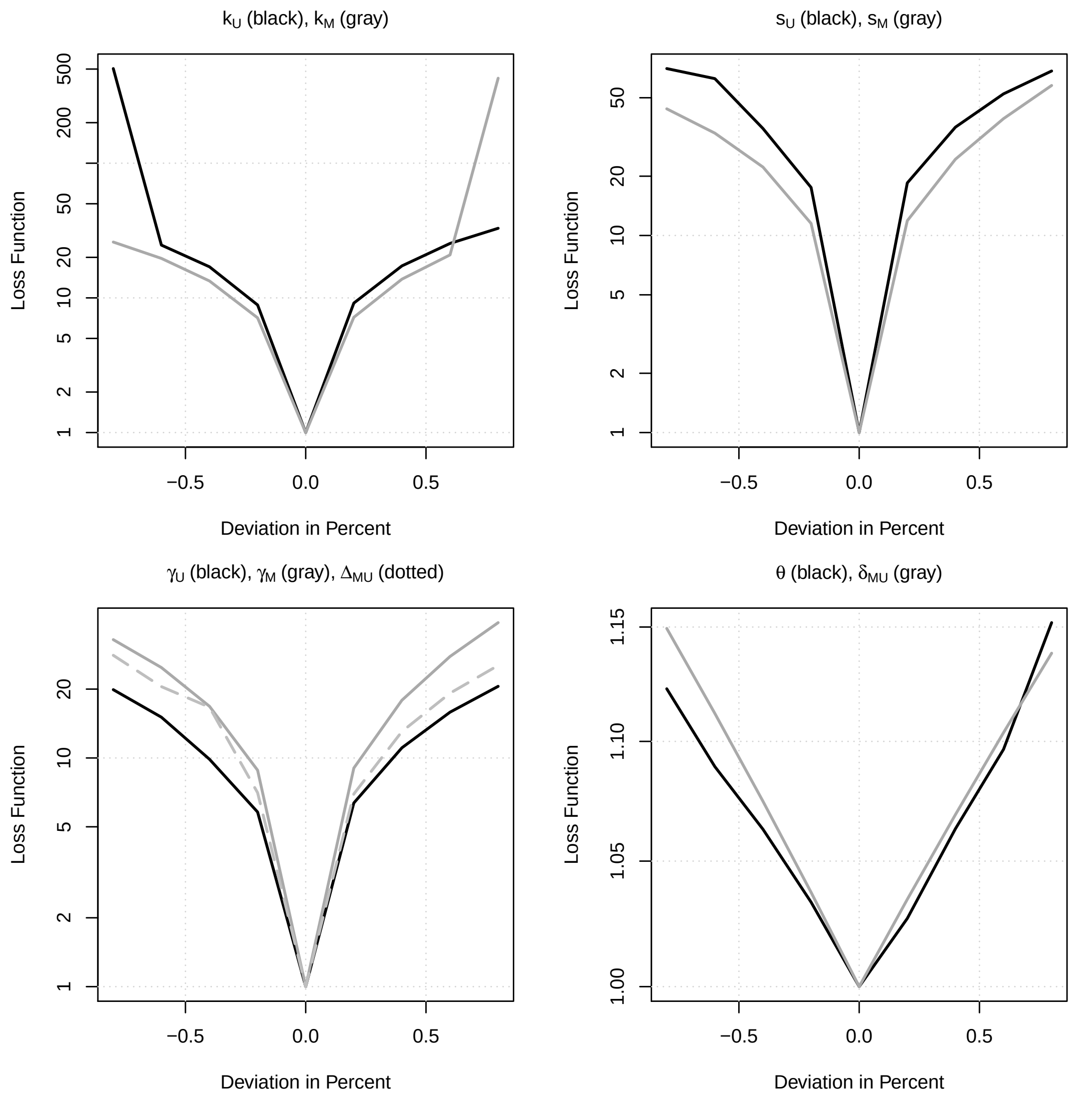


Figure C.1: Evaluation of the Best Parameter Vector

Figure C.1 presents the values of loss function $\zeta(\Xi)$, Eq. (C.4) in the neighborhood of the best parameter vector. Four panels represent one-dimensional marginal values with respect to 9 calibrated parameters: $k_U, k_M, s_U, s_M, \gamma_U, \gamma_M, \Delta_{MU}, \delta_{MU}, \theta$. Horizontal axes represent deviations in the value of respective parameters (calibrated value normalized to 0), while vertical axes depict values of loss function (minimized value normalized to 1).

$$\hat{S}_{UM}(x^c_{UM}) = \int_{x^c_{UM}}^{1} \partial/\partial x_U C\,(r, \phi(r))\,\mathrm{d}r, \tag{C.I6}$$

$$\left[-\int_{x^c_U}^{1} w_U(r)\mathrm{d}S_U(r)\right] \cdot \left[\int_0^1 \pi_U(r)\mathrm{d}r\right]^{-1} = \hat{w}_U^{share}/\hat{\pi}_U^{share}, \tag{C.I7}$$

$$\left[-\int_{x^c_M}^{1} w_M(r)\mathrm{d}S_M(r)\right] \cdot \left[\int_0^1 \pi_M(r)\mathrm{d}r\right]^{-1} = \hat{w}_M^{share}/\hat{\pi}_M^{share}, \tag{C.I8}$$

$$\sum_{x\in\{0,0.1,\ldots,1\}} \Big(\partial/\partial x_M C(x, x_M) - \hat{P}(x)\Big)^2 \to 0. \tag{C.I9}$$

Table C.4: Identification of Model Objects

Object Name	Symbol	Moment	Calibration	Data	Equation
Multiplicative Factors	k_U	$\hat{c}^f_U$	4,846	4,846	(C.I3)
in Production Function	k_M	$\hat{c}^f_M$	1,513	1,513	(C.I4)
Spreads of Mapping between	s_U	$\hat{S}_{UM}(0)$	0.137	0.137	(C.I6)
Skills and Output	s_M	$\hat{w}^c_M$	1,602	605	(C.I5)
Spreads of Mapping between	γ_U	$\hat{w}^{share}_U$	0.560	0.560	(C.I7)
Productivity and Output	γ_M	$\hat{w}^{share}_M$	0.521	0.520	(C.I8)
Utility Costs of Migration	Δ_{UM}	$\hat{w}^m_M$	51,793	45,834	(C.I2)
Monetary Costs of Migration	δ_{UM}	$\hat{w}^c_U$	3,997	4,133	(C.I1)
Correlation between Skills	θ	$\hat{P}(\cdot)$	distance: 0.129		(C.I9)
Distribution of Skills in U.S. Resident Population	$F(\cdot)$	$\hat{w}_{UU}(\cdot)$	forced perfect fit		-

The set of equations (C.I3)-(C.I6) jointly determines the production function parameters in both countries: k_U, s_U, k_M, s_M, for given values of γ_U, γ_M. Equations (C.I3)-(C.I4) indicate that for a given fixed costs c_i^f and minimal wages w_i^c, there exist a combination of k_i, s_i for $i \in \{U, M\}$ that imposes that the gross surplus produced by the worst match in economy i yields exactly the sum of minimal wage and the fixed production cost (zero profit at the cutoff). Equations (C.I5)-(C.I6) determine the spread of Mexican wage distribution and the total mass of Mexican migrants in the U.S., respectively. Note that equation (C.I5) has no counterpart in the U.S. economy. The spread of U.S. citizen wages gives only the range of admissible pairs of k_U, s_U, not the actual values of these two parameters, because the distribution of wage in the population of U.S. citizens is exploited to compute the U.S. skill distribution, $F(\cdot)$, using all degrees of freedom. In this way, one must find another source of identification of k_U, s_U. In our case, this job is done by the equation that characterizes the mass of Mexican immigrants, which depends on the separation function $\phi(\cdot)$, which in turn relates on production functions in both countries.[23]

[23] This module is close to what has been discussed regarding identification of the self-selection model by Roy (1951) in the paper by Heckman and Honore (1990). Parameters k_i relate to the dispersion of the wage distributions, parameters s_i alter the skewness of wage distributions, while migration costs determine location and log-location of the two wage distributions.

Equations (C.I7)-(C.I8) determine the ratios of aggregated wage bills to total profits earned by firms in both economies. Therefore, they directly relate to the moments that describe the structure of GDPs discussed in Table 1. For given parameters k_i, s_i, these two equations determine the magnitudes of γ_i in both countries, as they control the bargaining power of firms in the process of sharing the surplus with workers.[24]

Finally, equation (C.I9) allows us to select the value of copula parameter θ that yields the closest fit to empirically observable conditional probabilities of emigration, $P(\cdot)$, along the distribution of Mexican wages, computed using the MMP data. Our model is over-identified, as long as we fit continuous distributions with parametric approximations. Heckman and Honore (1990) prove that 3 moments per country wage distribution suffice to fully identify the log-normal self-selection model by Roy (1951).

Simulation algorithm In counterfactual simulations we manipulate the values of additive (and multiplicative) migration costs. We solve for the new equilibrium, keeping the set of parameters: $\{k_i, \gamma_i, s_i\}$ for $i \in \{U, M\}$ and θ constant. δ_{UM} (and Δ_{UM}) change by definition, while all the other variables and functions in the model become endogenous.

The algorithm solves for the new equilibrium following a sequential computation procedure. Taking a first guess on the total number of Mexican migrants to the U.S., $S_{UM}(0)$, it recomputes all characteristics of skill and wage distributions, for the new migration costs. This gives the workers' mobility equilibrium. Then, separately, the procedure computes the mass of firms in Mexico and in the U.S. by fixing the costs of entry, and solving for the potential number of firms, R_i^F. These steps allow to obtain country-specific labor market equilibria. Finally, the trade matrix is updated, price indexes are recomputed and new guess on the counterfactual number of Mexicans in the U.S. can be produced. This iterative procedure is continued as long as the aggregated deviation in all endogenous variables in consecutive steps is smaller than $1/K$. The whole procedure ends with determining the final distributions of net real wage changes (as a difference between the counterfactual and the benchmark values) in both countries for all

[24] This part of the identification strategy significantly differs from a standard self-selection model of Roy (1951), as in the latter there are no firms (or equivalently: all firms are homogeneous), which corresponds to a situation in which $\gamma_i = 0$. It is instructive to say that positive values of γ_i are the artifacts of our model that differentiate our approach from the classical self-selection model, and bring us closer to better understanding the importance of matching in real world labor markets.

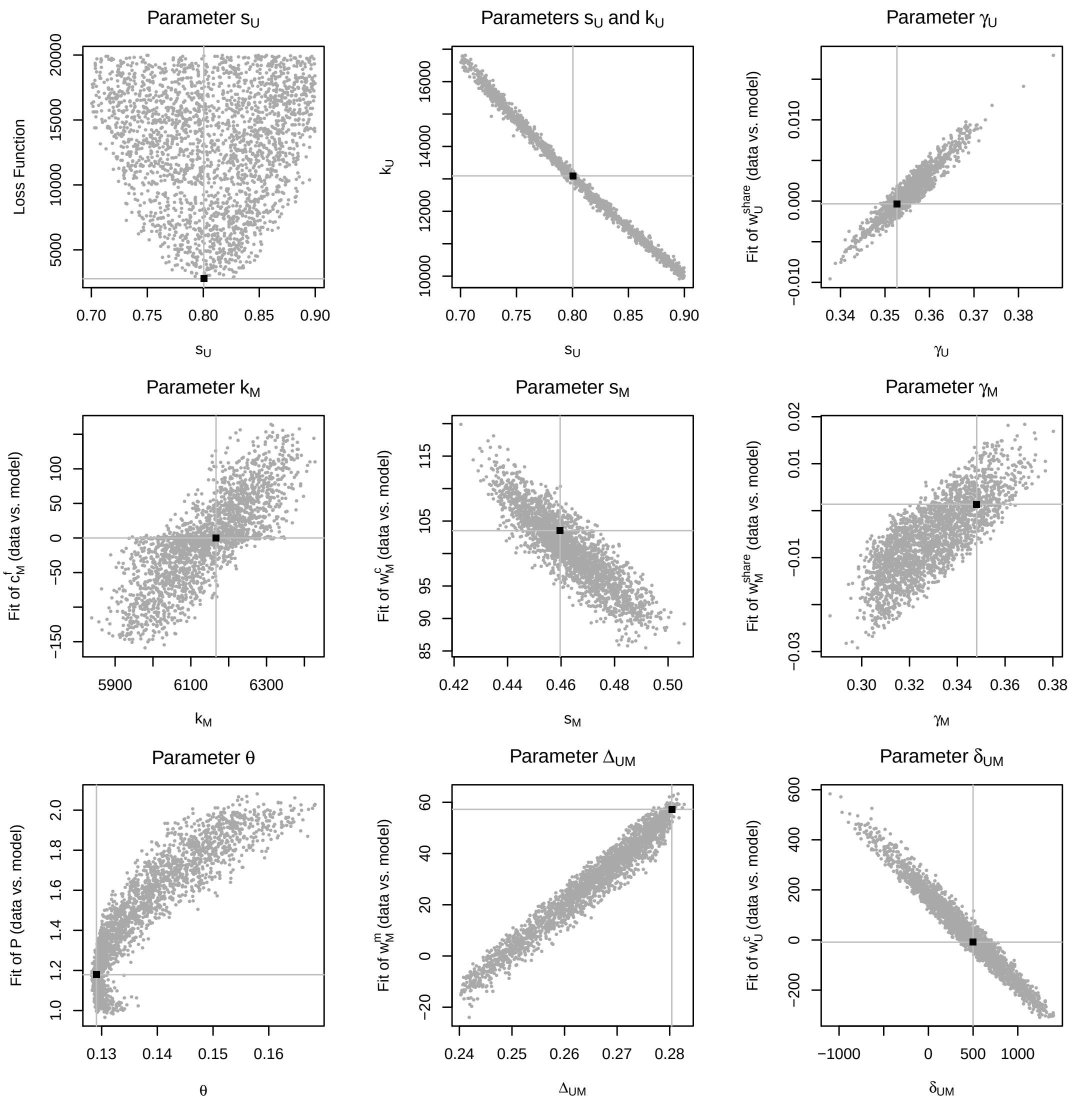
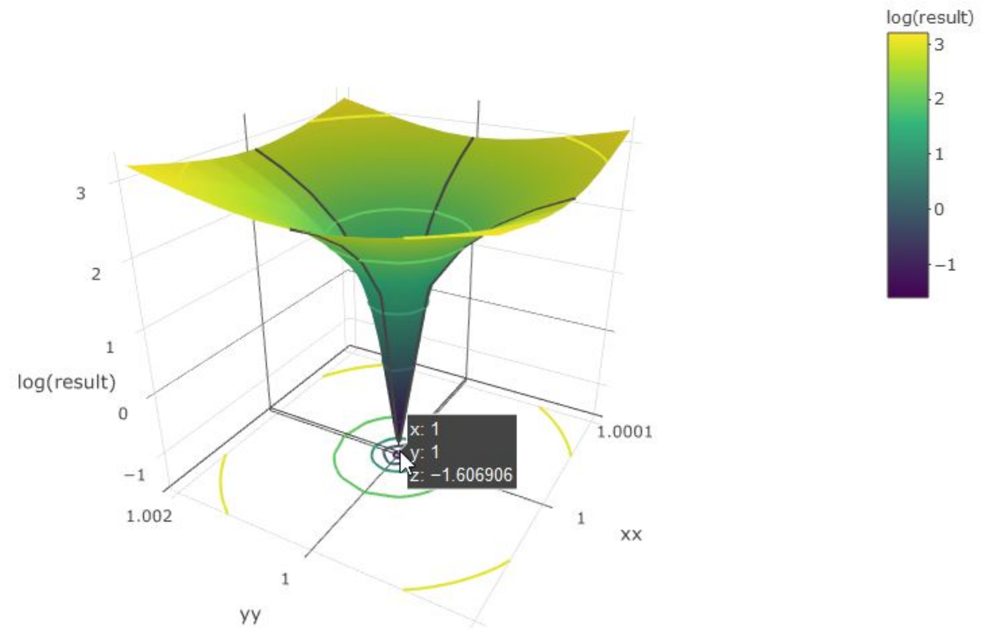


Figure C.2: Identification of Model Parameters

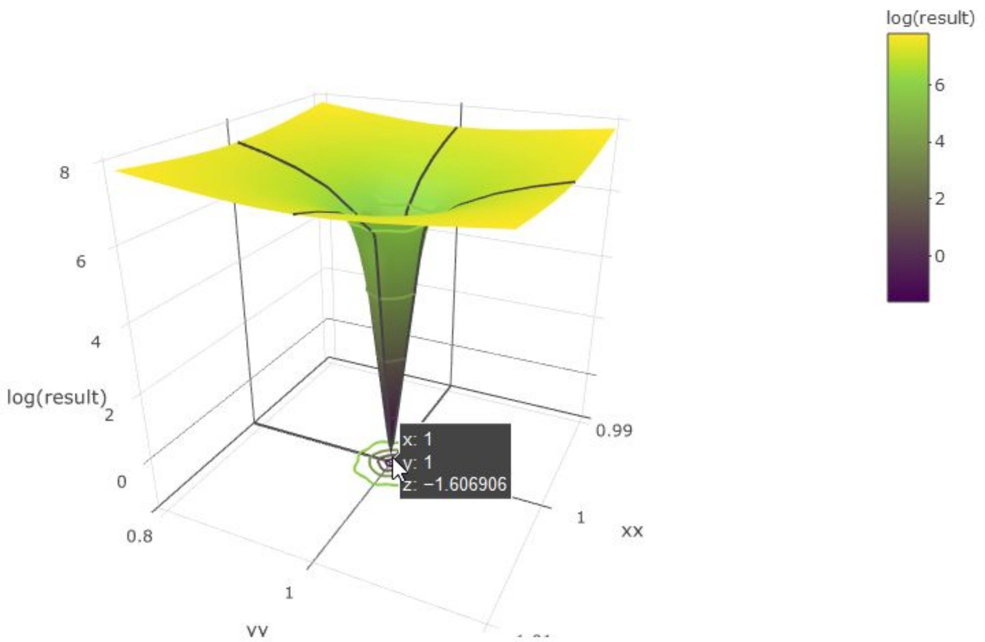
Figure C.2 presents the results of Monte Carlo calibrations for 9 fitted parameters and respective empirical moments matched (for the mapping between parameters and moments see Table C.4). Horizontal axes represent values of respective parameters, while vertical axes depict differences between observed and model values of matched moments. Gray points illustrate 2,000+ outcomes of repeated Monte Carlo search procedure, whereas the black square indicates the actual (best) calibration chosen.

groups of workers. In Figure C.3, we present deviations in values of GDPs after recomputing the labor market equilibrium for (non-)equilibrium initial values of GDPs (note that both figures depict the same 3-d function but in different scales). Only one point (the actual equilibrium) is mapped on itself; other starting points map to different points with positive distance from the initial ones. This indicates that the two-market general equilibrium necessarily has a unique solution which can be computed in an iterative procedure (first solve labor market, than solve

international goods market, repeat until convergence).



(a)



(b)

Figure C.3: Uniqueness of Model's General Equilibrium

Figure C.3 presents the numerical proof of uniqueness of the model's general equilibrium. The "xx" ("yy") axis represents initial deviations in Mexican (U.S.) GDP, while the vertical axis depict logarithms of Euclidean distances between initial and computed vectors of GDPs. General equilibrium GDP levels are normalized to 1.

Appendix D Robustness checks

We verify the robustness of our main results by performing several additional simulations, including alternative parameter values (for the market size effect and the structure of costs), functional forms of the distribution of unemployed and assumptions about the composition of the population of Mexican migrants in the U.S.

Alternative distributions of the skills of unemployed Any shock to the supply of skills in the U.S. affects workers' participation. More precisely, the presence of Mexicans discourages some previously employed Americans to quit the labor market. Importantly, we do not observe the wages (nor the skills) of these unemployed individuals; thus, we can only speculate about the distribution of their skills. In the benchmark, we assume that the skills of out-of-the-market individuals are distributed uniformly. In what follows, we verify this by taking exponential (strictly convex) and logarithmic (strictly concave) CDFs. Both have a negligible impact on the wage effect, as depicted in Figure D.1a.

Adding illegal Mexican immigrants Illegal migration from Mexico to the U.S. proves to be one of the key points in the overall discussion about American migration policy. Therefore, we assess the welfare effects of removing all Mexican immigrants by including undocumented Mexicans. Unfortunately, there is no official dataset that would give us an idea of the wage distribution of illegals. We overcome this difficulty by taking the necessary numbers from estimates available in the literature. First, the actual number of undocumented Mexicans in the U.S. is unknown, although a recent briefing by the Pew Research Center provides some trustworthy estimates of this figure.[25] The authors calculate that out of 11.7 million Mexican immigrants in the U.S. in 2014, there were approximately 5.8 million illegals. Our data consider 7 million working-age migrants (according to the crude estimates, one-third/one-fourth of illegals are included in the U.S. census); thus, we will increase their number to 10.5 million, which leaves us with $S_U^M(0) = 0.206$. Illegal migrants earn substantially lower wages than their legal peers. Due to the work of Caponi and Plesca (2014), we are able to compute the wage penalty for illegals along the wage distribution.[26] We find that the penalty is approximately

[25]For details, please consult *http://www.pewresearch.org/fact-tank/2017/03/02/what-we-know-about-illegal-immigration-from-mexico/*.

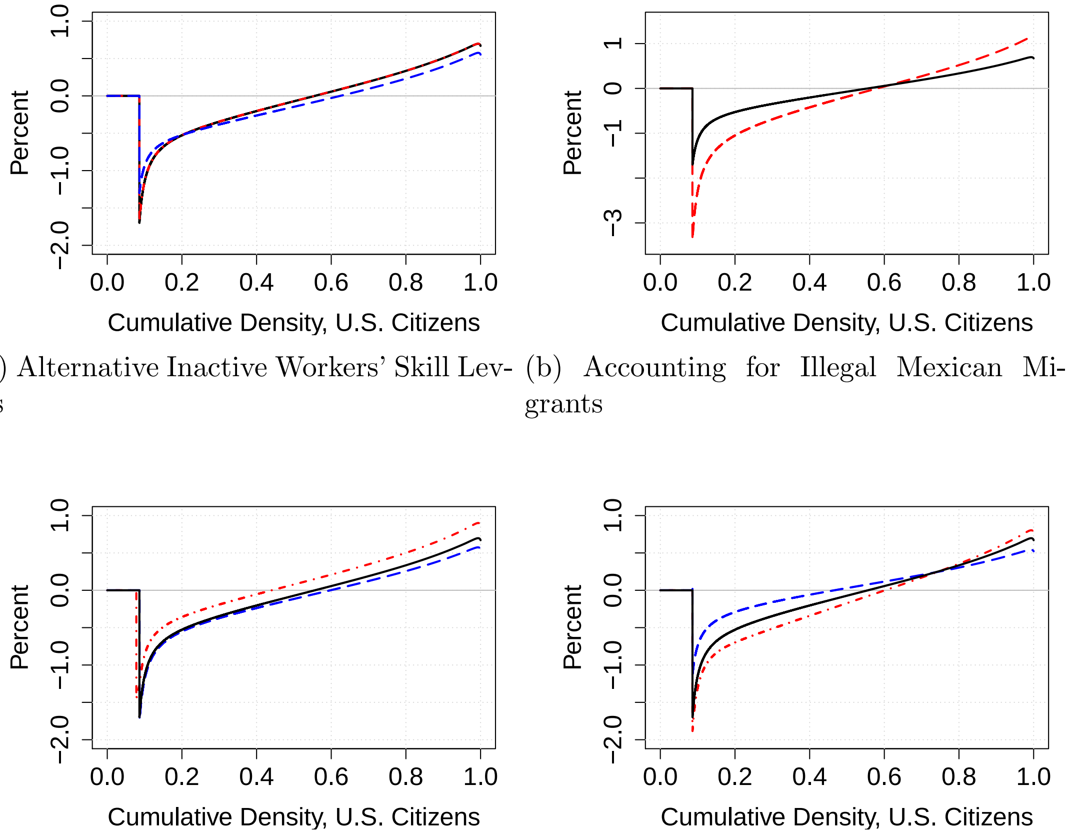
[26]Caponi and Plesca (2014) dispose of data describing illegal immigrants to the U.S. from all origins and compute the kernel distributions of gender-specific penalties; see their Figure

15-20%, in line with the findings of Massey and Gentsch (2014). The new distribution of immigrants' wages would now be an average of the distribution of legal migrants (taken with a weight of 2/3) and illegal migrants (weight of 1/3). While the former is used in the reference calibration, the latter includes a 20% reduction in wages across all worker types. The quantitative outcomes of including illegal Mexican immigrants in the no-migration scenario are depicted in the Figure D.1b. The magnitudes of the economic impacts become significantly more pronounced, while the measure of winners from this is virtually unchanged.

Modifying the market size effect The literature provides numerous estimates of the elasticity of trade flows with respect to trade costs (equivalent to the elasticity of substitution between varieties, ε, in our approach). The various model specifications and datasets used, however, allow us to formulate a convergent view on the magnitude of this particular variable. In the Melitz (2003) trade model with heterogeneous firms, Simonovska and Waugh (2014b) indicate that the 80% confidence interval is $[4.1, 6.2]$. Melitz and Redding (2015) use $\varepsilon = 4$ in their simulations. In the framework developed by Eaton and Kortum (2002), this elasticity is found to be in the range of $[3.8, 5.2]$ according to Bernard et al. (2003); Donaldson (2018); Burstein and Vogel (2010); Eaton et al. (2011); Parro (2013); Simonovska and Waugh (2014a); Caliendo and Parro (2015), although Eaton and Kortum (2002) estimate it at the level of 8. Therefore, we verify the consequences of alternative estimates of ε for our main results. Figure D.1c summarizes the effects of setting infinitely large immigration cots assuming different magnitudes of the market size effect. The black line indicates the reference value of $\varepsilon = 7$; in the blue line, we assume $\varepsilon = 9$; and the red line imposes $\varepsilon = 5$. Higher elasticities (lower market size effects) move the welfare effects very slightly downward. A stronger market size effect has a significantly positive impact on the gains from inviting immigrants, which increases the mass of winners to 60%.

Changing the structure of capital costs One degree of freedom in the calibration process is subject to a broad interpretation of underlying data. This problem concerns the division between variable and fixed costs of capital, that are necessary to pin down production costs: c_i^f and c_i^v. In the benchmark calibration, we assume that the consumption of fixed capital that relates to structures constitutes the fixed part of capital costs, while equipment and intellectual property

1. Using these estimates, we pool the sample and produce a single distribution of differences between legal and undocumented immigrants.



(a) Alternative Inactive Workers' Skill Levels

(b) Accounting for Illegal Mexican Migrants

(c) Robustness to Market Size Effect

(d) Robustness to Capital Costs Structure

Figure D.1: Robustness Checks

Note: Figure D.1 illustrates the economic effects of Mexican migration to the United States, with alternative assumptions about the structure of the model. Figure D.1a experiments with the distribution of skills of inactive workers. The reference scenario (black) assumes a linear CDF, the "convex scenario" (long-dashed blue) assumes exponential CDF, while the "concave scenario" (double-dashed red) assumes logarithmic CDF. Figure D.1b considers a model with illegal Mexican immigrants (long-dashed red line). Figure D.1c assumes alternative values for the elasticity of substitution between varieties (solid black benchmark: $\varepsilon = 7$, double-dashed red: $\varepsilon = 5$, long-dashed blue: $\varepsilon = 9$). Figure D.1d assumes alternative structure of capital and investment costs (solid black benchmark: fixed costs constitute 35% of capital costs, double-dashed red: 0%, long-dashed blue: 100%).

costs are ascribed to its variable part. In this robustness check, depicted in Figure D.1d, we verify the results of our migration scenario in two extreme cases of 100% of capital consumption being related to the fixed (variable) costs of production, illustrated by the blue (red) line. Higher fixed share of capital costs attenuates the magnitudes of the economic effect of Mexican immigration to the U.S. as the result function becomes twisted clockwise. Higher variable share of costs inflates the magnitudes of extreme effects, but keeps the indifferent individual at 60^{th}

percentile, close to our benchmark result.

Appendix E Additional Results

Reference to the canonical model of labor market and multiple-bins CES model In the introduction we discussed the conceptual differences between the matching-selection model developed in this paper and the labor market characterized by a standard constant elasticity of substitution (CES) production function. In this section, we perform a comparison of labor market effects generated by our model with two fundamental state-of-the-art models of labor markets: the "the canonical CES model" of labor markets by Acemoglu and Autor (2011) and the "multiple-bins CES model" by Dustmann et al. (2013). Before proceeding with results, let us first lay out the two theories to which we relate.

The labor market analyzed by Acemoglu and Autor (2011) is a k-skill extension of the standard two-skill CES model, in which the number of efficiency units of skill k is continuously distributed according to an exogenously given distribution $\mathcal{L}_k$. The aggregated production function can be expressed as a nested CES function (with the first tear referring to education, and the second to origin):

$$Y = \left[\sum_{k\in\mathcal{K}} \alpha_k L_k^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \quad L_k = \left[\alpha_k^N (L_k^N)^{\frac{\sigma_m-1}{\sigma_m}} + \alpha_k^M (L_k^M)^{\frac{\sigma_m-1}{\sigma_m}}\right]^{\frac{\sigma_m}{\sigma_m-1}}. \tag{E.1}$$

The total number of efficiency units supplied by a particular skill-origin cell equals $L_k^o = \int_{i\in\mathcal{L}_k^o} l_i d_i$, while σ and σ_m stand for a constant elasticity of substitution between inputs of different sort of skill/origin, respectively. The wage rate of any individual $i \in \mathcal{L}_k^o$ equals: $w_i = w_k^o l_i$, where w_k^o is the remuneration of a unit of skill k supplied by members of the origin group $o \in \{N, M\}$. Knowing the distribution of efficiency unit supplies, it is straightforward to compute the distribution of wage effects in the population of U.S. workers generated by removing all Mexicans workers from the U.S. economy. The downside of this approach consists in keeping the distributions of efficiency units constant. Therefore, the overall economic effects of migration can be computed as a linear combination of direct effects on k discrete groups weighted by exogenously given distribution of k types of efficiency units along the distribution of wages.

The model developed in Dustmann et al. (2013) builds on a different set of assumptions that allow to mimic a continuous distribution of wages in a CES model. The production process takes place using labor and capital (as we consider only

the long-term effects, we assume capital fully adjusts). Individuals are categorized into K groups according to their location in the wage distribution (migrants and natives being perfect substitutes within the so-defined skill group). We write:

$$Y = \left(\beta H^{\frac{s-1}{s}} + (1-\beta) K^{\frac{s-1}{s}}\right)^{\frac{s}{s-1}}, \quad H = \left(\sum_{K \in \mathcal{K}} \alpha_K (l_K)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}. \tag{E.2}$$

Assume that the size of group K equals the sum of native U.S. workers and immigrants Mexican workers that earn the same wage rate: $l_K = l_K^N + l_K^M$. While the allocation of workers in Acemoglu and Autor (2011) was done according to their observable characteristics (i.e. education level), in Dustmann et al. (2013) workers are classified into discrete groups according to their location in the observed wage distribution. In this way, the wage effects induced by immigrants are, as summarized by Dustmann et al. (2013), dependent upon their relative density along the distribution of wages.

Figure E.1 compares the economic effects for U.S. citizens of exogenous removals of different cohorts of Mexican migrants from the U.S. labor market. Black lines relate to the labor market effects (short-term nominal wage and long-term firms' entry and exit effects) generated by our selection and matching model, the red lines depict the wage effects produced by a $\mathcal{K} = 4$ group CES model in the vein of Acemoglu and Autor (2011), whereas the blue lines present the wage effects assuming the CES model by Dustmann et al. (2013) with $\mathcal{K} = 100$ groups of workers.[27] Our results resemble the outcomes from Acemoglu and Autor (2011) model only when the imposed migration shock is diversified enough regarding the skill structure of removed Mexican immigrants, and the similarity is the closest for the random sample shocks. Whenever we impose skill-specific shocks (especially removing 1% and 10% of top skilled Mexicans), the model by Acemoglu and Autor (2011) indicates almost no distribution across workers: Assuming exogenous weighting distributions of efficient skills makes the heterogeneous effects on 4 discrete skill groups cancel out along the distribution, which is visible in the second row of figures in its extreme case. The percentile-bins CES model by Dustmann et al. (2013) gives similar predictions only when considering random deletion of Mexican immigrants (i.e. when the shock is spread across all quantiles of wage distribution). In all other cases the model generates positive wage effects

[27] Four groups of workers in the implemented Acemoglu and Autor (2011) framework relate to the following education groups: high school dropouts, high school degree, some college education, completed college degree. 100 groups in the implementation of Dustmann et al. (2013) model assume percentiles of observed wage distribution in the U.S.

for the group which is hit by the negative supply shock, and uniform and negative effects for all other quantiles of wage distribution (as all percentiles are identically substitutable in the production function). The magnitudes of these effects are significantly higher (as measured on the right-hand-side axes).

Having considered these results, it is vital to highlight fundamental differences between the paradigms analyzed in this quantitative exercise. First, the matching model microfounds the distance-dependent elasticity of substitution (DIDES), as in Teulings (1995, 2005); Costrell and Loury (2004). The latter imposes that the elasticity of substitution is endogenous along the wage distribution and depends on the whole distributions of skills supplied and firms' productivities. This property is in sharp contrast with classical CES models, which assume a constant elasticity of substitution, in which case the structure of labor market interactions is fixed along the distribution of wages as in Acemoglu and Autor (2011) or is symmetric across wage quantiles as in Dustmann et al. (2013). Second, as workers are preassigned to discrete groups in the CES model in which they are perfect substitutes (classification being done according to their education level or their placement in wage distribution in the baseline), there is no room for heterogeneous effects of migration within these broadly defined groups. Moreover, a skill-specific migration shock cannot induce workers' mobility along the wage distribution, which is the main criticism of the multiple-bins CES model articulated by Dustmann et al. (2013). In contrast, our selection and matching model allows for rematching between incumbent workers and firms after any shock to the supply of labor. This, in turn, induced heterogeneous welfare effects along the distribution of wages, as long as the change in the density of skills is different from the change in the density of productivities. Our approach explicitly accounts for natives' mobility along the distribution of wages induced by an inflow/outflow of migrants, similar to the task-upgrading phenomenon discussed by Peri and Sparber (2009) and Foged and Peri (2016).

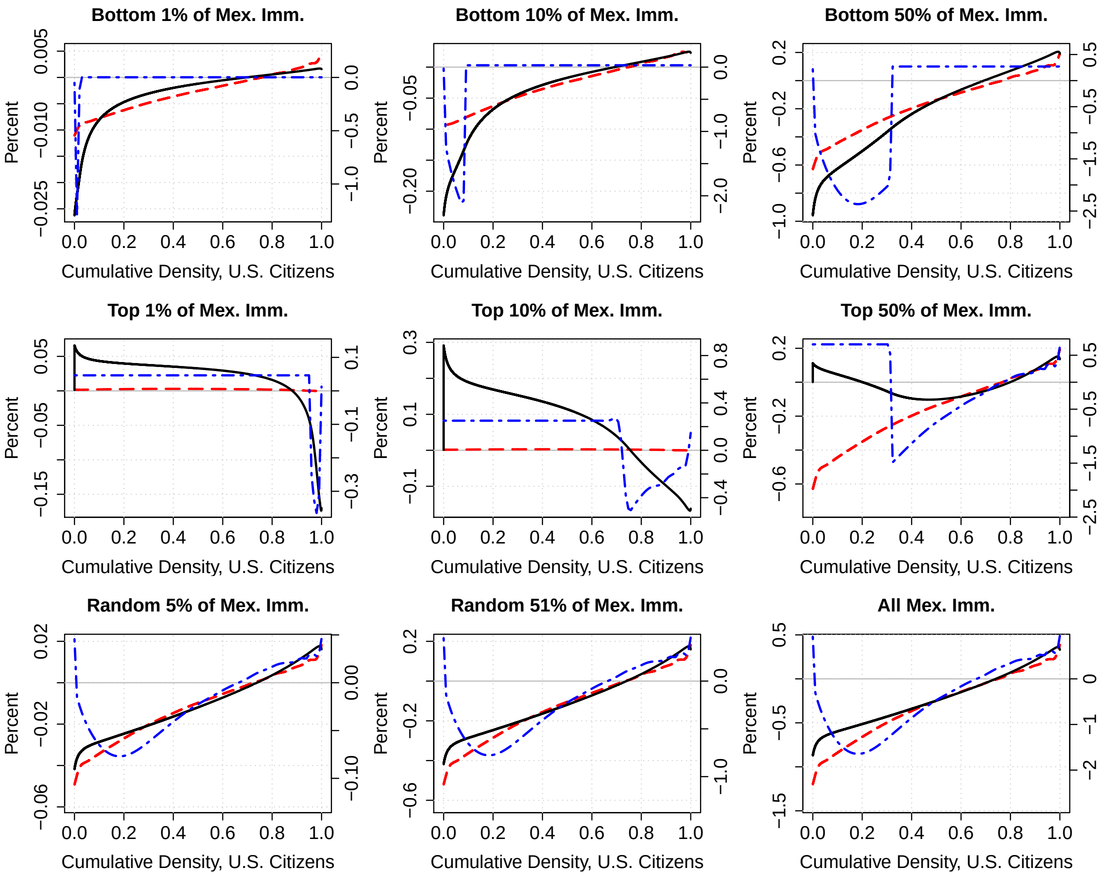


Figure E.1: Labor Market Effects in the Matching and CES models

Figure E.1 presents the quantifications of labor market effects for U.S. citizens from exogenous immigration shocks generated by: the selection and matching model of this paper (black solid curves), the canonical model of labor markets by Acemoglu and Autor (2011) (red long-dashed line) and the multi-bins CES model by Dustmann et al. (2013) (blue two-dashed line). The magnitude of the effects generated by the last model are summarized in the right-hand-side axes. The first (second) row of figures assumes removing 1%, 10% and 50% of the lowest (highest) earning Mexican immigrants in the United States, while the third row depicts the outcomes of a random deletion of 5%, 51% and 100% of Mexican migrants in the United States. All effects are in percentage points, horizontal axes represent quantiles of U.S. citizen wage distribution.

References

Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Berge, C. (1963). *Topological Spaces.* Macmillan.

Bernard, A. B., Eaton, J., Jensen, J. B., and Kortum, S. (2003). Plants and Productivity in International Trade. *American Economic Review*, 93(4):1268–1290.

Burstein, A. and Vogel, J. (2010). Globalization, Technology, and the Skill Premium: A Quantitative Analysis.

Caliendo, L. and Parro, F. (2015). Estimates of the Trade and Welfare Effects of

NAFTA. *The Review of Economic Studies*, 82(1):1–44.

Caponi, V. and Plesca, M. (2014). Empirical Characteristics of Legal and Illegal Immigrants in the USA. *Journal of Population Economics*, 27(4):923–960.

Costrell, R. M. and Loury, G. C. (2004). Distribution of Ability and Earnings in a Hierarchical Job Assignment Model. *Journal of Political Economy*, 112(6):1322–1363.

Donaldson, D. (2018). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review*, 108(4-5):899–934.

Dustmann, C., Frattini, T., and Preston, I. P. (2013). The Effect of Immigration Along the Distribution of Wages. *Review of Economic Studies*, 80(1):145–173.

Eaton, J. and Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica*, 70(5):1741–1779.

Eaton, J., Kortum, S., and Kramarz, F. (2011). An Anatomy of International Trade: Evidence from French Firms. *Econometrica*, 79(5):1453–1498.

Foged, M. and Peri, G. (2016). Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data. *American Economic Journal: Applied Economics*, 8(2):1–34.

Gola, P. (2018). Supply and Demand in a Two-Sector Matching Model.

Heckman, J. J. and Honore, B. E. (1990). The Empirical Content of the Roy Model. *Econometrica*, 58(5):1121.

Massey, D. S. and Gentsch, K. (2014). Undocumented Migration and the Wages of Mexican Immigrants. *The International Migration Review*, 48(2):482–499.

Melitz, M. (2003). The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations. *Econometrica*, 71(6):1695–1725.

Melitz, M. J. and Redding, S. J. (2015). New Trade Models, New Welfare Implications. *American Economic Review*, 105(3):1105–46.

Milgrom, P. and Segal, I. (2002). Envelope Theorems for Arbitrary Choice Sets. *Econometrica*.

Parro, F. (2013). Capital-Skill Complementarity and the Skill Premium in a Quantitative Model of Trade. *American Economic Journal: Macroeconomics*, 5(2):72–117.

Peri, G. and Sparber, C. (2009). Task Specialization, Immigration, and Wages. *American Economic Journal: Applied Economics*, 1(3):135–169.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3(2):135–146.

Simonovska, I. and Waugh, M. E. (2014a). The Elasticity of Trade: Estimates and Evidence. *Journal of International Economics*, 92(1):34–50.

Simonovska, I. and Waugh, M. E. (2014b). Trade Models, Trade Elasticities, and the Gains from Trade.

Tarski, A. (1955). A Lattice-Theoretical Fixpoint Theorem and its Applications. *Pacific Journal of Mathematics*.

Teulings, C. N. (1995). The Wage Distribution in a Model of the Assignment of Skills to Jobs. *Journal of Political Economy*, 103(2):280–315.

Teulings, C. N. (2005). Comparative Advantage, Relative Wages, and the Accumulation of Human Capital. *Journal of Political Economy*, 113(2):425–461.

LUXEMBOURG INSTITUTE OF
SOCIO-ECONOMIC RESEARCH

11, Porte des Sciences
Campus Belval
L-4366 Esch-sur-Alzette

T +352 58 58 55-1
F +352 58 58 55-700

www.liser.lu