Appendices for Online Publication for: "Mexican Migration to the United States: Selection, Assignment, and Welfare"*

Michał Burzyński

Luxembourg Institute of Socio-Economic Research (LISER)

Paweł Gola

University of Edinburgh

January 2025

Abstract

Online Appendix A provides all ommitted proofs, derivations and results of all formal statements. Online Appendix B provides a graphical comparison of the Clayton, Gaussian, and Gumbel copulas. Online Appendix C provides details of the calibration procedure, Online Appendix D reports the results of several robustness checks.

*Contact: Michał Burzyński: michal.burzynski@liser.lu; Paweł Gola: pawel.gola@ed.ac.uk.

A Ommitted Proofs and Derivations

Proof of Theorem 1

The proof should be read after (or concurently with) the sketch provided in the main text.

Wages

The *critical skill in country* i (i.e., the skill level of the least-skilled worker employed in country i) is defined as $x_i^c = \min\{x_{iU}^c, x_{iM}^c\}$. Given the critical skill level, the matching function, wage gradient and wage function are given by Equation (4), (5) and (6), respectively, and the wage paid to the worker with critical skill level x_i^c is determined by the outside option of that worker. In particular, if $x_i^c < 1$ then $w_M^c(x_M^c) = w_M^c$, whereas $w_U(x_U^c) = \min\{w_U^c, e^{\Delta_{UM}} P_U\left(w_M^c/P_M + \delta_{UM}\right)\}$, where $w_i^c \equiv P_i^c \bar{w}_i^c$. To see why, first suppose that $i = M$ (the argument for $i = U$ is analogous). It follows immediately from Equation (16) and the definitions of x_{MM}^c and U_{MM}^c 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.

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 (16) that if $x_{UM}^c < 1$, then $w_U(x_{UM}^c) = P_U\left(e^{\Delta_{UM}} w_M^c/P_M + \delta_{UM}\right)$.

Supply Functions

Given the definition of the separation function, Equation (19) follows immediately by the exact same reasoning as in the proof of Lemma 2 in Gola (2021).

Feasibe Allocations

Define the set $\mathbb{E}$ of *partial equilibrium allocations* as the set of allocations for which there exists a pair of wage functions w_U, w_M that induces both the supply and the demand functions to be equal to S_U, S_M and satisfies the zero-expected-profit-condition. Define also $x_M^s = \inf\{x_M \geq x_M^c : \psi(x_M) = 1\}$. Equations (18)–(21) and the requirement that labor markets clear put strong restrictions on the partial equilibrium allocations. In particular, the restrictions on allocations imply that for any $A \in \mathbb{E}$, it must be the case that (1) $S_{ij} : [0,1] \to [0, R_i^W]$ is non-decreasing, absolutely continuous, and semi-differentiable on the interior, with $S_{ij}(1) = 0$; (2) the function $\psi : [x_{MM}^c, 1] \to [x_{UM}^c, 1]$ 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 $x_M \in (x_M^c, x_M^s)$; (3) S_{MU}, S_{MM} satisfy Equation (20) and $-\partial_+ S_{Mi}(x_i)/\partial x_M \in (0,1)$; (4) $S_{UU}(x_U)$ satisfies (21), and (4) $S_i(0) \leq R_i^F$, where $S_i(x_i) = S_{ii}(x_i) + S_{ij}(x_i)$ and $S_{MU}(x_M) = 0$.[1] The allocations meeting conditions (1)–(4) will be called *feasible* and the set of all feasible allocations will be denoted by $\mathbb{A}$. Clearly, $\mathbb{E} \subset \mathbb{A}$.

Uniqueness

Proposition 1. A worker and firm allocation $A^* \in \mathbb{A}$ can hold in the partial labor market equilibrium if and only if it maximizes the weighted sum of (net) revenues:

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

Proof. 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 $A^* \neq A'$.[2] The tuple $w = (w_M, w_S)$ that clears markets for A^* is denoted as w^*.[3]

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:

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

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

$$\begin{aligned} \bar{w}^A(w, A) = \; & e^{-\Delta_{UM}} \left[\bar{w}^A_{UM}(w_U, S_{UM}) + \bar{w}^A_{UU}(w_U, S_{UU}) + \bar{w}^c_U F(x^c_U(S_{UU})) R^W_U \right] \\ & + \bar{w}^A_{MM}(w_M, S_{MM}) + \bar{w}^c_M C(x^c_{UM}(A), x^c_M(A)) - \delta_{UM} S_{UM}(0). \end{aligned}$$

As $S^*_{UM}, S^*_M, S^*_{UU}$ and w^* are the equilibrium supply and wage functions, respectively, it

[1]Condition (1) follows from Equations (18)–(21) and market clearing. Condition (2) follows from differentiating Equation (18) on (x^c_M, x^s_M) and noting that wages must be non-decreasing in equilibrium; (3) follows from Equations (19) and (20); and (4), obviously, from Equation (21). Note that R^W_i denotes the measure of workers born in country i, with R^W_M normalized to 1.

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

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

follows from the first equilibrium condition (Definition 2) 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 3.3. Suppose that $R_U^{F*}, R_M^{F*} > 0$, the other cases are considered in footnote 4. 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 $\frac{\pi_i^{E*}}{P_i} - c_i = 0$ 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).[4]

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

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

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

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, $S_i^* = S_i'$ 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

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

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 2) 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:

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

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 $\mathrm{int}(\mathbb{A}(R_U^F, R_M^F))$ 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.[5] 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}^*).[6] 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 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

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

---

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

[6]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. As long as $\underline{\epsilon}, \bar{\epsilon}$ are small enough, each $\psi(\cdot; t_M)$ is strictly increasing and thus 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 $t_M \in [0,1]$.

Thus it can be shown that:

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

Because $V(t_i)$ is a 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 $x_M \in (x_{MM}^{c*}, x_{MM}^{s*})$, then

$$o(x_M) \equiv e^{-\Delta_{UM}}\psi_{x_M}^*(x_M)\frac{\partial}{\partial x_U}\bar{r}_U(\psi^*(x_M), m_U^*(\psi^*(x_M))) - \frac{\partial}{\partial x_M}\bar{r}_M(x_M, m_M^*(x_M)) = 0, \qquad \text{(A.11)}$$

where $\psi_{x_M}^*(x_M) = \frac{\partial}{\partial x_M}\psi^*(x_M)$.

Proof. Consider such family $S_{MM}(\cdot;t_M)$ that $x_{MM}^c(t_M) = x_{MM}^{c*}$, $S_{MM}(0;t_M) = S_{MM}^*(0)$, $\frac{\partial}{\partial x_M}S_{MM}(x_{MM}^{c*};t_M) = \frac{\partial}{\partial x_M}S_{MM}^*(x_{MM}^{c*})$, and $S_{MM}(x_M;t_M) = S_{MM}^*(x_M)$ for all $x_M \geq x_{MM}^{s*}$.This implies that x_{MM}^c, x_{UM}^c and $S_{MM}(0)$ do not change with t_M, and thus neither does $S_{UM}(0)$, because $C(x_{UM}^{c*}, x_{MM}^{c*}) = 1 - S_{UM}(0) - S_{MM}(0)$ by the definition of $S_{UM}(S_{MM})$. It follows that Equation (A.9) reduces to

$$\begin{aligned}
\frac{\partial}{\partial t_M}V(t_M) = e^{-\Delta_{UM}}&\int_{x_{MM}^c}^{x_{MM}^s}\frac{\partial}{\partial t_M}S_{MM}(x_M)\psi'(x_M)\frac{\partial}{\partial x_U}\bar{r}_U(\psi(x_M), m_U(\psi(x_M)))\mathrm{d}x_M. \\
&- \int_{x_{MM}^c}^{x_{MM}^s}\frac{\partial}{\partial t_M}S_{MM}(x_M)\frac{\partial}{\partial x_M}\bar{r}_M(x_M, m_M(x_M))\mathrm{d}x_M \qquad \text{(A.12)}
\end{aligned}$$

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 $t''_M > t'_M$, $\text{sgn}(S_{MM}(x_M; t'_M) - S_{MM}(x_M; t''_M)) = \text{sgn}(o(x_M))$ 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! □

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. As discussed on page 2 of this Appendix, the wage functions w_U, w_M for which conditions (2) and (4) of equilibrium are satisfied, are given by Equation (6), where $\bar{w}^c_M(x^c_M) = \bar{w}^c_M$ and $\bar{w}_U(x^c_U) = \min\{\bar{w}^c_U, e^{\Delta_{UM}}(\bar{w}^c_M + \delta_{UM})\}$. For condition (1) to be satisfied, it must be the case that these w_U, w_M satisfy (i) Equation (18) as well as (ii) $\bar{w}_U(x^{c*}_{UU}) = \bar{w}^c_U$.

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

$$\bar{r}_U(x^{c*}_U, m_U(x^{c*}_U)) + \int_{x^{c*}_U}^{x^{c*}_{UU}} \frac{\partial}{\partial x_U} \bar{r}_U(x_U, m_U(x_U)) \mathrm{d}x_U = \bar{w}^c_U. \tag{A.13}$$

Let us focus on $\frac{\partial}{\partial t_M} V(t_M)$ and consider a family $S_{MM}(\cdot; t_M)$ such that $S_{MM}(x^{c*}_{MM}; t_M) = S^*_{MM}(x^{c*}_{MM})$, $x^c_{UM}(t_M) = x^{c*}_{UM}$, $S_{MM}(x_M; t_M) = S^*_{MM}(x_M)$ for all $x_M \geq x^{s*}_{MM}$ and $t^*_M \in (0,1)$. This implies that (a) $\frac{\mathrm{d}}{\mathrm{d}t_M} S_{UM}(0; t^*_M) = 0$ as well as (b) $\frac{\mathrm{d}}{\mathrm{d}t_M} S_{MM}(0) = \frac{\partial}{\partial t_M} x^c_{MM}(t_M) \frac{\partial}{\partial x_M} C(x^c_{MM}(t_M), x^{c^*}_{MU})$. Substituting this and Equation (A.11) into Equation (A.9) yields

$$\bar{r}_M(x^{c*}_M, m_M(x^{c*}_M)) = \bar{w}^c_M. \tag{A.14}$$

Consider such family $S_{MM}(\cdot; t_M)$ that $S_{UM}(0; t_M) \neq S^*_{UM}(0)$ for $t_M \neq t^*_M$, and $S_{MM}(x_M; t_M) = S^*_{MM}(x_M)$ for all $x_M \geq x^{s*}_{MM}$. Then subsituting Equations (A.11) and

[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

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

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 $\psi(x_M; t_M) = \psi^*(x(M))$.

(A.14) into Equation (A.9) yields

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

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

$$e^{\Delta_{UM}}\left(\bar{w}_M^c + \delta_{UM}\right) - \bar{w}_U^c = \bar{w}_U(x_{UM}^{c*}) - \bar{w}_U(x_{UU}^{c*}), \tag{A.16}$$

which implies that $x_{UM}^{c*} \geq x_{UU}^{c*}$ if and only if $e^{\Delta_{UM}}\left(\bar{w}_M^c + \delta_{UM}\right) \geq \bar{w}_U^c$.

Suppose that $x_{UM}^{c*} \geq x_{UU}^{c*}$. Then $\bar{w}_U(x_U^{c*}) = \bar{w}_U^c$ and condition (ii) follows immediately. Further, Equation (A.13) reduces to $\bar{r}_U(x_U^{c*}, m_U(x_U^{c*})) = \bar{w}_U^c$. Substituting this into Equation (A.15) ensures that $\bar{w}_U(x_{UM}^{c*}) = e^{\Delta_{UM}}\left(\bar{w}_M(x_{MM}^{c*}) + \delta_{UM}\right)$. As Equation (A.11) is the same as the first derivative of Equation (18) on $(x_{UM}^{c*}, x_{UM}^{s*})$, it follows that condition (ii) must be satisfied as well.

Suppose that $x_{UM}^{c*} < x_{UU}^{c*}$. Then $\bar{w}_U(x_{UM}^{c*}) = e^{\Delta_{UM}}\left(\bar{w}_M(x_{MM}^{c*}) + \delta_{UM}\right)$, which reduces Equation (A.16) to $\bar{w}_U(x_{UU}^{c*}) = \bar{w}_U^c$. Again, Equation (A.11) is the same as the first derivative of Equation (18) on $(x_{UM}^{c*}, x_{UM}^{s*})$, 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 OA.11 in Gola (2021). Consider some maximizer (R_U^{F*}, R_M^{F*}) of the outer problem and some $R_i^{F'}$. Define the function $R_i^F(t_R) = t_R R_i^{F*} + (1 - t_R) R_i^{F'}$. Note that

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

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

$$V^I(S_{UU}, S_{MM}, t_R) = \max_{S_{UM}\in\mathbb{S}_{UM}(S_{MM})} \bar{V}(S_{UU}, S_{UM}, S_{MM}, R_U^F(t_R), R_M^F(t_R)).$$

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 $R_i^F(t_R) = S_i(0)$), and also that whenever $V_t^I(S_U, S_M, t)$ does exist we have that

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

where

$$\pi^E_i(S_i,R^F_i) = \begin{cases} \int_0^1\int_0^h \frac{\partial}{\partial h} r_i(S_i^{-1}((1-p)R^F_i),p)\mathrm{d}p + r_i(S_i^{-1}(R^F_i),0)\,\mathrm{d}h & \text{for } R_i \in (0,S_i(0)), \\ \int_{1-\frac{S_i(0)}{R^F_i}}^1 \int_{1-\frac{S_i(0)}{R^F_i}}^h \frac{\partial}{\partial h} r_i(S_i^{-1}((1-p)R^F_i),p)\mathrm{d}p\,\mathrm{d}h & \text{for } R^F_i > S_i(0). \end{cases} \tag{A.17}$$

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

$$|V_{t_R}(S_{UU},S_{MM},t)| \leq (R^{F*}_U - R^{F'}_U)\max\{c^e_U,\bar{r}_U(1,1)\} + (R^{F*}_M - R^{F'}_M)\max\{c^e_M,\bar{r}_M(1,1)\}$$

which proves

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

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

Define $S_{UU}(t_R), S_{MM}(t_R) \in \arg\max_{(S_{UU},S_{MM})\in\mathbb{A}} V^I((S_{UU},S_{MM},t_R)$, $T(t_R) \equiv V(t_R) + c^e_U R^F_U(t_R) + c^e_M R^F_M(t_R)$ and pick any $t \in (0,1)$ for which $V(\cdot)$ is differentiable. Consider two $\tilde{c}^e_U, \tilde{c}^e_M \in \mathbf{R}_{\geq 0}$ such that $\tilde{c}^e_i = \pi^E_i(R^F_U(t),R^F_M(t))$. For entry costs $\tilde{c}^e_M, \tilde{c}^e_S$, the allocation $A(t) = (S_{UU}(t), S_{UM}(S_{MM}(t)), S_{MM}(t), R^F_U(t), R^F_M(t))$ is a partial labor market equilibrium, implying that it maximizes the function $\tilde{V}(t) = T(t) - \tilde{c}^e_U R^F_U(t) - \tilde{c}^e_M R^F_M(t)$. Clearly, both $\tilde{V}(\cdot)$ and $T(\cdot)$ are differentiable at t as well. It follows from first-order conditions that $\tilde{V}_{t_R}(t_R) = 0$ implying that

$$T_{t_R}(t_R) = (R^{F*}_U - R^{F'}_U)\tilde{c}^e_U + (R^{F*}_M - R^{F'}_M)\tilde{c}^e_M$$
$$= (R^{F*}_U - R^{F'}_U)\pi^E_U(R^F_U(t),R^F_M(t)) + (R^{F*}_M - R^{F'}_M)\pi^E_M(R^F_U(t),R^F_M(t)).$$

This proves that

$$V_{t_R}(t) = (R^{F*}_U - R^{F'}_U)(\pi^E_U(R^F_U(t),R^F_M(t)) - c^e_U) + (R^{F*}_M - R^{F'}_M)(\pi^E_M(R^F_U(t),R^F_M(t)) - c^e_M).$$

Note, by the way, that because we can induce an equilibrium for any values of (R^F_U, R^F_M) by an appropriate choice of $(\tilde{c}^e_U, \tilde{c}^e_M)$, 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 $R^{F*}_U > 0$ then $\pi^E_M \geq c^e_M P_M$. First, pick some $R^{F'}_M < R^{F*}_M$ 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 $\pi_i^E(t)$ is continuous even for $R_M^F = 0$, in the sense that the limit of the average 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. □

Existence

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 2 of this Appendix 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 OA.2 in Gola (2021) 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.

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 $p_W q_W = Y_W$ Equation (13) for $i = W$ can be rewritten as

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

[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.

[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 U.S. 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 (23) results in

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

where $a \in (0, \min_{i\in\{U,M,W\},j\in\{U,M\}}\{(\frac{\tau_{ji}}{p_W\tau_{iW}})^{1-\varepsilon}\})$. 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.18) must also satisfy Equation (13).[10] Therefore, it follows trivially that any such $\mathbf{P}$ solves also the system given by (23).

Lemma 2. Consider the set $\mathbb{P}$ of all $\mathbf{P} \in \mathbf{R^3}_{>0}$ that solve Equation (A.18) for all $i \in \{U, M, W\}$. The set $\mathbb{P}$ is non-empty.

Proof. Consider the interval $I_i = [\left[\frac{(\tau_{iU})^{1-\varepsilon}Y_U}{aY_W} + \frac{(\tau_{iM})^{1-\varepsilon}Y_M}{aY_W} + (\tau_{iW}p_W)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}, \tau_{iW}p_W]$ 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

$$\mathscr{T}_i(\mathbf{P}) \equiv \left[\frac{(\tau_{iU})^{1-\varepsilon}Y_U}{a\mathscr{F}(\mathbf{P}) + \sum_k Y_k\tau_{kU}^{1-\varepsilon}P_k^{\varepsilon-1}} + \frac{(\tau_{iM})^{1-\varepsilon}Y_M}{a\mathscr{F}(\mathbf{P}) + \sum_k Y_k\tau_{kM}^{1-\varepsilon}P_k^{\varepsilon-1}} + (\tau_{iW}p_W)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}.^{11} \tag{A.19}$$

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

□

The largest vector of price indexes solving Equation (A.18) 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 is 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 we can define the map $\mathscr{K} : \mathbb{R}^3_{\geq 0} \to \mathbb{R}^3_{\geq 0}$ such that

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

[10] Multiplying voth sides of (A.18) by $P_i^{\varepsilon-1}Y_i$, summing by i and rearranging results in

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

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 $\mathscr{T}_i(\mathbf{P}) \leq \tau^{Wi}p_W$.

Any fixed point of this map characterizes a general equilibrium of our 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] Thus 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 (23) is solved uniquely by $P_U = P_M = P_W$. This further implies that $\mathscr{B}_U(\mathbf{Y}) = \mathscr{B}_M(\mathbf{Y}) = \mathscr{B}_W(\mathbf{Y})$. $\mathscr{F}(\mathbf{P}) = 0$ gives that $\mathscr{B}_W(\mathbf{Y}) = p_W^\varepsilon q_W$, which pins down the unique equilibrium.

Equilibrium with Empty Mexico

We will now demonstrate that if the transportation costs either to or from Mexico are sufficiently large, then there must exist an equilibrium in which $S_M(0) = 0$. In what follows we use the notation from the proof of Theorem 2. Consider an auxiliary economy in which $r_M^A(x, h) = 0$ but the rest of the model is unchanged. Trivially, if there exists an equilibrium of the actual model in which $S_M(0) = 0$, then the vector of equilibrium total expenditures $\mathbf{Y}^A$ must satisfy $\mathscr{K}^A(\mathbf{Y}^A) = \mathbf{Y}^A$, where $\mathscr{K}^A$ is the map determining the equilibrium of the auxiliary economy. It follows that $\mathbf{Y}^A$ is independent of τ_{iM}, τ_{Mi} for all $i \in \{U, W\}$; by inspection of Equation (A.18) so are $\bar{P}_U(\mathbf{Y}^A), \bar{P}_W(\mathbf{Y}^A)$. Suppose that $\tau_{WM} = \tau_{MW} = \tau_{UM} = \tau_{MU} = a$. Then

$$B_M = a^{\frac{1-2\varepsilon}{\varepsilon}} \frac{\left(\sum_{k\in\{U,W\}} Y_k^A (\bar{P}_k(\mathbf{Y}^A))^{\varepsilon-1}\right)^{1/\varepsilon}}{\left[Y_U^A \left(\sum_{k\in\{U,W\}} Y_k^A \left(\bar{P}_k(\mathbf{Y}^A)/\tau_{Uk}\right)^{\varepsilon-1}\right)^{-1} + p_W^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}}$$

Thus, if

$$a \geq \left((\bar{U}_M + c_M^f) f_M(1,1)^{\frac{1-\varepsilon}{\varepsilon}} \frac{\left[Y_U^A \left(\sum_{k\in\{U,W\}} Y_k^A \left(\bar{P}_k(\mathbf{Y}^A)/\tau_{Uk}\right)^{\varepsilon-1}\right)^{-1} + p_W^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}}{\left(\sum_{k\in\{U,W\}} Y_k^A (\bar{P}_k(\mathbf{Y}^A))^{\varepsilon-1}\right)^{1/\varepsilon}} \right)^{\frac{\varepsilon}{2\varepsilon-1}}$$

then $\bar{w}_M(1) \leq r_M(1,1)/P_M \leq \bar{U}_M$ and thus $S_M(0) = 0$ as required.

[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 continuous by the same reasoning as that in footnote 8.

B Copula Functions

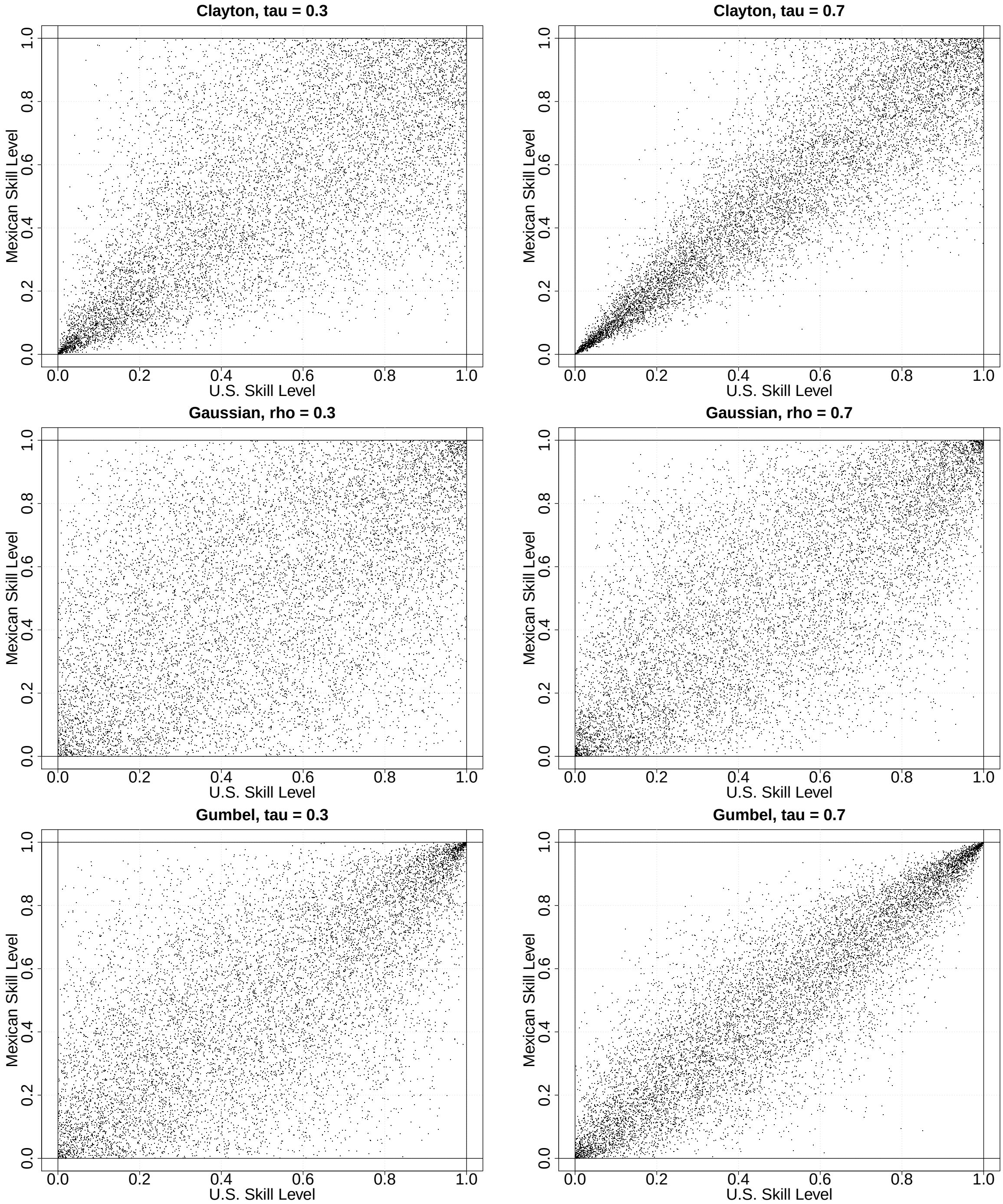


Figure B.1: Two-dimensional Distributions for Clayton, Gaussian and Gumbel Copulas

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.

C Further Calibration Details

Identification As discussed in the main text, our model requires an identification of nine unknown parameters. Our identification strategy relies on matching five discrete empirical moments, a set of conditional emigration probabilities from Mexico and two wage distributions. While in our calibration procedure all of these equations matter, the first 19 (highlighted 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 $\Xi = \{K_U, s_U, \gamma_U, K_M, s_M, \gamma_M, \theta, \delta_{UM}, \Delta_{UM}\}$, for $i \in \{U, M\}$: $K_i, s_i, \gamma_i, \theta, \Delta_{UM} > 0$; $\delta_{UM} \in \mathbb{R}$, as depicted in Figure C.1 and summarized in Table 3. Some parameters are precisely identified by respective model equations and data moments, other emerge as a solution to a subsystem of simultaneous equations.

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

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

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

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

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

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

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

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

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

As in every selection model, migration costs, Δ_{UM} and δ_{UM} define the shape (the skewness) of Mexicans' wage distributions in Mexico and in the United States. By construction, in our model, these variables also determine the minimal (maximal) wages received by Mexicans in Mexico (in the United States). Equations (C.I1) and (C.I2) are jointly solved by Δ_{UM} and δ_{UM} for given values of minimal and maximal wages received by Mexicans in Mexico and the United States. These two parameters determine the relative positioning of distributions of wages for Mexican stayers and emigrants and are identified by the extremes of wage distributions, as the no-arbitrage migration equation has to be fulfilled for the least and the most skilled Mexican worker. There exists a close relation between the multiplicative (additive) migration cost and the maximal (minimal) calibrated wage attainable in Mexico (the United States), as summarized in Table 3 and depicted in Figure C.1, graphs 8 and 9. However, we do not want to make our calibration strategy vulnerable to and dependent on subjective choices of wage distributions' cut-offs (imposed to keep our model Lipschitz continuous). Therefore, even though we retain the minimal and maximal wages in both countries as targeted empirical moments in the loss function, we assign a relatively low weight to these conditions and do not fit them exactly.

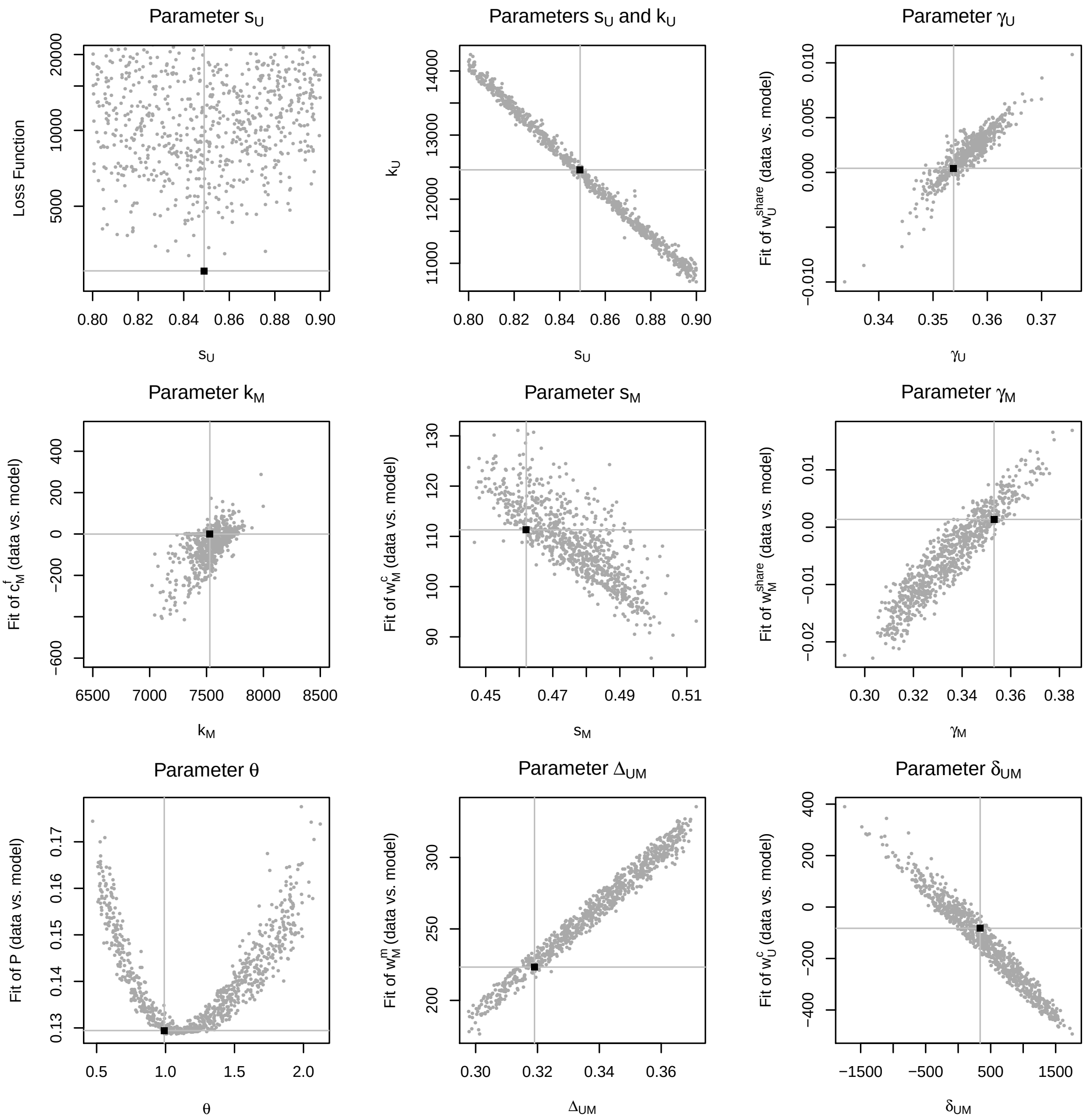


Figure C.1: Identification of Model Parameters

Figure C.1 shows the results of Monte Carlo calibrations for 9 fitted parameters and respective empirical moments matched (for the mapping between parameters and moments see Table 3). Horizontal axes represent values of respective parameters, while vertical axes depict differences between observed and model values of matched moments (exceptions are top-left figure, in which we plot the value of loss function, and top-middle figure, in which we plot values of K_U parameters). Gray points illustrate outcomes of 800 Monte Carlo calibrations, whereas the black square indicates the best calibration.

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. Since these four parameters affect not only country-specific moments (fixed costs and the dispersion of wage distributions) but also influence the location of the separation function, they cannot be individually determined. 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 United States, respectively. Note that equation (C.I5) has no counterpart in the U.S. economy. The spread of U.S. residents' 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. residents 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 and 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.[14] Even though Figure C.1 reveals a close relationship between these four parameters and particular moments, one has to bear in mind that K_U, s_U, K_M, s_M are a solution to a system of four simultaneous equation rather than an explicit one-to-one identification.

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.

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 Honoré (1990) prove that 3 moments per country wage distribution suffice to fully identify the log-normal self-selection model by Roy (1951).

Solution of the model For a given vector of parameter guesses (denoted by 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. Using Equation (5), we arrive

[14] 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 Honoré (1990). Parameters K_i relate to the location of the wage distributions, as they are identified by the fixed costs, c_i^f. Then, parameters s_i determine the dispersion of wage distributions, while migration costs determine the skewness of the two wage distributions, as in all self-selection models.

at the following differential equation:

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

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 of U.S. skills $F(\cdot)$, multiplied by the density of skills supplied by U.S. residents: $F'(x_U)$. 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.[15]

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 (18), and taking its first derivative:

$$\frac{\partial}{\partial x_U}\bar{w}_M(\phi(x_U)) = e^{-\Delta_{UM}}\frac{\partial}{\partial x_U}\bar{w}_U(x_U) \leftrightarrow$$
$$\frac{\partial}{\partial x_U}r_M(\phi(x_U), h_M(\phi(x_U)))\phi(x_U)' = \frac{P_M}{P_U}e^{-\Delta_{UM}}\frac{\partial}{\partial x_U}r_U(x_U, h_U(x_U)). \tag{C.2}$$

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: $\phi(1) = 1$.[16] For the given solution for selection pattern, determined by the separation function $\phi(\cdot)$, the mass of Mexican immigrants in the United States can be computed by using Equation (19), 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 skills 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 (5) to determine the wage distributions of Mexican workers in the United States and in Mexico. The final result of the calibration for a given guess of parameter values Ξ is a set of three wage distributions: U.S. residents, $w_U \equiv (w_U(x_U), F(x_U))$, Mexican immigrants in the United States, $w_{UM} \equiv (w_U(x_U), F_{UM}(x_U))$, and Mexican stayers, $w_M \equiv$

[15] 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: $y(x(t)) = y(x(t-1)) + (x(t) - x(t-1))f(x(t-1))$.

[16] 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 that retain this important feature of the model with (untruncated) log-normal skills.

$(w_M(x_M), F_M(x_M))$.[17]

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 fitting crucial moments in the data. The solution of the model requires finding functions and distributions, for which there exists no analytical solution, therefore to calibrate the model we cannot escape solving it for each guess 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.[18] Our implementation of the random search through the parameter space is in principle a variant of the Simulated Annealing Optimization method.

Each vector Ξ is evaluated using a subjective goal function:[19]

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

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.[20] The $P(\cdot)$ function computes the conditional probabilities of emigration from Mexico (see Equation C.I9), while functions $w_i(\cdot)$ represent the group-specific distributions of wages. The goal function aims at minimizing: (i) the distance between eight model variables 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 United States from the model and from the data (weighted by p_9) and the distances between model and

[17]The proposed notation includes skill CDFs in the analyzed groups of workers. $F_{UM}(x_U) = \left(S_{UM}(x_{UM}^c) - S_{UM}(x_U)\right)/S_{UM}(x_{UM}^c)$, while: $F_M(x_M) = \left(S_M(x_M^c) - S_M(x_M)\right)/S_M(x_M^c)$.

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

[19]$p_1 = p_2 = 100, p_3 = p_4 = p_5 = 1, p_6 = p_7 = 5 \cdot 10^5, p_8 = 10^4, p_9 = 4 \cdot 10^5, p_{10} = 50, p_{11} = 3, p_{12} = 2$.

[20]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.

data wage distributions in three populations (weighted by $p_{10}, ..., p_{12}$).

The proposed Monte Carlo search method assumes the following procedure:

1. Select a randomly drawn guess of parameters Ξ_0.

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

3. Search for a new vector of parameters in a close 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.

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.1. For a graphical analysis of the loss function minimum achieved by the best parameter vector consult Figure C.2, 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 higher dispersion in skills pricing on the American market comparing to Mexico. Then, firms' component in surplus appears to be almost identical in the United States and in Mexico. Interestingly enough, our best calibration returns a rather low value of the copula parameter θ. Its value close to 1 indicates that U.S. and Mexican skills are weakly related with an average rank correlation of 0.33. Migration costs take values in expected ranges: the multiplicative one equals $1 - \Delta_{UM} = 68\%$ of migrant's wage in the United States, while the additive one is $\delta_{UM} = 338$ USD. Trade costs, reported in Table C.2, take values ranging between 1 and 2.1, 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.[21]

Backward Recalibration Our calibration strategy considers a single, cross-sectional snapshot of U.S. and Mexican economies, by fitting nine model parameters to eight discrete moments and a set of conditional probabilities of emigration. This might raise concerns that (i) the calibrated parameters are unstable over time and (ii) that there is no

[21]Note that the fact the trade costs are smaller between the US and ROW than the US and Mexico does not constitute a failure of simple gravity, as the ROW includes many much larger economies than Mexico, one of which (Canada) is also a neighboring country.

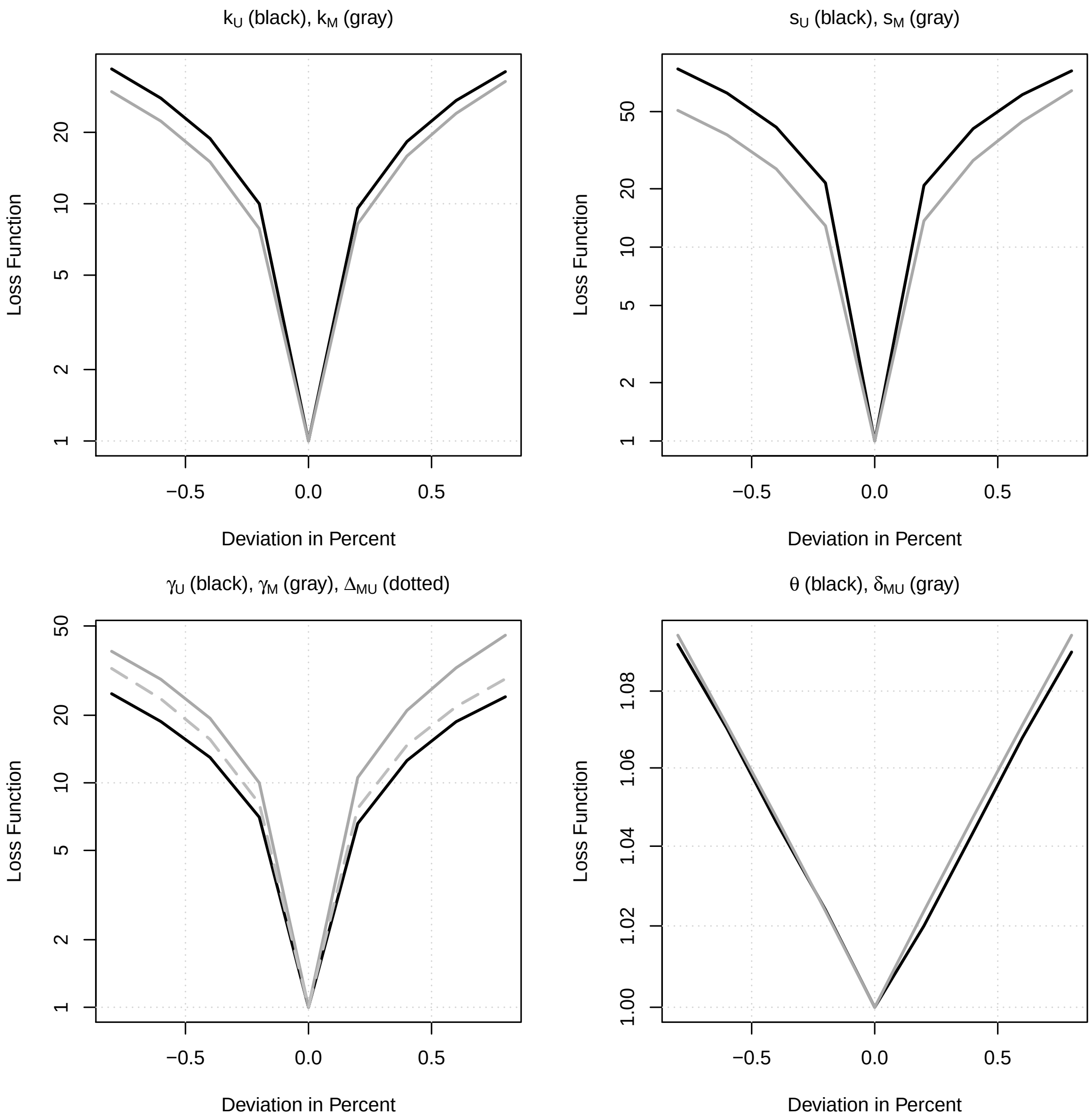


Figure C.2: Evaluation of the Best Parameter Vector

Figure C.2 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).

Table C.1: Calibrated values of parameters

US Market	MEX Market	Migration Parameters
$K_U = 12,457.9$	$K_M = 7,529.1$	$\theta = 0.990$
$s_U = 0.849$	$s_M = 0.462$	$\delta_{UM} = 338.2$
$\gamma_U = 0.354$	$\gamma_M = 0.356$	$\Delta_{UM} = 0.319$

Table C.2: Calibrated trade costs

To:\From:	ROW	MEX	US
ROW	1.00	1.82	1.14
MEX	1.96	1.00	1.39
US	1.99	2.09	1.00

natural external validation of our calibration. To dispel these worries, we investigate the fit of our 2015 model to 2010 data on labor markets, wages and migration. First, we argue that parameters that represent production technology and prices for skills $(s_U, s_M, \gamma_U, \gamma_M)$ are held constant throughout the course of five years. This assumption is motivated by the fact that the standard deviations of wage distributions are almost identical across the two waves, so as the share of firms' profits in GDP. Second, we recalibrate the model using only five model parameters $(k_U, k_M, \delta_{UM}, \Delta_{UM}, \theta)$, and show that we are able to fit the 2010 data with a reduced set of parameters. The outcomes of the calibration are depicted in Figure C.3.

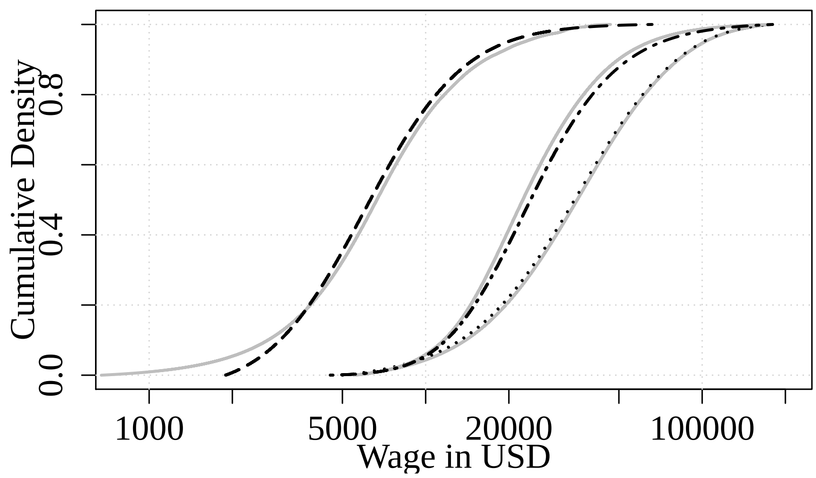


Figure C.3: Model Fit to 2010 data

We observe only slight changes in the values of recalibrated parameters. The multiplicative constant in production function represents total factor productivities in both countries, changes by -11% in the United States and remains roughly identical in Mexico. The utility costs of migration are lower in 2010 by 4 percentage point (of the wage lost when migrating), while monetary costs of migration are larger by approximately 100 USD in 2010. Recall that the Deferred Action for Childhood Arrivals (DACA) legislation was already active in 2015, but not present in 2010, which might explain these differences. Finally, the copula parameter equals 1.2 in 2010, which implies a small increase in the skill correlation among Mexicans by 4 percentage points.

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, while the remaining 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 US, $S_{UM}(0)$, it recomputes the skill and wage distributions, for the new migration costs. Then, separately for each economy, the procedure computes the mass of firms by setting expected profits equl to the fixed costs of entry. 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 United States can be produced. This iterative procedure is continued as long as the aggregated deviation in all endogenous variables in consecutive steps is smaller than $1/K$. In Figure C.4, we present deviations in values of GDPs after recomputing the labor market equilibrium for (non-) equilibrium initial values of GDPs. 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 (in a fairly large neighborhood of the initial equilibrium) which can be computed using an iterative procedure (first solve labor market, than solve international goods market, repeat until convergence).

D Additional Results

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), and functional forms of the distribution of inactive individuals.

Fiscal Effects Mexican migrants in the United States tend to locate in the left tail of wage distributions, consequently they are expected to have a net fiscal contribution different from the one of U.S. or Mexican residents. To quantify the extent to which the U.S. and Mexican fiscal balances change due to Mexican immigration, we add to our model the fiscal extension.[22] Double-dashed black lines in Figure D.1a illustrate that due to Mexican immigration, U.S. residents are forced to pay 94 USD of the budget-balancing lump-sum tax, while Mexican stayers benefit from a 46 USD lump-sum transfer

[22] 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. Regarding the data, we collect income and corporate tax rates and thresholds for the United States and Mexico from the OECD. Finally, we assume balanced governmental budgets, and choose that the lump-sum transfers adjust after shocking the economy.

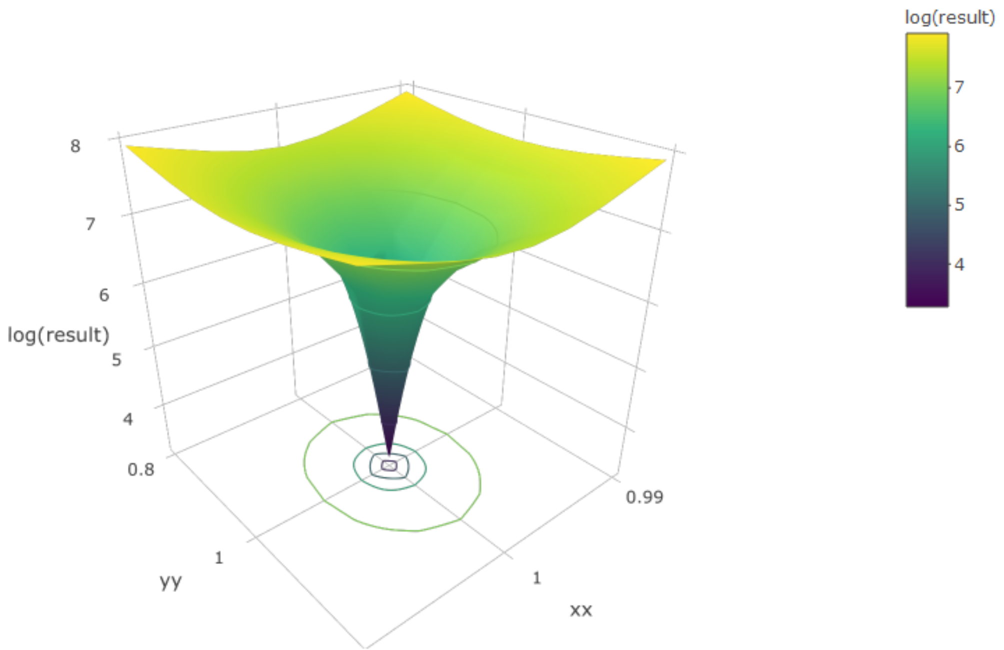


Figure C.4: Uniqueness of Model's General Equilibrium
Figure C.4 illustrates the numerical proof of uniqueness of the model's general equilibrium. The "xx" ("yy") axis represents initial deviations in Mexican (U.S.) GDP, while the vertical axis depict logarithms of Euclidean distances between initial and computed vectors of GDPs. General equilibrium GDP levels are normalized to 1.

(note that the solid light gray lines recall the benchmark results from Figure 3). Despite being quantitatively small, the change in net benefits received by incumbent residents sets the share of winners and losers to approximately 40:60 in the United States and 60:40 in Mexico.

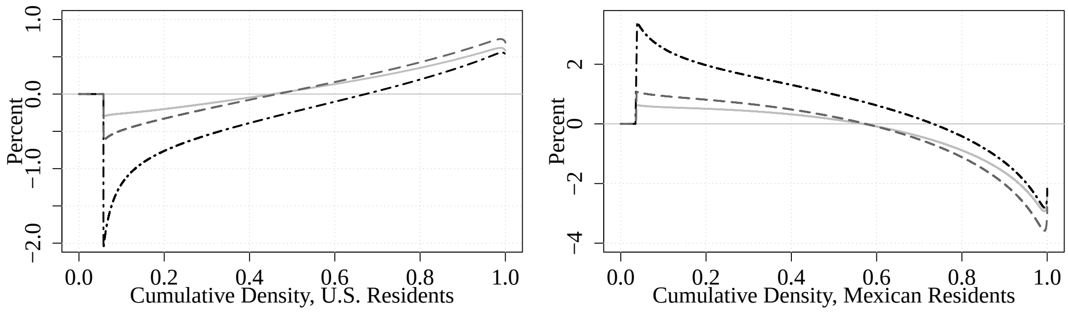
Illegal Mexican Immigration Illegal migration from Mexico to the United States proves to be one of the key points in the overall discussion about American migration policy. To analyze its economic importance, we include estimates of the number of undocumented Mexicans and their wage distribution in our quantifications.[23] The quantitative outcomes of including illegal Mexican immigrants in the no-migration scenario are depicted in Figure D.1a with long-dashed gray lines. The magnitudes of the economic impacts become significantly more pronounced (especially for the low-skilled U.S. residents), while the measures of losers and winners stay virtually unchanged.

[23]We take the number of undocumented Mexicans from the Pew Research Center. The authors calculate that out of 11.7 million Mexican immigrants in the United States 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, in this simulation we increase the number of Mexicans in the US to 10.5 million. Illegal migrants earn substantially lower wages than their legal peers. Caponi and Plesca (2014) compute the wage penalty for illegals along the wage distribution (see their Figure 1), which equals approximately 15-20%, in line with the findings of Massey and Gentsch (2014).

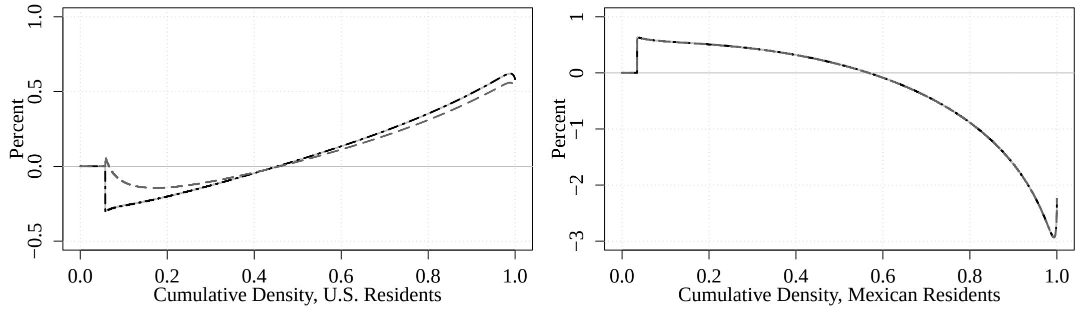
Alternative distributions of inactive individuals' skills Any shock to the supply of skills in the United States affects workers' participation. More precisely, the presence of Mexicans discourages some previously employed Americans to quit the labor market. Importantly, we do not observe the wages (nor the skills) of these inactive 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.1b.

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 model). 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 main results assuming different magnitudes of the market size effect. The solid gray line indicates the reference value of $\varepsilon = 7$; with the dark-gray line, we assume $\varepsilon = 9$; and the black 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 100%.

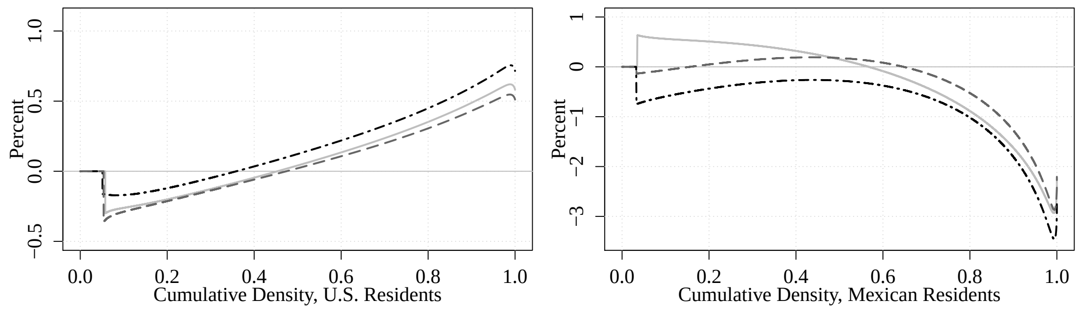
Changing the structure of capital costs One degree of freedom in the calibration process is subject to a broad interpretation of the underlying data. This problem concerns the division between variable and fixed costs of capital, that are necessary to pin down production costs. In the benchmark calibration, we assume that the consumption of fixed capital that relates to structures constitutes the fixed part of capital costs, while equipment and intellectual property 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 dark-gray (black) line. Higher fixed share of capital costs twists the gain distribution clockwise, while higher variable share of costs inflates



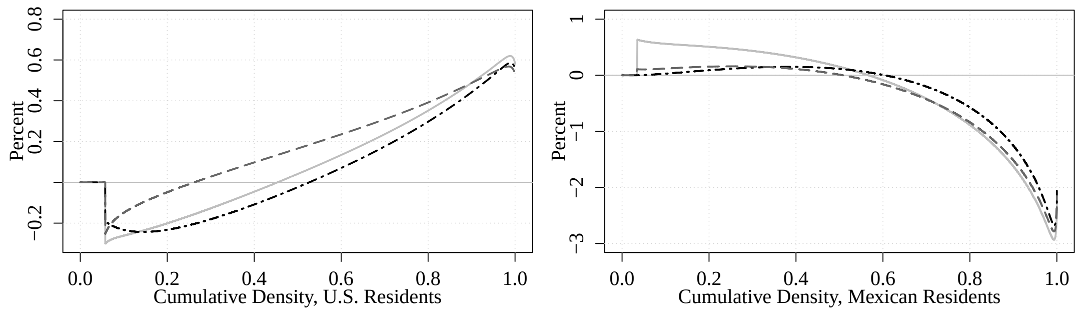
(a) Illegal Migration and Fiscal Effects



(b) Alternative Inactive Workers' Skill Levels



(c) Magnitude of the Market Size Effect



(d) Changes in 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 includes illegal immigrants (gray dashed line) and fiscal effects (black double-dashed line). Figure D.1b experiments with the distribution of skills of inactive workers. The reference scenario (solid gray) assumes a linear CDF, the "convex scenario" (long-dashed dark gray) assumes exponential CDF, while the "concave scenario" (double-dashed black) assumes logarithmic CDF. Figure D.1b assumes alternative values for the elasticity of substitution between varieties (solid gray benchmark: $\varepsilon = 7$, double-dashed black: $\varepsilon = 5$, long-dashed dark gray: $\varepsilon = 9$). Figure D.1c assumes alternative structure of capital and investment costs (solid gray benchmark: fixed costs constitute 35% of capital costs, double-dashed black: 0%, long-dashed gray: 100%).

the magnitudes of extreme effects, but keeps the indifferent individual at around 40th percentile, close to our benchmark result.

Applying redistribution among U.S. citizens Below, we complement the findings of Section 5.2 by deriving the tax rates imposed on all U.S. citizens that finance a lump-sum transfers designed to keep the variance of U.S. citizens wage distribution constant between the reference and the counterfactual scenarios. Note that this redistribution policy does not affect average wages among U.S. citizens. Figure D.2 depicts the outcomes with the long-dashed gray (solid black) line indicating the case of labor market effects (labor market and market size effects). The induced tax rates are almost linear in migration cost liberalizations. In order to maintain the variance of U.S. citizens wage distributions constant, every 100 USD reduction of visa costs should be followed by an increase in proportional income taxes by 0.027 percentage points. The respective number for the full general equilibrium model is 0.02.

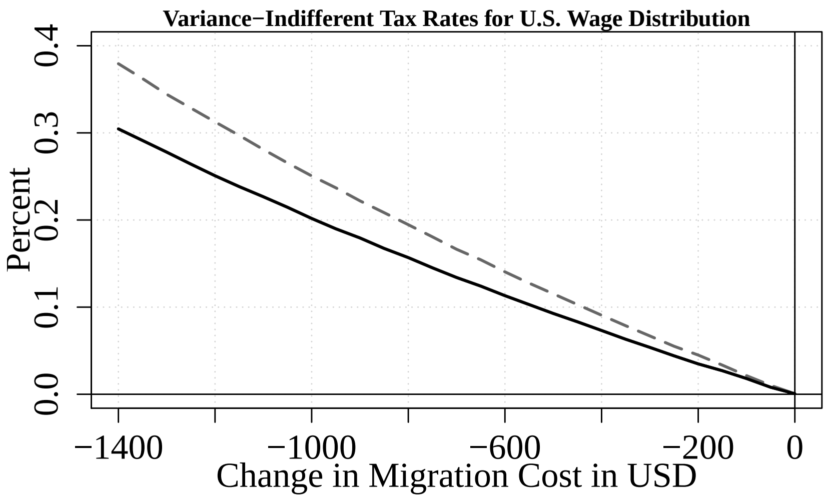


Figure D.2: Redistribution Among U.S. Citizens: Variance-Preserving Tax Rates

Note: Figure D.2 plots the tax rates that compensate for the change in wage inequality generated by more liberal Mexico-US immigration policies. The long-dashed gray line (the solid black line) represents labor market effects (labor and market size effects). Horizontal axes present deviations in monetary costs of migration, δ_{UM}, relative to the *status quo*

References

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.

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

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.

Gola, P. (2021). Supply and demand in a two-sector matching model. *Journal of Political Economy*, 129(3):940 – 978.

Heckman, J. J. and Honoré, 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):482â499.

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.

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.*