8 Appendix to "Intergenerational Mobility and the Political Economy of Immigration"

(Appendix not for publication but intended to be offered online)

by Henning Bohn and Armando R. Lopez-Velasco

8.1 Appendix to section 3.1 Foreign Born Population & Immigration

Using the CPS, the education levels of the foreign born population 25-years and older entering the US each decade since 1980 is shown in table 8. The data is also put in the form of percentages of the respective immigration flows. For example, for the decade of 2000-2009, out of the 8.66 millions of (net) people entering the US, 31.6% were individuals with less than a high school degree, 34.7% had a high-school degree or some college (either education in their home country or some of it in the US) and 33.7% had an advanced degree (BA and above). These flows remained relatively stable since 1980 (with an exception for the 2010-2013 period).

Table 9 displays the shares of the US population 25-years and older by education for the same period as the foreign born population. The percentage of medium-skilled agents is very stable (high school diploma or some college), fluctuating between 54% and 58% for the 1980-2013 period. The shares of people with less than a high-school diploma has been decreasing, with 26.8% of the population in 1980-89, while the share of the population without a high-school diploma is 12.4% in 2010-2013. The share of the US population with a BA degree or above increased from 19% in 1980-89 to 30.7% in 2010-2013.

Accurate estimates of the net annual flow of immigrants (without conditioning by education) to the US are presented in Blau and Mackie, editors (2017). Table 8 also present the estimates for the same periods discussed (estimates available until 2013). Using those estimates, together with the rest of the information in tables 7 and 8 we are able to compute the estimates by decade presented in table 1.

Table 8. **Number of Foreign people entering the US by decade, by education level (in 1000's), 1980-2013**

	2010-13^a	(%)	2000-09	(%)	1990-99	(%)	1980-89	(%)
Less than High School	1917.5	21.1	2503	31.6	2831	32.7	2204	31.5
High School & Some College	4160	45.8	2743	34.7	3519	40.6	2849	40.7
BA degree & Above	3002.5	33.1	2665	33.7	2314	26.7	1951	27.9
Total	9080	100	7911	100	8664	100	7004	100

Source: U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplement, 2009 and 2013. Table 2.5.

Notes (a) Flow put in the form of flow/decade

Table 9. **US average population shares by education and decade, 1980-2013**

	2010-13	2000-09	1990-99	1980-89
Less than High School	12.4	14.8	19.2	26.8
High School & Some College	56.9	57.5	58.0	54.2
BA degree & Above	30.7	27.7	22.8	19.0
Total	100	100	100	100
Annual immigration/1000 natives[a]	3.1	3.29	4.8	2.8

Source population shares: U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplement. Table A-1, 2013 . Source net annual immigration to the US: Blau, Francine and Christopher Mackie, editors (2017). The Economic and Fiscal Consequences of Immigration. Washington, DC:The National Academies Press. doi: https://doi.org/10.17226/23550.

8.2 Appendix to section 2.2. Ruling out negative wage premiums

All agents works at their own (highest) skill level if the resulting "unconstrained" wages satisfy $w_1 \leq w_2 \leq w_3$. Otherwise, it is efficient (both output maximizing and individually rational) for some agents to take jobs at a lower skill level.

(Note to readers: The appendix is lengthy because we feel obliged to cover all theoretically possible cases. Readers interested in the cases discussed in the text may focus on labor supply ratios Z in the unconstrained set Ω_Z^u, both defined below.)

To model job assignment in general, we distinguish labor supplies $L_i = N_i(1+\theta_i)$ from actual labor inputs $\hat{L}_i$, $i = 1, 2, 3$. Let $\hat{L}_{ij} \geq 0$ denote labor supplied by workers with skill i at skill level $j < i$. (Throughout, $i, j = 1, 2, 3$ unless noted.) Then $\hat{L}_1 = L_1 + \hat{L}_{21} + \hat{L}_{31}$, $\hat{L}_2 = L_2 - \hat{L}_{21} + \hat{L}_{32}$, and $\hat{L}_3 = L_3 - \hat{L}_{31} - \hat{L}_{32}$. Note that $\hat{L}_{31}$ is redundant because any set of labor inputs $(\hat{L}_1, \hat{L}_2, \hat{L}_3)$ obtained with $\hat{L}_{31} > 0$ can also be obtained with $\hat{L}_{31} = 0$ if $\hat{L}_{32}$ and $\hat{L}_{21}$ are increased by the original amount of $\hat{L}_{31}$. Hence we set $\hat{L}_{31} = 0$ without loss of generality.

Assuming CES production, output is $Y = [\sum_i \phi_i(\hat{L}_i)^\rho]^{\frac{1}{\rho}}$. The resulting wages can be written as

$$w_i = \frac{\partial Y}{\partial \hat{L}_i} = Y^{1-\rho} \cdot \left(\hat{L}_i/\phi_i^{\frac{1}{1-\rho}}\right)^{\rho-1} \tag{A.1}$$

For given (L_1, L_2, L_3), maximizing Y by choice of $(\hat{L}_{32}, \hat{L}_{21}) \geq 0$ implies the Kuhn-Tucker conditions $\hat{L}_{32} \cdot (w_3 - w_2) = 0$, $\hat{L}_{21} \cdot (w_2 - w_1) = 0$, and $w_3 \geq w_2 \geq w_1$. With two sets of Kuhn-Tucker conditions, the solutions divide into four cases: (1) $\hat{L}_{32} = 0$ and $\hat{L}_{21} = 0$, which is the case of "unconstrained" wages; (2) $\hat{L}_{32} > 0$ and $\hat{L}_{21} = 0$, which implies $w_2 = w_3$; (3) $\hat{L}_{32} = 0$ and $\hat{L}_{21} > 0$, which implies $w_2 = w_1$; (4) $\hat{L}_{32} > 0$ and $\hat{L}_{21} > 0$, which implies $w_1 = w_2 = w_3$.

Wages depend naturally on labor supplies but can be written more compactly in terms of the ratios $z_1 = L_1/L_2$ and $z_3 = L_3/L_2$, exploiting constant returns to scale. The latter is also useful because

$$z_1 = \frac{x_1\,(1+\theta_1)}{(1+\theta_2)} \text{ and } z_3 = \frac{x_3\,(1+\theta_3)}{(1+\theta_2)}$$

are functions of the state (X,θ), whereas L_i depends on native population N_i, which is not part of the Markov state.

Define the vector of labor supply ratios

$$Z = (z_1, z_3) = Z(X,\theta),$$

total labor $L = L_1 + L_2 + L_3$, constants $\phi_{12} = (\phi_1^{\frac{1}{1-\rho}} + \phi_2^{\frac{1}{1-\rho}})^{1-\rho}$, $\phi_{23} = (\phi_2^{\frac{1}{1-\rho}} + \phi_3^{\frac{1}{1-\rho}})^{1-\rho}$, and $\phi_{123} = (\phi_1^{\frac{1}{1-\rho}} + \phi_2^{\frac{1}{1-\rho}} + \phi_3^{\frac{1}{1-\rho}})^{1-\rho}$, the dummy variable $z_2 \equiv 1$, and the sets

$$\begin{aligned}
\Omega_Z^u &= \{Z : z_3 \le \bar{z}_3, z_1 \ge \underline{z}_1\}, \text{ where } \underline{z}_1 = (\phi_1/\phi_2)^{\frac{1}{1-\rho}}\,,\ \bar{z}_3 = (\phi_3/\phi_2)^{\frac{1}{1-\rho}} \\
\Omega_Z^{23} &= \{Z : z_3 > \bar{z}_3, z_1 \ge \underline{z}_{13}(1+z_3)\}, \text{ where } \underline{z}_{13} = (\phi_1/\phi_{23})^{\frac{1}{1-\rho}}\,, \\
\Omega_Z^{12} &= \{Z : z_3 \le \bar{z}_{31}(1+z_1), z_1 < \underline{z}_1\}, \text{ where } \bar{z}_{31} = (\phi_3/\phi_{12})^{\frac{1}{1-\rho}}\,, \\
\Omega_Z^{123} &= \{Z : z_3 > \bar{z}_{31}(1+z_1), z_1 < \underline{z}_{13}(1+z_3)\}.
\end{aligned}$$

Note that $\underline{z}_{13}(1+\bar{z}_3) = \underline{z}_1$ and $\bar{z}_{31}(1+\underline{z}_1) = \bar{z}_3$. Hence the closures of the four sets above intersect at $(\underline{z}_1, \bar{z}_3)$, where $L_i/\phi_i^{\frac{1}{1-\rho}} = L/\phi_{123}^{\frac{1}{1-\rho}}$ for all i, so unconstrainted wages satisfy $w_3 = w_2 = w_1$ without job reassignments. Wage differentials arise if high skills are relatively more scarce; job reassigments must occur if high skills are in relatively greater supply. Specifically:

Lemma 2.2A (Four Cases): Exactly one of the following four cases applies for all Z:

(1) If $Z \in \Omega_Z^u$, wages are unconstrained, $Y = [\sum_i \phi_i\,(L_i)^\rho]^{\frac{1}{\rho}}$, and $\hat{L}_{21} = \hat{L}_{32} = 0$. Moreover, $z_3 < \bar{z}_3$ implies $w_2 < w_3$, and $z_1 > \underline{z}_1$ implies $w_1 < w_2$.

(2) If $Z \in \Omega_Z^{23}$, then $w_2 = w_3$, $Y = [\phi_1\,(L_1)^\rho + \phi_{23}(L_2+L_3)^\rho]^{\frac{1}{\rho}}$, $\hat{L}_{21} = 0$, and $\hat{L}_{32} > 0$. Moreover, $z_1 > \underline{z}_{13}(1+z_3)$ implies $w_2 < w_3$.

(3) If $Z \in \Omega_Z^{12}$, then $w_1 = w_2$, $Y = [\phi_{12}(L_1+L_2)^\rho + \phi_3\,(L_3)^\rho]^{\frac{1}{\rho}}$, $\hat{L}_{21} > 0$ and $\hat{L}_{32} = 0$. Moreover, $z_3 < \bar{z}_{31}(1+z_1)$ implies $w_1 < w_2$.

(4) If $Z \in \Omega_Z^{123}$, then $w_1 = w_2 = w_3$, $Y = (\phi_{123})^{\frac{1}{\rho}}\,L$, $\hat{L}_{21} > 0$, and $\hat{L}_{32} > 0$.

Proof: Since the four sets partition the space $\{Z : z_1 \ge 0, z_3 \ge 0\}$, exactly one applies for any Z. (1) From (A.1), $w_i = w_j$ iff $\hat{L}_i/\phi_i^{\frac{1}{1-\rho}} = \hat{L}_j/\phi_j^{\frac{1}{1-\rho}}$. Since $\rho - 1 < 0$, $w_i \ge w_j$ iff $\hat{L}_i/\phi_i^{\frac{1}{1-\rho}} \le \hat{L}_j/\phi_j^{\frac{1}{1-\rho}}$. If $z_3 \le (\phi_3/\phi_2)^{\frac{1}{1-\rho}} = \bar{z}_3$ and $z_1 \ge (\phi_1/\phi_2)^{\frac{1}{1-\rho}} = \underline{z}_1$, then $L_3/\phi_3^{\frac{1}{1-\rho}} \le L_2/\phi_2^{\frac{1}{1-\rho}} \le L_1/\phi_1^{\frac{1}{1-\rho}}$, so $\hat{L}_i = L_i$ is consistent with $w_3 \ge w_2 \ge w_1$. By analogous reasoning, strict inequalities $z_3 < \bar{z}_3$ and/or $z_1 > \underline{z}_1$ imply the corresponding strict inequalities for wages.

(2) If $z_3 > \bar{z}_3$, then $L_3/\phi_3^{\frac{1}{1-\rho}} > L_2/\phi_2^{\frac{1}{1-\rho}}$ and $\hat{L}_3/\phi_3^{\frac{1}{1-\rho}} \le \hat{L}_2/\phi_2^{\frac{1}{1-\rho}}$ imply $\hat{L}_{32} > 0$, so $w_2 = w_3$. In turn, $w_2 = w_3$ implies $(L_3 - \hat{L}_{32})/\phi_3^{\frac{1}{1-\rho}} = (L_2 + \hat{L}_{32} - \hat{L}_{21})/\phi_2^{\frac{1}{1-\rho}}$, so $\frac{\hat{L}_{32}}{L_2} = \frac{\phi_2^{\frac{1}{1-\rho}} z_3 - \phi_3^{\frac{1}{1-\rho}} + \phi_3^{\frac{1}{1-\rho}} \frac{\hat{L}_{21}}{L_2}}{\phi_2^{\frac{1}{1-\rho}} + \phi_3^{\frac{1}{1-\rho}}} \ge (\frac{\phi_2}{\phi_{23}})^{\frac{1}{1-\rho}} \cdot z_3 - (\frac{\phi_3}{\phi_{23}})^{\frac{1}{1-\rho}}$ and $\frac{\hat{L}_3}{L_2} = z_3 - \frac{\hat{L}_{32}}{L_2} \le (\phi_3/\phi_{23})^{\frac{1}{1-\rho}} \cdot (1+z_3)$.

Suppose for contradiction that $\hat{L}_{21} > 0$. Then $\frac{\hat{L}_1}{L_2} = z_1 + \frac{\hat{L}_{21}}{L_2} > z_1$, and $\frac{\hat{L}_{32}}{L_2} \geq (\frac{\phi_2}{\phi_{23}})^{\frac{1}{1-\rho}} \cdot z_3 - (\frac{\phi_3}{\phi_{23}})^{\frac{1}{1-\rho}}$ would imply $\frac{\hat{L}_3}{\hat{L}_1} = \frac{\hat{L}_3}{L_2}/\frac{\hat{L}_1}{L_2} < (\phi_3/\phi_1)^{\frac{1}{1-\rho}}$ and hence $w_1 < w_3 = w_2$, violating the Kuhn-Tucker condition. By contradiction, $\hat{L}_{21} = 0$, which implies $\frac{\hat{L}_3}{L_2} = (\phi_3/\phi_{23})^{\frac{1}{1-\rho}}(1+z_3)$ and $\frac{\hat{L}_2}{L_2} = (\phi_2/\phi_{23})^{\frac{1}{1-\rho}}(1+z_3)$. With these labor inputs $\hat{L}_i$, $Y = [\sum_i \phi_i (L_i)^\rho]^{\frac{1}{\rho}} = [\phi_1 (L_1)^\rho + \phi_{23} (L_2 + L_3)^\rho]^{\frac{1}{\rho}}$.

(3) The proof for $z_1 < \underline{z}_1$ and $z_3 \leq \bar{z}_{31}(1+z_1)$ is analogous to case (2), with analogous steps, first showing that $z_1 < \underline{z}_1$ implies $\hat{L}_{21} > 0$, and then showing that $\hat{L}_{32} > 0$ would lead to a contradition, so $\hat{L}_{32} = 0$.

(4) Conditions $z_3 > \bar{z}_{31}(1+z_1)$ and $z_1 < \underline{z}_{13}(1+z_3)$ imply $z_3/z_1 = L_3/L_1 > \phi_3^{\frac{1}{1-\rho}}/\phi_1^{\frac{1}{1-\rho}}$. Since $\hat{L}_3/\phi_3^{\frac{1}{1-\rho}} \leq \hat{L}_1/\phi_1^{\frac{1}{1-\rho}}$ one cannot have $\hat{L}_{32} = \hat{L}_{21} = 0$. By similar reasoning $\frac{z_3}{1+z_1} = \frac{L_3}{L_1+L_2} > \bar{z}_{31}$ rules out $\hat{L}_{32} = 0$ and $\frac{z_1}{1+z_3} = \frac{L_1}{L_3+L_2} < \underline{z}_{13}$ rules out $\hat{L}_{12} = 0$. Thus $\hat{L}_{32} > 0$ and $\hat{L}_{21} > 0$, hence $w_1 = w_2 = w_3$. Equality of $\hat{L}_i/\phi_i^{\frac{1}{1-\rho}}$ for all i then implies $\hat{L}_i/\phi_i^{\frac{1}{1-\rho}} = L/\phi_{123}^{\frac{1}{1-\rho}}$, $\hat{L}_i = \phi_i^{\frac{1}{1-\rho}}/\phi_{123}^{\frac{1}{1-\rho}} \cdot L$, and $Y = (\phi_{123})^{\frac{1}{\rho}} L$. QED.

Lemma 2.2B (Wages): (1) For $Z \in \Omega_z^u$: $w_2 = \phi_2 \{\phi_1 z_1^\rho + \phi_2 + \phi_3 z_3^\rho\}^{\frac{1-\rho}{\rho}}$, $w_1 = w_2 \cdot \frac{\phi_1}{\phi_2} \cdot z_1^{\rho-1} \leq w_2$, and $w_3 = w_2 \cdot \frac{\phi_3}{\phi_2} \cdot z_3^{\rho-1} \geq w_2$.

(2) For $Z \in cl(\Omega_z^{23})$, where the closure covers cases with $w_2 = w_3$ in Ω_z^u: Define $z_{13} \equiv \frac{z_1}{1+z_3}$, then $w_2 = w_3 = \phi_{23} \{\phi_1 (z_{13})^\rho + \phi_{23}\}^{\frac{1-\rho}{\rho}}$ and $w_1 = w_2 \cdot \frac{\phi_1}{\phi_{23}} \cdot (z_{13})^{\rho-1} \leq w_2$.

(3) For $Z \in cl(\Omega_z^{12})$, where the closure covers cases with $w_1 = w_2$ in Ω_z^u: Define $z_{31} \equiv \frac{z_3}{1+z_1}$, then $w_2 = w_1 = \phi_{12} \{\phi_3 (z_{31})^\rho + \phi_{12}\}^{\frac{1-\rho}{\rho}}$ and $w_3 = w_2 \cdot \frac{\phi_3}{\phi_{12}} \cdot (z_{31})^{\rho-1} \geq w_2$.

(4) For $Z \in cl(\Omega_z^{123})$: $w_1 = w_2 = w_3 = (\phi_{123})^{\frac{1}{\rho}}$.

Proof: Follows from differentiating the equations for Y in Lemma 2.2A and verifying that wages are continous at boundaries between cases. QED.

8.3 Appendix to section 2.6. The model without dynamic effects. Proofs and generalizations

8.3.1 Cases discussed in the text

We prove the claims in section 2.6 first for $w_1 < w_2 < w_3$ and $w_2 \leq \bar{w}$, and then generalize. The wages w_2 and $\bar{w}$ that determines voter consumption c_2 have the following properties:

Lemma 2.6A (Derivatives of w_2): (1) For $Z \in int(\Omega_z^u)$: $\frac{\partial w_2}{\partial z_1} = (1-\rho)\frac{w_2 w_1}{y} > 0$ and $\frac{\partial w_2}{\partial z_3} = (1-\rho)\frac{w_2 w_3}{y} > 0$, where $y = Y/L_2 = \{\phi_1 z_1^\rho + \phi_2 + \phi_3 z_3^\rho\}^{\frac{1}{\rho}}$. Moreover, $\frac{\partial w_2}{\partial \theta_1}, \frac{\partial w_2}{\partial \theta_3} > 0$ and $\frac{\partial w_2}{\partial \theta_2} < 0$.

(2) For $Z \in int(\Omega_z^{23})$: $\frac{\partial w_2}{\partial z_{13}} = (1-\rho)\frac{w_2 w_1}{y} > 0$, $\frac{\partial w_2}{\partial \theta_1} > 0$, and $\frac{\partial w_2}{\partial \theta_2}, \frac{\partial w_2}{\partial \theta_3} < 0$.

(3) For $Z \in int(\Omega_z^{12})$: $\frac{\partial w_2}{\partial z_{31}} = (1-\rho)\frac{w_2 w_3}{y} > 0$, $\frac{\partial w_2}{\partial \theta_3} > 0$, and $\frac{\partial w_2}{\partial \theta_1}, \frac{\partial w_2}{\partial \theta_2} < 0$.

(4) For $Z \in int(\Omega_z^{123})$: $\frac{\partial w_2}{\partial z_1} = \frac{\partial w_2}{\partial z_3} = 0$ and $\frac{\partial w_2}{\partial \theta_i} = 0$ for all i.

Proof: Derivatives with respect to z_i follow from differentiating the equations for w_2 in Lemma 2.2B. Derivatives with respect to θ_i follow, using $\frac{\partial z_i}{\partial \theta_2} = -\frac{L_i}{L_2^2}\frac{\partial L_2}{\partial \theta_2} = -\frac{N_2 L_i}{L_2^2} < 0$, $\frac{\partial z_i}{\partial \theta_i} = \frac{1}{L_2}\frac{\partial L_i}{\partial \theta_i} = \frac{N_i}{L_2} > 0$

for $i=1,3, \frac{\partial z_{13}}{\partial z_1} = \frac{1}{1+z_3} > 0$, $\frac{\partial z_{13}}{\partial z_3} = -\frac{z_{13}}{1+z_3} < 0$, $\frac{\partial z_{31}}{\partial z_3} = \frac{1}{1+z_1} > 0$, $\frac{\partial z_{31}}{\partial z_1} = -\frac{z_{31}}{1+z_1} > 0$, and applying the chain rule. QED.

Remark: Since the derivatives differ across cases, w_2 is generally not differentable at the boundaries. For example, $\frac{\partial z_{13}}{\partial z_1} = \frac{1}{1+z_3}$ in Ω_z^{23} implies $\frac{\partial w_2}{\partial z_1} = (1-\rho)\frac{w_2 w_1}{y(1+z_3)}$ which differs by a scale factor from $\frac{\partial w_2}{\partial z_1}$in Ω_z^u.

Lemma 2.6B (Average wage $\bar{w}$): The average wage $\bar{w} = \sum_i \frac{L_i}{L} w_i$ can be expressed as

$$\bar{w} = \sum_i \frac{z_i}{1+z_1+z_3} w_i = \frac{y}{1+z_1+z_3} = \frac{1}{1+z_1+z_3}\left\{\phi_1 z_1^\rho + \phi_2 + \phi_3 z_3^\rho\right\}^{\frac{1}{\rho}},$$

and its derivatives are

$$\frac{\partial \bar{w}}{\partial z_i} = \frac{w_i - \bar{w}}{1+z_1+z_3} \text{ for } i=1,3, \text{ and } \frac{\partial \bar{w}}{\partial \theta_i} = \frac{N_i}{L}(w_i - \bar{w}) \text{ for all } i$$

Proof: $\bar{w} = \frac{1}{L}\sum_i L_i w_i = \frac{Y}{L}$ by Euler's law; $\bar{w} = \sum_i \frac{L_i/L_2}{L/L_2} w_i = \sum_i \frac{z_i}{1+z_1+z_3} w_i$ and $\frac{Y}{L} = \frac{y}{1+z_1+z_3}$ follow. Differentiating $\bar{w} = \frac{1}{L}\sum_j L_j w_j$ with respect to w_i yields $\frac{\partial \bar{w}}{\partial L_i} = \frac{w_i}{L} + \frac{1}{L}\sum_j L_j \frac{\partial w_j}{\partial L_i} - \frac{1}{L^2}\sum_j L_j w_j = \frac{1}{L}(w_i - \bar{w})$ because $\sum_j L_j \frac{\partial w_j}{\partial L_i} = 0$ by constant returns to scale and $\frac{1}{L^2}\sum_j L_j w_j = \frac{\bar{w}}{L}$. Differentiating $\frac{y}{1+z_1+z_3}$ yields $\frac{\partial \bar{w}}{\partial z_i} = \frac{1}{1+z_1+z_3}\frac{\partial y}{\partial z_i} - \frac{y}{(1+z_1+z_3)^2} = \frac{w_i - \bar{w}}{1+z_1+z_3}$ since $\frac{\partial y}{\partial z_i} = w_i$. Finally, $\frac{\partial \bar{w}}{\partial \theta_i} = N_i \frac{\partial \bar{w}}{\partial L_i}$ since $L_i = N_i(1+\theta_i)$. QED.

Remark: Lemma 2.6B shows that $\frac{\partial \bar{w}}{\partial \theta_i} > 0$ iff $w_i > \bar{w}$, as claimed in the text.

Lemma 2.6C: Suppose $w_1 < w_2 < w_3$ and $w_2 \le \bar{w}$. Then (a) $\frac{\partial c_2}{\partial \theta_3} > 0$, (b) $\frac{\partial c_2}{\partial \theta_2} < 0$, and (c) $\frac{\partial c_2}{\partial \theta_1} > 0$ for τ in a neighborhood of zero and $\frac{\partial c_2}{\partial \theta_1} < 0$ for τ in a neighborhood of one.

Proof: Differentiating $c_2(Z) = (1-\tau)w_2 + \tau\bar{w}$ as function of $Z = Z(X,\theta)$, one obtains

$$\frac{\partial c_2}{\partial \theta_i} = (1-\tau)\frac{\partial w_2}{\partial \theta_i} + \tau\frac{\partial \bar{w}}{\partial \theta_i},$$

where Lemma 2.6A implies $\frac{\partial w_2}{\partial \theta_1}, \frac{\partial w_2}{\partial \theta_3} > 0$ and $\frac{\partial w_2}{\partial \theta_2} < 0$, since $Z \in int(\Omega_z^u)$ from $w_1 < w_2 < w_3$. For $w_2 \le \bar{w}$, Lemma 2.6B implies $\frac{\partial \bar{w}}{\partial \theta_1} < 0$, $\frac{\partial \bar{w}}{\partial \theta_2} \le 0$ and $\frac{\partial \bar{w}}{\partial \theta_3} > 0$. Combining terms, (a) follows from $\frac{\partial w_2}{\partial \theta_3} > 0$ and $w_3 > \bar{w}$; (b) follows from $\frac{\partial w_2}{\partial \theta_2} < 0$ and assumption $w_2 \le \bar{w}$; and $\frac{\partial c_2}{\partial \theta_1}$ is a weighted average of $\frac{\partial w_2}{\partial \theta_1} > 0$ and $\frac{N_i}{L}(w_1 - \bar{w}) < 0$ with weights $(1-\tau, \tau)$. Hence $\frac{\partial c_2}{\partial \theta_1} > 0$ as $\tau \to 0$ and $\frac{\partial c_2}{\partial \theta_1} < 0$ as $\tau \to 1$, proving (c). QED.

Corollary to 2.6C: Suppose $w_1 < w_2 < w_3$ and $w_2 \le \bar{w}$ for given X and for all $\theta \in \Omega_\theta$, and $\beta = 0$. Then all optimal policies θ^* satisfy $\theta_2^* = 0$ and $\theta_3^* = \theta_3^{\max}$; and Ortega's (2005) tie-breaking convention to minimize population selects a unique policy (i.e., unique θ_1^*).

Proof: For $\beta = 0$, problem (11) reduces to finding $\theta^* \in \arg\max_{\Theta\in\Omega_\Theta}\{u_2(X,\theta)\}$, which further reduces to finding $\theta^* \in \arg\max_{\Theta\in\Omega_\Theta}\{c_2\}$. Since Lemma 2.6C implies $\frac{\partial c_2}{\partial \theta_3} > 0$ and $\frac{\partial c_2}{\partial \theta_2} < 0$, $\arg\max_{\Theta\in\Omega_\Theta}\{c_2\}$ is attained at the corners $\theta_2^* = 0$ and $\theta_3^* = \theta_3^{\max}$. Since optimal policies vary only by θ_1^*, Ortega's convention selects the unique minimum of θ_1^*. QED.

Remark: Since $\theta_3^* = \theta_3^{\max}$ and $w_1 < w_2 < w_3$ requires $z_3 < \bar{z}_3$, the corollary can only apply if

$x_3(1+\theta_3^{\max}) < \bar{z}_3$, i.e., if $\theta_3^{\max}$ is sufficiently small.

Lemma 2.6D: Suppose $w_1 < w_2 < w_3$ and $w_2 \leq \bar{w}$ for all $\theta \in \Omega_\theta$, given X and assuming $\beta = 0$. Consider a range of tax rates $(\tau_-, \tau_+) \subset [0,1]$ and the associated optimal policies $\theta^*(\tau)$, where in case of multiple solutions, $\theta^*(\tau)$ is selected by Ortega's (2001) tie-breaking convention to minimize population. Then $\theta_1^* = \theta_1^*(\tau)$ is decreasing in τ; and strictly decreasing at interior solutions $\theta_1^*(\tau) \in (0, \theta_1^{\max})$.

Proof: The Corollary to 2.6C implies $\theta_2^* = 0$ and $\theta_3^* = \theta_3^{\max}$ for all $\tau \in (\tau_-, \tau_+)$. Let $c_2(\theta_1, \tau)$ denote consumption implied by immigration $\theta = (\theta_1, 0, \theta_3^{\max})$. By the theorem of the maximum, $\tilde{\theta}_1^*(\tau) \equiv \arg\max_{\theta_1 \in [0, \theta_1^{\max}]} \{c_2(\theta_1, \tau)\}$ is a non-empty u.h.c. correspondence.

Consider any pair $\tau_a < \tau_b$ and any $\theta_{1a} \in \tilde{\theta}_1^*(\tau_a)$, $\theta_{1b} \in \tilde{\theta}_1^*(\tau_b)$. Optimality implies $c_2(\theta_{1b}, \tau_b) \geq c_2(\theta_{1a}, \tau_b)$ and $c_2(\theta_{1a}, \tau_a) \geq c_2(\theta_{1b}, \tau_a)$. Since c_2 is continuous and differentiable,

$$\begin{aligned} & c_2(\theta_{1b}, \tau_b) - c_2(\theta_{1a}, \tau_b) - c_2(\theta_{1b}, \tau_a) + c_2(\theta_{1a}, \tau_a) \\ & = \int_{\tau_a}^{\tau_b} \int_{\theta_{1a}}^{\theta_{1b}} \frac{\partial^2 c_2(\theta_1, \tau)}{\partial \tau \partial \theta_1} d\theta_1 d\tau = \frac{\partial^2 c_2(\theta_{1m}, \tau_m)}{\partial \tau \partial \theta_1} \cdot (\theta_{1b} - \theta_{1a}) \cdot (\tau_b - \tau_a) \geq 0 \end{aligned}$$

where $\tau_m \in (\tau_a, \tau_b)$ and θ_{1m} is between θ_{1a} and θ_{1b}, using the mean value theorem. Differentiating $\frac{\partial c_2}{\partial \theta_1}$ in Lemma 2.6C, $\frac{\partial^2 c_2(\theta_1, \tau)}{\partial \tau \partial \theta_1} = -\frac{\partial w_2}{\partial \theta_1} + \frac{\partial \bar{w}}{\partial \theta_1} < 0$ for all (θ_1, τ). Since $\tau_b - \tau_a > 0$, $\frac{\partial^2 c_2(\theta_{1m}, \tau_m)}{\partial \tau \partial \theta_1} < 0$ implies $\theta_{1b} \leq \theta_{1a}$, proving that $\theta_1^*(\tau)$ is decreasing.

Note that $\frac{\partial c_2}{\partial \theta}(\theta_{1b}, \tau_b) = \frac{\partial c_2}{\partial \theta}(\theta_{1b}, \tau_a) + \int_{\tau_a}^{\tau_b} \frac{\partial^2 c_2(\theta_{1b}, \tau)}{\partial \tau \partial \theta_1} d\tau < \frac{\partial c_2}{\partial \theta}(\theta_{1b}, \tau_a)$. Hence $\theta_{1b} = \theta_{1a}$ would be inconsistent with the first order condition $\frac{\partial c_2}{\partial \theta} = 0$ for interior solutions. Hence $\theta_{1b} < \theta_{1a}$ unless $\theta_{1a} = \theta_{1b} = 0$ or $\theta_{1a} = \theta_{1b} = \theta_1^{\max}$. Finally, uniqueness of $\theta_2^* = 0$ and $\theta_3^* = \theta_3^{\max}$ implies that if $\tilde{\theta}_1^*(\tau)$ has multiple values, population is minimized by setting $\theta_1^*(\tau) = \min \tilde{\theta}_1^*(\tau)$. QED.

Lemmas 2.6A-D prove claims (a)-(c) in Section 2.6. Note that Lemma 2.6D would be much easier to prove if $\partial^2 c_2 / \partial \theta_1^2 < 0$, using the implicit function theorem; but $\partial^2 c_2 / \partial \theta_1^2 < 0$ may not always hold. Note also that since the proof uses $\tilde{\theta}_1^*(\tau)$, the tie-breaking convention $\theta_1^*(\tau) = \min \tilde{\theta}_1^*(\tau)$ is without loss of generality.

8.3.2 Generalizations

Now consider the static model without assumptions $w_1 < w_2 < w_3$ and $w_2 \leq \bar{w}$. (Note to readers: We provide this analysis mainly to motivate our focus on small $\theta_3^{\max}$ in settings (II) and (III). That is, this appendix section is meant to show that while it is possible to prove results for high $\theta_3^{\max}$, the analysis is much more complicated and does not provide significant new insights. Readers may skip this subsection without loss of continuity.)

(1) Regarding $w_2 \lesseqgtr \bar{w}$, note that $w_2 \leq \bar{w}$ was invoked only in the proof of Lemma 2.6C to sign $\frac{\partial c_2}{\partial \theta_2} < 0$. If $w_2 > \bar{w}$ for some (X, θ), the sign of $\frac{\partial c_2}{\partial \theta_2}$ becomes ambiguous and hence $\theta_2^* > 0$ cannot be ruled out. However, θ_1^* and θ_2^* cannot both be positive, i.e., $\theta_1^* \theta_2^* = 0$. The proof is implied by a more general insight:

Lemma 2.6E: Derivatives of $c_i = (1-\tau)w_i + \tau\bar{w}$ satisfy $\sum_j L_j N_j \frac{\partial c_i}{\partial \theta_j} = 0$ whenever wages are differentiable.

Proof: Consider $\sum_j \frac{\partial c_i}{\partial L_j} L_j = (1-\tau)\sum_j \frac{\partial w_i}{\partial L_j} L_j + \tau \sum_j \frac{\partial \bar{w}}{\partial L_j} L_j$. Constant returns to scale imply $\sum_i \frac{\partial w_2}{\partial L_i} L_i = 0$. From Lemma 2.6B, $\sum_i \frac{\partial \bar{w}}{\partial L_i} L_i = \sum_i \frac{1}{L}(w_i - \bar{w})L_i = \sum_i \frac{L_i}{L} w_i - \bar{w} = 0$. Hence $\sum_j \frac{\partial c_i}{\partial L_j} L_j = 0$. Since $\frac{\partial c_i}{\partial L_j} = N_j \frac{\partial c_i}{\partial \theta_j}$, $\sum_j L_j N_j \frac{\partial c_i}{\partial \theta_j} = 0$. QED.

Corollary to 2.6E: Suppose $w_2 < w_3$ for given X and for all $\theta \in \Omega_\theta$, and $\beta = 0$. Then all $\theta^* = P(X)$ satisfty $\theta_3^* = \theta_3^{\max}$ and $\theta_1^* \theta_2^* = 0$.

Proof: Since $w_2 < w_3$ implies $w_3 > \bar{w}$, $\frac{\partial w_2}{\partial \theta_3} > 0$, $\frac{\partial \bar{w}}{\partial \theta_3} > 0$ follow from Lemmas 2.6A-B. Hence $\frac{\partial c_2}{\partial \theta_3} > 0$ and $\theta_3^* = \theta_3^{\max}$. Then from Lemma 2.6E, either $\frac{\partial c_2}{\partial \theta_1} < 0$, which implies $\theta_1^* = 0$, or $\frac{\partial c_2}{\partial \theta_2} < 0$ which implies $\theta_2^* = 0$, or both, which combines to $\theta_1^* \theta_2^* = 0$. QED.

Intuitively, $\theta_2^* > 0$ is possible for $w_2 > \bar{w}$ because medium skilled immigrants then generate net taxes that reduce the tax burden on native medium skilled agents. However, medium-skilled immigrants also reduce w_2. Hence $\theta_2^* > 0$ occurs only when the tax rate and initial value of x_1 are so high that the fiscal benefits outweigh the reduction in wages; and under these conditions, low-skilled immigrants would greatly reduce c_2. This provides an intuition why $\theta_1^* = 0$ when $\theta_2^* > 0$.

(2) Consider wage equalization. This is straightforward but requires multiple case distinctions and new notation, because immigration has wage effects that differs across cases in Lemma 2.6A. For given X, Z must be in the feasible set

$$\Omega_Z(X) \equiv \{Z : z_1 = \frac{x_1(1+\theta_1)}{1+\theta_2}, z_3 = \frac{x_3(1+\theta_3)}{1+\theta_2} \text{for some } \Theta \in \Omega_\Theta\}.$$

which may overlap with some or all of the sets in Lemma 2.6A. Note that $\Omega_Z(X)$ is compact and convex and that any $Z \in \Omega_Z(X)$ can be implemented by the immigration policy

$$\theta_2^o = \max\{\frac{x_1}{z_1} - 1, \frac{x_3}{z_3} - 1, 0\}, \theta_1^o = \frac{z_1(1+\theta_2^o)}{x_1} - 1, \theta_3^o = \frac{z_3(1+\theta_2^o)}{x_3}.$$

Moreover:

Lemma 2.6F: (a) The feasible set $\Omega_Z(X)$ is equivalently to $z_1 \in [z_1^{\min}, z_1^{\max}]$ and $z_3 \in \left[z_3^{\min}, z_3^{\max}(z_1)\right]$, where $z_1^{\min} = \frac{x_1}{1+\theta_2^{\max}}$, $z_1^{\max} = x_1 \cdot (1+\theta_1^{\max})$, $z_3^{\min} = \frac{x_3}{1+\theta_2^{\max}}$, and $z_3^{\max}(z_1) = \frac{x_3(1+\theta_3^{\max})}{\max(1, x_1/z_1)}$.

(b) For any $Z \in \Omega_Z(X)$, policy θ^o is the unique policy that satisfies Ortega's (2001) tie-breaking convention to minimize population. If $z_3 > x_3$, then $\theta_1^o \theta_2^o = 0$.

Proof: (a) By construction, $Z \in \Omega_Z(X)$ implies $z_1 \in \Omega_{z_1}$ and $z_3 \in \Omega_{z_3}(z_1)$. Conversely, for any $z_1 \in \Omega_{z_1}$ and $z_3 \in \Omega_{z_3}(z_1)$, the policy defined in (b) satisfies $\Theta \in \Omega_\Theta$, hence $Z \in \Omega_Z(X)$.

(b) Any policy θ' that implements Z must satisfy $(1+\theta_2')z_1 = x_1(1+\theta_1')$ and $(1+\theta_2')z_3 = x_3(1+\theta_3')$. Minimizing population $L = N_2 \sum_i x_i(1+\theta_i')$ subject to these constraints implies $\theta_2' = 0$ if $z_1 \geq x_1(1+\theta_1')$ and $z_3 \geq x_3(1+\theta_3')$, implies $\theta_2' = \frac{x_1}{z_1} - 1 > 0$ if $z_1 < x_1(1+\theta_1')$ and

$(1+\frac{x_1}{z_1})z_3 \geq x_3(1+\theta_3')$, and $\theta_2' = \frac{x_3}{z_3} - 1 > 0$ if $z_3 < x_3(1+\theta_3')$ and $(1+\theta_2')z_1 \geq x_1(1+\theta_1')$. Hence $\theta_2' = \theta_2^o$. Then θ_1^o, θ_3^o follow from the constraints. Moreover, $z_3 > x_3$ implies $\frac{x_3}{z_3} - 1 < 0$, hence $\theta_2^o = \max\{\frac{x_1}{z_1} - 1, 0\}$ and $\theta_1^o = \frac{z_1(1+\theta_2^o)}{x_1} - 1 = \max\{\frac{z_1}{x_1} - 1, 0\}$, which implies $\theta_1^o\theta_2^o = 0$. QED.

Remark: Part (a) facilitates sequential choice of z_3 followed by z_1, which useful because $\frac{\partial c_2}{\partial \theta_3}$ is easy to sign in all cases. Part (b) shows that Ortega's (2005) tie-breaking convention yields a unique policy in the static model whenever the set

$$\Omega_Z^* \equiv \arg\max\{c_2(Z) : Z \in \Omega_Z(X)\}$$

is single valued.

(3) To focus the analysis, suppose $w_1 < w_2 < w_3$ applies at least at X, the starting point without immigration; that is, $Z(X,0) = (x_1, x_3) \in int(\Omega_Z^u)$. Also assume $\theta_3^{\max} \leq \theta_2^{\max}$, as implied by the notion that high skilled immigrants are relatively scarce.[16] One finds:

Lemma 2.6G: Suppose $X \in int(\Omega_Z^u)$. Then:

(a) If $\theta_3^{\max} \leq \bar{z}_3/x_3 - 1$, all optimal policies satisfy $\theta_3^* = \theta_3^{\max}$, $\theta_1^*\theta_2^* = 0$, all imply unconstrained wages, and Ortega's (2001) tie-breaking convention generically selects a unique policy.

(b) If $\theta_3^{\max} > \bar{z}_3/x_3 - 1$, there exists an optimal policy θ^* with unconstrained wages ($Z^* \in \Omega_Z^u$). It is characterized by either:

(i) $\theta_3^* = \bar{z}_3/x_3 - 1$, $\theta_1^* \geq 0$, $\theta_2^* = 0$, and $w_1 < w_2 = w_3$, or

(ii) $\theta_3^* = \bar{z}_3/x_3 - 1$, $\theta_1^* = 0$, $\theta_2^* > 0$, and $w_1 < w_2 = w_3$, or

(iii) $\theta_3^* = \theta_3^{\max}$, $\theta_1^* = 0$, $\theta_2^* > \frac{x_3}{z_3}(1+\theta_3^{\max}) - 1$, and $w_1 < w_2 < w_3$.

Moreover, Ortega's (2001) tie-breaking convention generically selects a unique θ^* in Ω_Z^u.

(c) If cases (b-i) or (b-ii) apply, Ω_Z^* also includes a line segment in Ω_Z^{23} consisting of all $Z \in \Omega_Z(X)$ such that $z_{13}^* = \frac{z_1}{1+z_3}$ is the same as for $Z^* \in \Omega_Z^u$; all such policies imply the same wages. In case (b-i), all policies in Ω_Z^{23} are eliminated by Ortega's (2001) tie-breaking convention; in case (b-ii), policies in Ω_Z^{23} are not eliminated, but all are economically equivalent to $Z^* \in \Omega_Z^u$ in the sense that they imply the same wages and the same job assignments $\hat{L}_i$.

Proof: (a) If $X \in int(\Omega_Z^u)$ and $\theta_3^{\max} \leq \bar{z}_3/x_3 - 1$, then $\Omega_Z(X) \subset \Omega_Z^u \cup \Omega_Z^{12}$.

Suppose for contradiction that $Z \in \Omega_Z^{12}$ is optimal: Since $x_1 > \underline{z}_1$ whereas $z_1 < \underline{z}_1$, $\theta_2 > x_1/\underline{z}_1 - 1 > 0$ for any policy θ that implement Z. Note that $\frac{\partial c_2}{\partial \theta_2} < 0$ on $Z \in int(\Omega_Z^{12})$ because $\frac{\partial w_2}{\partial \theta_2} < 0$ from Lemma 2.6A and $\frac{\partial \bar{w}}{\partial \theta_2} \leq 0$ from Lemma 2.6B with $w_1 = w_2 < w_3$. Hence $c_2(Z)$ is strictly less than c_2 under the feasible policy θ' that replaces θ_2 by $\theta_2' = x_1/\underline{z}_1 - 1$ and implements $Z' \in \Omega_Z^u$ (on the boundary to Ω_Z^{12}), contradicting optimality of $Z \in \Omega_Z^{12}$ and proving $\Omega_Z^* \subset \Omega_Z^u$.

Considering $Z \in \Omega_Z(X) \cap \Omega_Z^u$: If $\theta_3^{\max} < \bar{z}_3/x_3 - 1$, then $w_2 < w_3$, so $\theta_3^* = \theta_3^{\max}$ and $\theta_1^*\theta_2^* = 0$ from Corollary 2.6E. If $\theta_3^{\max} = \bar{z}_3/x_3 - 1$, $\frac{\partial c_2}{\partial \theta_3} > 0$ for $Z \in int(\Omega_Z^u)$ implies $\theta_3^* = \theta_3^{\max}$, and hence $z_3^* =$

[16] Cases with $X \notin int(\Omega_Z^u)$ would require numerous case distinctions that are utterly irrelevant in the context of the main model, where X is generated by transition matrices that tend to have $x_3/\bar{z}_3 << 1$ and $x_1/\underline{z}_1 >> 1$, which means X is far from the boundaries of Ω_Z^u. If $\theta_2^{\max} < \theta_3^{\max}$, optimal policy could be technically in Ω_Z^{23} though economically equivalent to policies in Ω_Z^u, requiring needless elaboration.

$z_3^{\max}(z_1^*)$. Hence the problem of maximizing $c_2(Z)$ on Ω_Z^u reduces to maximizing $c_2(z_1, z_3^{\max}(z_1))$ by choice of z_1. Since $[\max\{z_1^{\min}, \underline{z}_1\}, z_1^{\max}]$ is compact, $z_1^* = \arg\max\{c_2(z_1, z_3^{\max}(z_1))\}$ is non-empty, showing existence of $Z^* \in \Omega_Z^u$. Moreover, $z_1^* = z_1^*(\tau)$ is strictly decreasing in τ for $z_1^* \in (\max\{z_1^{\min}, \underline{z}_1\}, z_1^{\max})$ by arguments analogous to the proof of Lemma 2.6D (hence details omitted), which implies uniqueness of (z_1^*, z_3^*) except for a countable number of tax rates, i.e. generic uniqueness. Then Lemma 2.6F implies generic uniqueness of θ^* with tie-breaking convention, and $\theta_1^*\theta_2^* = 0$.

(b) For $\theta_3^{\max} > \bar{z}_3/x_3 - 1$, $\Omega_Z(X)$ may overlap with Ω_Z^{123} and Ω_Z^{23}.

Suppose for contradiction that $Z \in cl(\Omega_Z^{123})$ is optimal: Since $\frac{x_1}{1+x_3} > \underline{z}_{13}$ for $X \in int(\Omega_Z^u)$, whereas $z_{13} \leq \underline{z}_{13}$ on $cl(\Omega_Z^{123})$, implementing Z requires $\theta_3 > 0$ and/or $\theta_2 > 0$ and implies existence of a feasible alternative θ' with $\theta_3 > \theta_3'$ and/or $\theta_2 > \theta_2'$ that implements $Z' \in \Omega_Z^{23} \cap cl(\Omega_Z^{123})$. As c_2 is constant on $cl(\Omega_Z^{123})$, $c_2(Z) = c_2(Z')$. Note that $\frac{\partial w_2}{\partial \theta_2}, \frac{\partial w_2}{\partial \theta_3} < 0$ in Ω_Z^{23} and that $\frac{\partial \bar{w}}{\partial \theta_i} = 0$ at Z'. Hence $\frac{\partial c_2}{\partial \theta_2}, \frac{\partial c_2}{\partial \theta_3} < 0$ a neighborhood of Z' inside Ω_Z^{23}, which means there is a feasible $Z" \in \Omega_Z^{23}$ for which $c_2(Z") > c_2(Z') = c_2(Z)$, contradicting optimality of $Z \in cl(\Omega_Z^{123})$. Analogous reasoning rules out optimal $Z \in \Omega_Z^{12}$, because there are superior alternatives either on the boundary between Ω_Z^u and Ω_Z^{12} or on the boundary between Ω_Z^{23} and Ω_Z^{123}.

Thus $\Omega_Z^* \subset \Omega_Z^u \cup \Omega_Z^{23}$.

For $Z \in cl(\Omega_Z^{23})$, Lemma 2.2B implies that w_2 and $\bar{w}$ are univariate functions of z_{13}. Note that $Z \in \Omega_Z(X) \cap cl(\Omega_Z^{23})$ implies $z_3 \geq \bar{z}_3$ and $z_1 \leq x_1 \cdot (1 + \theta_1^{\max})$, so $z_{13} \leq \frac{x_1}{1+\bar{z}_3} \cdot (1 + \theta_1^{\max}) \equiv z_{13}^{\max}$; also, , $z_{13} \geq \underline{z}_{13}$, and $z_3 \geq \bar{z}_3$ further implies $\frac{x_3(1+\theta_3)}{1+\theta_2} \geq \bar{z}_3$, $1 + \theta_2 \leq \frac{x_3}{\bar{z}_3}(1 + \theta_3^{\max})$, and $z_{13} \geq \frac{x_1}{1+\theta_2+x_3(1+\theta_3)} \geq \frac{x_1}{(1+\theta_3^{\max})(x_3/\bar{z}_3+x_3)}$. Thus, using $z_{13}^{\min} \equiv \max\{\underline{z}_{13}, \frac{x_1}{(1+\theta_3^{\max})(x_3/\bar{z}_3+x_3)}\}$, one finds $\Omega_Z(X) \cap cl(\Omega_Z^{23}) \subset \{Z : z_3 \geq \bar{z}_3, z_{13}^{\min} \leq z_{13} \leq z_{13}^{\max}\}$.

Conversely, note that $Z^u = (z_{13}(1 + \bar{z}_3), \bar{z}_3) \in \Omega_Z(X) \cap cl(\Omega_Z^{23})$ for all $z_{13} \in \left[z_{13}^{\min}, z_{13}^{\max}\right]$, and $Z^u \in \Omega_Z^u$. Define $\Omega_{13}^* \equiv \arg\max\{c_2(Z^u),\ z_{13} \in \left[z_{13}^{\min}, z_{13}^{\max}\right]\}$, which is non-empty because $\left[z_{13}^{\min}, z_{13}^{\max}\right]$ is compact, and generically single-valued by arguments analogous to the proof of Lemma 2.6D. Since w_2 and $\bar{w}$ are univariate functions of z_{13} on $cl(\Omega_Z^{23})$, $c_2(Z)$ is constant for all $Z \in cl(\Omega_Z^{23})$ with common ratio $\frac{z_1}{1+z_3} = z_{13}$. Hence for any $z_{13}^* \in \Omega_{13}^*$, $Z^{u*} = (z_{13}^*(1 + \bar{z}_3), \bar{z}_3)$ maximizes $c_2(Z)$ on $\{Z : z_3 \geq \bar{z}_3, z_{13}^{\min} \leq z_{13} \leq z_{13}^{\max}\}$, and hence on the subset $\Omega_Z(X) \cap cl(\Omega_Z^{23})$. The set of policies that maximize $c_2(Z)$ on $\Omega_Z(X) \cap cl(\Omega_Z^{23})$ therefore consist of Z^{u*} and the corresponding line segment(s) in Ω_Z^{23} with $z_{13} = z_{13}^*$.

Now consider all $Z \in \Omega_Z(X) \cap \Omega_Z^u$. Since $Z^{u*} \in \Omega_Z^u$, $\max\{c_2(Z) : Z \in \Omega_Z(X) \cap \Omega_Z^u\} \geq c_2(Z^{u*})$, which leaves two possibilities:

If $\max\{c_2(Z) : Z \in \Omega_Z(X) \cap \Omega_Z^u\} = c_2(Z^{u*})$, $Z^{u*} \in \Omega_Z^u$ and the line segments in Ω_Z^{23} with $z_{13} = z_{13}^*$ are optimal. In addition, $arg\max\{c_2(Z) : Z \in \Omega_Z(X) \cap \Omega_Z^u\}$ may include policies with $z_3 < \bar{z}_3$. Since $\frac{\partial c_2}{\partial \theta_3} > 0$ on $int(\Omega_Z^u)$, $z_3^* = \min\{z_3^{\max}(z_1^*), \bar{z}_3\}$, which maxizes z_3 within Ω_Z^u, and $\theta_1^*\theta_2^* = 0$ follows as in (a). This allows for three cases: (i) if $z_3^* = \bar{z}_3$ and $\theta_2^* = 0$, then $\theta_3^* = \bar{z}_3/x_3 - 1$ and $\theta_1^* \geq 0$; (ii) if $z_3^* = \bar{z}_3$ and $\theta_2^* > 0$, then $\theta_3^* = \bar{z}_3/x_3 - 1$ and $\theta_1^* = 0$; (iii) if $z_3^* < \bar{z}_3$, then $z_3^* = z_3^{\max}(z_1^*)$, so $\theta_3^* = \theta_3^{\max}$ and $\theta_2^* > \frac{x_3}{z_3}(1 + \theta_3^{\max}) - 1 > 0$; the latter implies $\theta_1^* = 0$. Also,

$z_3^* < \bar{z}_3$ implies $Z^* \in int(\Omega_Z^u)$, so $w_1 < w_2 < w_3$. This proves the properties claimed in (i-iii). Alternatively, if $\max\{c_2(Z) : Z \in \Omega_Z(X) \cap \Omega_Z^u\} > c_2(Z^{u*})$, then all optimal policies must satisfy $z_3 < \bar{z}_3$, and by the arguments above, they have properties (iii).

Note that all optimal policies in $\Omega_Z(X) \cap \Omega_Z^u$ must maximize c_2 by choice of $z_1 \in [\max\{\underline{z}_1, z_1^{\min}\}, z_1^{\max}]$ with implied $z_3 = \min\{\bar{z}_3, z_3^{\max}(z_1)\}$. For this univariate choice problem, arguments analogous to the proof of Lemma 2.6D imply that $z_1^*(\tau)$ is strictly decreasing in τ, hence generically unique, which means Ortega's (2001) tie-breaker selects a unique θ^*.

(c) The claimed line segments are defined in the proof of (b) above. In case (i), constant z_{13}^* on a line segment requires that θ_1^* and θ_3^* are increasing in z_3. Hence Ortega's tie breaker selects $(z_{13}^*(1+\bar{z}_3), \bar{z}_3)$, which is in Ω_Z^u. In case (ii), constant z_{13}^* on a line segment requires that θ_3^* is increasing in z_3 whereas θ_3^* is decreasing. Since $\theta_1^* = 0$, $L_1 = N_1$ is given, and constant $z_{13}^* = \frac{N_1}{L_2+L_3}$ implies common values for $L_2 + L_3$, for wages, and hence the same job assignments $\hat{L}_i$. Ortega's tie breaker is unhelpful because $N_1 + L_2 + L_3$ is common. In case (iii), $z_3 < \bar{z}_3$ rules out policies in Ω_Z^{23}. QED.

Remarks: Lemma 2.6G generalizes 2.6F and shows that there is always an optimal policy with unconstrained wages. The multiplicity of optimal policies is largely resolved by Ortega's tie-breaking convention, except for non-generic cases with multiple solutions and in subcase (b)(ii), when high-skilled immigrants are assigned to medium-skilled job. The latter is economically irrelevant (not affecting the skill mix actually used ($\hat{L}_i$)), an artifact of labeling immigrants by innate skills in a scenario that assignes them to the same job regardles of skill. A natural tie breaker is to assume that immigrants are admitted at the skill level at which they are employed–then all optimal policies have unconstrained wages.

8.4 Appendix to Section 3.2. Details on Mobility Matrices

We describe in more detail the analysis done in the text and present some robustness checks. We use individuals whose interview was obtained during the period 1977-2012, who were aged 25-55 and born at least since 1945. Then we consider some subsamples. The total number of observations used in the main text includes 18999 observations for children of natives, and 1447 observations for children of immigrants. Among natives, 8476 are men and 10523 are women. Among immigrants, 636 are men and 811 are women. In what follows, we label "children of immigrants" as "immigrants" for simplicity. We show in this section that the conclusions remain even if use only sons or daughters, as well as controlling by white race.

Table 10 shows the estimated intergenerational (transition) matrices for immigrants. Matrix #1′ is the transition matrix of men (sons of immigrants), matrix #2′ is the same concept for women, while matrix #3′ is for both men and women (used in main text). Table 10 shows the different matrices estimated for natives: matrix #1 is for men, matrix #2 is for white men, matrix #3 is for women, matrix #4 is for white women and matrix #5 is for all men and women (used in text).

Table 10. **Estimated Transition Matrix for Children of Immigrants**

	Transition Matrix ($\mathbf{Q}^I$)				Men/Women	Race	Age
	.240	.544	.216	N_1= 171			
$1'$	.068	.647	.285	N_2= 323	Men	All	25-55
	.021	.338	.641	N_3= 142			
	.192	.628	.180	N_1= 250			
$2'$	.067	.622	.311	N_2= 389	Women	All	25-55
	.023	.314	.663	N_3= 172			
	.211	.594	.195	N_1= 421			
$3'$	.067	.633	.299	N_2= 712	Both	All	25-55
	.022	.325	.653	N_3= 314			

We test whether the probability distributions for natives and children of immigrants are statistically the same, first at the row level and then at the matrix level. At the row level the null hypothesis is given by $H_0 : q_{ij} = q^I_{ij}$ for all j=1, 2, 3, given row $[i]$.

Let the parameters $\widetilde{q}_{ij}$ be the overall population parameters (for both natives and immigrants); when they are unknown (as in this case) they are estimated with the sample counterpart. Define also n_{im} be the total counts of row i (in transition matrix) of sample m. Under the null, the statistic $\sum\limits_{m=1}^{2} n_{im} \sum\limits_{j=1}^{k} \frac{(q_{ij}-\widetilde{q}_{ij})^2}{\widetilde{q}_{ij}}$ is distributed chi-square with $(k-1)$ degrees of freedom $(k=3)$, where n_{im} is the number of counts of row i of sample m. The statistic for each row is to be compared with $\chi^2_{(2)}$, which is 5.99 (9.21) at the 5% (1%) level of significance. For example, for the main matrices used in the paper ($Q[\#5]$ and $Q^I[\#3']$) the statistics are 58.74 for the first row, 18.63 for the second and 10.16 for the third, thus rejecting equality between natives and immigrants at each skill level.

We can also test for the equality of both matrices. The test (see Amemiya Pp. 417 and Mood-Graybill-Boes Pp. 449) is given by summing over rows, with the null hypothesis that $q_{ij} = q^I_{ij}$ for all $i = 1, 2, 3$ and $j = 1, 2, 3$. Then the null is rejected if $\sum\limits_{i=1}^{k} \sum\limits_{m=1}^{2} n_{im} \sum\limits_{j=1}^{k} \frac{(q_{ij}-\widetilde{q}_{ij})^2}{\widetilde{q}_{ij}} > \chi^2_{(k(k-1))}$ where the degrees of freedom are in this case $(k-1)\,k = 3\,(2) = 6$. The test produces a statistic of 87.52, which is significant at the 1% level (where thresholds are 16.81 at the 1% level and 12.59 at the 5% level). Hence, the test rejects that children of natives and those of first generation immigrants have the same transition matrix.

The results that low-skilled mobility of immigrants is higher for immigrants, as well as similar mobility in the high-skilled category are very robust across subsamples, while for medium-skilled men there's no statistical difference in the mobility of immigrants and natives (native matrix $Q\,[\#1]$ and immigrant matrix $Q^I\,[\#1']$, with a row statistic of 3.47, less than 5.99 which is the 5% significance level), and there's still significant difference for women ($Q\,[\#3]$ and $Q\,[\#2']$, with a statistic of $17.45 > 5.99$). For the test at the matrix level for these same subsamples of men or women, we reject that they are statistically equal at the 1% significance level. Similar results for white men vs immigrant men ($Q\,[\#2]$ and $Q^I\,[\#1']$) and white women vs immigrant women ($Q\,[\#4]$ and $Q^I\,[\#2']$). The tests are reported in table 12.

Table 11. **Estimated Transition Matrices for Natives**

#	*Transition Matrix* (**Q**)				Men/Women	Race	Age
	.281	.628	.091	N_1= 1538			
1	.068	.692	.240	N_2= 4985	Men	All	25-55
	.008	.401	.590	N_3= 1953			
	.282	.622	.096	N_1= 1094			
2	.064	.685	.251	N_2= 4318	Men	White	25-55
	.007	.384	.609	N_3= 1806			
	.239	.687	.075	N_1= 2317			
3	.057	.719	.224	N_2= 6091	Women	All	25-55
	.011	.394	.595	N_3= 2115			
	.232	.690	.078	N_1= 1465			
4	.047	.713	.231	N_2= 4984	Women	White	25-55
	.009	.381	.610	N_3= 1901			
	.256	.663	.081	N_1= 3855			
5	.062	.707	.231	N_2= 11076	Both	All	25-55
	.010	.397	.593	N_3= 4068			
	.253	.661	.086	N_1= 2559			
6	.055	.700	.245	N_2= 9302	Both	White	25-55
	.008	.382	.609	N_3= 3707			

Table 12. **Tests with 5% critical values** (χ^2)

H_0 (Null Hypothesis)	Row1 Test CV=5.99	Row2 Test CV=5.99	Row3 Test CV=5.99	Matrix Test CV=12.59
$Q[\#1] = Q^I[\#1]$	26.05**	3.47	5.80	33.87**
$Q[\#2] = Q^I[\#1]$	21.53**	2.12	5.21	27.69**
$Q[\#3] = Q^I[\#2]$	32.72**	17.45**	5.03	56.06**
$Q[\#4] = Q^I[\#2]$	26.78**	14.68**	4.26	46.89**
$Q[\#2] = Q[\#4]$	12.03**	19.83**	3.42	34.89**
$Q^I[\#1] = Q^I[\#2]$	3.00	0.58	0.55	3.79
$Q[\#1] = Q[\#3]$	14.28**	11.27**	1.95	26.49**
$Q[\#5] = Q^I[\#3]$	58.74**	18.63**	9.34**	87.52**

Notes: The 1% critical value (CV) of the test for equality of matrices is 16.81 (6 degrees of freedom). The 1% CV (**) for tests of row equality is 9.21.

We also look at differences between men and women transition matrices, for natives-only and immigrants-only. Comparing the transition matrices for native men and women (native matrix #1 vs native matrix #3), the first two rows suggest that men have slightly more extreme outcomes than women. The row tests show that the probability distributions for sons of low-skilled and medium-skilled parents are different to those of daughters, with statistics of 14.28 (low-skilled parents) and 11.27 (medium-skilled), while we cannot reject that the probability distribution of sons and daughters of high-skilled parents is the same (statistic=0.95 $<$ 5.99). The matrix test rejects that men and women have the same transition matrices, with an statistic of 26.49 $>$ 12.59. In the case of immigrants (immigrant matrix #1 and immigrant matrix #2), we cannot reject that the rows and the whole matrices are statistically the same. The differences are similar to those of natives, but in this case the lower number of observations is the cause that we can't reject the null of matrix equality. When we restrict the men-women comparison to native whites, the matrices and tests results remain essentially unchanged.

8.4.1 Transition Matrices under an Alternative Definition of 2nd Generation Immigrants

Table 13 presents the transition matrices when second generation immigrant are defined as individuals whose both parents were born outside the US under the same filters as described in the main text. The sample size across all skills categories is only 530 observations (as opposed to 1447 used in the main analysis).

Table 13. **Transition Matrices of children of immigrants: both parents born outside US**

	Transition Matrix ($\mathbf{Q}^I$)				Men/Women	Race	Age
	.286	.464	.250	N_1= 84			
$1''$	.085	.606	.309	N_2= 94	Men	All	25-55
	.043	.391	.565	N_3= 46			
	.165	.617	.218	N_1= 133			
$2''$	.058	.545	.397	N_2= 121	Women	All	25-55
	.019	.288	.692	N_3= 52			
	.212	.558	.230	N_1= 217			
$3''$	.070	.572	.358	N_2= 215	Both	All	25-55
	.031	.337	.633	N_3= 98			

With the new definition, we also test whether matrix #$3''$ is statistically different from matrix #$3'$, which considers whether the mobility matrix obtained for immigrants when these are defined as children of at least one foreign-born parent is statistically different from the same matrix obtained for children of two foreign-born parents. The row tests yield 1.89, 3.79 and 0.41 for low-skilled, medium-skilled and skilled parents respectively, which are all lower than the 5.99 threshold for rejection. The matrix test yields 6.10, lower than the 12.59 threshold at the 5% level. Thus confirming that these matrices are statistically the same.

8.4.2 Average Schooling Years by Cohort in the GSS

The average number of schooling years for individuals born in the US and aged between 25 and 55 years old at the time of the interview ranges from 11.03 for those born in the period 1915-1924 to 13.86 for the cohort 1975-1984. We use individuals born on or after 1945 because the average schooling years by cohort are roughly constant since then, while the average schooling years trend upward for previous cohorts.

Table 14. **Average schooling years by GSS cohort***

Cohort born	1915-24	1925-34	1935-44	1945-54	1955-64	1965-74	1975-84
Mean	11.03	12.01	12.93	13.65	13.64	13.87	13.86
Std Dev	3.33	3.20	2.95	2.74	2.52	2.55	2.74
Sample (N)	115	1193	3342	7157	7277	3561	996

*Individuals born in the US, age 25-55 at the time of interview. Men and Women.

8.5 Appendix to Section 3.3. Estimation of TFR's for Native and Foreign Born Women in the US

We estimate total fertility rates (TFR) by education level and nativity, starting from the fertility rates by education level that can be computed from birth and census data. Since birth data available at the VitalStats website doesn't distinguish whether the mother was US-born or foreign-born, we construct estimates with census information as well as with the help of the American Community Survey (ACS), which identifies the place where the mother was born and which can be used to estimate $TFR's$. The estimators are derived for each specific level of skill levels as defined in this paper. We first estimate TFR's for all US women for years 1990, 2000 and the CPS for 2005. The results are in table 15.

Table 15. **TFR's for all women in US by education**

Year	Less than HS	HS + Some College	BA+Beyond
2005	2.73	1.95	1.85
2000	2.26	2.00	1.84
1990	2.43	2.04	1.61
Average	2.47	2.00	1.77

Sources: Census 1990, 2000. CPS-2005.

Now we show how we estimate $TFR's$ by nativity from the levels in table 15 which are for all women. For the particular estimates, we keep notation simple by not including a particular skill level (but calculations are done for each education level defined in the paper). Define $\mathbf{X}_i^k =$

Number of children born during year to women type k in age-group i; $\mathbf{Y}_i^k$ =total # of women type k in age-group i; $\mathbf{k} = \{N, F\}$, where N=US-born and F=foreign-born; $\mathbf{n}$ = # of age-groups (5-year groups); TFR^k =Total fertility rate of women type k ; TFR =Total fertility rate of all women living in the US.

The goal is to obtain estimates for TFR^N and TFR^F as all data for their direct estimation is not available. So we construct estimators of TFR^N and TFR^F from our estimates for TFR for all women. The TFR can be written as $TFR = 5\Sigma_i^n \left[\left(X_i^N + X_i^F\right) / \left(Y_i^N + Y_i^F\right)\right]$. It can then be rewritten as $TFR = TFR^F - 5\Sigma_i^n \left(1 - w_i^F\right) m_i$, with weights $w_i^F = Y_i^F / \left(Y_i^N + Y_i^F\right)$, and where $m_i = \left[\left(X_i^F / Y_i^F\right) - \left(X_i^N / Y_i^N\right)\right]$ is the difference in births per foreign–born women and births per US-born women in age group i (for a given level of education). Hence an estimate of TFR^F is obtained by computing $TFR^F = TFR + 5\Sigma_i^n m_i \left(1 - w_i^F\right)$. Similar algebra yields native fertiliy as $TFR^N = TFR - 5\Sigma_i^n m_i w_i^F$.

Table 16. **Estimated weights ($\mathbf{w_i^F}$)**

Census 1990, 2000 and CPS-2005 by skill level (L, M, H)*

	1990			**2000**			**2005**		
Age Group	L	M	H	L	M	H	L	M	H
15-19	.082	.068	.250	.092	.083	.256	.078	.089	.209
20-24	.209	.079	.091	.296	.113	.135	.263	.109	.107
25-29	.229	.077	.106	.394	.127	.160	.383	.133	.157
30-34	.248	.075	.111	.400	.122	.161	.476	.148	.189
35-39	.257	.070	.096	.360	.107	.155	.452	.142	.177
40-44	.241	.075	.099	.339	.097	.135	.386	.117	.159
45-49	.208	.074	.101	.345	.087	.115	.348	.099	.141
50-54	.185	.077	.106	.303	.088	.106	.345	.090	.116

* L= Low-Skilled, M=Medium-Skilled, H=High-Skilled

The estimates for the weights w_i come from the census of 1990, 2000 and the 2005 CPS. However, data is not available for the differences m_i for these exact years. Thus we estimate the average difference m_i for each age-education group for years $2001 - 2008$ from ACS data. Using these numbers, we arrive at the fertility rates by skill level and nativity shown in the text.

Table 17. **Estimates of differences in births per women (foreign - natives : $m_i = \left[\left(X_i^F / Y_i^F\right) - \left(X_i^N / Y_i^N\right)\right]$)***

	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54
Low	.014	.002	.042	.051	.036	.015	.000	.000
Medium	.002	.000	.0016	.037	.030	.012	.003	.000
High	-.003	.010	.005	-.004	.012	.009	.003	.001

*averages for 2001-2008 (ACS).

8.6 Appendix to section 4.1. Steady State Composition of the Native Population in absence of Immigration

In the absence of immigration we have that $\mathbf{S_t}$ is time invariant, so we drop the time subscript. Thus in this case we have that $\mathbf{S} = \mathbf{Q}'\eta$ and the scalar given by $\left(\mathbf{S_{t[2]}} X_t\right) = \sum_i^3 x_{it}\eta_i q_{i2}$. Therefore the evolution of the native population is given by

$$X_{t+1} = \mathbf{Q}'\eta X_t / \left(\sum_i^3 x_{it}\eta_i q_{i2}\right).$$

Multiply both sides of the above expression by the scalar $\left(\sum_i^3 x_{it}\eta_i q_{i2}\right)$, and evaluate the product $\mathbf{Q}'\eta X_t$ in order to write

$$\begin{bmatrix} x_{1t+1}\left(\sum_i^3 x_{it}\eta_i q_{i2}\right) \\ \sum_i^3 x_{it}\eta_i q_{i2} \\ x_{3t+1}\left(\sum_i^3 x_{it}\eta_i q_{i2}\right) \end{bmatrix} \equiv X_{t+1}\left(\sum_i^3 x_{it}\eta_i q_{i2}\right) = \mathbf{Q}'\eta X_t \equiv \begin{bmatrix} \sum_i^3 x_{it}\eta_i q_{i1} \\ \sum_i^3 x_{it}\eta_i q_{i2} \\ \sum_i^3 x_{it}\eta_i q_{i3} \end{bmatrix}.$$

Substracting $\mathbf{Q}'\eta X_t$ on both sides yields the [3x1] system of equations

$$X_{t+1}\left(\sum_i^3 x_{it}\eta_i q_{i2}\right) - \mathbf{Q}'\eta X_t = \mathbf{0}.$$

At a steady state, this system is given by

$$\begin{bmatrix} x_1^0\left(\sum_i^3 x_i^0\eta_i q_{i2}\right) - \sum_i^3 x_i^0\eta_i q_{i1} \\ 0 \\ x_3^0\left(\sum_i^3 x_i^0\eta_i q_{i2}\right) - \sum_i^3 x_i^0\eta_i q_{i3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{A.2}$$

which is a quadratic system of 2 equations in the 2 unknowns (x_1^0, x_3^0), where one of the solutions is the steady state composition of the population in the absence of immigration.

8.7 Appendix to sections 4.4 - 6.2: Numerical Analysis and Optimal Policies

This section explains the algorithm used for the numerical analysis, including comments on the choice of technical parameters.

We define the steady state pair ratios in absence of immigration as $\{x_1^0, x_3^0\}$, which is the solution to (A.2) and turn out to be (.09782, .5429). A grid of N_{X_1} points in $((1-a_1)\,x_1^0, (1+a_2)\,x_1^0)$ and N_{X_2} points in $((1-b_1)\,x_3^0, (1+b_2)\,x_3^0)$ for a total of $N_{X_1} * N_{X_2}$ pairs in the resulting grid is used for the value function iteration procedure, where a_1, a_2, b_1 and b_2 are positive constants that represent the deviation from the no-immigration steady state. The specific nodes that are used for interpolation are chosen as the roots of Chebyshev polynomials in the considered state spaces. The specific constants that define the state grid depend on the policy space (i.e. big high-skilled quotas might imply an induced state out of the grid if b_2 is too small), which we now describe. The policy space considers N_{θ_1} points for θ_1 in the space $[0, \theta_1^{\max}]$, N_{θ_2} points for θ_2 in $[0, \theta_2^{\max}]$ and N_{θ_3} points for θ_3 in $[0, \theta_3^{\max}]$.

We first explain setting (II), which has a perfectly inelastic supply of high skill immigrants ($\theta_3^{\max} = 13\%$) in section 4.5 in detail as it provides our main results, then comment on the case with wage-elastic supply of high-skilled immigrants (section 6.1), and finally on Setting (I) in section 4.4. Comments on Settings (II) also apply to case (III), since allowing guest workers only adds to choice variables but does not expand the range of parameters.

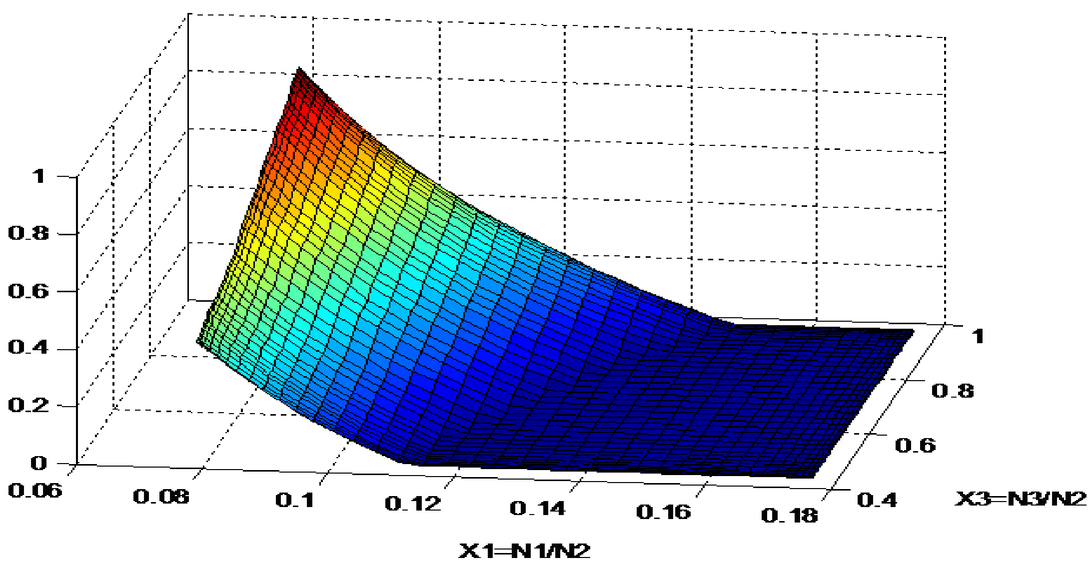

Figure 2. **Optimal Policy function for low-skilled immigration.** Case with $\theta_3^{\max} = 13\%$

In the case of a "small pool" of high-skilled immigrants ($\theta_3^{\max} = 13\%$), we use the particular grid given by points in $[.07823, .1761] \times [.4344, .9773]$, with $N_{X_1} = N_{X2} = 40$ nodes in each axis, for a total of 1600 nodes (points) in the state space. The considered immigration policies are elements $(\theta_1, \theta_2, \theta_3) \in [0, \theta_1^{\max}] \times [0, \theta_2^{\max}] \times [0, \theta_3^{\max}] \subseteq R_+^3$. For this version of the model, the number of policy points in θ_2 and the maximum immigration $\theta_2^{\max}$ quota turn out to be irrelevant since the model always predicts (given parameterization of model) that the majority chooses $\theta_2^* = 0$. Similarly, given that the maximum level $\theta_3^{\max} = 0.13$ is lower than the quota that would be freely chosen under setting (I), the majority chooses $\theta_3^* = \theta_3^{\max} = 0.13$. The relevant policy information in this version of the model (in addition to the state grid) are the number of policy points available for θ_1 and having a large enough $\theta_1^{\max}$ in order to allow for an interior solution, together with the specific level of $\theta_3^{\max}$ (set at 13% in this version). We show the case below (used as "baseline") where $\theta_1^{\max} = 1$, and $\theta_3^{\max} = 0.13$, with 801 equidistant points in the policy space for $\theta_1 \in [0, 1.25]$ (this case the number of points for high skill immigration is irrelevant). We use $\beta = .6325$ (this level chosen to get as close as possible to a target of $\theta_1^* = 18\%$). Solving the model with a tolerance of

.0000001 in the sup norm between the value functions of the current and the previous iteration, we obtain a quota of low-skilled workers of 18.01%, and the high-skilled immigration quota is optimally chosen at 13%. The induced steady state (x_1^{ss}, x_3^{ss}) is $(.1004, .589)$. Cases where there are guest worker quotas available typically involve increasing the size of the policy space.

As an example of the solution to the case with a wage-elastic supply as captured by (16), we use an elasticity of 10, while the rest of the parameters set as in the main text; with a state grid of 900 pairs of points in $[.097826, .176084] \times [.434353, .977295]$. In this case the $i-th$ immigration policy point $(\theta_{1i}, \theta_{2i}, \theta_{3i})$ at the state node (x_1, x_3) is feasible if $\theta_{3i} \leq \theta_3^{\max}(w_3(x_1, x_3, (\theta_{1i}, \theta_{2i}, \theta_{3i})))$. We use 90 points for $\theta_1 \in [0, 2]$ and 150 points in $\theta_3 \in [0, \theta_3^{\max}(w_3(\cdot))]$. Again $\theta_2^{\max}$ (and the number of policy points for θ_2) are irrelevant. At the no-immigration steady state the optimal policies are $(43.4\%, 0\%, 19.9\%)$. Below we show the immigration policies chosen in the state-grid considered. The shape of the low-skilled immigration quota is just like before: increasing in x_3 and decreasing in x_1, while the shape of the high-skilled immigration policy is decreasing in x_3 and slightly increasing in x_1. Other numerical cases look qualitatively identical and are therefore not discussed further.

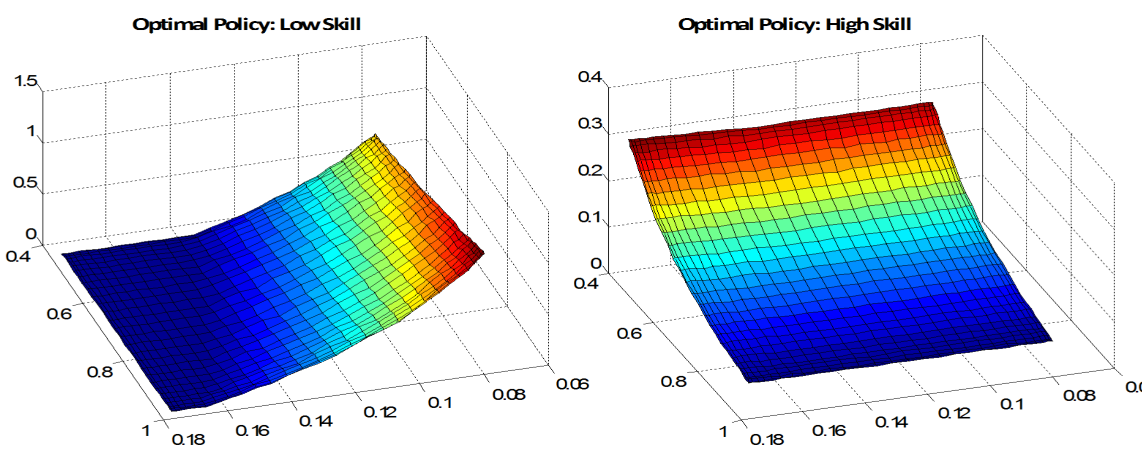

Figure 3. **Optimal policy functions with an elastic response of $\theta_3^{\max}$ (Elasticity=10)**

In setting (I), which is the case of a huge pool of high-skilled immigrants, we successively expanded the maximum immigration quotas until we found a policy space such that the resulting optimal functions were in the interior for all possible states. In the case of medium-skill immigration, $\theta_2^{\max} = 100\%$ was high enough, as the optimum policy is $\theta_2^* = 0$. For the other types, $\theta_1^{\max} = 4.5$ and $\theta_3^{\max} = 3$ turned out to be sufficient.

8.8 Appendix to section 6.2 on definition of skill groups

In section 6.2 we did some sensitivity analysis regarding the definition of the skill categories. We defined medium skilled workers as those with a high-school diploma, some college or having a college degree. For the computations of the immigration estimates, we use data from 2000 to 2013. Using the same estimates of net international migration (3.1 for 2010-2013 and 3.29 for 2000-2009), but using the information in table 18, we estimate that there were an average (given new definitions) of 6.55 annual low-skilled immigrants per thousand low-skilled natives, 2.45 medium skilled per thousand medium skilled natives, and 4.44 high-skilled immigrants per thousand high-skilled natives. This translates into approximately 20% quota for the low-skilled group and we obtain again 7% quota for the medium-skilled and 13% for the high-skilled.

Table 18. **Education Shares of Foreign Born and US Population (2000-2013)**

	2010-2013		2000-2009	
	Foreing Born Entering	US Pop	Foreign Born Entering	US Pop
LTHS Diploma	21.1	12.4	31.6	14.8
HS+Some College+College	63.5	76.6	55.4	75.7
Master & Above	15.4	11.0	13.0	9.5

Given the new skill definitions we obtain wage premia of $\frac{w_3}{w_2} = 1.727$ and $\frac{w_2}{w_1} = 1.4845$ for the same samples and filters as previously discussed (except education). Total fertility rates can be computed from the ACS, which yield fertility rates for native women of $(2.41, 2.04, 1.97)$ and $(3.23, 2.44, 2.08)$ for foreign-born women as the time-averages for the years 2001-2008. This in turn imply model parameters of $\eta = diag\,\{1.21, 1.02, .99\}$ for natives and $\eta^I = diag\,\{1.62, 1.22, 1.04\}$ for immigrants. Given $\rho = 1/2$, mobility and fertility rates, the production share parameters (ϕ_1, ϕ_2, ϕ_3) are calibrated as (.11209, .56996, .31795) following the steps described in the paper. The tax rate used is the same as in the main parameterization ($\tau = 30\%$). The model is studied under log utility ($\sigma = 1$) and we set $\beta = .985^{30}$ as in the initial experiments. In this case, under setting (II) there's no β that is able to produce a low-skilled quota of 20%, but the qualitative results to all experiments remain the same (shape of optimal policy, guest workers for low-skilled, immigration for the high-skilled while minimizing medium-skill immigration, etc.). Numerical results are summarized in table 7 in the paper.

References in Appendix

[1] Anderson, T. W. and Leo Goodman (1957) "Statistical inference about Markov chains". Annals of Mathematical Statistics, Pp. 89-109.

[2] Amemiya, Takeshi (1985). Advanced Econometrics. Harvard University Press.

[3] Mood, Graybill and Boes (1974) Introduction to the Theory of Statistics. McGraw-Hill, 3rd edition.

[4] National Academies of Sciences, Engineering, and Medicine (2017). The Economic and Fiscal Consequences of Immigration, Blau, F. D. & C. Mackie, (Eds.). Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/23550.