Intergenerational Mobility and the Political Economy of Immigration*

Henning Bohn† Armando R. Lopez-Velasco‡

September 2014

Abstract

Flows of US immigrants are concentrated at the extremes of the skill distribution. We develop a dynamic political economy model consistent with these observations. Individuals care about wages and the welfare of their children. Skill types are complementary in production. Voter support for immigration requires that the children of median-voter natives and of immigrants have sufficiently dissimilar skills. We estimate intergenerational transition matrices for skills, as measured by education, and find support for immigration at high and low skills, but not in the middle. In a version with guest worker programs, voters prefer high-skilled immigrants but low-skilled guest workers.

---

*We thank Finn Kydland, Peter Rupert, Giulio Zanella, Shawn Knabb, Stephen Trejo, Pia Orrenius, Carlos Zarazaga, Erick French and Julian Neyra for valuable comments and discussions. We also thank the comments by Nicola Pavoni as well as of two anonymous referees. All remaining errors are our own.

†Department of Economics, University of California Santa Barbara; Santa Barbara CA 93106; and CESifo. Phone (805) 893-4532. E-mail: bohn@econ.ucsb.edu. Homepage: http://www.econ.ucsb.edu/~bohn

‡Department of Economics, 253 Holden Hall. Texas Tech University. Lubbock, TX. Email: ar.lopez@ttu.edu.
1 Introduction

Stylized facts of international migration are that immigrants tend to be concentrated at the extremes of the skill distribution—high and low—and that high- and low-skilled immigrants are treated very differently. Many countries allow or even encourage immigration of high-skilled workers, but they prohibit low-skilled immigration or only accept low-skilled foreigners temporarily (e.g., as guest workers or by allowing illegal immigrants subject to instant deportation to stay).

We examine the political economy of immigration in a dynamic model in which natives care about their children and recognize that immigration influences the labor market for current and future generations. Skill types are complementary and the majority of natives is medium-skilled. Hence from a static perspective, the native majority benefits from foreign workers with skills far from the middle, both high and low.

The challenge is to explain the differential treatment of high and low skilled foreigners. A common argument is that natives worry about unskilled migrants relying on welfare, whereas the high-skilled pay more taxes. Our model includes a simple tax-transfer system to account for this, but we find the welfare argument unconvincing, at least for countries like the US that exclude migrants from most welfare benefits. Our main contribution is to provide an alternative explanation: Using U.S. data on generational mobility, we show that children of unskilled workers tend to compete in the labor market with the children of medium-skilled natives. In contrast, children of high-skilled workers have a skill distribution more complementary to the children of medium-skilled natives. Hence concern about children can provide a positive theory of why high-skilled workers are allowed to enter permanently, i.e., with their children, whereas low-skilled workers are not accepted permanently.\footnote{Undocumented workers are typically confined to low skilled work and cannot easily settle down as families, being under a constant threat of deportation. Hence it can be argued that de facto tolerance of illegal immigrants is similar to a guest worker program for unskilled workers.}

Our data analysis focuses on the US. Since 1970, about 60% of immigrants are either "high-skilled" (BA degree or more) or "unskilled" (less than High School degree), while the rest is "medium-skilled" (High School degree or some college). Because the share of medium-skilled individuals in the US population is more than 50%, the amount of high-skilled and low-skilled individuals as percentage of their native counterparts is larger than for the middle group. For example, for the decade of the 90’s, we estimate that there were 5.65 unskilled (net) immigrants per 1000 unskilled natives in the US, about 2.02 medium-skilled per 1000 native medium skilled, and 4.15 skilled immigrants per 1000 skilled US natives.\footnote{See appendix for details on these numbers. We use the terms "unskilled" and "low-skilled" interchangeably; and we use "skilled" and "high-skilled" in the same manner.} In a 30 year period these numbers would account for immigration quotas of roughly 18%, 6% and 13%, when compared to the composition of the native population.
A difficulty in interpreting these flow data is that control over immigration is highly imperfect. Observed immigration is arguably a combination of legal and illegal flows, and of job-related and other flows.\textsuperscript{3} To interpret the data, we set up a political-economy model to derive predictions about optimal immigration under alternative assumptions, and we examine under what conditions the model provides a positive theory.

The model has three types of labor inputs defined as low, medium and high-skilled; and two types of migrants, permanent immigrants and temporary guest workers. Each worker supplies one unit of their work-type to the production process, earns a wage, pays proportional taxes that are then redistributed via lump-sum, has children according to a fertility profile that depends on skill and place of birth. The model is calibrated to match the transition matrix of intergenerational skill transmission for natives and immigrants in the US. The calibrated demographic process for the US is such that with or without immigration, the medium-skilled type would be the absolute majority in each generation.

Immigration policy is defined by a set of quotas indexed by skill level and type of immigration permit (permanent vs guest-worker). Votes over immigration policy occur before the skill type of children is revealed. Immigrants don’t have the right to vote, but the children of immigrants are citizens identical in everything to natives, which in turn have the right to vote. We use Markov Perfect as the equilibrium concept, as it is now common in the literature of dynamic political economy.

To streamline the presentation, our analysis initially sets aside guest workers and focuses on the more challenging problem of modelling permanent immigration. Thereafter, the possibility of guest workers is added, which is straightforward in our setting because guest workers don’t raise intertemporal issues. We show that in the presence of guest worker and full immigration, the medium-skilled majority chooses a guest worker program for low-skilled individuals, no immigration (for economic reasons) for the medium-skilled and full immigration for the high-skilled individuals. The model is consistent with proposals in the US to create a guest worker program as part of immigration reform, and with a continuation of relatively easy immigration by very high skilled individuals and/or a push for allowing more high-skilled immigration.

Two important objects in the model and of independent interest in this paper are the matrices of intergenerational mobility for natives and for immigrants, which we estimate from the General Social Survey. This survey collects information on education data on the respondents and their parents, and it also identifies whether the parents are foreign born, among other variables. The data required for our purposes is available since 1977. We find that on average the children of unskilled and medium-skilled parents do better than their native counterparts, while there is no statistical difference for children of high-skilled

\textsuperscript{3}In the US, a significant share of immigration involves non-working family members of earlier immigrants who have attained citizenship. We do not model family unification, but the phenomenon is consistent with our premise that intergenerational issues are central to thinking about immigration.
parents. Consistent with this, Card, DiNardo and Estes (2000) find with regression analysis that children of immigrants (second generation immigrants) have on average higher schooling and wages than children of native parents with comparable education level. We now elaborate on findings of the paper.

Among other findings we show that the intergenerational mobility of children of low-skilled workers can be a major determinant for the political support of low-skill immigration by the majority (medium-skilled). In this case 2nd generation immigrants from low-skilled parents mostly "complement" the children of medium-skilled natives. Also the higher the probability that children of low-skilled immigrants become medium-skilled, the lower the low-skilled immigration quota that is politically chosen because 2nd generation immigrants from low-skilled parents are mostly "substitutes" of the children of medium-skilled natives. Since we estimate that on average children of unskilled immigrants have a higher probability of upward mobility, and a lower probability of becoming medium-skilled or staying unskilled than natives, these effects help explain why the US allows significant quantities of unskilled immigrants. Another way to put this is the following. If children of unskilled immigrants had on average the same probability of becoming skilled as the unskilled natives, the number of unskilled immigrants allowed in equilibrium would decrease considerably from current flows. In addition, the model suggests that intergenerational mobility of 2nd generation immigrants from skilled parents is not very important for the admission of skilled individuals. The high skilled individuals are so productive and the intergenerational mobility matrices such that the medium-skilled majority wouldn’t want to restrict their immigration from current flows, even though it could mean more competition for the children of medium skilled that were in the "skilled" state. Their desirability would decrease (their quota) only if they had an unrealistically high probability of having unskilled children.

This paper is closely related to Ortega’s politico-economic models on immigration with intergenerational mobility (2005, 2010). Ortega (2005) identifies political-economic trade-offs of immigration in a model with two skill levels and skill upgrading. In it he finds the Markov Perfect Equilibrium of the model in the presence of skill upgrading, while abstracting from redistribution issues. Ortega (2010) extends the previous model by allowing agents to vote over immigration and redistribution in a richer model with alternative immigration regimes. In our paper we extend the analysis from two-skill levels to three because of 1) the composition of what is traditionally defined as unskilled workers (education level lower than a college degree) is very different for natives and immigrants (a big percentage of immigrants have much lower schooling than natives in this category) and 2) because the demographic profiles of natives and immigrants are significantly different: intergenerational mobility is statistically different for some skill levels, and it has been documented that fertility profiles are also different. An important difference with Ortega’s models is that in ours there’s no political trade-off since the majority of voters in our model are of the medium-skilled type (consistent with definitions used throughout the paper) and the estimated demographic profiles are such that the majority will continue to be the medium-skilled. Another difference with Ortega (2010) is that in our model
the level of redistribution is exogenous. The model of this paper allows for the
study of questions regarding mobility and the resulting equilibrium immigration
policies in the presence of an exogenous welfare state.

Other related literature on equilibrium political-economic models of immi-
gration includes the seminal work of Benhabib (1996) on immigration policy
under heterogeneous agents. Dolmas and Huffman (2004) study the interaction
of immigration and redistribution. Sand and Razin (2007) study the joint de-
termination of immigration and social security with a Markov equilibrium con-
cept. Cohen, Razin and Sadka (2009), study the theory and empirical relation
between the size of the welfare state and the composition of the immigration
flows in a static model. Lopez-Velasco and Bohn (2008) study immigration
games when immigrants have a larger fertility rate than natives. On the dy-
namic fiscal effects of immigration there are papers by Storesletten (2000) and
related to capital accumulation. The paper is also related to the macroeconomic
literature where current voters foresee the consequences of their choices on the
future behavior of voters. Some examples include Krusell and Rios-Rull (1999)
and Hassler et.al (2003). Empirical studies on the intergenerational mobility
of immigrants include Borjas (1992) and Card, DiNardo and Estes (2000). On
the economics of guest worker programs Djaji´c (2014) has discussed welfare
implications of low skilled guest workers, the optimal design of temporary mi-
gration programs has been discussed by Schiff (2007) and Djaji´c, Michael and
Vinogradova (2013).

Regarding preferences over immigration, Dustmann and Preston (2005) find
empirically that competition in the labor market, the effects of immigration on
the public burden, and efficiency considerations of immigration shape immigra-
tion attitudes. Mayda (2006) concludes that both economic variables (mainly
labor market outcomes in her analysis) and non-economic variables shape the
attitudes of natives toward immigration, and non-economic variables do not
alter the labor market results. The paper by Card, Dustmann and Preston
(2009) concludes that immigration preferences are shaped not only by the eco-
nomic effects of immigration, but by externalities associated to the changes on
the composition of the population induced by immigration policies. This is one
of the main arguments in our paper: that the current and future composition
of the population is affected by the composition of the immigration flows and
their different-from-natives demographic profiles (mobility and fertility), which
in addition to the welfare state and labor market effects would then shape views
on immigration, and that ultimately can dictate policy.

The paper is organized as follows. In section 2 we present the model and
the equilibrium concept. Section 3 discusses the data, the choice of functional
forms and the estimation/calibration of parameters of the model with particular
emphasis on the estimation of intergenerational transition matrices for US na-
tives and immigrants from the General Social Survey. In section 4 we do some
experiments with the model and the main results are presented. Section 5 does
sensitivity analysis. Section 6 concludes.
2 The Model

2.1 Demographics

Consider a model with 3 types of agents: low-skilled workers (also called unskilled, labeled type 1), medium-skilled (type 2) and high-skilled (type 3) which for brevity we also call "skilled". Individuals work full-time when adult, vote over immigration policy and also have children whose skill type is unknown at the time of the vote but whose probability distribution over skill types is stationary. Children are assumed not to take any economic decisions.

Natives have $i$ children, while immigrants have $0$ for skill levels $i=1,2,3$, and these fertility profiles are later estimated by skill levels and nativity.

There is intergenerational mobility across skill types, where each type will earn the current wage that their labor skill level commands. Define $q_{ij}$ as the probability that the children of a parent of skill type $i$ will be of skill type $j$ ($i=1,2,3$ and $j=1,2,3$). For example, $q_{13}$ denotes the probability that the children of an unskilled parent will be skilled. With this notation we define the intergenerational ("transition") matrix for the children of native parents as $Q_n$ given by

$$Q_n = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} = \begin{bmatrix} -Q_{n[1]}^- \\ -Q_{n[2]}^- \\ -Q_{n[3]}^- \end{bmatrix}. \quad (1)$$

The rows of the transition matrix add up to 1 as they are probability distributions. We denote these $[1 \times 3]$ row vectors (and probability distributions) from the transition matrix as $Q_n[i]$ for $i=1,2,3$. Similarly, the children of first generation immigrants have a transition matrix $Q_m$ which is allowed to be different of $Q_n$

$$Q_m = \begin{bmatrix} q'_{11} & q'_{12} & q'_{13} \\ q'_{21} & q'_{22} & q'_{23} \\ q'_{31} & q'_{32} & q'_{33} \end{bmatrix} = \begin{bmatrix} -Q_{m[1]}^- \\ -Q_{m[2]}^- \\ -Q_{m[3]}^- \end{bmatrix}. \quad (2)$$

Define the $[3 \times 1]$ vectors of skill-types native population and the vector of immigrants at time $[t]$ as $N_t = [N_{1,t} \ N_{2,t} \ N_{3,t}]'$ and $M_t = [M_{1,t} \ M_{2,t} \ M_{3,t}]'$ respectively. Also define immigration quotas of each type (chosen by the median voter) as $\theta_{i,t} = \frac{M_{i,t}}{N_{i,t}}$.

Given the assumptions on intergenerational mobility and fertility, the evolution of the native population is given by

$$N_{i,t+1} = \sum_{j=1}^{3} (\eta_j q_{ji} + \eta'_j q'_{ji} \theta_{j,t}) N_{j,t} \quad (3)$$

for $i = 1,2,3$

which in matrix notation can be written as
\[ N_{t+1} = (Q^T_n \eta + Q^T_m \eta_m \Theta_t) N_t, \]  

(4)

where the fertility profiles and immigration quotas are written for conformity as [3 \times 3] diagonal matrices given by \( \eta = diag \{ \eta_1, \eta_2, \eta_3 \} \) and \( \eta_m = diag \{ \eta'_1, \eta'_2, \eta'_3 \} \), and \( \Theta_t = diag \{ \theta_{1t}, \theta_{2t}, \theta_{3t} \} \). Thus the quantity of immigrants is given by \( M_t = \Theta_t N_t \).

The evolution of the demographic process is assumed to be such that the native medium skilled is always the majority, at any point in time, irrespective of the immigration policy. This assumption will be justified in the calibration of the model. Immigrants don’t have the right to vote, but the children of immigrants are citizens identical in every respect to natives, and as such they have the right to vote. For analysis on immigration policies and different voting regimes for the children of immigrants, see Ortega (2010).

In what follows, it will be convenient to define the evolution of the composition of the population as ratios of unskilled (skilled) to medium-skilled natives. That is, we define the variables

\[ x_{1t} = \frac{N_{1t}}{N_{2t}}, \quad x_{2t} = 1, \quad x_{3t} = \frac{N_{3t}}{N_{2t}}. \]

Define the [3 \times 1] vector stacking these ratios as

\[ X_t = [x_{1t} \quad 1 \quad x_{3t}]^T. \]

In the appendix, we show that the evolution of these ratios is given by

\[ X_{t+1} = S(\Theta_t) X_t S_2(\Theta_t) X_t^{-1}, \]

(5)

\[ S(\Theta_t) = (Q^T_n \eta + Q^T_m \eta_m \Theta_t) \]

(6)

where \( S_2(\Theta_t) \) is the second row-vector from the matrix \( S(\Theta_t) \), while the term \( [S_2(\Theta_t) X_t]^{-1} \) is just the scalar \( \frac{N_{2t}}{N_{2t+1}} \) which can be written as

\[ [S_2(\Theta_t) X_t]^{-1} = \frac{1}{\sum_{i=1}^3 x_{it}(\eta'_i q_{i2} + \eta_i q_{i2})}. \]

(7)

2.2 Production

There exist three inputs of production that correspond to the skill types of agents: unskilled, medium-skilled and skilled labor. Each worker supplies one unit of their labor-type inelastically and all labor inputs are used to produce a final good via a production function \( F(\bullet) \)

\[ Y_t = F(L_{1t}, L_{2t}, L_{3t}), \]

(8)

where \( F(\bullet) \) is constant returns to scale, \( L_{i,t} \) is the total amount of input of type \( i \) at time \( t \). Thus \( L_{1,t} \) is the quantity of unskilled workers, \( L_{2,t} \) the quantity of medium-skilled workers and \( L_{3,t} \) the number of skilled workers. We assume that the production function satisfies \( F_i > 0 \) and \( F_{ii} < 0 \), \( F_{ij} > 0 \) for \( i=1,2,3 \) and \( j=1,2,3 \). In particular, we assume that the production function is of the constant-elasticity-of-substitution form:

\[ Y_t = [\phi_1 (L_{1,t})^p + \phi_2 (L_{2,t})^p + \phi_3 (L_{3,t})^p]^{\frac{1}{p}}, \]

(9)
where for simplicity the share parameters are normalized so that

\[ 1 = \phi_1 + \phi_2 + \phi_3. \tag{10} \]

The elasticity of substitution between labor types is defined as \( \sigma = \frac{1}{1+\rho} \). This production function encompasses among other cases, Cobb-Douglas (\( \sigma = 1 \) if \( \rho = 0 \)), Leontief (\( \sigma = 0 \) if \( \rho = -\infty \)) and perfect substitutes (\( \sigma = \infty \) if \( \rho = 1 \)) production functions. The domain of the parameter \( \rho \) is \((-\infty, 1]\). We discuss calibration of the production function later and abstract from the case where \( \rho = 1 \).

Wages are paid their marginal products

\[ w_{i,t} = \frac{dF}{dL_{i,t}} = \phi_i \left( \frac{Y}{L_i} \right)^{1-\rho}, \text{ for } i=1, 2, 3. \tag{11} \]

Immigrants and natives are assumed to be perfect substitutes in each skill category\(^4\), which defines \( L_{i,t} \) as

\[ L_{i,t} = N_{i,t} + M_{i,t} = N_{i,t} (1 + \theta_{i,t}) \text{ for } i=1, 2, 3. \tag{12} \]

Since the production is constant returns to scale, wages depend on the relative quantity of labor-types and therefore can be written as functions of the the vector \( X_t \), as well as of the immigration quotas \( \Theta_t \),

\[ w_{i,t} = w_i (X_t, \Theta_t), \text{ for } i=1, 2, 3. \tag{13} \]

### 2.3 Government

The role of the government is to proportionally tax income and redistribute the proceeds via lump-sum. The level of taxation/redistribution is exogenous as we focus on immigration for a given fiscal policy.

Because of the higher wages of skilled workers, they are net contributors to the system, while unskilled workers would receive a net benefit. Depending on wages, medium-skilled workers can be net contributors or net beneficiaries.

Defining \( \tau \) as the tax rate, then the transfer \( b_t \) that each worker in the economy receives is given by

\[ b_t = \tau \left[ \frac{\sum_{i=1}^{3} x_{i,t} (1 + \theta_{i,t}) w_{i,t}}{\sum_{i=1}^{3} x_{i,t} (1 + \theta_{i,t})} \right]. \tag{14} \]

\(^4\)There is evidence by Ottaviano and Peri (2006) that immigrants and natives for a same level of schooling/experience are not perfect substitutes. Borjas (2009) argues that for all practical purpose they can be considered perfect substitutes. For the present purpose out of simplicity we assume that they are perfect substitutes.
2.4 Preferences

2.4.1 Static Preferences

Agents receive utility from consumption, which is given by net-wages plus transfers \((b_t)\):

\[
c_{i,t} = (1 - \tau) w_{i,t} + b_t, \quad i = 1, 2, 3, \tag{15}
\]

We assume that the medium-skilled is the majority and therefore choose the immigration policies that maximize the medium-skilled consumption. In the immigration policies considered we assume that the medium-skilled remains the majority and that vote is not given to immigrants as we want to build on this insight for the dynamic case. The exogenous tax rate satisfies \(\tau \in (0, 1)\).

Define the space of feasible policies as elements in \([0, \theta_1^{\text{max}}] \times [0, \theta_2^{\text{max}}] \times [0, \theta_3^{\text{max}}]\), and the optimal (static) policies as \(\theta_i^*\), for \(i = 1, 2, 3\).

We assume that the supply side of high-skilled immigrants is not large enough as to equalize the wages of skilled and medium-skilled agents. This can be considered a reduced-form assumption in order to avoid some unnecessary complications.\(^5\) We do however, discuss the alternative case where \(\theta_3^{\text{max}}\) is such that \(w_3(\theta_3^{\text{max}}) = w_2(\theta_3^{\text{max}})\) in the appendix, where it is assumed that high-skilled workers are able to work as medium-skilled in order to avoid a negative wage premium (see Ortega (2005)).

The conditions that maximize medium-skilled consumption, whose signs we discuss later are

\[
\left[\text{foc} \theta_i \right] \frac{\partial c_2}{\partial \theta_i} = (1 - \tau) \frac{\partial w_2}{\partial \theta_i} + \tau \frac{\partial \pi}{\partial \theta_i} \geq 0; \quad i = 1, 2, 3. \tag{16}
\]

The terms in the above conditions affecting consumption are (1) the net-of-taxes change in medium-skilled-wage due to immigration of agents with skill level \(i\), and (2) the change in the average transfer due to immigration of agents type \(i\). The effect on wages is positive if immigration is of a distinct skill type and negative if same \((\frac{\partial w_2}{\partial \theta_1} = F_{21}, \text{ with } F_{21} > 0, F_{23} > 0 \text{ and } F_{22} < 0)\). In the appendix it is shown that the effect on transfers is positive if the wage that commands the skill type of the immigrant is above the average wage, and negative otherwise \((\frac{\partial \pi}{\partial \theta_1} > 0 \text{ iff } w_i > \overline{w})\). Now we describe the optimal choice for the medium-skilled majority.

With respect to high-skilled immigration we have \(\frac{\partial w_2}{\partial \theta_3} > 0\) for all \(\theta_3 \in [0, \theta_3^{\text{max}}]\) since in this case skilled immigration increases both wages and the

\(^5\) For a small country the assumption of an exogenous limit in the supply of skilled immigrants is probably incorrect. However, since we are discussing immigration to the US (a large country), the assumption makes sense. There are other better reasons why this represents a good first approximation. First, unrealistically large immigration flows would depress the skilled premium in such a way that many agents would find it a better alternative to migrate to other countries rather than to the US. Second, the acquisition of human capital for many professions (lawyers for example) is intrinsically different in different countries. Finally, the order of magnitude of the immigration needed to equalize wages between categories can be very large, which would be incompatible with the supply of land, housing and infrastructure in a country, all these factors are omitted in this model for simplicity.
average transfer. Hence $\theta_3^* = \theta_3^{\text{max}}$. For the other types of immigration the level of redistribution matters. In addition, we show in the appendix that the optimal medium-skilled and unskilled immigration quotas satisfy the condition $\theta_1^* \theta_2^* = 0$. So at least one of those two immigration quotas is 0. We now explain the possible cases for medium and unskilled immigration.

If (the exogenous level of) redistribution is low, the medium-skilled prefers to maximize skilled immigration and chooses the quota of unskilled immigration such that the gains in production complementarities compensate the cost in the form of lower transfers. If redistribution is high, the medium-skilled shuts down medium and unskilled immigration. The medium-skilled might allow some medium-skilled immigration (and no unskilled immigration) if $w_2 (\theta_3^{\text{max}}) > w (\theta_3^{\text{max}})$ in the presence of a large welfare state and after maximizing skilled immigration. In other words, it is theoretically possible for the medium skilled to find in its best interest to allow some medium-skilled immigration to help pay for an expensive welfare state, while necessarily shutting unskilled immigration. We now discuss these possibilities.

Define the auxiliary level $\bar{\theta}_1$ that yields an interior solution (possibly negative) to the unskilled immigration condition

$$\bar{\theta}_1 = \left\{ \theta_1 : (1 - \tau) \frac{\partial w_2}{\partial \theta_1} + \tau \frac{\partial w}{\partial \theta_1} = 0 ; \theta_2 = 0, \theta_3 = \theta_3^{\text{max}} \right\}.$$

Then define $\theta_1^* = \max \left[ 0, \bar{\theta}_1 \right]$. If $w_2 (\theta_3^{\text{max}}, \theta_1^*) \leq w (\theta_3^{\text{max}}, \theta_1^*)$ then the optimal choice for medium skilled is just $\theta_2^* = 0$, while for unskilled immigration is $\theta_1^*$. Notice that if $w_2 (\theta_3^{\text{max}}, \theta_1^*) \leq w (\theta_3^{\text{max}}, \theta_1^*)$ it is the case that $\frac{\partial \bar{\theta}_2}{\partial \theta_2} < 0$ since both terms in (17) would be negative: there would be no complementarity gains and a negative effect in transfers. Regarding unskilled immigration and the size of the welfare state, in the appendix it is shown that $\frac{\partial \bar{\theta}_2}{\partial \theta_2} < 0$ for $0 < \rho < \frac{1 - a + a \left( \frac{L_3}{L_1} \right)}{(1 + a) - 2a \left( \frac{L_3}{L_1} \right) + a \left( \frac{L_3}{L_1} \right)} < 1$ and $a = \left( \frac{L_1}{L_2 + L_3} \right)^6$. In words, the assumptions of this section imply that a bigger welfare state decreases the optimal immigration quota of unskilled workers, everything else constant.

For completeness, if we assume (the empirically irrelevant) case where $w_2 (\theta_3^{\text{max}}, \theta_1 = 0) > w (\theta_3^{\text{max}}, \theta_1 = 0)$ then we can define the auxiliary level $\bar{\theta}_2$ that yields an interior solution (possibly negative) to medium-skilled immigration as

$$\bar{\theta}_2 = \left\{ \theta_2 : (1 - \tau) \frac{\partial w_2}{\partial \theta_2} + \tau \frac{\partial w}{\partial \theta_2} = 0 ; \theta_1 = 0, \theta_3 = \theta_3^{\text{max}} \right\}.$$

Then define $\theta_2^* = \max \left[ 0, \bar{\theta}_2 \right]$. When $\theta_2^* > 0$ and $\theta_1^* = 0$, then this is the optimal choice. This is the case where some medium skilled immigration can

---

6The constraint is sufficient to obtain $\frac{\partial \bar{\theta}_2}{\partial \theta_2} < 0$. Empirically the constraint is satisfied when numbers calibrated to the US. The ratio $\frac{L_1}{L_2 + L_3} = a$ is necessarily less than one as we assume that the medium skilled is the absolute majority.
help pay for an expensive welfare state. Finally, if at the same time \( \theta_1 > 0 \) (defined for \( \theta_2 = 0 \)) and \( \theta_2 > 0 \) (defined for \( \theta_1 = 0 \)), then the policy that yields the higher consumption is optimal, where one of the two immigration quotas is necessarily zero as these options are mutually exclusive. In summary, the optimal policy vectors in this case are one of the following 3 critical points \((0, 0, \theta_3^{\text{max}})\), \((\theta_1^*, 0, \theta_3^{\text{max}})\), \((0, \theta_2^*, \theta_3^{\text{max}})\).

### 2.4.2 Preferences if agents care about their offspring

Following Ortega (2005, 2010), we assume that agents consume (and derive utility \( u(c_{i,t}) \) from it) and vote over immigration policy before the skill-type of their children is revealed. Hence we assume that they value the welfare of their offspring in an expected-utility form. The lifetime utilities of each agent-type in this model are given by

\[
v_{i,t}(c_{i,t}, E[v_{t+1}|i]) = u(c_{i,t}) + \beta E[v_{t+1}|i] \quad \text{for } i = 1, 2, 3. \tag{18}
\]

Where \( \beta \) is a scalar that governs the strength of the altruism motive, the function \( v_{i,t}(c_{i,t}, E[v_{t+1}|i]) \) is the lifetime utility of an agent of skill \( [i] \) at time \( [t] \) and the expectation is conditional on the skill type of the parent. Given the transition matrix, the lifetime utilities can be written as

\[
v_{i,t} = u(c_{i,t}) + \beta [q_{i1}v_{1,t+1} + q_{i2}v_{2,t+1} + q_{i3}v_{3,t+1}] \quad \text{for } i = 1, 2, 3, \tag{19}
\]

where for simplicity we omitted the arguments in \( v_{i,t} \) but discuss this more thoroughly in the next section. In matrix notation we can write

\[
V_t = U(C_t) + \beta Q_n V_{t+1} \quad \tag{20}
\]

where the vectors \( V_t \) and \( U(C_t) \) are defined as \( V_t = [v_{1,t} \ v_{2,t} \ v_{3,t}]^T \) and \( U(C_t) = [u(c_{1,t}) \ u(c_{2,t}) \ u(c_{3,t})]^T. \)

Since wages depend on the relative supply of each type of input, immigration affects utility of natives directly by its current effect on wages, but also affects the wages and welfare in future periods by altering the composition of the future native labor force. In addition, an immigration policy changes \( b_t. \)

### 2.5 Dynamic Equilibrium

The transition matrices that we consider and fertility rates are such that the medium-skilled are the majority, irrespective of the immigration quotas (see the next section for the empirical justification). Hence they choose policy every period.

Since at every point in time the problem is repeated for each generation, where the majority is the medium-skilled and the state of the economy is given by the composition of the population (vector \( X_t \)), we use Markov perfect equilibrium (MPE) as the equilibrium concept. In the MPE the equilibrium strategies
are a function of the state but not of the history of the game. This concept is typically used in dynamic political macroeconomic games.

Under this equilibrium concept, agents fully incorporate the expected induced effects on future policy due to changes in the current policy. In other words, when making a voting decision agents consider how future generations will respond to changes in current policy. For the medium skilled workers, who are the majority, a strategy maps the state $X_t$ into an immigration quota for each type of worker.

Since the production function is constant returns to scale, the variables in the vector $X$ are sufficient to characterize wages, and those ratios are also sufficient to characterize the optimal policies since they characterize the state. We now provide a definition of the political economic equilibrium.

Definition. A politico-economic equilibrium is a policy rule $g : X \mapsto \Theta \in \mathbb{R}_+^3$ and 3 value functions $(v_1^*, v_2^*, v_3^*)$ such that

i) Given the policy rule $g$, continuation values are given by

\[
\begin{align*}
v_1^* (X) & = u_1 (X, g (X)) + \beta \sum_j q_{1j} v_j^* (\Psi (X, g (X))) \\
v_2^* (X) & = u_2 (X, g (X)) + \beta \sum_j q_{2j} v_j^* (\Psi (X, g (X))) \\
v_3^* (X) & = u_3 (X, g (X)) + \beta \sum_j q_{3j} v_j^* (\Psi (X, g (X)))
\end{align*}
\] (21)

where

\[
\Psi (X, g (X)) = (Q_n^T \eta + Q_m^T \eta_m g (X)) X [S_2 (g (X)) X]^{-1}
\] (22)

\[
[S_2 (g (X)) X]^{-1} = \frac{1}{\sum_i x_i (\eta_i \theta_i q_{i1} + \eta_i q_{i2})} ; \quad x_2 = 1
\] (23)

ii) the optimal policy rule solves the problem

\[
g (X) = \arg \max_{\Theta \in H(X)} \left\{ u_2 (X, \Theta) + \beta \sum_j q_{2j} v_j^* (\Psi (X, \Theta)) \right\}
\] (24)

What the definition above says is the following. Given a policy rule $g$, the value functions $v_i^*$ ($i=1, 2, 3$) are the expected lifetime utilities of agents that are consistent with the population process and the respective state of the economy induced by $g$. The policy $g$ in turn is the policy that maximizes the expected lifetime utility of the medium skilled agent and takes into account that their offspring might be in the minorities in the next generation, as well as the response of the next generation to the induced state $(\Psi (X, g (X)))$. The value functions of the unskilled and skilled types are given by the expected lifetime utilities that result from the current majority (the medium-skilled worker) choosing the immigration quotas, taking into account that their offspring could be choosing policy in the next period (if in state "medium-skilled").
2.5.1 Are the optimal policies for the static and dynamic versions the same?

Ortega (2010) shows that with 2 skill levels and political choices over immigration and redistribution, the optimal policies for the static and the dynamic versions of the model are the same given that the choice set is not state-dependent. This is in general not the case in our model that has 3 skill levels and where for some versions of the model the choice set is also not state-dependent. We discuss only the case where the choice is not state-dependent, but the main ideas apply to the case where there exists a positive-sloping supply of high-skill immigrants.

Notice that the first order conditions for the static and dynamic problems are different. The dynamic problem has first order conditions that consist of the trade-offs of the static problem (complementarities vs fiscal effects), plus terms that depend on the future composition of the population induced by the chosen immigration policy.

Define the first order conditions of the dynamic problem by

\[
fo_{ci} = \left( \frac{\partial u_2}{\partial c_2} \frac{\partial c_2}{\partial \theta_i} + \beta \sum_{j} q_{2j} \frac{\partial v_j^*}{\partial \theta_i} \right); \quad (25)
\]

where we assume that the value functions \( v_i^* \) (i=1, 2, 3) are known. Then for the system of first order conditions (i=1,2,3) either of the following sets of conditions hold:

\[
fo_{ci} \leq 0 \quad \& \quad fo_{ci} * \theta_i^* = 0 \quad \& \quad fo_{ci} * (\theta_i^* - \theta_i^{\text{max}}) < 0, \quad \text{or}
\]

\[
fo_{ci} \geq 0 \quad \& \quad fo_{ci} * \theta_i^* > 0 \quad \& \quad fo_{ci} * (\theta_i^* - \theta_i^{\text{max}}) = 0, \quad \text{with}
\]

\[
fo_{ci} = 0 \quad \text{if } 0 < \theta_i^* < \theta_i^{\text{max}}, \quad fo_{ci} \leq 0 \quad \text{if } \theta_i^* = 0 \quad \text{and} \quad fo_{ci} \geq 0 \quad \text{if } \theta_i^* = \theta_i^{\text{max}}.
\]

Where the inequalities are the Kuhn-Tucker conditions to accommodate the possibilities of corner solutions due to immigration quotas of zero, or quotas hitting the maximum \( \theta_i^{\text{max}} \).

Notice that the term \( \frac{\partial c_2}{\partial \theta_i} \) in the dynamic first order conditions is precisely the term obtained from the static problem (agents don’t care about their offspring). Thus the first order condition in (25) can be written as

\[
fo_{ci} = \frac{\partial u_2}{\partial c_2} \text{[static } fo_{ci}] + \beta \sum_{j} q_{2j} \frac{\partial v_j^*}{\partial \theta_i}.
\]

There’s no reason a priori to expect for the first order conditions of the static and dynamic problems to be the same. In the computational part of this paper we use a value function iteration algorithm and the first iteration is typically given by the solution to the static problem, but more iterations

\footnote{For the first iteration, if for the initial value functions we use the value 0 for all state points in order to initialize the iterative procedure leads to a first iteration that finds the solution to the static problem.}
are required in order for a solution to be found, and the numerical solutions are indeed different. Cases in which the solution would be the same typically involve parameterizations where in the dynamic problem results in $\theta^*_1 = 0$, $\theta^*_2 = 0$ and $\theta^*_3 = \theta_{3\text{max}}$. Hence, provided the problem yields at least one solution that is interior (for at least one of the immigration quotas) the policies won’t be the same. $^8$

3 Data and Calibration

3.1 Intergenerational Mobility Matrices

Many studies of intergenerational mobility use transition matrices to analyze the relative position of people in the income distribution compared to that of their parents. Such studies typically suffer from possible life-cycle effects and a small number of observations (PSID, NLS for example). Since we are interested in the intergenerational mobility of individuals based on skills, we focus on the education levels of individuals and their parents in studying intergenerational mobility.

We use the General Social Survey (GSS) for this matter, which is an annual survey since 1972. Starting in 1977, this survey captures, in addition to the schooling level of the respondent, information on the schooling level of the respondent’s parents. The survey also identifies whether the respondents and their parents were born in the US. Hence, we can estimate transition matrices and perform some statistical tests for natives and children of first generation immigrants.

We consider individuals who were born on or after 1945 and whose age at the time of the interview was between 25 and 55 years old. Some individuals were born before, but since the earliest GSS wave that has the information used in this paper is in 1977, that implies using individuals that might be too old. We cap it at 55 because of a possible relationship between mortality and education level. The education variable used to classify individuals is based on whether the individuals (either respondent or his/her parents) obtained any of the following degrees: less than high school, high school (HS), junior college, college and grad school. We classify individuals as 2nd generation immigrants, whose transition matrix we are interested in, if the respondent was born in the US but any of the parents were born outside the US.$^9$ Natives in turn are individuals whose

$^8$In particular, if we use a version of the model with high-enough redistribution (and conditional on the rest of parameters) and a small pool of high skill immigrants (implying a perfectly inelastic supply at the level $\theta_{3\text{max}}$), we can obtain the same optimal policy under the static and dynamic problems given by $\theta^*_1 = \theta^*_2 = 0$ and $\theta^*_3 = \theta_{3\text{max}}$. However, if there is an interior solution to the unskilled-immigration first order condition in the dynamic problem and even assuming $\theta^*_2 = 0$ and $\theta^*_3 = \theta_{3\text{max}}$, then the result would be clearly different for $\theta^*_1$ in the dynamic and static versions as the static first order condition contains some of the terms included in the first order condition of the dynamic problem.

$^9$We could classify second generation immigrants as those respondents born in the US whose both parents (rather than only one) were born outside the US. We don’t do this because 1) the number of observations greatly decreases for second generation immigrants (from 1259 to
parents were born in the US.

For the transition matrices we define "unskilled" as an individual with less than a high school degree, "medium-skilled" as either having a high-school or a junior college degree, and a "skilled" individual if he has either a college or graduate degree. For individuals with information on both parents, we use the maximum degree obtained by any of them. The details on the construction of these matrices, including the criteria for data selection, matrices estimated under alternative assumptions as well as statistical tests on them are found in the appendix.

The transition matrices $Q_n$ and $Q_m$ are estimated from the GSS as follows. For each element $q_{ij}$ in (1) and $q'_{ij}$ in (2), the estimate is the number of children with corresponding education level $i$ and parents with education level $j$ divided by the total number of parents with educational level $j$. The empirical frequencies differ somewhat across subsamples in the GSS (e.g. for men and women; see appendix); but since they do not differ greatly -and we cannot reject that they are different for native men and women, our main calibration uses results for both all men and women, given by

$$Q_n = \begin{bmatrix} .256 & .663 & .081 \\ .062 & .707 & .231 \\ .010 & .397 & .593 \end{bmatrix}, \quad Q_m = \begin{bmatrix} .211 & .594 & .195 \\ .067 & .633 & .299 \\ .022 & .325 & .653 \end{bmatrix}.$$

The first two rows of the transition matrices show that the children of unskilled and medium skilled immigrants to the US appear to be more "successful" than natives. In the case of unskilled parents, children of immigrants have a lower probability of staying unskilled, and a higher probability of upward mobility. Indeed, children of unskilled natives have probability of 8% of becoming skilled, while the children of unskilled immigrants have a probability of almost 20%. Given medium-skilled parents, the differences are not as marked as in the unskilled case but their odds seem to be somewhat better.

We test formally whether the probability distributions for natives and children of immigrants are statistically the same, conditional on the skill of the parents. This is just a test on whether the rows in both matrix $Q_m$ and $Q_n$ are statistically the same ($H_0 : q_{ij}^1 = q_{ij}^2$ for all $j = 1, 2, 3$, given row $i$). Under the null, the statistic

$$\sum_{m=1}^{2} n_{im} \sum_{j=1}^{k} \frac{(q_{ij} - \hat{q}_{ij})^2}{\hat{q}_{ij}}$$

is distributed as a chi-square with $(k - 1)$ degrees of freedom, where $n_{im}$ is the number of counts of row $i$ of sample $m$. Since $k = 3$, the statistics for each row are to be compared with $\chi^2_2$ , which are 5.99 and 9.21 for the 5% and 1% significance levels, respectively. For more details on these tests see the appendix.

The statistics are 58.74 for the first row, 18.63 for the second and 10.16 for the third. Hence the data rejects that the probability distribution of each type of worker is the same at the 1% level.

434] and 2) because the numbers don’t appear to be significantly different according to the unskilled, medium-skilled, skilled classification used in this paper. See the appendix.
We also test for the equality of both matrices, with a test given by summing over the statistic of each row, with the null hypothesis that \( q_{ij}^1 = q_{ij}^2 \) for all \( i = 1, 2, 3 \) and \( j = 1, 2, 3 \). Then the null is rejected if
\[
\sum_{i=1}^{k} \sum_{m=1}^{2} n_{im} \sum_{j=1}^{k} \frac{(q_{ij} - \bar{q}_{ij})^2}{q_{ij}} > \chi^2_{(k(k-1))}
\]
where the degrees of freedom are in this case \( k (k - 1) = 3 (2) = 6 \). The test produces a statistic of 87.52, which is to be compared with a \( \chi^2_{(6)} \) critical value at 1% significance of 16.81. Hence, the test rejects that children of natives and those of first generation immigrants have the same transition matrix. In the appendix we perform more tests for several subcategories of the data. In general, the findings are that 1) the unskilled immigrant category has a higher upward mobility \( (q_{13}^1) \) as well as lower probabilities for the unskilled and medium-skilled state, 2) for most partitions of the data the probability distribution for children of skilled parents is statistically the same for both natives and immigrants; 3) for medium skilled workers there is only statistical difference for women and 4) all tests of equality of matrices between natives and immigrants rejected that the intergenerational transition matrices are statistically the same, mostly due to differences in intergenerational mobility of the unskilled.

Another way to characterize the differences in population dynamics is by looking at the steady state distributions induced by the estimated transition matrices and fertility rate profiles. We do this after the following section on fertility rates.

### 3.2 Fertility Rates

In order to calibrate the number of children that agents have, we construct total fertility rates by education level (TFR) for 3 different years: 1990, 2000 and 2005. The concept of total fertility rate measures the expected number of children that a woman would have in her lifetime if she was subject to the current (cross-section) age-specific fertility profiles. \(^{10}\)

Total fertility rates can be accurately calculated with birth data from the National Center for Health Statistics (downloaded from their VitalStats system), and age-education groups from Census data (1990, 2000) and the Current Population Survey (2005). However, the available data on births doesn’t detail whether the mothers are US-born, or foreign-born. In order to arrive at total fertility rates by place of births of mothers we also use information from several years of the American Community Survey (ACS). Details on the construction of these estimates are in the appendix.

\(^{10}\)Total Fertility rates are constructed as follows. First divide the total number of births whose mothers are in a specific age-education group (i.e. 20-24 year old mothers with 11 or less years of education) by the number of women in that age-education group. Since the age groups are for 5 year increments, the resulting number is multiplied by 5. Finally, those quantities are added for a given education level. The result is a cross-section measure of expected number of children a woman would have in her lifetime.
Table 1. Total fertility rates by skill level. Years 1990, 2000 & 2005

<table>
<thead>
<tr>
<th>Year</th>
<th>US-Born</th>
<th>Foreign-Born</th>
<th>All Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Med</td>
<td>Hi</td>
</tr>
<tr>
<td>2005</td>
<td>1.98</td>
<td>1.94</td>
<td>1.82</td>
</tr>
<tr>
<td>2000</td>
<td>2.24</td>
<td>2.00</td>
<td>1.59</td>
</tr>
<tr>
<td>1990</td>
<td>2.41</td>
<td>1.89</td>
<td>1.82</td>
</tr>
<tr>
<td>Average</td>
<td>2.21</td>
<td>1.94</td>
<td>1.75</td>
</tr>
</tbody>
</table>

The estimated fertility rates show the well-known negative relationship between education and fertility, and also display that foreign-born women have higher fertility rates than US-born women of the same skill level. Given these estimated TFR’s, the implied model parameters for fertility by skill level are \( \eta_1 = 1.1, \eta_2 = 0.97, \eta_3 = 0.87 \) and \( \eta_1' = 1.5, \eta_2' = 1.22, \eta_3' = 0.96 \).

3.3 Population Composition at Steady State

A population process is described in this paper by an intergenerational transition matrix \((Q)\) and a (diagonal) matrix of fertility rates \((\eta)\), which yield a specific composition of the population at steady state \((X)\). Because we allow for different population processes for natives and first-generation immigrants and in some respects comparing those objects for natives and immigrants can be uninformative, we also look at the steady state distributions induced by those objects \((Q_i, \eta_i)\). Notice however, that when considering immigration we only assume that the first-generation of immigrants is subject to the different intergenerational mobility and fertility process, as \(2^{nd}, 3^{rd}\) and \(n\)-th generation immigrants are considered natives.

Under identical fertility but different transition matrices the steady state distribution induced by \((Q, \eta)\) yields an unskilled to medium-skilled ratio \(x_1\) of 10.10% and high-skilled to medium-skilled \(x_3\) of 53.32%. The pair \((Q_m, \eta)\) in turn yields \(x_1 = 11.83\%\) and \(x_3 = 85.09\%\). Hence, assuming identical (estimated) fertility the pair \((Q_m, \eta)\) yields a distribution with more weight on the extremes of the skill distribution: about 18% more unskilled workers than natives and about 60% more skilled individuals. When allowing for differential fertility (higher fertility for immigrants), the composition is such that there are 24% more unskilled workers and 47% more skilled workers, and this in turn involves less medium-skilled workers as total percentage of the population.

Table 2. Steady State Composition of Population.

<table>
<thead>
<tr>
<th></th>
<th>Identical Fertility</th>
<th>Differential Fertility</th>
<th>(\left( Q_{n[i]}, \eta_{n[i]} \right)) rows in place for (\left( Q_{n[i]}, \eta_{n[i]} \right))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_1)</td>
<td>(Q_n, \eta)</td>
<td>(Q_m, \eta)</td>
<td>(Q_n, \eta_n)</td>
</tr>
<tr>
<td></td>
<td>.1010</td>
<td>.1187</td>
<td>.0978</td>
</tr>
<tr>
<td>(x_3)</td>
<td>.5332</td>
<td>.8499</td>
<td>.5429</td>
</tr>
</tbody>
</table>
We also see the effect of replacing the demographic profile of natives by that of immigrants, one agent-type at the time. In all cases the steady state composition has more low-skilled and high-skilled agents relative to the medium skilled. When replacing the demographic profile of low-skilled natives by those of low-skilled immigrants, the result is about 1.4% more low-skilled workers and 2.7% more high-skilled workers at steady state. Doing this exercise for medium skilled workers results in 15% more unskilled and 22% more skilled workers. Finally, in the case of skilled agents, there is 13% more unskilled and 19.5% more skilled agents. Table 2 summarizes the steady state ratios $x_1 = \frac{N_1}{N_2}$ and $x_3 = \frac{N_3}{N_2}$ under the different assumptions.

### 3.4 Production

Production parameters are $\phi_1, \phi_2, \phi_3$ and $\rho$. In the labor literature, the elasticity of substitution $\left(\frac{1}{\rho}\right)$ is reported to be between 1.4 and 2.5 in several studies.\(^{11}\) We use $\rho = \frac{1}{2}$, which implies an elasticity of substitution of 2, but later we do sensitivity analysis.

Given $\rho$, we can calibrate the stylized production function to match average skill premia so that the model replicates the wage premia at the steady state of the demographic process. The equations relating wage premia to the production parameters are

$$\left(\frac{\phi_i}{\phi_j}\right) = \frac{w_{i,t}}{w_{j,t}} \left(\frac{L_{j,t}}{L_{i,t}}\right)^{\rho-1} \quad \text{for } i \neq j=1,2,3,$$

and the normalized shares equation given by (10).

For the wage premia we use census data of the mean hourly wage of workers that are between 25 and 65 years old by educational attainment; that work at least 40 hours per week, and that worked at least 40 weeks in the previous year for census years 1990, 2000 and 2005. The skill categories are defined in terms of schooling under the same definition as for the intergenerational mobility matrices. The average wage ratios obtained are $\left\{\frac{w_2}{w_1}, \frac{w_3}{w_2}\right\} = \{1.315, 1.67\}$.

The demographic process given by the intergenerational mobility matrix $Q_n$ and estimated fertility rates for US native women yield steady state native ratios of $x_1^* = \frac{N_1}{N_2} = 0.09782$ and $x_3^* = \frac{N_3}{N_2} = 0.5429$. Taking into account immigration quotas of 18% for the unskilled group, 6% for the medium skilled and 13% for the high skilled group yield labor ratios given by $\left\{\frac{L_1}{L_2}, \frac{L_3}{L_2}\right\} = \{.1089, .5788\}$. Using $\rho = \frac{1}{2}$ for a baseline parameterization, we obtain estimates $\left\{\widehat{\phi}_1, \widehat{\phi}_2, \widehat{\phi}_3\right\}$ given by $\{.0995, .3966, .5039\}$.

---

\(^{11}\)See for example Ottaviano and Peri (2006).
3.5 Preferences

We assume that the period utility function \( u(x) \) is constant-relative-risk-aversion

\[
 u(x) = \frac{x^{1-\sigma} - 1}{1 - \sigma},
\]

(27)

Where \( \sigma \) is set at a baseline level of 1 (log-utility). For the discount factor \( \beta \), we follow Ortega (2005) in setting an annual value of \( \beta = .985 \), which for 30 year periods yield a model parameter of \( \beta = .985^{30} = .63546 \). We perform sensitivity analysis later.

3.6 Supply of Immigrants and Guest Workers

In most exercises we assume that agents choose immigration policies in the space \([0, \theta_1^{\text{max}}] \times [0, \theta_2^{\text{max}}] \times [0, \theta_3^{\text{max}}] \). For the upper limit on unskilled immigrants (\( \theta_1^{\text{max}} \)), we assume that \( \theta_1^{\text{max}} \) is a huge number so that any unskilled number desired by the majority is available as for all practical purposes that seems to be the case for a country like the US. In the case of the medium skilled, a priori we could argue that it is a big number that implies a bigger pool for medium skilled workers than the pool for high skilled (reflecting relative scarcity of high-skilled), but this parameter turns out to be irrelevant in the numerical exercises. Modelling the pool of high skilled immigrants in turn is not straightforward for reasons that we detail later. Since the quantitative exercises will depend on the specific assumption on the pool of high skilled immigrants we proceed by analyzing the model under different assumptions: (i) an initial case where the supply of high skilled immigrants is unrealistically "large", where the maximum skilled immigration quota (\( \theta_3^{\text{max}} \)) is large enough as to equalize the wages of the medium and high-skilled workers if the majority chooses it; (ii) a case with a "smaller" limit on the supply of high skilled immigration; (iii) cases where guest workers are available for each skill-type in addition to immigration; and (iv) cases with a supply of high-skilled immigration that is wage-dependent elastic. We discuss the rationale for each of these cases below.

(i) An unrealistically "large" pool of skilled immigrants. The initial parameterization allows for a very large pool of skilled immigrants in the set of available immigration policies in \([0, \theta_1^{\text{max}}] \times [0, \theta_2^{\text{max}}] \times [0, \theta_3^{\text{max}}] \). These policies never imply a negative wage premia as it is assumed that high-skilled individuals can work as medium-skilled if they choose to do so, which would happen only if the immigration quota of the high-skilled was (unrealistically) very large.\(^{12}\) The same applies to the medium-skilled if they choose to work as unskilled workers.

Note that policies that yield the same wage structure in the current period produce a different future composition of the population, and hence policies that yield the same wage structure represent distinct immigration policy points due to differing dynamic effects. We later show that this assumption on the

\(^{12}\)In the main calibration of the model an immigration quota of about 200% would be needed in order to equalize the wages of the medium skilled and the high skilled at the initial steady state.
supply side leads to quantitative results that are unreasonable, but we include it because it can help in understanding the results of the other (more relevant) cases.

(ii) A "small" pool of high skilled immigration. In this case we assume that the pool of high-skill immigration is constrained by some number $\theta_3^{\text{max}}$ that is not large enough as to equalize wages between medium and high skilled workers. In terms of a supply curve for high-skill immigration, the supply is perfectly inelastic at the level $\theta_3^{\text{max}}$. The political majority in the host economy chooses their most preferred level of high skilled immigration between 0 and $\theta_3^{\text{max}}$. We offer 3 possible justifications for this assumption. First, since immigration laws in the US typically allow high-skilled individuals easier access to coming/staying in the country (specially those with advanced degrees), the observed skilled immigration quota could represent a supply-side constraint. An argument against this possibility is that HB visas in the US typically exhaust very quickly.\footnote{We thank an anonymous referee for this observation.} Second, it is possible that the current immigration policy of the US with only $\theta_3^* = 13\%$ is suboptimal (thus not at an interior solution) and understanding the effects of an exogenous limit on high-skilled immigration is of independent interest on the immigration policies that would result from removing this constraint since immigration reform can be observed in a "piece-meal" approach. Hence a reform over high-skilled immigration is likely to bring changes to the other categories over time, and vice-versa. The final justification is as follows. First assume that the world supply of high skilled immigration is perfectly inelastic and that the home economy is competing with other countries for the world pool of high skilled workers. Provided that the wages of the home economy (including immigration) are higher than in the rest of the world then the world supply of high skilled immigrants will also be the supply of high skilled immigrants for the discussed economy. By assuming a more realistic level of $\theta_3^{\text{max}}$ than in the previous case we capture some of this intuition and later show that practically the same effects are obtained with a wage-elastic supply of high skilled immigration.

(iii) The supply of guest workers. Full-immigration in this model has dynamic effects because immigrants have children, which affect the composition of the (future) native population. A guest worker program in this model has only static effects since those agents have a fertility rate of 0 (they return to their home country after their working-period). We assume that the hard maximum of immigrants+guest workers of each type is given by $\theta_i^{\text{max}}$ ($i = 1, 2, 3$). Agents are willing to come to the host economy as either full-immigrants who get to stay in the host country with their children, or as guest workers. Define the quota of guest workers chosen as $\theta_{iB}$ and the quota of immigrants as $\theta_i$ for $i = 1, 2, 3$, then it is assumed that $\theta_{iB} \leq \theta_i^{\text{max}}$, $\theta_i \leq \theta_i^{\text{max}}$ and $\theta_{iB} + \theta_i \leq \theta_i^{\text{max}}$.

(iv) Wage-dependent elastic supply of high skilled immigrants. The previous specifications have policy spaces that are not state-dependent. We relax this assumption and make the findings more robust by considering a wage-elastic supply of high skilled immigration. This assumption captures a stylized
constraint where the host economy competes with other countries for the pool of high skilled immigrants in the world. We leave the discussion of this case to the sensitivity section.

3.7 The Model Under an Initial Parameterization and a Very Large Pool of High-Skilled Immigrants

In this section we use an initial parameterization of the model in order to understand its properties when the pool of skilled immigrants is (unrealistically) very large. That is, we investigate which immigration policies would be chosen by the majority in the absence of any meaningful constraint from the supply side of immigration.

We interpret a model-period as 30 years, which is roughly the average age of mothers. Given the 5.15 immigrants per 1000 natives annual estimate for unskilled workers, and 4.15 skilled immigrants per 1000 skilled natives, the quotas for a 30 year equivalent are roughly \( \theta_1 = 18\% \) and \( \theta_3 = 13\% \). For the tax rate, we use an average tax rate of 30\%, approximately the current average from the series computed in McDaniel (2012) for labor taxes and consumption taxes for the US.

We use a value function iteration algorithm in order to solve for the Markov perfect equilibrium (MPE) of the model, with discretized state and policy spaces and bilinear interpolation in the evaluation of the value function.\(^\text{14}\) The state variables of the model are the ratios low–skilled-to-medium-skilled natives \( x_1 = \frac{N_1}{N_2} \) and high-skilled-to-medium-skilled natives \( x_3 = \frac{N_3}{N_2} \). Given the demographic profile of natives \((Q, \eta)\), we first compute the steady state distribution of skill types of the native population in absence of immigration, which we denote as \( \{x_1^{SS}, x_3^{SS}\} = \{0.09782, 0.5429\} \). Then given the demographic profile of immigrants \((Q_m, \eta_m)\) we consider a grid in the state space that contains both the initial steady state and any possible future steady state induced by the space of immigration policies. The policy space is given by elements in \([0, \theta_1^{\max}] \times [0, \theta_2^{\max}] \times [0, \theta_3^{\max}]\).\(^\text{15, 16}\)

We document the results when the maximum immigration quotas are \( \theta_1^{\max} = \)

\(^{14}\)Spline interpolation was also used. The results are identical to bilinear interpolation, but the computational cost is much higher.

\(^{15}\)In general, using an appropriate grid and policy space requires some experimentation since choosing a very small policy space might produce results that involve the upper bound in the policy space for unskilled or high-skilled immigration; and in general bigger policy spaces require (somewhat) bigger state spaces so that (i) the induced states are part of the state-grid considered and (ii) the immigration policies chosen by the medium skilled are in the interior of the policy space.

\(^{16}\)Multiple policy points might produce the same payoffs, only when \( \beta = 0 \). When \( \beta = 0 \) it can be the case that for some policy given by \( \{\theta_1, \theta_2, \theta_3\} \) we have that \( u(\theta_1, \theta_2, \theta_3) = u(k\theta_1, k\theta_2, k\theta_3) \) for any positive constant \( k \) for policy points that don’t include a zero immigration quota (i.e. \( \theta_1 = 0 \) not included). However, we proved in a previous section that for the static case \( (\beta = 0) \) the agent optimally chooses \( \theta_1^* \theta_2^* = 0 \). When \( \beta > 0 \) (the dynamic model) the expected utility of any 2 policies that yield the same wages and transfers in the first period produce different composition of the population in the future, which implies different payoffs (expected utility).
As explained before, these are very large numbers in order to understand optimal choices in the absence of constraints on the supply side (as opposed to using a small policy space where the solution could imply hitting one or more of the maximum immigration quotas). Holding everything else constant, it makes no difference on the results of the model to consider a bigger/smaller unskilled policy space provided that the optimal policy function is in the interior of that space. In the case of medium-skilled immigration, we show that the optimal policy of equilibrium under any numerical exercise will always dictate \( \theta_2^* = 0\% \) and hence the maximum immigration quota \( (\theta_2^{\text{max}}) \) turns out to be irrelevant as the optimal choice for any point in the state space is \( \theta_2^* = 0\% \). The specific quota for high skilled immigration has to be at least 1.98 (198%), which is the number needed for an equalization of wages of the skilled and medium skilled at the initial steady state. The specific number used (2.5) is large enough to accommodate the optimal choices at state points \((x_1, x_3)\) other than at the initial steady state.

Given this parameterization, the model predicts immigration of the extremes in the skill distribution, with a high immigration quota for the unskilled and somewhat smaller quota (but higher absolute numbers) for the high-skilled. At the no-immigration steady state the quotas chosen are \( \theta_1^* = 281\% \) and \( \theta_3^* = 200\% \). In the induced steady state the immigration quotas are 189% and 101% respectively. The assumption of a huge pool of skilled immigrants, in addition to being hard to justify due to the relative scarcity of high skilled immigrants, leads also to unreasonable results and hence we modify the model in subsequent sections.

The optimal policy functions in this exercise show a very intuitive behavior. The quota of high-skilled immigration is decreasing in \( x_3 \) (the higher the share of native high-skilled relative to the medium-skilled, the lower the demand for high skilled immigration), and is also increasing in \( x_1 \) as more low-skilled immigration also benefits the high-skilled immigration by increasing wages of the high skilled. The opposite is observed for immigration of the unskilled: it is increasing in \( x_3 \) as more high skilled immigration make room for more unskilled immigration by not making unskilled immigration as expensive in terms of fiscal costs, and it is decreasing in \( x_1 \) which means that the higher the share of unskilled natives as percentage of medium skilled, the lower the demand for unskilled immigrants. Given the assumptions, the trade-offs are such that medium-skilled immigration is not allowed.

One way in which these numbers can diminish somewhat is by assuming a small constant marginal cost per immigrant that is paid out of the transfers to

---

17 In terms of people, an immigration quota of 100% for the medium skilled is a very large number in absolute terms since the majority in the population is of that type, though using a much larger number (or smaller) has no effect on the findings.

18 Higher maximum immigration quotas than those used do not affect the solution. Indeed since \( \theta_2^* = 0 \) is chosen, the size of \( \theta_2^{\text{max}} \) turned out to be irrelevant. In the case of \( \theta_1^{\text{max}} \) and \( \theta_3^{\text{max}} \) it is important to choose a space "big enough" as to allow for the interior solution so that the optimal policy is in the interior of the policy space. Choosing other maximum immigration quotas do not affect the solution once the policy space is big enough as to produce the interior solutions.
all agents. This could be justified in terms of externalities associated to the absorption of very big immigration flows that are not captured in this parsimonious model; like congestion effects. In such a case, a very small cost (i.e. a small cost per-immigrant chosen so that the lump-sum transfer decreases by a very small amount that in the presence of huge immigration quotas represents about 1% of the lump-sum transfer) can produce a reduction in immigration quotas. This reduces the skilled quota slightly, and more significantly the unskilled quota. The numbers are still very large, given by 

\[ \frac{1}{244\%} \] and 

\[ \frac{3}{198\%} \]; and

\[ \frac{1}{245\%} \] and 

\[ \frac{3}{100\%} \]. Increasing this cost leads to lower overall immigration, with \( \theta_1 \) decreasing faster than \( \theta_3 \).

**Table 3. Initial Parameterization**

<table>
<thead>
<tr>
<th>Demographic Profiles</th>
<th>Production and Taxes</th>
<th>Immigration Pool and Prefs</th>
</tr>
</thead>
</table>
| \( Q_n \) = \[
\begin{bmatrix}
0.256 & 0.663 & 0.081 \\
0.062 & 0.707 & 0.231 \\
0.100 & 0.397 & 0.593
\end{bmatrix}
\] | \( L_2 \ \frac{L_1}{L_3} \) = \[
\begin{bmatrix}
0.1089 \\
0.5788
\end{bmatrix}
\] | \( \theta_1^{\text{max}} = 4.5 \) |
| \( Q_m \) = \[
\begin{bmatrix}
0.211 & 0.594 & 0.195 \\
0.067 & 0.633 & 0.299 \\
0.022 & 0.325 & 0.653
\end{bmatrix}
\] | \( \frac{w}{w_1} \ \frac{w_2}{w_3} \) = \[
\begin{bmatrix}
1.315 \\
1.670
\end{bmatrix}
\] | \( \beta = 0.985^{30} \) |
| \( \eta \) = diag\{1.1, 0.97, 0.87\} | \( \tau = 0.30 \) |
| \( \eta_m \) = diag\{1.5, 1.22, 0.96\} | |

We entertain some justifications as to why this parameterization of the model overpredicts immigration. First, the observed flows of high-skilled immigration might be suboptimal as more firms, politicians and academics call for more high-skilled immigration than the current status-quo. Under this interpretation, the observed high-skilled flows could be an exogenous limit on immigration. Second, the policy space is invariant to the state space (and thus invariant to the wage premia). For the case of the skilled workers this is far-fetched as it implies that even if the wage premia is 0 (cannot be negative in our model), the available supply of high skilled workers would still be the same (\( \theta_3^{\text{max}} \)). This suggests that a supply of high-skilled immigrants that respond to wages could change the results of how many skilled immigrants are allowed, and this in turn would influence the chosen amount of low-skilled and perhaps affect medium-skilled immigration. This is explored later.

\[ ^{19} \text{In this case transfers can be shown to be given by } \tilde{b} = b - \text{cost} \* \left[ \frac{\Sigma^3_{i=1} \Theta_i}{\Sigma^3_{i=1} \pi_i (1+\theta_i)} \right], \text{ which reduces to the equation shown in the main text when cost = 0.} \]
3.8 The Model with a Perfectly Inelastic Supply of Skilled Immigrants (A Small Pool)

In this section we study the model under a perfectly inelastic supply of high-skilled immigration where the pool of skilled immigrants $\theta_3^{\text{max}}$ is not large enough as to equalize wages between medium skilled and skilled workers.\(^{20}\)

We perform experiments with 2 different levels of $\theta_3^{\text{max}}$, the first one is $\theta_3^{\text{max}} = 13\%$ which is the level that has been observed in the US for the analyzed period and that could represent either the supply side constraint (in light of the results of the above section) but more likely represents some other exogenous constraint. Another level considered is $\theta_3^{\text{max}} = 30\%$, which is a very large number in relation to historical levels of US immigration and that represents a reference that helps seeing the effects of relaxing the exogenous constraint under the interpretation that 13% is suboptimal. These levels matter only for the magnitudes of the chosen policies, but not for the qualitative effects as other levels of $\theta_3^{\text{max}}$ also yield the same (qualitative) predictions provided that high skilled immigration is relatively scarce (as opposed to the assumption in the previous section).

The parameter $\beta$ can be calibrated as to yield $\theta_1^* = 18\%$ as it is found that $\theta_1^*$ and $\beta$ are inversely related (i.e. for the static problem where $\beta = 0$, we obtain much higher unskilled immigration than when $\beta > 0$). We calibrate this parameter as $\hat{\beta} = 0.6312$. With a higher $\beta$, more weight is given on the expected utility of children, and agents pay attention to the bad state "unskilled" and optimally decide to allow less unskilled immigration than otherwise. We label this case with $\theta_3^{\text{max}} = 13\%$ as the "baseline" as this model reasonably and parsimoniously allows for the analysis of many issues in the subsequent sections; while the qualitative predictions are robust to changes that are later discussed.

The optimal policy function consists of maximizing high skilled immigration, minimizing medium skilled immigration and choosing a quota of unskilled workers that is decreasing in $x_1 \left( x_1 = \frac{N_1}{N_2} \right)$ and increasing in $x_3 \left( x_3 = \frac{N_3}{N_2} \right)$. In words, the higher the number of low-skilled natives relative to medium-skilled, the lower demand there is for unskilled immigration as ceteris-paribus there is less "room" for new unskilled workers whose positive effects on medium-skilled wages would be smaller but their net benefit from transfers possibly larger. Similarly, the higher the number of high-skilled agents relative to medium-skilled, the higher the equilibrium quota of low-skilled immigrants since the wages of low and medium-skilled are higher and this in turn allows for more immigration from a redistributive standpoint. Medium-skilled workers are not accepted in equilibrium as the "cost" of allowing some of them (lower current medium-skilled wages) is higher than future benefits. Even though in reality we do observe some medium-skilled immigration, in general it is "small" when compared to the same category in the host economy and hence the model captures the stylized fact of mostly observing immigration of the extremes in the skill distribution.

\(^{20}\)Depending on the size of the pool of skilled immigrants, it can be the case that at some points in the grid in $(x_1, x_3)$ it could be possible for some immigration policies to imply equalization of wages between the high skilled and the medium skilled. However, given the parameterization of the model and the pool of skilled immigrants, that wasn’t the case.
These optimal policies induce a steady state where the native share of low-skilled and high-skilled relative to the medium skilled majority are higher than at the initial steady state, with $x_1 = 0.10$ and $x_3 = 0.589$ (as opposed to initial steady state values of 0.0978 and 0.5429 respectively). The slightly higher share of low-skilled natives induces less low-skill immigration, but the higher share of high skilled natives induces the opposite effect on low-skill immigration, for a total effect at the induced steady state of $\theta_1 = 17.6\%$ and $\theta_3 = 13\%$.

Regarding the effect of $\theta_3^{\text{max}}$ on the immigration quotas of the other types in equilibrium, we find that a higher level of $\theta_3^{\text{max}}$ allows the medium skilled to choose a higher quota of low-skilled workers. When $\theta_3^{\text{max}} = 13\%$ the unskilled quota chosen is $\theta_1^* = 18\%$, while the case when $\theta_3^{\text{max}} = 30\%$ yields $\theta_1^* = 21.6\%$ (and a new steady state induced quota of $\theta_1 = 29\%$ as opposed to 17.6%, due to the higher share of high skilled agents in the native population). Hence, a "piece-meal" approach that relaxes first the amount of high-skilled immigrants would also lead to a politically motivated higher demand for low-skilled immigrants.\footnote{We also analyze a case where we completely shut down high-skill immigration ($\theta_3^{\text{max}} = 0$). In this case the unskilled quota would be slightly lower in the initial steady state ($\theta_1^* = 17\%$ as opposed to 18% when $\theta_3^{\text{max}} = 13\%$), and because in the long run there wouldn’t be as many skilled individuals as in the baseline, the long run quota of unskilled individuals is even lower (10%).}

The politically chosen immigration policy induces a unique steady state irrespective of initial conditions. In general, the policy implies a faster convergence toward the new steady state starting from any point in the state grid than convergence to the old steady state in absence of immigration. The optimal policies are "state-contingent". That is, if the country is in a state with a very small share of low-skill natives, there would be (higher) demand for unskilled immigration. The "bracero" program that the US signed with Mexico in 1942 could be an example of this. Similarly, illegal immigration (typically confined to unskilled jobs) could also be interpreted as a demand side factor (at least partially) as the native unskilled population in the US as percentage of the labor force has been shrinking over time.

### 3.9 Immigration Policy if Guest-Workers are Available

In this section we enlarge the policy space to allow for the possibility of guest worker quotas, in addition to immigration quotas. In the model, the only difference between guest workers (or braceros) and immigrants is that immigrants affect the future composition of the native population because they have children, while guest workers have a zero fertility rate (they return to their home country). We perform this exercise under the exogenous limit on skilled immigration, and in the sensitivity section we repeat the analysis with a wage-elastic supply for high skill immigration.

Using the same parameterization as in the previous section (the baseline) but allowing for the 3 extra choice variables which are guest worker quotas ($\theta_{1B}, \theta_{2B}, \theta_{3B}$), we find that the median voter would choose full immigration to the available pool of high skilled immigrants (no guest worker for them),
no immigration/guest worker for the medium skilled, and for the unskilled a
positive quota of guest workers (without immigration). We discuss the reasons
below.

The median voter would offer full immigration to skilled individuals rather
than a guest worker program because in equilibrium the median voter would like
more skilled workers than the current available supply of skilled immigants.
By granting them full immigration, the native medium-skilled can affect the
composition of the population in a more advantageous way to their children,
who have a large probability of being medium skilled as their parents.

In the case of the medium skilled, allowing for guest workers doesn’t change
the results since there would only be "costs" of allowing braceros of the medium-
skill type (lower wages), while there would be no benefits (no possibly advan-
tageous change in the future composition of native workers) for the medium
skilled native.

The main changes are observed in the low-skill category as the medium
skilled majority prefers the guest worker program to immigration. When com-
paring across regimes, the quota of low-skill guest workers is higher than the
low-skill immigration quota in absence of the possibility of a guest worker pro-
gram ($\theta_{1B} = 87\% > 18\% = \theta_1$ if no guest worker program is available under
parameterization of previous section). There are two reasons for this. First,
low-skill immigrants have a much higher fertility rate than natives. Hence, al-
lowing unskilled immigrants can affect more easily the size of the future native
population (and the composition) than allowing the same number of individ-
uals of a different type; and second, low-skill immigrants have a majority of
children that become medium-skilled, which in turn would most likely compete
with native children of medium-skilled. When the dynamic effects of unskilled
immigration are removed (via allowing guest workers), the only effects that are
considered by the natives are the contemporaneous wage effects: the medium-
skilled voter allows unskilled guest workers until the marginal benefit (higher
wages for medium-skilled) equals the marginal cost (redistribution) to these
workers, everything else constant. Another way to see this is the following:
if a guest worker program is not feasible, then the immigration quota for un-
skilled workers would be lower than the quota of guest workers due to the future
competition that these children represent for the medium-skilled.

For the interpretation, note that admitting unskilled guest workers is eco-
nomically equivalent to a policy of tacitly tolerating unskilled illegal immi-
grants/undocumented workers. One way to implement such a policy is by ne-
eglecting border controls combined with measures that exclude these individuals
from medium-and high-skilled jobs, e.g., background checks of licensing require-
ments. Thus the voting equilibrium in this section resembles US immigration
policy, which permits high-skilled immigration and seems to tolerate illegal im-
migrants doing unskilled work.

Also note that admitting a succession of unskilled guest workers (without
children) would be equivalent to admitting unskilled immigrants if all their
descendants remained unskilled; for example, if their children were given low
quality education. It may be an interesting issue for future research to examine
if the changing educational opportunities of immigrant children (e.g., the end of bilingual education in California) have influenced voter support for immigration. Our analysis implies, for example, that voters will be less tolerant of immigration if a greater proportion of immigrant children is medium-skilled rather than low-skilled.

4 Using the Model

4.1 How does the Intergenerational Mobility of Immigrants affect Immigration Policy?

In this section we study the effects of changing entries in the transition matrix of immigrants $Q_m$ in the baseline case ($\theta_3 \leq 13\% = \theta_3^{\text{max}}$). Since we cannot change individual entries in $Q_m$ without affecting other entries in that row (since for example $\Delta q_{11}' + \Delta q_{12}' + \Delta q_{13}' = 0$), then in this section whenever we change an entry, the ratio of the other two probabilities in that row is left unchanged. We use a 5 percentage point increase in each entry of the baseline $Q_m$ matrix and document the results.

The most significant changes in equilibrium unskilled immigration come from changes in the probability distribution of unskilled agents. In particular, an increase of 5 percentage points in $q_{13}'$ (which represents the probability that an unskilled parent has a high skilled child) increases $\theta_1'$ to 42% in the initial steady state (from a baseline of 18%), with a long run quota (at the new induced steady state) of 34.5%. In turn, modifying $q_{11}'$ changes the equilibrium unskilled quota marginally (15%), while an increase in $q_{12}'$ would result in a much smaller quota of unskilled immigrants allowed of 5% (8% in the long run). We explain these results. A higher $q_{12}'$ leads to lower immigration of unskilled immigrants because their children would compete in the most likely scenario with the children of the native medium-skilled (there is a 70% probability that children of medium skilled continue being medium skilled). In turn, a higher $q_{13}'$ leads to higher immigration as the possible complementarity of the children of immigrants with natives who are in states "unskilled" or "medium-skilled" outweighs the cost of the competition if the children of natives were in the "skilled" state.

Changing the probability distribution of high-skill workers almost doesn’t affect the immigration quotas at the steady state without immigration. However, the steady state induced by the preferred immigration policy has a significantly lower low-skill quota when $q_{31}'$ is increased ($\theta_1' = 12\%$), which happens because the high-skill immigrants have a higher percentage of unskilled children than natives, which yields a lower demand for future low-skill immigration.

The changes to any of the probabilities for medium-skill parents don’t affect the quotas of low and high-skill quotas because the equilibrium quotas for medium-skilled are zero in the model and the "small" changes in probabilities considered here are not enough to produce positive medium-skill quotas.

It is possible to generate cases in which the demand for high-skill immigration is lower than the observed 13% (where the chosen immigration quota is
optimally $\theta^*_3 < \theta^{\text{max}}_3$). However, given the parameterization of the model, very unrealistic probability distributions for the children of high-skill immigrants would be required (i.e. adding 60 percentage points to $q'_{31}$ so that the children of high skilled immigrants are mostly unskilled). If $\theta^{\text{max}}_3$ was larger, then it would be relatively easier to generate an interior solution in skilled immigration. For example, adding 40 percentage points to $q'_{31}$ when $\theta^{\text{max}}_3 = 30\%$ would produce $\theta^*_3 = 24\% < 30\%$. We speculate that lack of information on intergenerational mobility of immigrants could generate an interior solution even in the case with the "small" pool of high-skill immigrants. That is, if the medium-skill worker has beliefs that high-skill individuals might have low mobility (or if there is uncertainty in the immigrants matrix $Q_m$) then their demand for skilled immigration would be lower. Another possibility is that there can be uncertainty about the "true" skill level of immigrants and so if there exist costs to verify the "true" skill level of immigrants, this would naturally lead to lower demand for high-skill immigration. We leave these issues for future research.\footnote{We do not pursue extensions to uncertainty due to a curse of dimensionality present in value function iteration algorithms. Because the value function iteration required to obtain the MPE requires optimization over a grid of states indexed by the entire distribution of types, computational effort is growing exponentially with the number of possible types.}

We discuss the results under a wage-elastic supply of high-skill immigrants in the sensitivity section.

4.2 What if Immigrants Had Identical Demographic Profiles to Natives?

From the previous section, it is clear that our baseline results are driven in part by differences between the transition matrices $Q_m$ and $Q_n$. In this section we study the effects in equilibrium immigration when we impose demographic profiles of natives to immigrants, given the calibration in the previous section. Holding fertility rates constant at their estimated levels ($\eta, \eta_m$) if we set $Q_m = Q_n$, equilibrium unskilled immigration is adversely affected because immigrants seem to have better upward mobility odds than natives (though this might not be true for every immigration group within a skill level, it is true in average\footnote{The information available from the GSS is too small as to make inferences of different groups.}). Hence with lower mobility more children of unskilled immigrants would be in states that turn out to be undesirable for the medium-skilled majority. In terms of magnitudes, the unskilled quota is shutdown (0\%) at both the initial and induced steady states. Repeating this experiment but with identical fertility ($\eta_m = \eta$) doesn’t change the qualitative results, but that model requires a different calibration for $\beta$ since the calibration is contingent on $\eta_m$ and $\eta$. We don’t discuss this further as there’s not any extra insight gained from this. The main point in this section is that a high upward mobility of unskilled immigrants implies a higher level of unskilled immigration in equilibrium.

In order to better understand these results we also examine the effects of replacing one row at a time in $Q_m$ by the respective row in $Q_n$. In words,
we study the effect of imposing the probability of intergenerational mobility of natives instead of the respective immigrant’s probability of the same type. When low-skill immigrants have a skill distribution of their children identical to that of natives ($Q_{m[1]} = Q_{n[1]}$, while $Q_{m[j]} \neq Q_{n[j]}$ for j=2,3), then the low-skill immigration quota is 0% in both the short and long run. This result is robust to either using identical fertility for natives and immigrants, or using the estimates for differential fertility. The fact that natives have a higher proportion of low-skilled children than immigrants helps explain this. Hence, when we set $Q_{m[1]} = Q_{n[1]}$, the medium-skill demands less low-skill immigration as that would imply more competition in the case that their children turn out to be low-skill, and at the same time less complementarity of the children as upward mobility of immigrants would be lower. Alternatively, the amount of low-skill immigration that can be supported decreases as in this case low-skill immigrants have more low-skill children and there is therefore less room for low-skill immigration every period.

Changing the medium-skilled distribution doesn’t change the unskilled immigration quota because in equilibrium there is no medium-skilled immigration in our models. And finally, if children of high-skill immigrants have the same transition probabilities as natives, the medium-skilled would demand slightly more low-skill immigration ($\theta_1^* = 20\%$ as opposed to 18%). This is so because the estimated probability distribution for immigrants has a slightly higher probability (but still small) that their children will become low-skilled. Hence using the same probabilities as natives implies lower number of unskilled natives over time (as well as less skilled natives) and the overall effect is to demand slightly more low-skilled immigration.

From this exercise we conclude that more successful children of low-skilled immigrants lends political support to a bigger low-skill quota, and the opposite is also true.

Holding constant the intergenerational mobility matrices at their estimated levels, we can also study the effects of modifying the fertility rates of each type of agent, first one-by-one, and then all at the same time when we impose fertility profiles to be identical to those of natives. The effect of fertility on low-skill immigration is significant. If the low-skill immigrants had children at the same level as those of natives, equilibrium immigration of the low-skill types would increase to a quota of 33% rather than 18%. The fertility of the medium skilled immigrant is not relevant, while the fertility of the high-skilled does indeed affect the equilibrium immigration of the low-skilled slightly (but not the high-skill immigration). Hence higher fertility rates for the low-skill decrease the demand for low-skill immigrants.
5 Sensitivity Analysis

5.1 The Model Under a Wage-Elastic Supply of High-Skill Immigration

In the initial parameterization of the model a huge pool of high-skill immigrants was assumed, and so the equalization of the wages between medium and high-skilled was among the possibilities that the majority could choose. Then the bulk of the analysis was performed with a version of the model where the pool of high skilled immigrants was capped at levels that don’t equalize the wages of the medium and high skilled \( \theta_3^{\text{max}} = 13\% \) but the supply side was still unresponsive to the wage. In this section we investigate the predictions of the model under a supply of high skilled immigrants that depends on the wage of the skilled agents. In general, we obtain the same qualitative predictions as those obtained by the baseline model \( \theta_3^* \leq \theta_3^{\text{max}} = 13\% \).

The particular functional form used for this exercise is given by

\[
\theta_3^{\text{max}}(w_3) = \gamma_1 \left( \frac{w_3}{w_3^*} \right)^{\gamma_2}, \tag{28}
\]

where \( \gamma_2 \) is the elasticity of \( \theta_3^{\text{max}} \) with respect to \( w_3 \), the constant \( w_3^* \) is a reference point that helps fixing the supply of high-skill immigrants to go through the point \( \theta_3^{\text{max}}(w_3 = w_3^*) = \gamma_1 \) in the space of \( w_3 \) and \( \theta_3^{\text{max}} \). This ensures that when considering alternative elasticity levels \( \gamma_2 \), the supply goes through the same reference point \((w_3^*, \theta_3^{\text{max}}(w_3^*))\), but the response to the left and right of that point depends on the elasticity parameter \( \gamma_2 \).

For the numerical choices of the parameters, it is not possible to calibrate them when it is assumed that the observed immigration choices are suboptimal. Hence we consider several elasticity levels and investigate the response of the medium skilled choices of immigration given each elasticity level. The considered elasticity levels ranges from 0 to 10.

The level of reference \( w_3^* \) used is the steady state wage of the high-skill agents in the absence of immigration.\(^{24}\) Finally, the parameter \( \gamma_1 \) has to be higher than the observed \( 13\% \) (which is observed inclusive of immigration) which we set at 30\%. Broadly speaking, the results still show immigration of the extremes as the trade-offs are not large enough as to justify immigration of the medium skilled. What is different now are the combinations of low and high-skill immigration quotas that would be chosen by the majority. Under the "high" elasticity of supply scenario we obtain less high-skilled immigration and more low-skill immigration than in the other cases: the higher the elasticity, the bigger the response of quantity of immigrants to the decrease in wages. Thus in the interest of keeping high-skill immigration at a higher level the majority allows

\(^{24}\)This number is \( w_3^* = .5424 \). It could be possible to back up \( w_3^* \) from data as a reference wage in the source countries, where it would be the average wage that high skilled immigrants to the US would earn in their respective countries. But doing this is a moot point since the level \( \gamma_1 \) can’t be identified when \( \theta_3 \) is suboptimal and for any level \( w_3^* \) and given an elasticity \( \gamma_2 \), the constant \( \gamma_1 \) can be normalized as to give the same supply with a different level \( w_3^* \).
even more low-skill immigrants than when the supply side is less responsive. Table 4 summarizes the numerical results.

<table>
<thead>
<tr>
<th>Elasticity Supply:</th>
<th>$\epsilon = 0$</th>
<th>$\epsilon = .25$</th>
<th>$\epsilon = 1$</th>
<th>$\epsilon = 4$</th>
<th>$\epsilon = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial steady state</td>
<td>$\theta_1^*$ = 26.2</td>
<td>27.8</td>
<td>29.0</td>
<td>35.7</td>
<td>42.7</td>
</tr>
<tr>
<td></td>
<td>$\theta_3^*$ = 30</td>
<td>29.5</td>
<td>28.2</td>
<td>24.4</td>
<td>20</td>
</tr>
<tr>
<td>Induced steady state</td>
<td>$\theta_1^*$ = 30.2</td>
<td>31</td>
<td>31.9</td>
<td>33.5</td>
<td>34.8</td>
</tr>
<tr>
<td></td>
<td>$\theta_3^*$ = 30</td>
<td>29.2</td>
<td>27.2</td>
<td>22.0</td>
<td>16.9</td>
</tr>
</tbody>
</table>

At the induced steady state the high-skill immigration quota is lowest in the high elasticity scenario. Since high-skill immigrants have a very high proportion of skilled children, allowing immigrants depress future wages more pronouncedly in the high-elasticity case. The low-skill immigration in turn has the opposite effect: the higher the elasticity, the higher the steady state low-skill immigration quota that is accepted in order to attract a higher supply of high-skill immigrants.

As previously found, the optimal policy function for low-skill immigration is decreasing in $x_1$ and increasing in $x_3$, but now the optimal policy for high-skill immigration is decreasing in $x_3$, and slightly increasing in $x_1$. In words, the higher the native share of skilled agents (as percentage of the medium skilled), the higher the low-skill immigration and the lower the equilibrium high-skill immigration quota. Similarly, the higher the native share of low-skilled agents, the lower the demand for low-skill immigration, and the higher the demand for high-skill immigration.

We also perform the experiments that replace the intergenerational mobility of immigrants by those of natives. We find similar qualitative effects as before since replacing the probability distribution of low-skill immigrants (holding fertility at its estimated levels) by those of natives shuts down low-skill immigration. Alternatively, when we use the fertility rates of natives in place of those of immigrants, there’s just a small change in high-skill immigration, but the effects on low-skill immigration would be more significant (from 29% to 42% in the unit elasticity case, and from 42% to 51% in the "high" elasticity case of 10). Using the whole demographic profile of natives instead of the one estimated for immigrants affects mostly low-skill immigration, by reducing it (from 29% to 15% in the unit elasticity case, and from 42% to 29% in the high elasticity case of 10).

When guest-worker programs are available, the results are still similar: full-immigration-only offered to high-skill immigrants, while low-skilled workers are only offered guest worker permits. For example, in the case with a supply-elasticity of 1 (10), the quotas chosen are $\theta_{1B}^* = 99\%$ (93%) for the low skilled-guest worker type, with a high-skill full-immigration quota of $\theta_3^* = 28.4\%$ (21%). The induced steady state has $\theta_{1B}^* = 134\%$ (109%) and $\theta_3^* = 27.2\%$ (16.7%). Notice that the low-skill guest-worker quota is much higher than the low-skill quota chosen when guest worker programs are not available (similar to what was found before when a perfectly inelastic supply of high skilled immigration was assumed). Very similar numbers are found for the other elasticities considered.
We also studied another specification of the supply for high skilled immigrants where it is assumed that \( \theta_3^{\text{max}} \) depends on the ratio \( \frac{w_3}{w_2} \), where \( \theta_3^{\text{max}} \) is positive provided that the ratio \( \frac{w_3}{w_2} \) is at least a specific exogenous level \( \left( \frac{w_3}{w_2} \right) \). Since the ratio \( \frac{w_3}{w_2} \) depends negatively on \( \theta_3 \) (raising \( w_2 \) and at the same time decreasing \( w_3 \)) the agent would choose a level \( \theta_3 \) such that \( \frac{w_3}{w_2} = \left( \frac{w_3}{w_2} \right) \). However, since there may be many policies \((\theta_1, \theta_2, \theta_3)\) that yield the static condition \( \frac{w_3}{w_2} = \left( \frac{w_3}{w_2} \right) \), the agent would choose the specific policy that maximizes utility according to the dynamic effects of the model. Results were qualitatively similar to the case discussed in this section, but we chose not to discuss this more thoroughly for this same reason, and also because the supply of immigrants would presumably respond to the level of high-skilled wages \( (w_3) \) rather than to the ratio \( \frac{w_3}{w_2} \).

5.2 Effect of \( \sigma \) (CRRA Preferences)

In our experiments we used logarithmic utility \( (\sigma = 1) \). Another number typically used in the literature is \( \sigma = 2 \), which leaves the results largely intact, with a few small numerical differences.

In the model with the huge high-skill supply the chosen immigration quotas are almost identical \( \theta_1^*(\sigma = 2) = 277\% \) and \( \theta_3^*(\sigma = 2) = 200\% \) as opposed to \( \theta_1^*(\sigma = 1) = 281\% \) and \( \theta_3^*(\sigma = 1) = 200\% \).

In the baseline version of the model \( (\theta_3^{\text{max}} = 13\%) \) and using \( \sigma = 2 \), we calibrate \( \beta = .5625 \) in order to obtain \( \theta_1^*(\sigma = 2) = 18\% \). We obtain identical comparative statics as before, with only small differences in magnitudes. If the policy space is enlarged to include guest-worker programs, we again obtain full immigration for the high-skill and guest worker for the low-skilled, with practically identical numbers as before \( \left( \theta_1^B = 87\%, \theta_1 = 0\% \right) \). Other realistic levels of risk aversion were also used \( (1 \leq \sigma \leq 4) \) and the qualitative behavior of the model is the same with some small difference in magnitudes.

5.3 Different degree of substitution in the Production Function \( (\frac{1}{1-\rho}) \)

We perform sensitivity analysis in the value of the elasticity of substitution between labor inputs, given by \( \varepsilon = \frac{1}{1-\rho} \). We do it for \( \rho = \frac{2}{5} \) and for \( \rho = \frac{3}{5} \). Using the same wage premia and labor ratios as before, we obtain share parameters \( (\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3) \) given by \( (.0836, .4160, .5004) \) in the case of \( \rho = \frac{2}{5} \) and \( (.1180, .3766, .5054) \) when \( \rho = \frac{3}{5} \). Just like before, in the absence of any

\[ \text{Under this specification, the predictions are similar (immigrations of the extremes), but when considering bracero vs guest worker we find a positive level of bracero and immigration at the same time for high skilled immigration.} \]
immigration cost, the unconstrained model with a "huge" $\theta_3^{\text{max}}$ yields quotas that are very high. Indeed, when using the version of the model where the policy space is state-invariant, there’s no combination of $(\beta, \sigma, \rho)$ that might yield magnitudes that are in line with what is observed in the data.

When $\rho = \frac{2}{5}$ we obtain $\theta_1^* (\rho = \frac{2}{5}) = 352\%$ and $\theta_3^* (\rho = \frac{2}{5}) = 227\%$. Alternatively, when $\rho = \frac{3}{5}$ we obtain $\theta_1^* (\rho = \frac{3}{5}) = 261\%$ and $\theta_3^* (\rho = \frac{3}{5}) = 325\%$. Introducing a small cost of immigration in the same magnitude as previously done (that overall represents 1% of transfer once all immigrants are taken into account), has the same effect as before since the magnitudes decrease significantly to $\theta_1^* (\rho = \frac{2}{5}) = 256\%$ and $\theta_3^* (\rho = \frac{3}{5}) = 153\%$; and in the case where $\rho = \frac{3}{5}$ we obtain $\theta_1^* (\rho = \frac{3}{5}) = 214\%$ and $\theta_3^* (\rho = \frac{3}{5}) = 286\%$. Summarizing this information, the unskilled quota is larger when $\rho$ is lower, while the opposite is true for high-skilled immigration. A higher level of $\rho$ represents a higher elasticity of substitution between the different skill types and hence low-skill immigration has less of a complementarity effect with the other labor types when $\rho$ is high, resulting in less low-skill immigration but more high-skill immigration.

For the main parameterization of the model where the supply of high skilled immigrants is perfectly inelastic at $\theta_3^{\text{max}} = 13\%$, we calibrate $\beta (\rho = \frac{2}{5}) = .875$ and $\beta (\rho = \frac{3}{5}) = .385$. The comparative statics are as before, but as expected the specific magnitude of the effects differ. Replacing the mobility of immigrants by those of natives shuts-down the demand for low-skill immigration. In both cases ($\rho = \frac{2}{5}$ and $\rho = \frac{3}{5}$) we obtain $\theta_1^* = 0$. This is driven by the mobility of the unskilled immigrants (first row in $Q_m$) as those results are obtained whenever we replace their probability distribution for the one from natives. The comparative statics for changes of 5 percentage points in each entry as performed in the main text were also performed, with results qualitatively identical to the case discussed in the main text. The case with $\rho = \frac{3}{5}$ produces low-skill immigration quotas that are more sensitive in terms of magnitudes to the changes in the entries in $Q_m$, relative to the baseline case.

As in previous sections, we find the result of a guest-worker program for low-skill immigrants, and full immigration to the high-skilled ones. The effects of the model are robust to this parameter.

5.4 An Alternative Definition of the Medium-Skill Group

In this section we consider a different definition of the medium-skill labor type. We put the high-school graduates and those with junior college or bachelor degrees in the middle category, with low-skilled defined as having less than a high school degree, while high-skilled is defined as those with a master/beyond degree.

The intergenerational mobility matrices in this case using the same filters as before yield
Under the new definitions we compute that during the 90’s there were 2.43 medium-skilled immigrants per 1000 medium-skilled natives, 4.3 skilled per 1000 skilled natives, and the unskilled remains the same. These numbers yield \( \frac{w_3}{w_1} = 1.727 \) and \( \frac{w_3}{w_2} = 1.4845 \) for the same samples and filters as previously discussed (except education). Using \( \rho = 1/2 \), the production share parameters (\( \phi_1, \phi_2, \phi_3 \)) are in this case given by (.1110, .5704, .3185). The fertility profiles (total fertility rates) can only be estimated at the level of disaggregation desired from the ACS for the years 2001-2008, which are \{2.41, 2.04, 1.97\} for native women and \{3.23, 2.44, 2.08\} for foreign-born women. This implies model parameters of \( \eta = diag\{1.21, 1.02, .99\} \) for natives and \( \eta_m = diag\{1.62, 1.22, 1.04\} \) for immigrants. Given the demographic parameters of natives, the steady state ratios in absence of immigration are given by \( \{x_1^{ss}, x_3^{ss}\} = \{0.0760, 0.0988\} \). Naturally, in this case there is a much bigger share of medium-skilled agents than in the case discussed in the main text.

We use the same tax rate used as before (\( \tau = 30\% \)) and evaluate the model for preference parameters given by \( \gamma = 1 \) and \( \beta = .98530 \). In the unconstrained version of the model, we obtain immigration of the extremes with very large immigration quotas given by \( \theta_1^* = 222\% \) and \( \theta_3^* = 231\% \), and induced quotas of 143\% and 132\% respectively at the new steady state. A small immigration cost as previously assumed yields quotas given by \( \theta_1^* = 207\% \) and \( \theta_3^* = 219\% \).

In the version of the model where we impose a perfectly inelastic supply of high skilled immigrants at the observed level (\( \theta_3^{max} = 13.75\% \)) and the rest of the parameters unchanged, the model produces qualitatively similar results as in the main parameterization in the text, but with a larger level of unskilled immigration (\( \theta_1^* = 141\% \) at initial steady state and 86\% in the induced one).

The comparative statics work in the same direction as previously (i.e. the low-skill quota depends highly on the intergenerational mobility of unskilled immigrants). When replacing probability distributions of immigrant types by those of natives we also obtain similar qualitative results. For example, using the probability distribution of the low-skill natives rather than the one for immigrants yields less demand for low-skill immigration. In the guest-worker vs immigration case we find again full-immigration for high-skill individuals, and guest-worker programs for the low-skilled, with a somewhat higher quota of low-skill immigration than when guest-worker programs are available.
6 Conclusions

We study a dynamic macroeconomic model of intergenerational mobility and immigration with three types of labor. In it, the demographic process is such that the medium-skill type is always the majority. We parameterize several versions of the model and study the results.

We calibrate the model for the US and among other things find that children of low-skill immigrants and medium-skill immigrants seem to be more "successful" than the children of natives (there is a higher probability for their children to become high-skilled), using data from the General Social Survey. We find the Markov Perfect equilibrium of the model and the optimal policy shows that if the children of unskilled immigrants are more successful than those of unskilled natives, there is more political support for unskilled immigration by the majority (the medium-skilled). The most preferred policy for the majority involves maximizing high-skill immigration, minimizing medium-skill immigration and for low-skill immigration it is increasing in the share of high-skilled natives and decreasing in the share of low-skilled natives.

In general, the current effect on wages and transfers from the high-skill agents are very important for the welfare of the medium-skilled and thus their intergenerational mobility is relatively unimportant for the political support of skilled immigration.

Under the interpretation that the observed flow of high-skill immigration is suboptimal, a "piece-meal" approach to immigration reform allowing more high-skill immigrants would also lead to an increase in the demand for low-skill immigrants.

If in addition to full immigration there are guest workers quotas available as policy tools, the medium-skilled would choose a guest-worker program for low-skill immigrants, and full-immigration for the high-skill immigrants.

Finally, a version of the model with a positive-sloped supply of high-skill immigrants produces similar results, and it is found that the higher the elasticity of supply, the higher the level of low-skill immigration chosen optimally. Sensitivity analysis shows that the effects of the model are robust.

References


[33] US Census Bureau at http://www.census.gov/
Appendix

A Open Economy Interpretation of the Production Function

Assume that the production function of the host country is given by

\[ Y = AK^\alpha L^{1-\alpha} \]

where \( A \) is technology level, \( K \) is aggregate capital and \( L \) is a composite labor input. Assuming perfect capital mobility with an exogenous world interest rate \( r \) then capital adjusts so that the return on domestic investment is the same as the world return. Equalizing the marginal product of capital to the world interest rate and substituting into the production function yields

\[ Y = A' L \]

with

\[ A' = \left( \frac{\alpha A}{r} \right)^{\frac{\alpha}{1-\alpha}} \]

where \( r \) is the return on capital. Since \( A' \) is just a multiplicative constant, without loss of generality we just assume that \( A' = 1 \). The resulting production function is a constant returns to scale in the composite labor input.

B Proofs on Static Preferences

B.1 Proof that \( \sum_i F_{ij} L_i = 0 \)

Assume \( F(L_1, L_2, L_3) \) is constant returns to scale, where the labor input is given by \( L_i = N_i (1 + \theta_i) \). Then we can write

\[ Y = F_1 L_1 + F_2 L_2 + F_3 L_3, \]

where \( F_i = \frac{\partial F}{\partial L_i} = w_i \).

Then differentiating \( Y \) w.r.t \( \theta_j \) we have

\[ \frac{\partial Y}{\partial \theta_j} \frac{\partial L_i}{\partial \theta_j} = \sum_{i=1}^3 F_{ij} L_i + F_j \frac{\partial L_i}{\partial \theta_j}, \]

where we define \( F_{ij} = \frac{\partial^2 F}{\partial L_i \partial L_j} = \frac{\partial w_i}{\partial L_j} \)

since \( \frac{\partial L_j}{\partial \theta_j} = N_j \) and \( \frac{\partial Y}{\partial \theta_j} = F_j \) we have

\[ F_j N_j = \sum_{i=1}^3 F_{ij} L_i + F_j N_j \]

which in turn implies that it must be the case that

\[ \sum_{i=1}^3 F_{ij} L_i = 0 \]

\[ \blacksquare \]
B.2 Proof that \( \frac{\partial \bar{w}}{\partial \theta_j} = \frac{N_j(w_j - \bar{w})}{\Sigma_i^j N_i(1+\theta_i)} \)

Define \( b = \tau \bar{w} \), where \( \bar{w} \) is just the average compensation in the economy. Thus \( \bar{w} = \frac{\Sigma_i^j N_i(1+\theta_i)w_i}{\Sigma_i^j N_i(1+\theta_i)} \). Replacing \( w_i \) with \( F_i \), and applying logs we have

\[
\ln \bar{w} = \ln \left[ \Sigma_i^j N_i (1 + \theta_i) F_i \right] - \ln \left[ \Sigma_i^j N_i (1 + \theta_i) \right]
\]

Then

\[
\frac{\partial \ln \bar{w}}{\partial \theta_j} = \frac{\left[ w_i N_j + \Sigma_i^j N_i (1 + \theta_i) F_{ij} \right] - N_j}{\Sigma_i^j N_i (1 + \theta_i) w_i} - \frac{N_j}{\Sigma_i^j N_i (1 + \theta_i)} \left\{ \frac{w_j - \bar{w}}{\bar{w}} \right\}
\]

Since \( \Sigma_{i=1}^j N_i (1 + \theta_i) F_{ij} = 0 \), we can write

\[
\frac{\partial \ln \bar{w}}{\partial \theta_j} = \frac{\Sigma_i^j N_i(1+\theta_i)w_i}{\Sigma_i^j N_i(1+\theta_i)} - \frac{N_j}{\Sigma_i^j N_i (1 + \theta_i)} \left\{ \frac{w_j - \bar{w}}{\bar{w}} \right\}
\]

hence

\[
\frac{\partial \bar{w}}{\partial \theta_j} = \frac{N_j (w_j - \bar{w})}{\Sigma_i^j N_i (1 + \theta_i)}
\]

Hence the average wage increases in \( \theta_j \) if \( w_j > \bar{w} \).

B.3 Proof that \( \theta^*_2 > 0 \& \theta^*_1 > 0 \) is not optimal

If, given the optimal skilled immigration policy involves \( w_2 (\theta^*_2) < \bar{w} \) then the proof is trivial since in that case necessarily we have that \( \frac{\partial w_2}{\partial \theta_2} < 0 \) as medium-skilled immigration decreases both the wage of the medium skilled and the average transfer. Thus in that case \( \theta^*_2 = 0 \) and we therefore have \( \theta^*_1 \geq 0 \) from the remaining first order condition. But more generally, now we show that even if \( w_2 (\theta^*_2) \geq \bar{w} \) it is the case that we can’t have an interior solution in both \( \theta_1 \) and \( \theta_2 \). In the algebra that follows, we allow the levels of immigration to be negative, just to show that those conditions that would define the solution to each of the equation cannot hold at the same time. Since positive immigration would be a subset of all of the values allowed in the first order conditions, this also implies that simultaneous interior solutions in \( \theta_1 \) and \( \theta_2 \) don’t exist. Hence at least one of the two immigration quotas is a corner solution of 0. We start by writing the interior solutions to the first order conditions as

\[
[foc \theta_i] \quad (\tau - 1) \frac{\partial w_2}{\partial \theta_i} = \tau \frac{\partial \bar{w}}{\partial \theta_i} \quad \text{for} \quad i = 1, 2
\]

since each element in the equations is different from 0, divide \([foc \theta_1]\) by \([foc \theta_2]\) and obtain

\[
(\tau - 1) \frac{\partial w_2}{\partial \theta_2} = \tau \frac{\partial \bar{w}}{\partial \theta_2} \quad \text{for} \quad i = 1, 2
\]

the following terms will be used

\[
\frac{\partial w_2}{\partial \theta_1} = \phi_2 (1 - \rho) \left( \frac{Y}{L_2} \right)^{-\rho} \left( \frac{1}{L_2} \right) \frac{\partial Y}{\partial L_1} \frac{\partial L_1}{\partial \theta_1} \frac{\partial \bar{w}}{\partial \theta_2} = \frac{(1 - \rho w_2 w_1 N_1)}{Y} > 0
\]

\[
\frac{\partial w_2}{\partial \theta_2} = \phi_2 (1 - \rho) \left( \frac{Y}{L_2} \right)^{-\rho} \left[ \frac{L_2 \frac{\partial \bar{w}}{\partial \theta_2} \frac{\partial L_2}{\partial \theta_2} - Y \frac{\partial L_2}{\partial \theta_2}}{(L_2)^2} \right] = (1 - \rho) w_2 \left[ \frac{L_2 w_2 - Y}{Y} \right] \frac{N_2}{L_2} < 0
\]

\[
\frac{\partial \bar{w}}{\partial \theta_1} = \frac{N_2 (w_2 - \bar{w})}{\Sigma_i^1 L_i} \quad \text{;} \quad \frac{\partial \bar{w}}{\partial \theta_2} = \frac{N_2 (w_2 - \bar{w})}{\Sigma_i^1 L_i}
\]

Inserting these terms obtain the equation

\[
(1 - \rho) w_2 \left( \frac{L_2 w_2}{Y} - \frac{1}{L_2} \right) \frac{N_2}{w_2} = \frac{N_2 (w_2 - \bar{w})}{\Sigma_i^1 L_i}
\]

simplifying
\[
\frac{w_1 N_1}{w_2 L_2 - Y} \Rightarrow \frac{N_2 (w_1 - \bar{w})}{w_2 L_2 - Y} = \frac{(w_2 - \bar{w}) L_2}{w_2 L_2 - Y} = \frac{(w_2 - \bar{w})}{w_1}
\]

where we assume that \( w_2 - \bar{w} > 0 \) (otherwise \( \frac{\partial u_2}{\partial w_2} < 0 \) necessarily if \( w_2 < \bar{w} \), which yields \( \theta_3^* = 0 \)). Multiply both sides by \(-1\) and rearrange
\[
\frac{w_2 L_2 - \bar{w} L_2}{Y - w_2 L_2} = \frac{\bar{w} - w_1}{Y - w_2 L_2} \Rightarrow \frac{w_2 L_2 - \bar{w} L_2}{Y - w_2 L_2} = \frac{\bar{w} - w_1}{Y - w_2 L_2} \Rightarrow \frac{Y - \bar{w} L_2}{Y - w_2 L_2} = \frac{\bar{w}}{w_1},
\]

then we can write
\[
(Y - \bar{w} L_2) w_1 = (Y - w_2 L_2) \bar{w}
\]

and after some algebra obtain
\[
(w_2 - w_1) L_2 = (\bar{w} - w_1) Y
\]

since \( \frac{\bar{w}}{w_1} = \Sigma_i L_i \) we can write
\[
w_2 L_2 - w_1 L_2 = \bar{w} \Sigma_i L_i - w_1 \Sigma_i L_i
\]

substituting \( \bar{w} \Sigma_i L_i \) by \( \Sigma_i w_i L_i \) and cancelling terms we finally obtain that
\[
w_3 L_3 = L_3 w_1 \text{ or } w_3 = w_1\]

which is false. Therefore both first order conditions can’t hold at the same time. Hence there is no case in which the medium skilled chooses \( \theta_1^* > 0 \) and \( \theta_2^* > 0 \). \[\blacksquare\]

**B.4 Proof that \( \frac{\partial u_3}{\partial \tau} < 0 \) for empirically relevant levels of \( \rho \)**

Given that \( \theta_1^* > 0 \) it must be the case that \( \theta_2^* = 0 \). For interior solutions (allowing for negative values for \( \theta_1 \) and \( \theta_2 \)) the first order condition (foc) with respect to \( \theta_1 \) holds with equality:
\[
\{\text{foc} \theta_1\} : \quad (\tau - 1) \frac{\partial m_2}{\partial \theta_1} = \tau \frac{\partial m_1}{\partial \theta_1}
\]

the following derivatives and substitutions will be used throughout this proof
\[
\frac{\partial m_2}{\partial \theta_1} = \frac{N_1 (w_1 - \bar{w})}{\bar{w} \Sigma_i L_i (1 + \theta_1)} ; \quad w_1 = \phi_1 \left( \frac{Y}{L_1} \right)^{1-\rho} ; \quad w_2 = w_1 \phi_2 \phi_1 \left( \frac{L_1}{L_2} \right)^{1-\rho}
\]

\[
w_3 = w_1 \phi_1 \left( \frac{L_1}{L_2} \right)^{1-\rho} ; \quad \bar{w} = \frac{w_1 L_1}{\Sigma_i L_i} \phi_1 (L_1)^{\rho}; \quad Y^\rho = \left[ \Sigma \phi_1 (L_1)^{\rho} \right]
\]

substituting \( \frac{\partial m_2}{\partial \theta_1} \) and \( \frac{\partial m_1}{\partial \theta_1} \) in \{foc \theta_1\} and rearranging terms we can write the implicit equation defining \( \theta_1^* \) as
\[
\theta_1 = \{\theta_1 : (1 - \tau) (1 - \rho) w_2 w_1 = \tau \bar{w} (\bar{w} - w_1)\}
\]

which can be written as
\[
(1 - \tau) (1 - \rho) \phi_1 \left( \frac{L_1}{L_2} \right)^{1-\rho} (w_1)^2 = \tau \frac{w_1 L_1}{\Sigma_i L_i} \left[ \frac{\Sigma \phi_1 (L_1)^{\rho}}{\phi_1 (L_1)^{\rho}} \right] (\bar{w} - w_1)
\]

eliminate \( \frac{w_1 L_1}{\phi_1} \) in both sides to obtain
\[
(1 - \tau) (1 - \rho) \phi_2 \left( \frac{1}{L_2} \right)^{1-\rho} w_1 = \tau \left[ \frac{\phi_2 (L_2)^{\rho} (L_1)^{1-\rho} + \phi_3 (L_3)^{\rho} (L_1)^{1-\rho} - \phi_1 L_2 - \phi_1 L_3} {\phi_1 [\Sigma L_i]} \right] (\bar{w} - w_1)
\]

It can be proven that the term \( (\bar{w} - w_1) \) can be written as
\[
(\bar{w} - w_1) = \frac{w_1 \left\{ \phi_2 (L_2)^{\rho} (L_1)^{1-\rho} + \phi_3 (L_3)^{\rho} (L_1)^{1-\rho} - \phi_1 L_2 - \phi_1 L_3 \right\}} {\phi_1 [\Sigma L_i]}
\]

and so we can eliminate \( w_1 \) from both sides to obtain
\[
(1 - \tau) (1 - \rho) \phi_2 \left( \frac{1}{L_2} \right)^{1-\rho} = \tau \left[ \frac{\phi_2 (L_2)^{\rho} (L_1)^{1-\rho} + \phi_3 (L_3)^{\rho} (L_1)^{1-\rho} - \phi_1 L_2 - \phi_1 L_3} {\phi_1 [\Sigma L_i]} \right]
\]

Since everything is positive in the above expression, we can take logs to this equation and totally differentiate. Solving for \( \frac{\partial \theta_1}{\partial \tau} \) obtain
whose sign is determined by the denominator since the numerator is positive. Below we derive a condition that yields a negative denominator, and this condition is very likely to be satisfied for the empirical range of $\rho$. 

In the denominator we can simplify one of the terms as

$$\frac{(1-\rho)(L_1)^{\rho-\rho}}{\phi_3(L_3)^{\rho}} \left[ \phi_2(L_2)^{\rho} + \phi_3(L_3)^{\rho} \right]$$

since $Y^\rho = \sum \phi_i(L_1)^\rho$, then

$$\frac{(1-\rho)(L_1)^{\rho-\rho}}{\phi_3(L_3)^{\rho}} \left[ \phi_2(L_2)^{\rho} + \phi_3(L_3)^{\rho} \right] = \frac{(1-\rho)\phi_2(L_2)^{\rho} + \phi_3(L_3)^{\rho}}{(1-\rho)[Y^\rho - \phi_3(L_3)^{\rho}]}.$$ 

and since $\bar{w} = \frac{w_1L_1}{[L_2L_3]} \frac{Y^\rho}{\phi_3(L_3)^{\rho}}$ then

$$\frac{w_1 \bar{w}}{[L_2L_3]} = \frac{Y^\rho}{\phi_3(L_3)^{\rho}} = \frac{Y^\rho}{\phi_3(L_3)^{\rho}} = \frac{Y^\rho}{\phi_3(L_3)^{\rho}} = \frac{Y^\rho}{\phi_3(L_3)^{\rho}} = \bar{w}.$$ 

Therefore we can write

$$\frac{(1-\rho)[\phi_2(L_2)^{\rho} + \phi_3(L_3)^{\rho}]}{\phi_3(L_3)^{\rho}} = \frac{(1-\rho)\bar{w}^{\rho-\rho}}{\bar{w}^{\rho-\rho}} = \frac{(1-\rho)[Y^\rho - \phi_3(L_3)^{\rho}]}{(1-\rho)[Y^\rho - \phi_3(L_3)^{\rho}]}.$$

Hence the sign of $\frac{d\theta_1}{d\tau}$ is determined by

$$\text{sign} \left[ \frac{d\theta_1}{d\tau} \right] = \text{sign} \left[ \frac{2}{[L_2L_3]} - \frac{\rho \phi_3(L_3)^{\rho}}{Y^\rho} - \frac{(1-\rho)[Y^\rho - \phi_3(L_3)^{\rho}]}{(1-\rho)[Y^\rho - \phi_3(L_3)^{\rho}]} \right]$$

and since $\phi_3(L_3)^{\rho-\rho} = \frac{w_1}{Y}$ we can write

$$\frac{\phi_3(L_3)^{\rho-\rho}}{[L_2L_3]} = \frac{w_1}{Y} - \frac{(1-\rho)[Y - w_1L_1]}{L_1[1 - w_1/L_1]}.$$ 

Now we show a condition that yields the above negative number, implying that $\frac{d\theta_1}{d\tau} < 0$. After some algebraic steps, the inequality can be written as

$$\frac{d\theta_1}{d\tau} < 0 \quad \text{if} \quad 2 < \frac{w_1}{\bar{w}} + \frac{(1-\rho)\bar{w}^{\rho-\rho}}{(\bar{w}^{\rho}-\bar{w}^{\rho})} - \frac{(1-\rho)[Y^\rho - \phi_3(L_3)^{\rho}]}{(1-\rho)[Y^\rho - \phi_3(L_3)^{\rho}]}.$$

Multiply both sides by the positive term $\left[ \frac{w_1}{\bar{w}} \right] \left[ \bar{w} - 1 \right]$ and simplify. Then

$$\frac{d\theta_1}{d\tau} < 0 \quad \text{if} \quad 2 < \frac{w_1}{\bar{w}} + \rho \bar{w} \left[ 2 - \frac{w_1}{\bar{w}} \right] + (1 - \rho) \left( \frac{\Sigma L_3}{L_1} \right).$$

Consider the empirically relevant case where $\rho \geq 0$ (for example when $\rho = 0$, function is Cobb-Douglas). The inequality would be satisfied for the case $\rho = 0$ provided that $L_2 + L_3 > 1$, but it is never satisfied in the limiting case of inputs that are perfect substitutes with $\rho = 1$. Below we describe a condition for the cut-off value of $\rho$ that yields $\frac{d\theta_1}{d\tau} < 0$.

For simplicity, define $a = \frac{L_1}{L_2 + L_3}$. The quantity $\left( \frac{\Sigma L_3}{L_1} \right)$ can then be written as $1 + \frac{a}{1 - a}$. Then we can write

$$\frac{d\theta_1}{d\tau} < 0 \quad \text{if} \quad 2 < \frac{w_1}{\bar{w}} + \rho \bar{w} \left[ 2 - \frac{w_1}{\bar{w}} \right] + 1 + \frac{a}{1 - a} - \rho \left( 1 + \frac{a}{1 - a} \right).$$

this inequality can be simplified in order to obtain the following condition

$$\frac{d\theta_1}{d\tau} < 0 \quad \text{if} \quad \rho < \frac{[1 - a] + a \left( \frac{w_1}{\bar{w}} \right)}{1 + a - 2a \left( \frac{w_1}{\bar{w}} \right) + a \left( \frac{w_1}{\bar{w}} \right)^2} \quad ; \quad a = \frac{L_1}{L_2 + L_3}$$

42
where it can be easily shown that \[
\frac{[1-a]+a(w_1)}{(1+a)-2a(w_1)+a(w_2)^2} < 1
\] and necessarily \(0 < \frac{w_1}{w_2} < 1\).

### B.5 Static Preferences if \(\theta_3^{\text{max}}\) is such that \(w_3 = w_2\)

We discuss this case for completeness. The level of \(\theta_3\) that yields \(w_3(\theta_3) = w_2(\theta_3)\) is given by

\[
\tilde{\theta}_3 = \frac{1}{x_3} \left( \frac{\phi_3}{\phi_2} \right)^{\frac{1}{x_3^*}} - 1. \tag{29}
\]

Under the assumption that skilled immigration is big enough to equalize the wages of the skilled and medium-skilled categories \((\theta_3^{\text{max}} > \tilde{\theta}_3)\) the model becomes one in which there are essentially 2 skills levels: unskilled and medium+high skilled which are assumed to be the majority. In the case of only 2 (effective) inputs, the optimal choice involves choosing the optimal ratio \(\frac{L_1}{L_2+L_3}\) that maximizes consumption \(c_2\). Given the assumptions \(\theta_1, \theta_2, \theta_3 \geq 0\), any optimal ratio \(\frac{L_1}{L_2+L_3}\) can be obtained by means of increasing either \(\theta_1\) and setting \(\theta_3^* = \tilde{\theta}_3\) or setting \(\theta_3^* > \tilde{\theta}_3\) with \(\theta_1^* = 0\).\(^{26}\) Indeed any point \((\theta_1', \theta_3')\) that satisfies \(\frac{N_1(1+\theta_1')}{N_2+N_3(1+\theta_3')} = \left(\frac{L_1}{L_2+L_3}\right)\) also maximizes \(c_2\). See Ortega (2010) for a richer analysis of the 2 input case.

### C Model Dynamics in terms of native ratios \((X_t)\)

Start from the evolution of the native population, given by

\[
N_{i,t+1} = \sum_{j=1}^{3} \left( \eta_j q_{ji} + \eta_{j'} q_{ji'} \theta_{j',t} \right) N_{j,t} \tag{30}
\]

for \(i = 1, 2, 3\)

divide both sides of each equation \((i=1,2,3)\) by \(N_{2t}\) and obtain the 3 equations

\[
\frac{N_{i,t+1}}{N_{2,t+1}} \frac{N_{2,t+1}}{N_{2,t}} = \sum_{j=1}^{3} \left( \eta_j q_{ji} + \eta_{j'} q_{ji'} \theta_{j',t} \right) \frac{N_{j,t}}{N_{2,t}} \tag{31}
\]

for \(i = 1, 2, 3\)

\(^{26}\)For simplicity, we don’t discuss in the text the specific composition of \(\theta_3^*\) when \(\theta_3^* > \tilde{\theta}_3\). Since in this case \(w_2 = w_3\), it can be possible to allow some medium skilled immigrants as part of the total medium+skilled labor force. In other words, since for \(\theta_3^* > \tilde{\theta}_3\) and \(N_2(1+\theta_2) \leq \left[ \frac{\phi_2}{\phi_3} \right]^{\frac{1}{x_3}} N_3(1+\theta_3)\) then the medium and skilled labor force would be perfect substitutes, there are many combinations of \(\theta_2\) and \(\theta_3\) that yield the same \(L_2 + L_3\). Hence the same allocation.
where on the left hand side we also multiply by \( \frac{N_{2,t+1}}{N_{2,t}} \) and rearranged. Using the substitution of variables \( \frac{N_{1,t}}{N_{2,t}} = x_{1,t} \), \( \frac{N_{3,t}}{N_{2,t}} = 1 \), \( \frac{N_{1,t}}{N_{2,t}} = x_{3,t} \), the system evolves as

\[
x_{i,t+1} \left( \frac{N_{2,t+1}}{N_{2,t}} \right) = \sum_{j=1}^{3} \left( \eta_{j} q_{ji} + \eta'_{j} q'_{ji} \theta_{ji,t} \right) x_{j,t} \quad \text{for } i = 1, 2, 3, \text{ and } x_{2,t} = 1
\]

(32)

Now define the vector \( X_t \) as \( X_t = [x_{1,t}, 1, x_{3,t}]^T \). We can then write the evolution of the vector \( X_t \) as

\[
X_{t+1} \left( \frac{N_{2,t+1}}{N_{2,t}} \right) = S(\Theta_t) X_t,
\]

\[
S(\Theta) = (Q^T \eta + Q^T_m \eta_m \Theta)
\]

where \( \left( \frac{N_{2,t+1}}{N_{2,t}} \right) \) is a scalar. Since the second component of \( X_{t+1} \) is just the scalar 1, hence \( \left( \frac{N_{2,t+1}}{N_{2,t}} \right) \) is just the second row in the \([3 \times 3]\) matrix \( S(\Theta_t) \), multiplied by the vector \( X_t \). Thus we can write the ratio \( \left( \frac{N_{2,t+1}}{N_{2,t}} \right) \) as a product given by

\[
\left( \frac{N_{2,t+1}}{N_{2,t}} \right) = S_2(\Theta_t) X_t
\]

(33)

where \( S_i \) is the \([1 \times 3]\) i-th row in \( S(\Theta_t) \):

\[
S(\Theta_t) = \begin{bmatrix} -S_1(\Theta_t) & -S_2(\Theta_t) & -S_3(\Theta_t) \end{bmatrix}
\]

(34)

Hence we can write

\[
X_{t+1} S_2(\Theta_t) X_t = S(\Theta_t) X_t,
\]

\[
S(\Theta_t) = (Q^T \eta + Q^T_m \eta_m \Theta)
\]

(35)

Since \( S_2(\Theta_t) X_t \) is a scalar, we can finally write the evolution of the composition of the population in the following way

\[
X_{t+1} = S(\Theta_t) X_t \left[ S_2(\Theta_t) X_t \right]^{-1},
\]

\[
S(\Theta_t) = (Q^T \eta + Q^T_m \eta_m \Theta)
\]

(36)

where \( [S_2(\Theta_t) X_t]^{-1} \) is just the scalar \( \frac{N_{3,t+1}}{N_{2,t+1}} \) (see equation 7).
C.1 Steady State Composition of the Native Population in absence of Immigration

In the absence of immigration the evolution of the native population is given by

\[ X_{t+1} = SX_t [S_2 X_t]^{-1}, \]  
\[ S = Q_n^T \eta. \]  

(37)

At steady state and after multiplying both sides by the scalar \( S_2 X_t \) we obtain

\[ XS_2 X = SX. \]

Hence we can write the system as

\[ (XS_2 - S) X = 0, \]

(38)

which is a quadratic system of equations in \( X \), with one of the solutions being the steady state distribution of \( X \) in the absence of immigration. Since \( X = [x_1 \ 1 \ x_3]^T \), the 2\(^{nd}\) row of the matrix given by the product \( XS_2 \) is just the \( S_2 \) row vector. Hence, the 2\(^{nd}\) row of the matrix given by \((XS_2 - S)\) is identically the \( 0 \) row vector. So we can finally write

\[ (XS_2 - S) X = \begin{bmatrix} x_1 S_2 - S_1 \\ 0^T \\ x_3 S_2 - S_3 \end{bmatrix} \begin{bmatrix} x_1 \\ 1 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \]

(39)

where \( x_1, 1 \) and \( x_3 \) are scalars; \( S_i \) is \([1 \times 3]\) and \( 0^T \) is the vector \([0 \ 0 \ 0]\). Hence the scalars \( x_1 \) and \( x_3 \) that solve this system of 2 quadratic equations in 2 unknowns is the steady state of the system.

D Foreign Born Population

The next table shows numbers of gross immigration to the US by education levels.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number (1000's)</td>
<td>3434</td>
<td>8679</td>
<td>6969</td>
<td>4494</td>
</tr>
<tr>
<td>Percentages of Corresponding Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>by Completed Schooling Levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than High School</td>
<td>30.7</td>
<td>34.9</td>
<td>34.9</td>
<td>30.5</td>
</tr>
<tr>
<td>High School+Some College</td>
<td>35.0</td>
<td>37.5</td>
<td>40.7</td>
<td>40.6</td>
</tr>
<tr>
<td>Bachelor Degree &amp; Beyond</td>
<td>34.3</td>
<td>27.6</td>
<td>24.4</td>
<td>28.9</td>
</tr>
</tbody>
</table>

In order to compute net immigration by skill level, we use the estimates in Ben-Gad (2004) of 3.2 immigrants per 1000 natives for the 1990-99 decade. Estimates of population by education level for individuals 25 years and older are 19.1% for LTHS, 58% for HS+SC and 22.8% for BA and beyond. With these numbers and some algebraic manipulation we can obtain estimates of 5.65 unskilled immigrants per 1000 unskilled natives, 2.025 medium skilled immigrants per 1000 natives and 4.14 skilled immigrants per 1000 skilled natives.

\[\text{Average Schooling by Cohort in the GSS}\]

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>
<thead>
<tr>
<th>Cohort</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1915 - 24</td>
<td>11.03</td>
<td>3.33</td>
<td>115</td>
</tr>
<tr>
<td>1925 - 34</td>
<td>12.01</td>
<td>3.20</td>
<td>1,193</td>
</tr>
<tr>
<td>1935 - 44</td>
<td>12.93</td>
<td>2.95</td>
<td>3,342</td>
</tr>
<tr>
<td>1945 - 54</td>
<td>13.65</td>
<td>2.74</td>
<td>7,157</td>
</tr>
<tr>
<td>1955 - 64</td>
<td>13.64</td>
<td>2.52</td>
<td>7,277</td>
</tr>
<tr>
<td>1965 - 74</td>
<td>13.87</td>
<td>2.55</td>
<td>3,561</td>
</tr>
<tr>
<td>1975 - 84</td>
<td>13.86</td>
<td>2.55</td>
<td>996</td>
</tr>
<tr>
<td>All</td>
<td>13.49</td>
<td>2.74</td>
<td>23,641</td>
</tr>
</tbody>
</table>

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

\[\text{Intergenerational Mobility Matrices}\]

\[\text{F.1 Estimation and tests of transition matrices}\]

Given the filters for the data described in the main text, define \(n_{ij}\) (i=1, 2, 3; j = 1, 2, 3) as the number of cases of a parent with education degree [i] with sons/daughters of degree [j]. Also define \(N_i\) (i=1, 2, 3) as the total number of observations whose parents have degree [i]. Then \(N_i = \sum_{j=1}^{3} n_{ij}\). The maximum likelihood estimates (MLE) of the probabilities describing a first order Markov process are given by \(\hat{q}_{ij} = \frac{n_{ij}}{N_i}\) (Anderson and Goodman (1957)).

Using data for men in the GSS, aged 25-55 at the time of the interview for individuals born at least since 1945, for interviews taken during the period 1977-
2012, we obtain 8476 observations for natives and 636 observations for sons of first generation immigrants. The estimated matrices are given by

\[
\hat{Q}_n = \begin{bmatrix}
  .281 & .628 & .091 \\
  .068 & .692 & .240 \\
  .008 & .401 & .590
\end{bmatrix}, \quad \hat{Q}_m = \begin{bmatrix}
  .240 & .544 & .216 \\
  .068 & .647 & .285 \\
  .021 & .338 & .641
\end{bmatrix}.
\]

All thee rows of the transition matrices show that the children of immigrants to the US appear to be more "successful" than natives. That is, they have a lower probability to become low or medium-skilled but a higher probability of becoming skilled. The only exception are the children of skilled immigrants that appear to have a slightly higher probability of being unskilled. The most dramatic difference between immigrants and natives are for unskilled parents as their children have a lower probability of becoming unskilled or medium skilled, but a higher probability of becoming skilled (21.6% for children of unskilled immigrants, and only 9.1% for children of unskilled natives).

We can formally test whether the probability distributions for natives and children of immigrants are statistically the same, conditional on the skill of the parents. In other words, we test whether the row \( [i] \) of matrix \( \hat{Q}_n \) is statistically different from the same row in the matrix \( \hat{Q}_m \). The null hypothesis is given by

\[ H_0 : q_{ij} = q_{ij}^2 \text{ for all } j=1,2,3, \text{ given row } [i]. \]

If the parameters \( q_{ij} \) of the population are unknown, then they are estimated via maximum likelihood (ML), which just considers one big sample of all immigrants and natives cases. Denote the ML estimator under the null hypothesis as \( \hat{q}_{ij} \). Then \( H_0 \) is rejected if the statistic

\[
\sum_{m=1}^{2} S_{im} > \chi^2_{(k-1)},
\]

where

\[
S_{im} = n_{im} \sum_{j=1}^{k} \left( \frac{(q_{ij} - \hat{q}_{ij})^2}{\hat{q}_{ij}} \right), \quad \text{and } n_{im} \text{ is the total counts of row } i \text{ (in transition matrix) of sample } m.
\]

This test can be generalized for more than two populations, in which case \( H_0 \) is rejected if

\[ \sum_{m=1}^{b} S_{im} > \chi^2_{(b-1)(k-1)} \]

where \( b \) is the number of samples that we are comparing under a null of identical population.

With this \( \chi^2_{(k-1)} \) test we obtain for each row the following statistics;

\[
\sum_{m=1}^{2} S_{1m} = 26.05, \quad \sum_{m=1}^{2} S_{2m} = 3.47 \text{ and } \sum_{m=1}^{2} S_{3m} = 4.35.
\]

Each one is to be compared with a \( \chi^2_{(3-1)} \) with \( 3-1 = 2 \) degrees of freedom. The \( \chi^2_{2} \) at the 5% and 1% significance levels are respectively 5.99 and 9.21. Hence with these tests we can only reject the null hypothesis for unskilled parents.

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 \( q_{ij} = q_{ij}^2 \) for all \( i = 1, 2, 3 \) and \( j = 1, 2, 3 \). Then the null
is rejected if \( \sum_{i=1}^{k} \sum_{m=1,2} S_{im} > \chi^2_{(k(k-1))} \) where the degrees of freedom are in this case \((k-1)k = 3(2) = 6\). The test produces a statistic of 33.87, which is to be compared with a \( \chi^2 \) critical value at 1% significance with 6 degrees of freedom, which is 16.81. Hence, the test rejects that children of natives and those of first generation immigrants have the same transition matrix.

In the case of women under the same filters we have the following estimated transition matrices

\[
\tilde{Q}^{women}_n = \begin{bmatrix} .239 & .687 & .075 \\ .057 & .719 & .224 \\ .011 & .394 & .595 \end{bmatrix}, \quad \tilde{Q}^{women}_m = \begin{bmatrix} .192 & .628 & .180 \\ .067 & .622 & .311 \\ .023 & .314 & .663 \end{bmatrix},
\]

where the total number of observations for the estimation of natives is 10523, and 811 for daughters of immigrants. In general we observe the same pattern as with men: the probabilities of upward mobility (according to schooling level) are higher for children of immigrants. In this case there are statistically significant differences for the first and second rows (daughters of medium and unskilled parents). The statistics are 32.72 for children of unskilled parents, and 17.45 for medium-skilled parents, which is higher than the 9.21 value of a \( \chi^2 \) with 2 degrees of freedom at the 1% level of significance. There is not statistical difference in the probability distribution of daughters of skilled parents, as the test statistic = 5.89 is slightly lower than the critical value of the test of 5.99 (at 5% level). The test for the whole matrix testing whether \( q_{ij}^{women} \) for all \( i,j \) are identical for both populations is rejected. The test statistic is 56.06 > 16.81.

F.2 Further partitions of the data

Men Vs Women. Comparing the transition matrices for native men and women, 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 unskilled and medium-skilled parents are different to those of daughters, with statistics of 14.28 (unskilled parents) and 11.27 (medium-skilled), while we cannot reject that the probability distribution of sons and daughters of skilled parents are the same (statistic=0.95). The matrix test rejects that men and women have the same transition matrices, with an statistic of 26.49 > 16.81.

In the case of immigrants, the matrix test produces an statistic of 3.79, which cannot reject that the matrices are equal. 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 white men are compared to white women (natives), the results remain essentially unchanged as the matrices almost don’t change, nor the results of the tests.

Using only white-men natives for comparison with sons of immigrants. If for native immigrants we use only data of white men, the results
are practically identical to our previous results. The row tests again show probability distributions significantly different for unskilled (statistic = 21.53) and statistically the same for children of medium skilled parents (statistic = 2.12), as well as those of skilled parents (4.05). The test for the whole matrix again rejects that they are identical (statistic = 27.69).

When using native women vs daughters of first generation immigrants, again the natives matrix for all races and for whites are almost identical. The row tests again rejects that the probabilities distributions are the same for unskilled and medium-skilled parents, while for skilled parents we can’t reject the null of equality.

F.3 Estimated Intergenerational Transition Matrices

### Transition Matrix for Children of Immigrants

<table>
<thead>
<tr>
<th>Transition Matrix (Qm)</th>
<th>Men/Women</th>
<th>Race</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.240 .544 .216</td>
<td>N1 = 171</td>
<td></td>
</tr>
<tr>
<td>0.068 .647 .285</td>
<td>N2 = 323</td>
<td>Men</td>
<td>All</td>
</tr>
<tr>
<td>0.021 .338 .641</td>
<td>N3 = 142</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.192 .628 .180</td>
<td>N1 = 250</td>
<td></td>
</tr>
<tr>
<td>0.067 .622 .311</td>
<td>N2 = 389</td>
<td>Women</td>
<td>All</td>
</tr>
<tr>
<td>0.023 .314 .663</td>
<td>N3 = 172</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.211 .594 .195</td>
<td>N1 = 421</td>
<td></td>
</tr>
<tr>
<td>0.067 .633 .299</td>
<td>N2 = 712</td>
<td>Both</td>
<td>All</td>
</tr>
<tr>
<td>0.022 .325 .653</td>
<td>N3 = 314</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Transition Matrices for Natives

<table>
<thead>
<tr>
<th>#</th>
<th>Transition Matrix (Qn)</th>
<th>Men/Women</th>
<th>Race</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.281 .628 .091</td>
<td>N1 = 1538</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.068 .692 .240</td>
<td>N2 = 4985</td>
<td>Men</td>
<td>All</td>
<td>25-55</td>
</tr>
<tr>
<td>0.008 .401 .590</td>
<td>N3 = 1953</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.282 .622 .096</td>
<td>N1 = 1094</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.064 .685 .251</td>
<td>N2 = 4318</td>
<td>Men</td>
<td>White</td>
<td>25-55</td>
</tr>
<tr>
<td>0.007 .384 .609</td>
<td>N3 = 1806</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.239 .687 .075</td>
<td>N1 = 2317</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.057 .719 .224</td>
<td>N2 = 6091</td>
<td>Women</td>
<td>All</td>
<td>25-55</td>
</tr>
<tr>
<td>0.011 .394 .595</td>
<td>N3 = 2115</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.232 .690 .078</td>
<td>N1 = 1465</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.047 .713 .231</td>
<td>N2 = 4984</td>
<td>Women</td>
<td>White</td>
<td>25-55</td>
</tr>
<tr>
<td>0.009 .381 .610</td>
<td>N3 = 1901</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.256 .663 .081</td>
<td>N1 = 3855</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.062 .707 .231</td>
<td>N2 = 11076</td>
<td>Both</td>
<td>All</td>
<td>25-55</td>
</tr>
<tr>
<td>0.010 .397 .593</td>
<td>N3 = 4068</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.253 .661 .086</td>
<td>N1 = 2559</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.055 .700 .245</td>
<td>N2 = 9302</td>
<td>Both</td>
<td>White</td>
<td>25-55</td>
</tr>
<tr>
<td>0.008 .382 .609</td>
<td>N3 = 3707</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$ Null Hypothesis</td>
<td>Matrix Test CV=12.59</td>
<td>Row1 Test CV=5.99</td>
<td>Row2 Test CV=5.99</td>
<td>Row3 Test CV=5.99</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>$Q_{n,[#1]} = Q_{m,[#1]}$</td>
<td>33.87**</td>
<td>26.05**</td>
<td>3.47</td>
<td>5.80</td>
</tr>
<tr>
<td>$Q_{n,[#2]} = Q_{m,[#1]}$</td>
<td>27.69**</td>
<td>21.53**</td>
<td>2.12</td>
<td>5.21</td>
</tr>
<tr>
<td>$Q_{n,[#3]} = Q_{m,[#2]}$</td>
<td>56.06**</td>
<td>32.72**</td>
<td>17.45**</td>
<td>5.03</td>
</tr>
<tr>
<td>$Q_{n,[#4]} = Q_{m,[#2]}$</td>
<td>46.89**</td>
<td>26.78**</td>
<td>14.68**</td>
<td>4.26</td>
</tr>
<tr>
<td>$Q_{n,[#2]} = Q_{m,[#4]}$</td>
<td>34.89**</td>
<td>12.03**</td>
<td>19.83**</td>
<td>3.42</td>
</tr>
<tr>
<td>$Q_{m,[#1]} = Q_{m,[#2]}$</td>
<td>3.79</td>
<td>3.00</td>
<td>0.58</td>
<td>0.55</td>
</tr>
<tr>
<td>$Q_{n,[#1]} = Q_{m,[#3]}$</td>
<td>26.49**</td>
<td>14.28**</td>
<td>11.27**</td>
<td>1.95</td>
</tr>
<tr>
<td>$Q_{n,[#5]} = Q_{m,[#3]}$</td>
<td>87.52**</td>
<td>58.74**</td>
<td>18.63**</td>
<td>9.34**</td>
</tr>
</tbody>
</table>

Notes: The 1% critical value of the test for equality of matrices is 16.81, from a $\chi^2$ distribution with 6 degrees of freedom. The 1% critical value (***) for the tests of equality of rows is 9.21.

F.4 On the Definition of Second Generation Immigrants

The following transition matrices are constructed when the definition of second generation immigrant is given to 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.

Transition Matrices of children of immigrants: both parents born outside US

<table>
<thead>
<tr>
<th>Transition Matrix ($Q_m$)</th>
<th>Men/Women</th>
<th>Race</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1'$</td>
<td>.286 .464 .250</td>
<td>$N_1 = 84$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.085 .606 .309</td>
<td>$N_2 = 94$</td>
<td>Men</td>
</tr>
<tr>
<td></td>
<td>.043 .391 .565</td>
<td>$N_3 = 46$</td>
<td></td>
</tr>
<tr>
<td>$2'$</td>
<td>.165 .617 .218</td>
<td>$N_1 = 133$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.058 .545 .397</td>
<td>$N_2 = 121$</td>
<td>Women</td>
</tr>
<tr>
<td></td>
<td>.019 .288 .692</td>
<td>$N_3 = 52$</td>
<td></td>
</tr>
<tr>
<td>$3'$</td>
<td>.212 .558 .230</td>
<td>$N_1 = 217$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.070 .572 .358</td>
<td>$N_2 = 215$</td>
<td>Both</td>
</tr>
<tr>
<td></td>
<td>.031 .337 .633</td>
<td>$N_3 = 98$</td>
<td></td>
</tr>
</tbody>
</table>

G Estimation of Fertility Rates of US and Foreign Born Women

Here we detail how we estimate fertility rates 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 the help of the American Community Survey (ACS), which identifies the place where the mother was born and can be used to estimate total fertility rates. The estimators are derived for a specific level of
education at a point in time, whose subscripts we don’t include in order to keep notation simple. Define

\[ X_{ik} = \text{# of children born to women type } k \text{ in age-group } i \text{ during the year}\]

\[ Z_{ik} = \text{# of mothers type } k \text{ in age-group } i \text{ during the year}\]

\[ Y_{ik} = \text{total # of women type } k \text{ in age-group } i \]

\[ k = \{N, F\} \text{ where } N=\text{US-born and } F=\text{foreign-born}\]

\[ i = \# \text{ of age-groups (each group includes 5-year categories)}\]

\[ F^k = \text{Total fertility rate of women type } k \]

\[ F = \text{Total fertility rate of all women living in the US} \]

**Total Fertility rates. Years 1990, 2000 and 2005**

<table>
<thead>
<tr>
<th>Year</th>
<th>less than HS</th>
<th>HS + Some College</th>
<th>BA+</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>2.73</td>
<td>1.95</td>
<td>1.85</td>
</tr>
<tr>
<td>2000</td>
<td>2.26</td>
<td>2.00</td>
<td>1.84</td>
</tr>
<tr>
<td>1990</td>
<td>2.43</td>
<td>2.04</td>
<td>1.61</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>2.47</strong></td>
<td><strong>2.00</strong></td>
<td><strong>1.77</strong></td>
</tr>
</tbody>
</table>

Then the formula for the total fertility rate is by definition

\[ F = \sum_{i=1}^{n} \left[ \frac{X^N_{i} w^N_{i} + X^F_{i} (1 - w^N_{i})}{Y^N_{i} + Y^F_{i}} \right] ; \]

with weights \( w^N_{i} \) given by \( w^N_{i} = \frac{Y^N_{i}}{Y^N_{i} + Y^F_{i}} \), which are just the US-born women of age-group \( i \) as a share of all women of age-group \( i \), for a given level of education. It can be shown that we can rewrite the total fertility rates of foreign women \( F^F \) and US-born women \( F^N \) as

\[ F^F = F + \sum_{i=1}^{n} m_i w^N_{i} \quad \text{and} \quad F^N = F - \sum_{i=1}^{n} m_i (1 - w^N_{i}) . \]

**Estimates of \( m_i \), averages for 2001-2008 ACS**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Unskilled</th>
<th>Med-Skilled</th>
<th>Skilled</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 – 19</td>
<td>.014</td>
<td>.002</td>
<td>-.003</td>
</tr>
<tr>
<td>20 – 24</td>
<td>.002</td>
<td>.000</td>
<td>.010</td>
</tr>
<tr>
<td>25 – 29</td>
<td>.042</td>
<td>.0016</td>
<td>.005</td>
</tr>
<tr>
<td>30 – 34</td>
<td>.051</td>
<td>.037</td>
<td>-.004</td>
</tr>
<tr>
<td>35 – 39</td>
<td>.036</td>
<td>.030</td>
<td>.012</td>
</tr>
<tr>
<td>40 – 44</td>
<td>.015</td>
<td>.012</td>
<td>.009</td>
</tr>
<tr>
<td>45 – 49</td>
<td>.000</td>
<td>.003</td>
<td>.003</td>
</tr>
<tr>
<td>50 – 54</td>
<td>.000</td>
<td>.000</td>
<td>.001</td>
</tr>
</tbody>
</table>

where \( m_i = \left( \frac{X^F_{i}}{Y^F_{i}} - \frac{X^N_{i}}{Y^N_{i}} \right) \) 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). This shows that we can estimate \( F^F \) and \( F^N \) from the total fertility rate \( F \) as long as we have data on both \( m_i \)’s and \( w^N_{i} \)’s. We can back up exact data on weights \( w_i \) from the census of 1990, 2000 and the 2005 CPS. However, data is not available for the differences \( m_i \) for these years. Hence, we estimate the average
difference $m_i$ for each age-education group for years 2001-2008 for which ACS data is available and use fertility rates $F$ to estimate fertility rates by nativity. Fertility rates, weights ($w_i$) and differences ($m_i$) are shown below.

Estimates of the $w_i$ weights, census 1990, 2000 and CPS-2005

<table>
<thead>
<tr>
<th>Age Group</th>
<th>1990</th>
<th>2000</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unskilled</td>
<td>Medium</td>
<td>Unskilled</td>
</tr>
<tr>
<td>15-19</td>
<td>0.082</td>
<td>0.068</td>
<td>0.250</td>
</tr>
<tr>
<td>20-24</td>
<td>0.209</td>
<td>0.079</td>
<td>0.091</td>
</tr>
<tr>
<td>25-29</td>
<td>0.229</td>
<td>0.077</td>
<td>0.106</td>
</tr>
<tr>
<td>30-39</td>
<td>0.257</td>
<td>0.070</td>
<td>0.096</td>
</tr>
<tr>
<td>35-39</td>
<td>0.241</td>
<td>0.075</td>
<td>0.099</td>
</tr>
<tr>
<td>40-44</td>
<td>0.208</td>
<td>0.074</td>
<td>0.101</td>
</tr>
<tr>
<td>50-54</td>
<td>0.185</td>
<td>0.077</td>
<td>0.106</td>
</tr>
</tbody>
</table>

H The Solution Algorithm and Optimal Policies

We explain the algorithm used for 2 cases: the case with an exogenous limit on immigration given by $\theta_3^{\text{max}} = .13$, and then the case where there’s a supply of high skilled immigrants as described in the text. Define the steady state pair ratios in absence of immigration as $\{x_1^{SS}, x_3^{SS}\}$.

These numbers are the solution to (38). A grid of $N_1$ points in $[(1-a_1) x_1^{SS}, (1+a_2) x_1^{SS}]$ and $N_2$ points in $[(1-b_1) x_3^{SS}, (1+b_2) x_3^{SS}]$ for a total of $N_1 \times N_2$ 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 low), which we now describe.

The policy space considers $N_{\theta_1}$ points for $\theta_1$ in the space $[0, \theta_1^{\text{max}}]$, $N_{\theta_2}$ points for $\theta_2$ in $[0, \theta_2^{\text{max}}]$ and $N_{\theta_3}$ points for $\theta_3$ in $[0, \theta_3^{\text{max}}]$. Since for the numbers calibrated to the US economy we find zero immigration of medium-skilled individuals, the number of points and space for $\theta_2$ are unimportant (optimally zero medium skilled immigration is chosen). For $\theta_3$ in turn we find that for all points in the state space the immigration quota hits the maximum ($\theta_3 = \theta_3^{\text{max}}$) in the version where there’s no possible equalization of wages ("small" pool).

The steady state ratios $\{x_1^{SS}, x_3^{SS}\}$ in absence of immigration given the demographic process characterized by $(Q_n, \eta_n)$ are obtained as the solutions to the equations (39), which yield $\{.09782, .5429\}$. The size of the state grid depends on the policy space considered. Here we describe the grids of the fixed-supply case, as well as the grid used for the positive-slope supply of high skilled immigration.

In the case of a maximum supply of high skilled immigrants at a level of $\theta_3^{\text{max}} = 13\%$, the grid parameters for the baseline case are $\{a_1, a_2, b_1, b_2\}$ are
given by \{2, 8, 2, 8\}, with \(N_{X_1} = 30\) points and \(N_{X_2} = 33\) points for a total of 990 state points in the resulting grid in \([.07826,.176084] \times [.434353,.977295]\). The considered immigration policies are elements \((\theta_1, \theta_2, \theta_3) \in [0, \theta_1^{\text{max}}] \times [0, \theta_2^{\text{max}}] \times [0, \theta_3^{\text{max}}] \subseteq R^3_+\). For this version of the model, the number of policy points in \(\theta_2\) and the maximum immigration \(\theta_2^{\text{max}}\) 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 \(\theta_3^{\text{max}} = 0.13\) is lower than what the majority would choose in the unconstrained version of the model (which results in very high immigration quotas - see baseline parameterization in text) the majority also chooses the maximum immigration quota for high skilled agents \(\theta_3^* = \theta_3^{\text{max}} = 0.13\). The only relevant policy information in this version of the model are the number of policy points available for \(\theta_1\), having a large enough \(\theta_1^{\text{max}}\) in order to allow for an interior solution, and the specific level of \(\theta_3^{\text{max}}\) (set at 13% in this version). Below we show the case where \(\theta_1^{\text{max}} = 1.25\), and \(\theta_3^{\text{max}} = 0.13\), with 401 equidistant points in the policy space for \(\theta_1 \in [0, 1.25]\). Solving the model with a tolerance of \(0.0000001\) of the sup norm between the value functions of the current and the previous iteration, we obtain the following optimal policy function for the preferred quota of unskilled workers for the median voter shown below. The unskilled immigration quota is 18.09%, and the high skilled immigration quota is chosen at 13%. The induced steady state is \((.1006,.5886)\).

**Optimal Policy function for unskilled immigration.**

Cases where there are guest worker quotas available involve increasing the size of the policy space.

For the case in which there exists a supply of high skilled immigration as captured by equation (28), we illustrate the case in which the elasticity of supply of the high skilled is 10 and the rest of the parameters as in the main text, the state grid is given by 900 pairs of points in \([.09478,.176084] \times [.434353,.7492]\). In this case ith immigration policy \((\theta_{1i}, \theta_{2i}, \theta_{3i}) \in \Theta\) at the state node \((x_1, x_3)\)
is feasible if $\theta_{3i} \leq \theta_{3i}^{\text{max}} (w_3(x_1, x_3, (\theta_{1i}, \theta_{2i}, \theta_{3i})))$. Using 90 points for $\theta_1 \in [0, 2]$ and 150 points in $\theta_3 \in [0, 0.4]$ where only the feasible policies due to the supply constraint on high skilled immigration are considered. Again $\theta_{3i}^{\text{max}}$ turned out to be irrelevant. At the no-immigration steady state the optimal policies are $(42.74\%, 0, 19.88\%)$. Below we show the immigration policies in the state grid considered. The shape of the unskilled immigration quota is just like before: increasing in $x_3$ and decreasing in $x_1$, while the shape of the skilled immigration policy is decreasing in $x_3$ and slightly increasing in $x_1$. Other cases look qualitatively identical.

Optimal policy functions $\theta_1$ and $\theta_3$ with an elastic response of $\theta_3^{\text{max}}$

Elasticity = 10