THE UNIVERSITY OF CHICAGO

QUANTITATIVE IMPACT OF REDUCING BARRIERS TO SKILLED LABOR IMMIGRATION: THE CASE OF THE US H-1B VISA

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECONOMICS

BY
HYUN LEE

CHICAGO, ILLINOIS
JUNE 2016

Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I am particularly grateful to my four thesis committee members: Robert E. Lucas, Jr., Nancy Stokey, Robert Shimer, and Joseph Vavra. Thank you for your incredible wisdom, support, and encouragement. It was truly my honor to learn from the best and be able to call you my adviser and mentor.

I would like to thank some of the professors who generously gave their time and advice in discussing elements of this dissertation, including Fernando Alvarez, Yongsung Chang, Mikhail Golosov, Aspen Gorry, Jeffrey Grogger, Benjamin Moll, Alexander Monge-Naranjo, Randy Wright, and many other people I was fortunate to meet on the academic job market this year. I would also like to thank the faculty and staff at the University of Chicago for providing such a supportive environment, and for granting me this opportunity. In addition, I gratefully acknowledge the financial support I have received from the Samsung Scholarship Foundation and the University of Chicago.

Many of my colleagues provided invaluable comments and support, including BongGeun Choi, Jung Sakong, Philip Barrett, Munseob Lee, Sara Moreira, and participants at the University of Chicago's Capital Theory Working Group, Western Economic Association Annual Conference, Economic Graduate Student Conference at the Washington University in St. Louis, and Fall 2015 Midwest Macro Meeting. I also want to give a shout out to the members of my study group from my "original" cohort—Aditya Bhave, Tim Kautz, Eric Nielsen, Marshall Steinbaum, Armin Rick, Enghin Atalay, Alice Chen, and Christina Jenq.

Finally, this dissertation would not have been possible without the love and support from my family—Dad, Mom, my brother Chan, and my in-laws—and, of course, the two women that I love the most, my wife Seung Hee "Christina" Kang and my daughter Sophia Soyoun Lee.

Abstract

In this dissertation, I develop a novel two-country general equilibrium model of immigration with heterogeneous skilled and unskilled labor. I use the model to study the short-run and long-run impacts of permanently doubling the US H-1B visa quota on the skill distribution, wages, output, and welfare in the United States and the source countries. I find that once the policy change takes place, less talented skilled foreigners who would not have applied for the visa at the old steady state now apply en masse such that the probability of obtaining the H-1B visa increases by only 11 percentage points. In the long-run, the United States experiences a 0.79% gain in output per capita once the economy completes transitioning to its new steady state. However, this aggregate gain hides wide heterogeneity in the effects on wages and welfare both across and within skill groups and age cohorts. In addition, I show that models that ignore return migration underestimate the average skill of an immigrant skilled worker, but overestimate the long-run increase in immigrant stock in response to reducing barriers to labor mobility. Because the latter effect dominates the former, I find that models with no return migration overestimate the changes in welfare by more than sixfold for certain cohorts compared to the benchmark model.

Chapter 1

Introduction

Skilled immigration is globally becoming ever more important. Docquier and Marfouk (2006) document that between 1990 and 2000, the total number of skilled immigrants legally residing in the OECD countries has multiplied by 1.64, whereas the total number of legally residing unskilled immigrants has multiplied by 1.14. Despite the relative increase in skilled immigration, the majority of the economic literature has focused on unskilled labor immigration and its impact on the natives' wages and employment.[1] Only in recent years have the researchers started to scratch the surface of skilled immigration, but most of these studies have focused on the link between skilled immigration and scientific innovation. So far, the literature has abstracted from skill heterogeneity and has thus been silent on the welfare effects of skilled immigration on both the source and destination countries. This paper fills that gap.

I develop a novel two-country life-cycle model with skilled and unskilled workers, in which skilled workers are heterogeneous in terms of their level of human capital and are able to switch their country of residence. I apply the model to the case of the US H-1B visa. Established in 1990, the H-1B visa is the largest US visa program for temporary skilled labor immigration to the United States. Since fiscal year 2004, the number of H-1B

[1] Borjas (1994), Friedberg and Hunt (1995), Okkerse (2008), and Kerr and Kerr (2011) provide good surveys of this literature.

visa applicants has been greater than the general quota of 65,000 per year, and the US government has resorted to a lottery to determine who receives the visa. The probability of obtaining the H-1B visa has been lower than 50% for the past few years, and there have been several bipartisan bills introduced in the US Congress to lower this barrier by doubling or even tripling the H-1B visa quota.[2]

Using the calibrated model, I study the impact on the United States and the source countries if the US were to pass the aforementioned bill and permanently double the US H-1B visa quota. I calculate the transition path from the old steady state to the new one for both countries, and analyze the short-run and long-run gains and losses in wages, output, and welfare. I find that once the US H-1B visa quota is doubled, the flow of skilled immigrants immediately doubles to its new steady state level, but the return migration flow, the immigrant stock, wages, and output per capita all converge gradually to their new steady states. The increased inflow of foreign skilled immigrants is mostly composed of less talented skilled workers who would not have applied for the visa at the old steady state. At the end of the transition path, the output per capita in the United States increases by 0.79% compared to the old steady state. However, this aggregate gain hides wide heterogeneity in the effects on wages and welfare both across and within skill groups and age cohorts. The average change in welfare for US skilled workers is negative since the increased inflow of skilled immigrants depresses the US skilled wage. However, there are US skilled workers who experience positive gain in welfare depending on their level of capital, human capital, and age. On average, the magnitude of the changes in welfare is greater the younger a worker is at the time of the policy change, and future generations that enter the model after the policy change experience even greater magnitudes of gains or losses.

There are three main contributions of this paper. First, to my knowledge, the framework developed in this paper is the first to divide labor into skilled labor and unskilled labor within a structural general equilibrium model of immigration. The majority of the immi-

[2]Immigration Innovation Act of 2013 and Immigration Innovation Act of 2015 are some examples.

gration literature uses identified regressions to analyze the economic impact of increasing immigration. However, this technique can neither answer the question of welfare gains and losses, nor conduct counterfactual analysis and predict how the economy transitions from one steady state to another after a change in immigration policy. Moreover, in addition to quantifying the impact of the increased inflow of foreign skilled workers on the US skilled population, my model allows for quantifying the general equilibrium effect on the wages and welfare of US unskilled workers. In fact, I find that for certain age cohorts, the magnitude of the average gain in welfare for US unskilled workers is greater than the magnitude of the average loss in welfare for US skilled workers.

Second, I show that models that ignore return migration significantly overestimate the impact of reducing barriers to labor mobility. The recent estimates of the return migration rates of immigrants going back to their place of origin have been non-negligible. Using the 2000 Census and the 2005 American Community Survey, the OECD (2008) estimates that the overall exit rate—after five years of residence—of immigrants who entered the United States in 1999 was 19%, and the five-year exit rate of immigrants with at least a college degree was even higher at 23.5%.[3] However, recent papers such as Klein and Ventura (2009) and Kennan (2013), which structurally estimated the magnitude of efficiency loss caused by international labor mobility barriers, ignore return migration in their models. By comparing my benchmark model to an alternate model that ignores return migration, I find that shutting down return migration results in the underestimation of the average skill level of skilled immigrants which underestimates the magnitude of the impact of lowering immigration barriers, and the overestimation of the long-run increase in immigrant stock which overestimates the magnitude of the policy impact. I find that the latter effect dominates the former, and consequently, the magnitude of the changes in welfare calculated from the alternate model

[3] Borjas and Bratsberg (1996) calculate—using the 1980 Census data—the re-emigration rate of immigrants after five years of residence or less to be 17.5% and after six to ten years of residence to be 21.5%. Similar patterns are observed for foreign doctoral students: Finn (2014) calculates that 34% of all foreign students who received doctorates in science and engineering (including social science) in 2006 from a US institution left the US by 2011.

is up to six times greater than the estimates from the benchmark model.

Third, by introducing skill heterogeneity within skilled workers, I assess the impact of the change in immigration policy on the skill distribution. The immigration literature—led by the seminal paper of Borjas and Bratsberg (1996) and empirically corroborated by studies such as Lubotsky (2007) and Grogger and Hanson (2015)—argues that return migration reinforces the original selection that characterized the immigrant population. That is, if immigrants are positively selected in terms of their skills, then the marginal immigrant most likely to become a return migrant is the least skilled worker of this batch of positively selected immigrants.[4] H-1B visas require sponsorship by US employers who are willing to hire the candidate, which implies that the H-1B workers are positively selected, and those who return should be negatively selected. I build a model with skill heterogeneity in which return migrants possess a lower level of human capital on average compared to those who stay in the United States in equilibrium, and show that a doubling of the quota in this setting leads to a massive increase in visa application of less talented skilled workers who now find it optimal to apply. As a result, the change in policy does little to lessen the bottleneck caused by the visa quota, while the average human capital of skilled immigrant workers decrease.

This paper relates to the extensive empirical literature on the effects of immigration on labor market outcomes in the destination country (e.g. Card (1990), LaLonde and Topel (1992), Borjas, Freeman, and Katz (1997), Card (2001), Borjas (2003), Ottaviano and Peri (2012)). My estimated impact of doubling the H-1B visa quota on equilibrium skilled labor wages is well within the range of the estimated elasticity of native wage to immigration found in this literature using identified regressions, as can be seen in the Kerr and Kerr (2011) survey article. It also relates to the literature on the impact of the H-1B program and/or skilled labor immigration on innovation, such as Kerr and Lincoln (2010), Hunt and Gauthier-Loiselle (2010), Kerr, Kerr, and Lincoln (2015), Peri, Shih, and Sparber (2015), Doran,

[4]Lubotsky (2007) finds that selective emigration by immigrants with poor earnings in the United States led to a systematic overstatement of assimilation and Grogger and Hanson (2015) find that science and engineering PhD's with the strongest academic potential, measured in terms of their attributes and performance at the time they enter graduate school, are those most intent on staying in the US after graduation.

Gelber, and Isen (2015), and Bound, Braga, Golden, and Khanna (2015). Whereas other papers focus on estimating the impact of the H-1B program on innovation and employment through regression studies, Bound, Braga, Golden, and Khanna (2015) use a calibrated partial equilibrium dynamic model to study the impact of skilled immigration on the labor market for computer scientists in the United States during the internet boom of the 1990s and the subsequent slump in the early 2000s. However, their method precludes any analysis of the overall welfare impact, which my paper achieves using a general equilibrium framework.

This paper is also related to Kennan (2013) and Klein and Ventura (2009)—two aforementioned papers that use structural models to estimate the loss in output and welfare associated with immigration barriers. Kennan (2013) uses a simple static model of migration costs with location-specific labor-augmenting productivity to estimate the gains from the complete elimination of labor mobility barriers. In a paper closely related to this one, Klein and Ventura (2009) use a dynamic life-cycle model with cross-country differences in total factor productivity and quantity of available land to assess the effects of removing barriers to labor mobility on output, capital accumulation, and welfare. Both papers find a sizable gain to completely eliminating the barriers to international labor mobility. My paper differs from Klein and Ventura (2009) by dividing labor into skilled and unskilled labor in order to focus on the effects of increased skilled immigration. Furthermore, my model has skill heterogeneity and allows for return migration. Even though these features substantially increase the complexity of the model, my results show that incorporating these features is not only crucial for matching the data, but also have significant impact on welfare estimations.

Methodologically, this paper is related to the literature on dynamic general equilibrium models with heterogeneous agents, especially the so-called Aiyagari (1994) models. The model developed in this paper is a two-country version of the Aiyagari (1994) economy populated by finitely lived heterogeneous skilled workers subject to uninsurable idiosyncratic human capital shock. Since the histories of human capital shocks differ across workers, the model generates equilibrium cross-sectional distributions of asset, human capital, and consumption. These cross-sectional distributions, especially that of human capital, are calibrated to fit their empirical counterparts in the data, and I analyze how they respond to a change in the immigration policy.

The rest of the paper is as follows. In Chapter 2, I describe the H-1B visa program in detail. In Chapter 3, I describe the model and its recursive stationary equilibrium. In Chapter 4, I calibrate the model, with one country representing the United States and the other representing a block consisting of low- and middle-income countries as classified by the World Bank, to match key moments from the data. I then solve for the steady state of the model. In Chapter 5, I conduct a policy experiment in which the H-1B visa quota is permanently doubled and analyze how the economy transitions from the old steady state to the new one. In Chapter 6, to assess the quantitative importance of return migration, I compare the baseline model against an alternate version of the model in which return migration is shut down. The last chapter concludes.

Chapter 2

H-1B Visa Program

The H-1B visa program was established by the Immigration Act of 1990 for temporary employment of foreign college-educated workers in the United States in "specialty occupations." Section 1184(i)(1) of Title 8 of the United States Code defines a specialty occupation as "an occupation that requires theoretical and practical application of a body of highly specialized knowledge, and attainment of a bachelor's or higher degree in the specific specialty (or its equivalent) as a minimum for entry into the occupation in the United States."

The H-1B is an employer-sponsored visa, which means that a firm applies for the visa on behalf of a foreign skilled worker it wants to hire. Firms that wish to hire foreign workers on H-1B visas must first file a Labor Condition Application (LCA) with the US Department of Labor. In the LCAs, the employer must attest that the firm will pay the foreign worker the greater of the actual compensation paid to other employees in the same job or the prevailing compensation for that occupation, and that the firm will provide working conditions for the foreign worker that do not adversely affect the working conditions of the other employees. Once the LCA is approved, the employer files Form I-129 (Petition for a Nonimmigrant Worker) requesting H-1B classification for the worker to the US Citizenship and Immigration Services Bureau (USCIS), accompanied by supporting documents demonstrating that the individual possesses the necessary credentials and all the required fees.[1] The USCIS then

[1]Base filing fee for H-1B is $325. In addition, employers with 1 to 25 full-time equivalent employees must

can approve the petition for the H-1B nonimmigrant for a period of up to three years, which can later be extended by an additional three years.[2] Once the work has started, the foreign skilled worker is effectively tied to the firm until he or she obtains permanent residency or obtains another visa.

Since its establishment in 1990, the H-1B visa program has had a cap on the number of new three-year H-1B visas issued per fiscal year. Initially, the annual quota was set at 65,000 that was rarely reached in the early 1990s. The quota was changed in 1998 by the American Competitiveness and Workforce Improvement Act, which temporarily increased the cap to 115,000 for fiscal years 1999 and 2000. The American Competitiveness in the 21st Century Act passed in 2000 increased the quota even further to 195,000 for fiscal years 2001, 2002, and 2003, while creating an uncapped category for non-profit research institutions and universities. In fiscal year 2004, the H-1B quota reverted back to the original 65,000, but the H-1B Visa Reform Act of 2004 added 20,000 to the cap for applicants with US postgraduate degrees. The H-1B non-immigrants who work at (but not necessarily for) universities, non-profit research facilities associated with universities, and government research facilities are still excluded from the cap.

The H-1B petitions can be filed up to six months before the day work starts. Because the US government's fiscal year starts on October 1, the first date on which one can apply for the H-1B visa for any given fiscal year is April 1. In recent years, the USCIS has received more applications than allowed for by the quota within the first few weeks of April. As a result, the agency has resorted to a lottery to select petitions. The lottery procedure is as follows. First, they randomly select petitions for the advanced degree exemption, which has a cap of 20,000 per fiscal year. Then, all unselected advanced degree petitions become part

pay \$750 while employers with 26 or more full-time equivalent employees must pay \$1,500. An additional fraud prevention and detection fee of \$500 is required to request an initial H-1B status but not for an extension. Another \$2,000 must be submitted by a petitioner that employs 50 or more employees in the United States, if more than half of those employees have an H-1B or L-1 nonimmigrant status. Firms that wish to have their petition processed within 15 calendar days can choose to do so by paying an additional premium processing fee of \$1,225.

[2] The firm can always choose to sponsor an H-1B visa holder for permanent residency before the expiration of the H-1B visa.

of the random selection process for the 65,000 general limit.

In fiscal year (FY) 2015 (October 2014 to September 2015), the USCIS received about 172,500 H-1B petitions from April 1, 2014 to April 7, 2014 (five business days).[3] The USCIS then conducted a computer-generated lottery on April 10, 2014, to pick 85,000 petitions (65,000 general pool + 20,000 US advanced degree pool) to be processed for a work start date of October 1, 2015. This implies that only 49% of the petitions were approved overall. After 20,000 US advanced degree cases were approved by random selection, approximately 152,500 cases—including an unspecified number of US advanced degree cases that did not get chosen in the first lottery for the 20,000—were left in the general lottery for the 65,000 spots. This implies that a non-advanced degree H-1B petition faced a 43% probability of visa approval.

Table 2.1: H-1B by Nationality (Top 6): Fiscal Year 2010

Low- and Middle-Income Countries	Percentage	High-Income Countries	Percentage
India	49.97%	South Korea	3.11%
China (Mainland)	9.58%	United Kingdom	3.03%
Philippines	2.7%	Japan	2.07%
Mexico	2.12%	Taiwan	1.76%
Brazil	1.37%	France	1.66%
Turkey	1.21%	Germany	1.4%
All	78.94%	All	21.06%

Note: Raw data can be accessed at http://travel.state.gov/content/dam/visas/Statistics/Non-Immigrant-Statistics/NIVDetailTables/FY10NIVDetailTable.xls

Table 2.1 shows the proportion of the 117,409 H-1B visas that were approved in FY 2010 by nationality.[4] Following the classification provided by the World Bank, I first divide countries into two categories: high-income countries and low- and middle-income countries.

[3]This number excludes petitions requesting to work at universities, non-profit research facilities associated with universities, and government research facilities, since they do not count toward the cap.

[4]This number is above the 85,000 quota because it includes non-capped workers who work at universities, non-profit research facilities associated with universities, and government research facilities.

In each category, I list the top six countries by the number of H-1B approvals. Indians take up almost half of all of the approved H-1B skilled workers, while mainland Chinese come in second with 9.58%. Overall, the absolute majority of H-1B workers are from low- and middle-income countries.

Chapter 3

Model

In this chapter, I describe my baseline model. I first describe the model environment, followed by the decision problems that workers face. Then, I define the recursive stationary equilibrium of the model.

3.1 Model Environment

There are two countries N and S. In both countries there are two types of workers: skilled and unskilled. In order to focus on the impact of skilled labor immigration and return migration, only the skilled workers are able to move between the two countries; unskilled workers are stuck in the country they are born in and inelastically supply labor. Capital moves costlessly across the two countries. Time is discrete and all workers live for 40 periods, entering the work force at age 25 and exiting at age 65. There is no population growth and each age cohort consists of the two types of workers from the two countries. Let $\pi^x_{i,a}$ denote the fraction (mass) of age a workers of skill type $i \in \{h, l\}$ born in country $x \in \{N, S\}$. Then,

$$1 = \pi^N_{h,a} + \pi^N_{l,a} + \pi^S_{h,a} + \pi^S_{l,a}, \forall a \in \{25, \ldots, 64\} \tag{3.1}$$

where the mass of each age cohort is normalized to one. For simplicity, I assume that the workers who exit the work force after 40 periods are replaced by newborn workers of the same type and same nationality in the exact same proportion. In other words, $\pi^x_{i,a} = \pi^x_i, \forall a \in \{25, \ldots, 64\}$.

3.2 Skilled Worker's Problem

Skilled workers own capital k and human capital h, neither of which are location-specific. When skilled workers enter the work force, they have no physical capital, and they draw initial human capital h from an exogenous lognormal distribution $\ln \mathcal{N}\left(\mu^x_h, (\sigma^x_h)^2\right)$, with $x \in \{N, S\}$ that indicates the country of birth. In each subsequent period until they exit the work force, the workers must decide how much to consume and save, how to allocate time between work and human capital accumulation, and whether or not to switch location.

At the beginning of the period, the skilled worker must first decide whether or not to apply for migration to a different location. If workers start the period in their birth country but wish to immigrate to a foreign country, they must finance a fixed cost κ out of their capital holdings in order to enter the visa lottery.[1] I assume that markets are incomplete such that workers cannot borrow.[2] All skilled workers face the same endogenous probability p of "winning" the visa lottery and starting the next period in the foreign country. In contrast, workers that start the period in a foreign country but desire to return migrate back to their birth country face no fixed cost and no visa lottery. In other words, return migration entails no cost, and return migrants are assured of starting the next period in their birth country. Any skilled worker who previously obtained the visa but return migrated must apply for the visa again if the worker wants to move to the foreign country again. In both immigration and return migration, the decision to apply for relocation takes place in the beginning of

[1] This fixed cost can be thought of as the total cost necessary to obtain the H-1B visa, which would include all the fees mentioned in Chapter 2 and the cost of matching with a potential employer (not explicitly modeled in this paper). Since this paper is not explicitly modeling decision problems of individual firms, I assume that the employee bears these costs.

[2] I do this mainly to avoid the complication of thinking about the borrowing contract and its enforceability.

the period, but the actual relocation—if approved—takes place at the end of the period. Because k and h are not location-specific, the migrating worker brings his or her savings and human capital intact to the new location.

Once the migration decisions are made, skilled workers work, accumulate human capital, consume, and save. Workers have a constant relative risk aversion utility of the form

$$u(c) = \frac{c^{1-\epsilon}}{1-\epsilon}$$

Time in each period is normalized to one so that the worker allocates a fraction $s \in [0, 1]$ to human capital accumulation and $1 - s$ to working. Skilled workers residing in country x receive equilibrium effective wage rate w_h^x per unit of h. Human capital production takes the Ben-Porath class in which human capital in the next period, h', is a function of the human capital in the current period, h, and the amount of time allocated to human capital accumulation, s, as in the following equation:

$$h' = e^{z_x} \left(h + \theta(hs)^{\nu}\right) \tag{3.2}$$

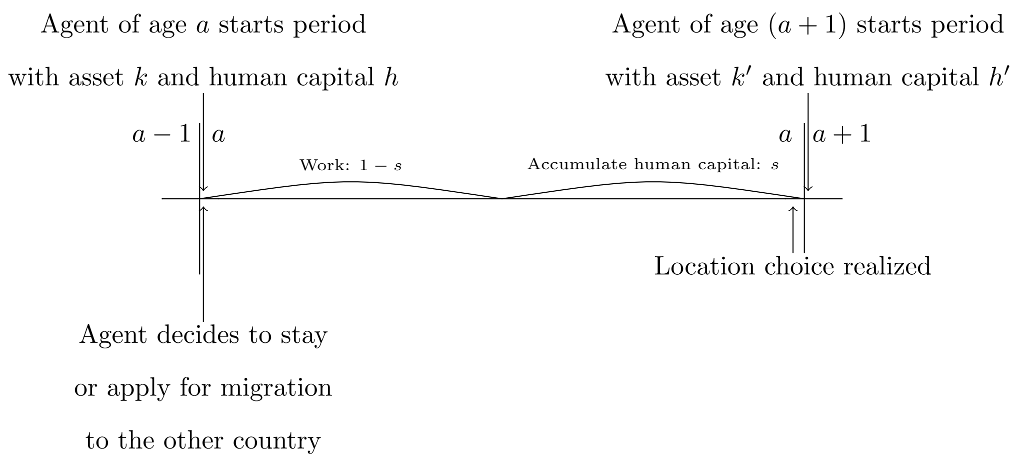
In Equation (3.2), θ and ν are the parameters of human capital technology in which θ controls the scale of production, and ν represents the elasticity of human capital with respect to input time. Human capital is subject to an idiosyncratic I.I.D. shock z each period where the distribution of z depends on the current location $x \in \{N, S\}$ in which human capital accumulation takes place. The $z_x \sim \mathcal{N}\left(\mu_x, \sigma_x^2\right)$, so e^{z_x} has a lognormal distribution.

Workers currently residing in a foreign country are subject to a utility flow cost[3] of $\frac{\lambda}{h}$ per period in which the cost is lower for skilled workers who possess a higher level of human capital. The utility flow cost takes this functional form for two reasons. First, utility flow cost has to be stochastic in order to generate the magnitude of return migration observed in the data. If the utility flow cost were a constant λ, the only uncertainty faced

[3]This cost is also widely known as the "psychic cost" in the immigration literature.

by skilled immigrants is their stock of human capital, and that alone generates too little return migration. Second, utility flow cost of $\frac{\lambda}{h}$, where λ is a constant parameter, is the most parsimonious way of having a stochastic utility flow cost, yet ensuring that return migrants are negatively selected with respect to their level of human capital.[4]

The following timeline sums up the sequence of events that occur within a period for a skilled worker.



There are four sets of Bellman equations for skilled workers: $V_a^{NN}(k,h)$, $V_a^{NS}(k,h)$, $V_a^{SN}(k,h)$, and $V_a^{SS}(k,h)$. The first superscript $x \in \{N,S\}$ indicates the country of birth. The second superscript $y \in \{N,S\}$ indicates the current country of residence. Subscript $a \in \{25,\ldots,64\}$ indicates age.

The Bellman equation for skilled workers who start the period in their birth country $(x = y)$ is as follows:

[4] Setting λ itself to be a stochastic state variable that transitions according to some probability is another way of adding uncertainty to utility cost, and I have solved that version of the model for robustness. However, I prefer my baseline model because the model with stochastic λ adds another state variable to my system, which not only lengthens the computation time needed to solve for the equilibrium of the model, but also complicates the calibration process by adding initial distributions and transition probabilities of λ that are difficult to pin down, even though the main results of the model do not change much.

$$V_a^{xx}(k,h) = \max \left\{ \begin{array}{l} \max\limits_{k' \geq 0, s \in [0,1]} u(c) + \beta \mathbb{E}\left[V_{a+1}^{xx}(k',h')\right], \\ \quad \text{s.t. } c+k'=k(1+r)+hw_h^x(1-s) \\ \max\limits_{k' \geq 0, s \in [0,1]} u(c) + \beta \left\{ p\mathbb{E}\left[V_{a+1}^{xy}(k',h')\right] + (1-p)\,\mathbb{E}\left[V_{a+1}^{xx}(k',h')\right] \right\} \\ \quad \text{s.t. } c+k'=(k-\kappa)(1+r)+hw_h^x(1-s) \end{array} \right\} \tag{3.3}$$

subject to

$$h' = e^{z_x}\left(h + \theta(hs)^{\nu}\right), \quad z_x \sim \mathcal{N}\left(\mu_x, \sigma_x^2\right)$$
$$V_{65}^{xx}(k,h) = \psi \cdot k, \quad k \geq 0$$
$$\forall x \in \{N, S\}$$

The worker must first decide whether to enter the visa lottery or not. The first line of the Bellman equation (3.3) represents the decision problem if the worker decides not to enter the visa lottery. In this case, the Bellman equation becomes a standard one in which the worker maximizes the current period utility and the next period's continuation value subject to the budget constraint and the human capital accumulation equation. There are two sources of income for the skilled worker: capital that they rent out to the capital market at the equilibrium rental rate r and the total labor income earned based on the market equilibrium effective wage (w_h^x), the amount of human capital the worker possesses (h), and the amount of time allocated to labor market ($1-s$). The expectation operator in the value function is due to the fact that human capital next period is subject to a shock z_x. The β is the discount factor. The terminal value at age 65 depends linearly on the amount of capital held at age 65. This can be interpreted as a bequest motive, in which the calibrated parameter ψ determines the marginal utility of capital at the exit age of 65.[5]

[5]Without bequest motive, workers decumulate capital as they near the age of 65 and display a hump-shaped wealth profile that does not match the data. Adding few periods of retirement in which workers simply consume without working does not alleviate this counterfactual wealth profile. Figure (B.2) shows

The second line of the Bellman equation (3.3) represents the decision problem if the worker decides to enter the visa lottery. Here, the worker faces an endogenous probability p of visa approval and moving to a foreign country and a probability $1-p$ of visa denial and staying in the birth country. In this case, the budget constraint is different because the worker must finance a fixed cost κ out of his or her capital holding in order to apply for the visa.

The Bellman equation for skilled workers who start the period in the foreign country $(x \neq y)$ are as follows:

$$V_a^{xy}(k,h) = \max \Bigg\{ \max_{k' \geq 0, s \in [0,1]} u(c) - \frac{\lambda}{h} + \beta \mathbb{E}\left[V_{a+1}^{xy}(k',h')\right], \\ \max_{k' \geq 0, s \in [0,1]} u(c) - \frac{\lambda}{h} + \beta \mathbb{E}\left[V_{a+1}^{xx}(k',h')\right] \Bigg\} \quad (3.4)$$

subject to

$$c + k' = k(1+r) + h w_h^y (1-s)$$
$$h' = e^{z_y}\left(h + \theta (hs)^{\nu}\right), \quad z_y \sim \mathcal{N}\left(\mu_y, \sigma_y^2\right)$$
$$V_{65}^{xy}(k,h) = \psi \cdot k, \quad k \geq 0$$

The worker must first decide whether to return migrate or not. The first line of the Bellman equation (3.4) represents the decision problem of a skilled worker born in $x \in \{N, S\}$ but residing in $y \neq x$, who decides to stay in the foreign country y. In this case, the equilibrium wages and human capital production function that the worker face are those of country y. Also, because work and human capital accumulation take place in the foreign country, the psychic cost $\frac{\lambda}{h}$ is incurred. The terminal value is the same as in the Bellman equation for skilled workers who start the period in the home country.

how well the model with bequest motive matches the data.

The second line of the Bellman equation (3.4) represents the decision problem of a skilled worker who decides to return migrate back to the birth country x. Because return migration requires no fixed cost, the budget constraint is exactly the same as if the worker stayed in the foreign country. Even though the return migration decision has been made at the beginning of the period, because the actual relocation takes place at the end of the period, the prices and human capital production technology the agent faces this period are those of the foreign country y.

3.3 Unskilled Worker's Problem

Unskilled workers own physical capital k but do not possess human capital and cannot relocate to a different country. Each period, they rent out their capital and supply inelastic labor to receive market wage w_l^x. Utility function $u(c)$ is the same as that of the skilled worker. Hence, the value function for the unskilled worker of country x with capital k and age a is as follows:

$$W_a^x(k) = \max_{k' \geq 0} \left\{ u(c) + \beta \left[W_{a+1}^x(k') \right] \right\} \tag{3.5}$$

$$\text{subject to} \qquad c + k' = k(1+r) + w_l^x$$

$$W_{65}^x(k) = \psi \cdot k, \quad k \geq 0$$

The terminal value is the same as that of a skilled worker and it can be interpreted as a bequest motive.

3.4 Aggregate Production

The production side is given by a two-level CES production function by Krusell et al. (2000) that is widely used in the literature. This production function allows for complementarity between skilled and unskilled labor. That is, for all $x \in \{N, S\}$, output Y_x is given by

$$Y_x = A_x F(K_x, H_x, L_x) = A_x \cdot \left(\alpha_x L_x^{\gamma} + (1-\alpha_x)\left[\rho_x K_x^{\eta} + (1-\rho_x) H_x^{\eta}\right]^{\frac{\gamma}{\eta}}\right)^{\frac{1}{\gamma}}$$

where A_x is the total factor productivity of country x. The K_x is the amount of aggregate capital used in country x's production, H_x is the aggregate amount of skilled labor provided by skilled workers residing in country x, and L_x is the aggregate unskilled labor of country x. The parameters $\gamma < 1$ and $\eta < 1$ determine the factor elasticities: $\frac{1}{1-\gamma}$ is the elasticity of substitution between capital and unskilled labor and between skilled labor and unskilled labor, whereas $\frac{1}{1-\eta}$ is the elasticity of substitution between capital and skilled labor. The parameters $0 < \alpha_x, \rho_x < 1$ control the factor shares.

3.5 Recursive Stationary Equilibrium

For the purpose of aggregation and defining the recursive stationary equilibrium of the model, it is necessary to describe the position of individuals across states. Let $\Omega_a^{xx}(k,h)$ represent the mass of skilled workers of age a residing in their birth country x with capital k and human capital h. Let $\Omega_a^{xy}(k,h)$ represent the mass of skilled workers of age a residing in foreign country y with capital k and human capital h. Also, let $\widetilde{\Omega}_a^{x}(k)$ represent the mass of unskilled workers of age a residing in country x with capital k.

The recursive stationary equilibrium of the model consists of the following.

1. **Skilled workers' optimization** : Given the probability p of visa approval and prices $\left(r, w_h^N, w_h^S\right)$, the value functions $V_a^{NN}(k,h)$, $V_a^{NS}(k,h)$, $V_a^{SN}(k,h)$, and $V_a^{SS}(k,h)$ are solutions to the skilled worker's optimization problem described in equations (3.3) and (3.4). $g_k(\cdot)$, $g_c(\cdot)$, $g_s(\cdot)$, and $g_m(\cdot)$ are the associated optimal decision rules with respect to capital in the next period, consumption, time allocated to human capital accumulation, and movement (location choice).

2. **Unskilled workers' optimization** : Given prices $\left(r, w_l^N, w_l^S\right)$, the value functions $W_a^N(k)$ and $W_a^S(k)$ are solutions to the unskilled worker's optimization problem described in equation (3.5). $\widetilde{g}_k(\cdot)$ and $\widetilde{g}_c(\cdot)$ are the associated optimal decision rules with respect to asset and consumption.

3. **Firm's optimization** : Prices $\left(r, w_h^N, w_h^S, w_l^N, w_l^S\right)$ satisfy the following marginal conditions for all $x \in \{N, S\}$.[6]

$$\begin{aligned} r &= A_x F_K(K_x, H_x, L_x) - \delta \\ w_h^x &= A_x F_H(K_x, H_x, L_x) \\ w_l^x &= A_x F_L(K_x, H_x, L_x) \end{aligned}$$

4. **Consistency (1)** : $\Omega_a^{xx}(k, h)$, $\Omega_a^{xy}(k, h)$, and $\widetilde{\Omega}_a^x(k)$ are respectively the stationary distribution of the skilled workers in their home country, skilled workers in the foreign country, and unskilled workers.

5. **Consistency (2)** : The probability of winning the visa lottery p is consistent with the aggregation of individual optimal movement choices, such that

$$p = \frac{\text{quota}}{\sum_{a=25}^{64} \int_{(k,h)\in(\mathbb{R}_+ \times \mathbb{R}_+)} g_{m,a}^{SS}(k, h)\, d\Omega_a^{SS}(k, h)}$$

where $g_{m,a}^{SS}(k, h)$ is the optimal relocation decision rule of a S-born skilled worker who started the period in his or her home country S with capital k, human capital h, and age a. The $g_{m,a}^{SS}(k, h) = 0$ indicates the said worker is staying in home country S, whereas $g_{m,a}^{SS}(k, h) = 1$ indicates that the worker will be entering the H-1B visa lottery this period.[7]

[6] δ is the depreciation rate of capital.

[7] There is also the probability of winning the visa lottery for N-born skilled workers that wish to move to S. However, I will omit the details of how that probability is calculated since the model will be calibrated such that no skilled workers born in N will ever move to S in equilibrium.

6. **Aggregation** : Capital supplied by heterogeneous individuals are aggregated appropriately as follows:

$$
\begin{aligned}
K_N + K_S = \sum_{x\in\{N,S\}} \sum_{a=25}^{64} \Bigg[\int_{(k,h)\in(\mathbb{R}_+\times\mathbb{R}_+)} & g_{k,a}^{xx}(k,h)\, d\Omega_a^{xx}(k,h) \\
& + \sum_{y\neq x} \int_{(k,h)\in(\mathbb{R}_+\times\mathbb{R}_+)} g_{k,a}^{xy}(k,h)\, d\Omega_a^{xy}(k,h) \\
& + \int_{k\in\mathbb{R}_+} \widetilde{g}_{k,a}^{x}(k)\, d\widetilde{\Omega}_a^{x}(k) \Bigg]
\end{aligned}
$$

Skilled and unskilled labor supplied by heterogeneous workers are aggregated appropriately as follows for all $x \in \{N, S\}$:

$$
\begin{aligned}
H_x &= \sum_{a=25}^{64} \int_{(k,h)\in(\mathbb{R}_+\times\mathbb{R}_+)} h\left(1 - g_{s,a}^{xx}(k,h)\right) d\Omega_a^{xx}(k,h) \\
&\quad + \sum_{y\neq x} \sum_{a=25}^{64} \int_{(k,h)\in(\mathbb{R}_+\times\mathbb{R}_+)} h\left(1 - g_{s,a}^{xy}(k,h)\right) d\Omega_a^{yx}(k,h) \\
L_x &= \sum_{a=25}^{64} \int d\widetilde{\Omega}_a^{x}(k) = \sum_{a=25}^{64} \pi_{l,a}^{x}
\end{aligned}
$$

Chapter 4

Computation of Recursive Competitive Equilibrium

In this chapter, I describe the numerical solution of the recursive competitive equilibrium of the model, using a two-country version of the computational algorithm in the Aiyagari (1994) heterogeneous agent model. Details of this algorithm can be seen in Appendix A.1. Here, I first discuss the calibration of the model and then show various moments generated at the steady state.

4.1 Calibration

Because this paper focuses on US skilled immigration policy, I calibrate the model such that country N represents the United States, and S represents a block of low- and middle-income countries as defined by the World Bank. There are two reasons why S represents low- and middle-income countries as a block and excludes non-US high-income countries, such as those in the European Union. First, the magnitude of return migration is different by an order of magnitude based on whether the skilled worker is from a developing or developed country. Using the 2000 Census 5% sample and the 2005 American Community Survey (ACS) 1% sample from Ruggles et al. (2015), I calculate a five-year re-emigration rate for

Table 4.1: Parameter Values: Baseline Model

Category	Symbol	Parameter Value
Preferences	(β, ϵ, ψ)	$(0.98, 2, 0.05)$
Aggregate Production Technology	$\left(\frac{1}{1-\gamma}, \frac{1}{1-\eta}, \delta\right)$	$(1.669, 0.669, 0.13)$
	$(\alpha_N, \alpha_S, \rho_N, \rho_S)$	$(0.675, 0.74, 0.36, 0.41)$
Migration Costs	(κ, λ)	$(1.7, 28)$
Total Factor Productivity	(A_N, A_S)	$(1, 0.49)$
Demographics	$\left(\pi_h^N, \pi_h^S, \pi_l^N, \pi_l^S\right)$	$(0.015, 0.051, 0.035, 0.899)$
Human Capital Technology	$h' = e^{z_x}\left(h + \theta(hs)^{\nu}\right)$	$(\theta, \nu) = (0.4, 0.7)$
Human Capital Shocks	$z_N \sim \mathcal{N}\left(\mu_N, \sigma_N^2\right)$	$(\mu_N, \sigma_N) = (-0.029, 0.111)$
	$z_S \sim \mathcal{N}\left(\mu_S, \sigma_S^2\right)$	$(\mu_S, \sigma_S) = (-0.037, 0.112)$
Initial Condition	$h_{25}^N \sim \ln\mathcal{N}\left(\mu_h^N, \left(\sigma_h^N\right)^2\right)$	$\left(\mu_h^N, \sigma_h^N\right) = (4.66, 0.213)$
	$h_{25}^S \sim \ln\mathcal{N}\left(\mu_h^S, \left(\sigma_h^S\right)^2\right)$	$\left(\mu_h^S, \sigma_h^S\right) = (4.18, 0.328)$

skilled workers aged 25 to 60 who entered the United States in 1999. Whereas only about 5% of skilled immigrant workers from low- and middle-income countries left the United States after five years, about 50% of the skilled immigrants from high-income countries (as classified by the World Bank) left the United States after the same number of years. Second, as described in Chapter 2, close to 80% of the newly issued H-1B are for skilled workers from low- and middle-income countries.

Table 4.1 summarizes the parameter value choices for the baseline model. For total factor productivity (TFP), I use the latest release of the Penn World Table (Feenstra et al. (2015)), which provides cross-country TFP values with the US value normalized to one. I set A_S to be 0.49, the population-weighted average of the 2010 TFP values of all low-income and middle-income countries. For demographic parameters, I use the 2010 US Census and World Bank data to estimate that about 5% of the world population resides in the United States.[1] 30.94% of the US population and 5.36% of the total population in all low- and middle-

[1] This calculation excludes the non-US high-income countries, since S represents a block of low- and middle-income countries as defined by the World Bank instead of the "rest of the world."

income countries are estimated to be skilled workers, which I define as those with a college degree. These values are used to calculate the proportion of skilled and unskilled workers from both countries in each age cohort. Specifically, $\pi_h^N = 0.015$, $\pi_l^N = 0.035$, $\pi_h^S = 0.051$, and $\pi_l^S = 0.899$. The period utility function has a CRRA property with parameter $\epsilon = 2$, which is well within the range of estimates used in the literature.

For the US human capital shock and technology parameters, I use the values estimated in Huggett, Ventura, and Yaron (2011).[2] Specifically, I set ν—the elasticity of human capital produced with respect to time—to 0.7, and θ—the scale parameter of human capital production—to 0.4. The initial distribution of human capital in the United States, $h_{25}^N \sim \ln \mathcal{N}\left(\mu_h^N, \left(\sigma_h^N\right)^2\right)$, from which newborn skilled workers draw their initial human capital, is calibrated to $\mu_h^N = 4.66$ and $\sigma_h^N = 0.213$. The blue curve in Figure (B.1) shows the initial distribution of human capital for US skilled workers. The idiosyncratic shock $z_N \sim \mathcal{N}\left(\mu_N, \sigma_N^2\right)$ to human capital in the United States is calibrated to $\mu_N = -0.029$ and $\sigma_N = 0.111$. Here, $\mu_N < 0$ embeds human capital depreciation and signifies that skilled workers must constantly accumulate human capital in order to stave off automatic decumulation.

I assume that the two human capital production technology parameters—θ and ν—take the same values in S as in N. To calibrate $h_{25}^S \sim \ln \mathcal{N}\left(\mu_h^S, \left(\sigma_h^S\right)^2\right)$, the initial distribution of human capital for skilled workers in country S, I use the index of human capital per person calculated in Feenstra et al. (2015): they find that the US had a human capital index of 3.62, whereas the population-weighted human capital index for all low-income and middle-income countries was 2.32 in 2010. That means the average person in the US has 56.38% higher level of human capital than his counterpart from S, so I set $\mu_h^S = 4.18$ and $\sigma_h^S = 0.328$ such that the mean of the lognormal distribution is 56.38% higher in the US than in S while

[2] They view lifetime inequality through the lens of a risky human capital model, very similar to this paper. They posit that two forces account for the increase in earnings dispersion: the initial condition and idiosyncratic human capital shocks received over the lifetime. They use the US earnings data to estimate both the shock process and the parameters governing the initial distribution of human capital.

keeping the variance of the lognormal distribution the same.[3] For the idiosyncratic shock $z_S \sim \mathcal{N}(\mu_S, \sigma_S^2)$, I pick the two parameters such that the age-earnings profile in S obtains a hump-shaped profile while maintaining the same level of variance between the lognormal distributions of e^{z_N} and e^{z_S}.

That leaves nine parameters yet to be calibrated: the annual discount factor β; factor share parameters in the aggregate production function α_N, α_S, ρ_N, and ρ_S; capital depreciation rate δ; fixed migration cost κ; the numerator of the period utility flow cost λ; and the parameter ψ from the bequest function. I calibrate these parameters such that the equilibrium properties of the model best match the nine target moments from the data: the rental rate of capital r, the aggregate capital to output ratio in the United States, the labor share in production, the labor earnings ratio between skilled and unskilled workers, the ratio of effective wages of skilled labor in the two countries $\frac{w_h^N}{w_h^S}$, the probability p of H-1B visa acceptance rate, the ratio of the flow of skilled workers from S to the total stock of US-born skilled workers, the return migration rate of S-born skilled workers within five years of arrival, and the average wealth profile of US workers.

For the first three target moments—r, $\frac{K_N}{Y_N}$, and labor share in the US—I use the standard estimates used in the literature of 0.042, 2, and 0.65, respectively. I use Goldin and Katz (2007) to target the labor earnings ratio between skilled and unskilled workers in the United States of 1.65. For the ratio of the effective wages of skilled labor in the two countries, I use the findings in Lagakos et al. (2015) that one log point difference in GDP per capita across two countries is associated with 20% higher return to experience. Because human capital is a proxy for experience in my model, I calculate the log of population-weighted GDP per capita for S, and compare that to the log of US GDP per capita to obtain that the US skilled wage rate should be about 55% higher than that of country S. For p, I use the value estimated in Chapter 2 of 0.49. I use the population data from the US Census and the visa data from the US State Department to calculate the ratio of the flow of H-1B skilled

[3] I do not have sufficient data to estimate the variance of the initial distribution in S. Hence, I focus on the cross-country differences in the mean of human capital.

workers from S to the total stock of native US skilled workers to obtain the value of 0.00174. For the return migration rate within five years of arrival, I use the value of 5% calculated earlier in this chapter. For the US profile of average wealth by age, I compare the wealth data published by the US Census—which uses Survey of Income and Program Participation to calculate various wealth moments by 10-year age bins[4]—against the model-generated age profile of average capital. The resulting calibrated parameter values are in Table 4.1, and the comparison of the nine target moments against the values calculated from the model can be seen in Table 4.2.

Table 4.2: Target Moments

Moment	Target	Model
r	0.042	0.044
$\frac{K_N}{Y_N}$	2	2.08
labor share in US	0.65	0.64
average labor earnings ratio (skilled vs. unskilled) in US	1.65	1.62
$\frac{w_h^N}{w_h^S}$	1.55	1.50
p	0.49	0.49
$\frac{\text{flow of skilled immigrants from } S}{\text{native skilled workers of } N}$	0.00174	0.00177
return migration rate after 5 years	0.05	0.052
average US wealth profile	Figure (B.2)	

4.2 Moments from Steady State

Figure (B.3) shows the average profiles of k, h, s, and the yearly labor earnings of skilled workers by their nationality. The model succeeds in producing the hump-shaped mean earnings profile that is well documented in the literature. Earnings at the early stage of life cycle are low because the initial human capital is low. Earnings rise as human capital

[4] Data can be accessed at http://www.census.gov/people/wealth/data/dtables.html

accumulates and as a greater fraction of time is devoted to market work. Earnings fall later in life because human capital depreciates and little time is put into producing new human capital. Therefore, the rise and fall of human capital drive the hump-shaped earnings profile in this model. The average time spent on human capital accumulation is initially higher in the US than in S, but that reverses at about age 40. Because idiosyncratic shocks to human capital and the initial human capital distributions are relatively more favorable in N, human capital and period earnings are higher in N than S. That also leads to higher average capital holdings in the US.

At the calibrated steady state, no skilled worker born in the US ever relocates to S. This is not surprising given that one of the targeted moments in the calibration showed that the skilled worker's effective wage rate w_h^x is about 50% higher in N than in S. Naturally, this creates a large incentive for skilled workers born in S to relocate to N. However, there are three forces that prevent a mass exodus. First, because the workers must pay the fixed migration cost κ in order to be eligible for the visa lottery, workers have to accumulate assets. Because newborn workers enter the model with zero physical capital, it takes a few years for them to build up enough savings to cover the fixed cost, whose calibrated value of $\kappa = 1.7$ is higher than the average annual earnings of all skilled workers born in S, as can be seen in the bottom right plot of Figure (B.3).[5] Further, if workers receive a series of bad human capital shocks, they have to spend more time accumulating human capital than working, which even further delays the relocation. Second, because the visa is awarded based on a lottery with endogenous success probability p, skilled workers who accumulate enough capital to cover the fixed cost κ still face a non-negligible probability of visa rejection. Third, even those who succeed in moving to N can be hit with a bad idiosyncratic human capital shock that may lead them to return migrate back to S.

Figure (B.4) shows the model-generated immigration and return migration pattern for

[5]The calibrated value of fixed cost κ may seem high, but this value certainly is not implausible if all the fees associated with applying for the visa and the implicit matching cost with a potential employer—which has not been explicitly modeled in this paper—are considered together.

S-born skilled workers. The top plot of Figure (B.4) compares the model-generated age distribution of skilled immigrant workers residing in N at the steady state to the data.[6] It shows that the model does a decent job of matching the age distribution of immigrants for the older cohorts but not so much for the younger. There are two reasons for this. First, the OECD data used to generate the black curve in the figure includes all college-educated individuals who were born outside of the US. That means the figure includes graduate students on student visas, who tend to be in their late 20s and early 30s. Second, because I have assumed that the skilled workers enter the model with zero capital for simplicity, those who would like to immigrate immediately are forced to delay their relocation until they have enough asset to pay κ. This leads to the model understating the mass of skilled immigrants in their late 20s and early 30s while overstating those in late 30s and early 40s. However, the model does a decent job overall of producing the age pattern in which the mass of skilled immigrants initially increases and then starts to decrease with the peak near the late 30s and early 40s.

The bottom plot of Figure (B.4) shows the flow of immigrant workers and return migrants generated from the model. The blue line represents the flow of skilled immigrants to N, and the red line represents the flow of skilled return migrants going back to S. The figure shows that much of the immigration takes place earlier in the life cycle as up to 0.2% of the S-born skilled workers in their early 30s decide to relocate to N. In the first half of the life cycle, return migration is close to nonexistent but as the workers age, return migration peaks—especially in their mid-40s—while immigration flow gradually decreases. The reason the return migration is relatively scarce in the early life cycle is because the immigrants that are hit with an adverse shock to their human capital tend to wait in case the human capital shock in the next period is favorable. However, as the immigrants age and start to invest

[6] I use the OECD's Database on Immigrants in OECD Countries (DIOC) and sum across all skilled immigrants from low- and middle-income countries living in the United States by age group. Since the model does not have population growth or death, I adjust these sums by the annual population growth rate of 1.4% and the death probabilities published in the US Social Security Administration's Actuarial Life Table.

less of their time on human capital production, they do not have as strong an incentive to stay and incur the per-period psychic cost.

Figure (B.5) shows the selection of immigrant workers and return migrants at the steady state. The top plot of Figure (B.5) shows the selection among skilled foreign workers who apply for the H-1B visa. It compares the cumulative distribution functions of human capital of skilled workers who do not apply for the visa (in red) and those who do apply (in blue). The dotted lines represent the mean of the distribution for each group. In equilibrium, the distribution function of those that apply for the visa stochastically dominates the distribution of those who do not apply, which means that the model succeeds in obtaining the positive selection of skilled immigrants, consistent with the selective process through which the US firms select their H-1B candidates. This result is driven mainly by the fact that workers start their life cycle with zero capital and that they have to pay fixed cost κ in order to enter the visa lottery. The higher the human capital that a worker possesses, the easier it is for the worker to pay κ out of his or her savings.

The bottom plot of (B.5) shows the selection among skilled workers who return migrate back to S. It compares the cumulative distribution function of human capital of the return migrants (in red) to that of immigrants who stay in N (in blue) at the steady state. The dotted line represents the mean of the distribution for each group. This figure shows that the equilibrium of the model displays positive (negative) selection for skilled immigrants who stay (return migrate) in the US. This result shows that the model behaves in accordance with the finding in literature that return migrants are negatively selected if the immigrants were positively selected in the first place.[7]

[7]Refer to Borjas and Bratsberg (1996), Lubotsky (2007), and Grogger and Hanson (2015).

Chapter 5

Policy Experiment

In this chapter, I conduct a counterfactual policy experiment of reducing the immigration barrier for skilled foreign workers. Specifically, I double the H-1B visa cap permanently. I solve for the new steady state and then compute the transition path from the old steady state to the new. Once the transition path is calculated, the welfare changes along the transition path—in addition to fluctuations in prices and aggregate variables—are calculated. I first compare the new steady state against the old and then describe the transition path and changes in welfare. Details of the computation algorithm can be seen in Appendix A.2.

5.1 Comparison of Steady States

Table 5.1 shows how the equilibrium prices and aggregate quantities differ across the two steady states. First, the table shows that the probability of visa approval has increased in the long-run by only 11 percentage points, which indicates that the visa bottleneck is still significant. This is equivalent to a 63% increase in the number of visa applicants, and Figure (B.6) shows how the human capital cumulative distribution function for the visa applicants at the new steady state is stochastically dominated by the distribution at the old steady state, which implies that the majority of the increase in visa applicants are by less talented skilled workers who would not have applied for the visa at the old steady state. As a result,

Table 5.1: Comparison of Steady States

Equilibrium Prices	Change	Equilibrium Aggregates	Change
r	0.01% ↑	$K_N + K_S$	0.36% ↑
w_h^N	0.79% ↓	H_N	1.87% ↑
w_h^S	0.47% ↑	H_S	0.6% ↓
w_l^N	0.59% ↑	output/capita in N	0.79% ↑
w_l^S	0.07% ↓	output/capita in S	0.14% ↓
p	Increases from 0.49 to 0.6		

the average human capital possessed by the visa applicant is much lower in the new steady state. Figure (B.7) compares the population mass of skilled immigrants at the two steady states by their level of human capital. The mass of workers near the peak of the distribution has doubled. However, the magnitude of the changes away from the peak is asymmetric: the mass barely increases the further one goes to the right of the distribution, unlike the left side of the peak.

Figure (B.8) shows the difference in the age distribution of immigrants across the two steady states. The majority of immigrants are still in their late 30s and early 40s, but the stock of immigrants in this age range has more than doubled. In contrast, the stock of immigrants in the older age cohort changes very little, which means the return migration rate for the older cohorts have increased a lot also. This is because the decrease in the labor mobility barrier led to an increased inflow of less talented foreigners who are much more likely to return migrate than their more talented peers.

Of all the equilibrium prices and aggregate variables, H_N—the aggregate human capital supplied by skilled workers residing in N—increases the most by 1.87%. This change is not surprising given the higher inflow of immigrants at the new steady state. However, the magnitude of the decrease in H_S is only about a third of the magnitude of the increase in H_N. There are two reasons for such asymmetry. First, the calibration for demographics was such that the mass of skilled workers born in S is more than double the mass of native US skilled workers. Second, return migration mitigates the loss of human capital from immigration since return migrants possess a higher level of human capital on average than those who never left S. In other words, even though S is experiencing a "brain drain" in which the country is losing its skilled workers, there is some "brain gain" that lessens the blow.

Basic economic theory of supply and demand predicts that an increase in labor supply will decrease wages.[1] In the current calibration, the US skilled wage decreases in the long-run by 0.79% across the two steady states. The elasticity implied by this result is well within the range of elasticities found in the empirical literature as documented in Kerr and Kerr (2011). Due to the increased outflow of skilled workers, the skilled wage in S increases by 0.47%. The complementarity of skilled and unskilled labor in the aggregate production function signifies that the increase in the inflow (outflow) of skilled workers leads to an increase (decrease) in the unskilled worker's wage. Hence, we observe that w_l^N increases by 0.59% whereas w_l^S decreases by 0.07%. I do not observe significant changes in the total capital in the economy and the price of capital across the two steady states.[2]

Just as the magnitude of the impact on human capital is asymmetric across the two countries, the magnitude of the impact on output is also quite asymmetric. US aggregate output increases by 1.52% in the long-run while aggregate output in S decreases by only 0.17%. Because the doubling of the inflow of immigrants causes the equilibrium population to increase in the US and decrease in S, the magnitude of the changes in output per capita are dampened relative to the changes in output. The output per capita of N increases by 0.79%, while it decreases by merely 0.14% in S in the long-run.

[1]This argument assumes that total factor productivity is not affected by the increased inflow of skilled immigrants. The results would be different if I assumed that skilled labor had a direct impact on productivity.

[2]However, there will be interesting changes to capital along the transition path, which will have significant impact on welfare evaluations.

5.2 Transition Path

Figures (B.9) and (B.10) show how various moments generated by the model behave along the transition path. Figure (B.9) shows the flow and stock of immigrant workers and return migrants. The top plot shows the flow of skilled workers born in S immigrating to N, the second plot shows the flow of skilled workers who return migrate back to S from N, and the last plot shows the stock of skilled immigrant workers in the US. In each plot, the first data point (transition period = 0) represents the value at the old steady state that is normalized to 100, and the last data point represents the new steady state. Transition period 1 in the x-axis represents the time at which the H-1B visa quota was doubled permanently. Figure (B.9) shows that the flow of immigrants immediately doubles at the time of the policy change and then stays that way for the rest of the transition path. However, the flow of return migrants stays around its old steady state level for six to seven years and then starts to pick up dramatically after that. In other words, the increase in the quota has an immediate impact on immigration, but a delayed impact on return migration. As a result, the total stock of skilled immigrant workers present in the US increases at an almost linear pace for the first ten years or so, and then gradually tapers off to its new steady state level. At the end of the transition path, the flow of skilled immigrants has doubled, the flow of return migrants has increased by 112%, and the the total stock of immigrants has increased by 94% compared to the old steady state.

Top left plot of Figure (B.10) shows how the probability p changes over the transition path. At the time of the policy change (transition period = 1), p increases modestly from 0.49 to 0.57, which means that the number of H-1B visa applicants immediately increases by 72%. This implies that many less talented foreign workers who would not have applied for the visa at the old steady state immediately apply for the visa en masse such that doubling of the quota does little to alleviate the friction caused by the H-1B visa lottery. After the initial spike, p gradually converges toward its new steady state level of 0.6, which implies that the number of H-1B visa applicants increases by 63% in the long-run. About two-thirds of the total change in p across the two steady states occurs in the initial period of the policy change and then the rest is spread out over the transition path.

Top right plot of Figure (B.10) shows how the total capital in the economy $K^* = K_N + K_S$ and the price of capital r change along the transition path. At the time of the policy change, K^* falls by 0.29% because of the sudden 72% increase in the number of skilled workers paying κ to apply for the visa. This causes r to increase by 1.2% immediately after the policy change. Eventually, K^* increases and r decreases back to their steady state levels, and in the long-run, there is a moderate increase in K^* but almost no change in r when compared to the old steady state.

Bottom two plots of Figure (B.10) show how the aggregate human capital supplied, the effective skilled wage, and the output per capita in each of the two countries change over the transition path.[3] Because the increase (decrease) in human capital of country N (S) depends solely on the inflow (outflow) of immigrants, the aggregate human capital's speed of convergence matches that of the total stock of immigrants. The effective wage of skilled labor w_h moves in the opposite direction as the aggregate human capital. US output per capita steadily and gradually increases to its new higher steady state, whereas the output per capita for S reaches its new lower steady state within 8 years after the policy change.

5.3 Welfare

For welfare calculations along the transition path, I use the following procedure. For each unique type of skilled worker and unskilled worker—where uniqueness is defined by their state variables—I calculate the one-time percentage increase in consumption the worker has to be compensated for at the old steady state in order to have the same level of total lifetime discounted utility along the transition path. For example, let $\omega_a^{xx}(k, h)$ denote the one-time welfare change for a skilled worker of age a in his or her home country $x \in \{N, S\}$ with asset

[3]Some of the figures display a sudden spike or drop at the end of the transition path. This is most likely caused by a discretization error associated with representing a function of a continuous variable by a finite number of evaluations in the computer.

k and human capital h. Also, let $\hat{V}_a^{xx}(k,h;t)$ denote the value function at transition period $t \in \{0,1,\ldots,\mathcal{T}\}$ of the aforementioned skilled worker, where $\hat{V}_a^{xx}(k,h;0)$, $\hat{V}_a^{xx}(k,h;1)$, and $\hat{V}_a^{xx}(k,h;\mathcal{T})$ are respectively the total discounted utility of this worker at the old steady state ($t=0$), the period in which the visa quota was doubled ($t=1$), and the new steady state ($t=\mathcal{T}$).[4] Suppose that the optimal relocation decision rule for this skilled worker at the old steady state is to not apply for the visa, which means $g_{m,a}^{xx}(k,h)=0$. Then, $\omega_a^{xx}(k,h)$ is calculated using the following equation.

$$\begin{aligned} V_a^{xx}(k,h;\omega) &\equiv u\left((1+\omega_a^{xx}(k,h))\cdot g_{c,a}^{xx}(k,h)\right)+\beta\mathbb{E}\left[V_{a+1}^{xx}\left(g_{k,a}^{xx}(k,h),g_{h,a}^{xx}(k,h)\right)\right] \\ &= \hat{V}_a^{xx}(k,h;t=1) \end{aligned} \tag{5.1}$$

where $V_a^{xx}(k,h;\omega)$ is the total discounted utility for this skilled worker at the old steady state after adjusting its current period's consumption by a factor of $(1+\omega_a^{xx}(k,h))$; $g_{c,a}^{xx}(k,h)$ and $g_{k,a}^{xx}(k,h)$ are respectively the optimal consumption and optimal capital next period; and $g_{h,a}^{xx}(k,h) = e^{z_x}\left(h+\theta\left(h\cdot g_{s,a}^{xx}(k,h)\right)^{\nu}\right)$ is the realized value of human capital next period given shock z_x and optimal time allocation $g_{s,a}^{xx}(k,h)$.[5] The welfare change for skilled workers in foreign country $\omega_a^{xy}(k,h)$ and for unskilled workers $\widetilde{\omega}_a^{x}(k)$ are calculated in a similar way.

Figure (B.11) shows the welfare calculation for three different age cohorts of US skilled workers: top, middle, and bottom plots represent the cohort of US skilled workers who were 35 $\left(\omega_{35}^{NN}(k,h)\right)$, 45 $\left(\omega_{45}^{NN}(k,h)\right)$, and 55 $\left(\omega_{55}^{NN}(k,h)\right)$, respectively, at the time of the policy change. In each of the three plots, x-axis denotes the percentile of human capital, and y-axis denotes the percentile of capital (or assets) within that age cohort of US skilled workers. For

[4] This uses the same notation as in the Appendix A.2, which describes the algorithm used to calculate the entire transition path.

[5] If the optimal relocation decision rule for this skilled worker at the old steady state is to apply for the visa such that $g_{m,a}^{xx}(k,h)=1$, then Equation (5.1) would have $\beta\left\{p\mathbb{E}\left[V_{a+1}^{xy}\left(g_{k,a}^{xx}(k,h),g_{h,a}^{xx}(k,h)\right)\right]+(1-p)\mathbb{E}\left[V_{a+1}^{xx}\left(g_{k,a}^{xx}(k,h),g_{h,a}^{xx}(k,h)\right)\right]\right\}$ instead of the $\beta\mathbb{E}\left[V_{a+1}^{xx}\left(g_{k,a}^{xx}(k,h),g_{h,a}^{xx}(k,h)\right)\right]$ term.

example, the top plot of Figure (B.11) shows that a US skilled worker with median levels of both human capital and capital within his cohort experiences about 11 percent decrease in lifetime utility due to doubling of the skilled visa quota. It also shows that all the US skilled workers who were 35 years old at the time of policy change lose welfare. However, the magnitude of the loss depends on how much capital and human capital a worker owns. Because the US skilled wage decreases but the price of capital increases at the time of policy change, as seen in Figure (B.10) earlier, US skilled workers who possess relatively more assets compared to their peers experience an increase in capital income that somewhat offsets the decrease in labor income. This effect is more pronounced for older cohorts as can be seen in the middle and the bottom plots of Figure (B.11) since older cohorts have accumulated greater amount of assets than the younger cohorts. In fact, a significant portion of older US skilled workers actually experience positive welfare gains from this policy.

Figure (B.12) shows the welfare calculation for S-born skilled workers whose age were 45 at the time of the policy change: the top plot is for those who were in S at the time of the change in policy $\left(\omega_{45}^{SS}(k,h)\right)$, while the bottom plot is for those who were in N $\left(\omega_{45}^{SN}(k,h)\right)$. As in Figure (B.11), x-axis denotes the percentile of human capital, and y-axis denotes the percentile of capital (or assets) within that age cohort. In the top plot, we see that all the skilled workers who were in S has positive welfare gain of about 6%, as both w_H^S and r increased. The changes in welfare are much more heterogeneous for the skilled immigrants, as can be observed in the bottom plot of Figure (B.12). Those who are less talented gain from this policy change, whereas immigrants who are relatively more talented lose. This is caused by the fact that return migrants are negatively selected as observed in the bottom plot of Figure (B.5). Less talented immigrants are very likely to return migrate soon, so they are affected in the same way as those who were in S at the time of the policy change. Talented immigrants are more likely to stay in N, so they are affected by the policy in a manner similar to the US skilled workers. In other words, $\omega_{45}^{SN}(k,h)$ is very similar to $\omega_{45}^{SS}(k,h)$ for low values of h and $\omega_{45}^{NN}(k,h)$ for high values of h.

This heterogeneity among skilled immigrants can be better observed in Figure (B.13), which contains three plots representing welfare changes of US skilled workers, S-born skilled workers in S, and S-born skilled immigrants, respectively, who were 45 years old at the time of the policy change. In contrast to earlier figures, the x-axis and the y-axis are no longer in terms of percentiles, but are in absolute model units. In the bottom plot, welfare changes of skilled immigrants whose level of human capital is less than 200 units resemble those of S-born skilled workers in S, whereas welfare changes of skilled immigrants whose level of human capital is more than 200 units resemble those of US skilled workers. In other words, the negative selection of return migrants creates the heterogeneity in welfare changes observed among skilled immigrants.

The top plot of Figure (B.14) shows skilled and unskilled workers' average changes in welfare at the time of the policy change for each age cohort. The blue line plots the average change in welfare for US skilled workers of age cohort a $\left(\int_{k,h} \omega_a^{NN}(k,h)\, d\Omega_a^{NN}(k,h)\right)$. The red line represents S-born skilled workers who were in S at the time of the policy change $\left(\int_{k,h} \omega_a^{SS}(k,h)\, d\Omega_a^{SS}(k,h)\right)$, while the green line represents S-born skilled immigrant workers $\left(\int_{k,h} \omega_a^{SN}(k,h)\, d\Omega_a^{SN}(k,h)\right)$. The black and purple respectively represent the average change in welfare for US unskilled workers $\left(\int_{k} \widetilde{\omega}_a^{N}(k)\, d\widetilde{\Omega}_a^{x}(k)\right)$ and unskilled workers in country S $\left(\int_{k} \widetilde{\omega}_a^{S}(k)\, d\widetilde{\Omega}_a^{x}(k)\right)$.

Three results stand out. First, even though welfare changes are quite heterogeneous even within the same cohort of a given skill type, the average change in welfare is relatively straightforward. The average US unskilled workers and average S-born skilled workers who were in S at the time of policy change gain from doubling the visa quota, while the opposite is true for US native skilled workers and unskilled workers in S. This is not surprising given that the change in policy immediately increases the inflow of skilled workers to N, which causes changes in wages that benefit the unskilled workers in N and the skilled workers living in S.

Second, the average change in welfare for skilled immigrants show a more complicated pattern. Because the bulk of return migration occurs in the later half of the lifecycle, the younger immigrants are mostly composed of those with a high level of human capital who tend to stay in the US, whereas the older immigrants have greater proportion of those with lower level of human capital that tend to return migrate. For that reason, the younger immigrants are closer to the US native skilled workers while the older cohorts are relatively closer to their compatriots in their home country. Hence, we observe that the average welfare changes are negative for younger immigrants but positive for older generations.

Third, as mentioned before, the younger the worker is at the time of the policy change, the greater the magnitude of the changes in welfare. In fact, the gain or loss in welfare is even greater for future generations who are not even alive at the time of the policy change, as can be seen in the bottom plot of Figure (B.14), which shows the average change for each age cohort—including those that enter the model after the policy change—over the entire transition path. The later the workers enter the model, the greater the magnitude of the average gain or loss in welfare compared to the old steady state.

Chapter 6

Shutting Down the Return Migration Channel

This chapter examines the importance of including return migration when structurally estimating the impact of relaxing immigration barriers. Specifically, this chapter solves and simulates an alternate version of the model in which return migration is shut down and then compares the steady state and transition path of that version to the baseline model.

Solving this alternate version of the model is very similar to solving the baseline model, except for the fact that the Bellman equation for skilled workers who start the period in foreign country changes from equation (3.4) to the following in which return migration is no longer a possibility.

$$V_a^{xy}(k,h) = \max_{k' \geq 0, s \in [0,1]} u(c) - \frac{\lambda}{h} + \beta \mathbb{E}\left[V_{a+1}^{xy}(k',h')\right] \tag{3.4$'$}$$

subject to

$$c + k' = k\,(1 + r) + h w_h^y (1 - s)$$

$$h' = e^{z_y}\left(h + \theta(hs)^{\nu}\right), \quad z_y \sim \mathcal{N}\left(\mu_y, \sigma_y^2\right)$$

$$V_{65}^{xx}\,(k, h) = \psi \cdot k, \quad k \geq 0$$

$$\forall x \in \{N, S\}$$

$$y \neq x$$

The return migration rate after five years—one of the nine target moments in the baseline model—is zero in this alternate model. In fact, the elimination of the option to return migrate—ceteris paribus—renders immigrants worse off, which lowers the incentive of the immigrants to apply for the H-1B visa. In order to match the target moment of the 49% probability of obtaining the visa, the parameters of the model must be adjusted. It turns out that recalibrating λ—the numerator of the psychic cost—to a lower value of 21.36 is sufficient. The comparison of the model moments calculated in this alternate model to those of the baseline model can be seen in Table 6.1.

Table 6.1: Target Moments: Baseline vs Alternate

Moment	Target	Baseline	Alternate
r	0.042	0.044	0.044
$\frac{K_N}{Y_N}$	2	2.08	2.09
labor share in US	0.65	0.64	0.64
average labor earnings ratio (skilled vs unskilled) in US	1.65	1.62	1.61
$\frac{w_h^N}{w_h^S}$	1.55	1.50	1.48
p	0.49	0.49	0.49
$\frac{\text{flow of skilled immigrants from } S}{\text{native skilled workers of } N}$	0.00174	0.00177	0.00175
return migration rate after 5 years	0.05	0.052	0

Figure (B.15) shows the differences in the transition of the probability of obtaining the visa across the two models. In the baseline model, p increases from 0.49 to 0.57 at the time of the policy change and eventually converges to 0.6, which means that the number of H-1B visa applicants immediately increases by 72% and then converges to a level 63% higher than at the old steady state. In the alternate model with no return migration, p increases from 0.49 to 0.64 at the time of the policy change and eventually converges to 0.79, which means that the number of H-1B visa applicants immediately increases by 53% and then converges to a level 24% higher than at the old steady state. In other words, the model with no return migration underestimates the elasticity of visa application with respect to the reduction in the labor mobility barrier, thereby overestimating the policy's impact on reducing the friction caused by the visa lottery.

The baseline model showed that doubling the visa quota results in a massive increase in visa applications of less talented skilled workers who would not have applied for the visa at the old steady state. In the alternate model, this effect intensifies. The top plot of Figure (B.16) shows that the human capital distribution of those who apply for the visa shows little difference across the two models at the old steady state. However, at the new steady state, the cumulative distribution function of the visa applicants in the alternate model is stochastically dominated by that of the baseline model, which means the average skill of a visa applicant is incorrectly estimated to be much lower than the estimate from the baseline model. The underestimation of the skill levels of visa applicants, reinforced by the fact that the alternate model has no return migration to send less skilled immigrants back to their home country, results in the underestimation of the overall skill level of immigrants as can be seen in the bottom plot of Figure (B.16).

Figure (B.17) shows how the flow and stock of immigrants compare across the two models as the economy transitions from the old steady state to a new one after the H-1B visa quota is permanently doubled. The top plot shows the flow of skilled workers from S to N, while the second plot shows the stock of skilled immigrant workers present in N. In these two plots, the baseline model's old steady-state value is normalized to 100. The top plot shows that in the alternate model with no return migration, it is still the case that doubling the quota immediately causes the number of applicants to jump and all the extra visa spots are instantly filled up. However, because return migration is not possible in the alternate model, the second plot of Figure (B.17) shows that the stock of skilled immigrants at the old steady state is incorrectly estimated to be 96% higher than in the baseline model. Not only is the stock level overestimated, the magnitude of the changes across the two steady states is overestimated as can be see in the bottom plot of Figure (B.17), which is a renormalized version of the second plot with the alternate model's old steady state also normalized to 100. In the baseline model, the stock of immigrants at the new steady state was 94% higher than the old steady-state level; without return migration, it is estimated to be 133% higher. In other words, shutting down return migration results in overestimation of both levels and changes of immigrant stock.

The alternate model has two forces working in the opposite direction: underestimation of the level of human capital of skilled immigrants underestimates the magnitude of the impact of lowering immigration barriers, but the overestimation of the increase in stock overestimates the magnitude. In the end, the latter effect dominates the former and results in the overestimation of the increase in aggregate human capital in N and the decrease in aggregate human capital in S, as top left plot of Figure (B.18) shows.[1] In the baseline model, I estimate that the doubling of the H-1B visa quota would increase the aggregate human capital in the US by 1.87% but lower that of S by 0.6% across the two steady states. In the alternate model—as Table 6.2 shows—these changes are amplified to a 7.51% increase in the US and a 3.21% decrease in S.

The rest of Figure (B.18) show not only how the levels of these prices and aggregate variables are over- or underestimated in the alternate model, but also how the entire transition path is amplified compared to the baseline model. In all of these figures, if the equilibrium

[1] It is worth emphasizing that the level of aggregate human capital is overestimated for country N but underestimated for country S throughout the entire transition path.

Table 6.2: Comparison of New Steady State Against Old Steady State

Equilibrium Prices	Baseline Model	Alternate Model
r	0.01% ↑	1.76% ↑
w_h^N	0.79% ↓	3.45% ↓
w_h^S	0.47% ↑	2.22% ↑
w_l^N	0.59% ↑	2.21% ↑
w_l^S	0.07% ↓	0.42% ↓
p	0.49 to 0.6	0.49 to 0.79
Equilibrium Aggregates	Baseline Model	Alternate Model
$K_N + K_S$	0.36% ↑	0.67% ↑
H_N	1.87% ↑	7.51% ↑
H_S	0.6% ↓	3.21% ↓
output/capita in N	0.79% ↑	1.50% ↑
output/capita in S	0.14% ↓	0.54% ↓

object decreases as the economy transitions from the old steady state to the new, then the alternate model estimates the initial steady state to be lower than it is in the baseline model, and the magnitude with which it decreases toward its new steady state level to be much greater. If the equilibrium object increases as the economy transitions from the old steady state to the new, then the alternate model estimates the initial steady state to be higher than it is in the baseline model, and the magnitude with which it increases toward its new steady state level to be amplified.

The overestimation of the policy impact caused by shutting down return migration also can be seen in the welfare calculations. Figure (B.19) recalculates the bottom plot of Figure (B.14) for the alternate model and compares it to the baseline model; that is, this figure compares across the two models the average gain or loss in consumption by each cohort after the US H-1B visa quota doubles. The US unskilled workers and S-born skilled workers still gain on average from doubling the visa quota, while the opposite is true for US skilled workers and S-born unskilled workers. The gain or loss in welfare is still greater for workers who are younger at the time of the policy change, and they are even greater for future generations who are not alive at the time. However, similar to what we observed for prices and aggregate variables, the magnitude of the welfare gain or loss is amplified when return migration is shut down. The magnitude of the estimation error is greater for the cohorts that are young or have yet to enter the model at the time of the policy change. For example, the average gain in welfare for a US unskilled worker that enters the model at the end of the transition path (the left-most point along the black curves) under the alternate model is 6.78 times that of the baseline model.

Chapter 7

Conclusion

Although the regulation of international labor mobility is one of the most severe distortions facing the world today, there are only a handful of recent papers that quantitatively estimate the magnitude of this distortion. Not only that, the vast majority of the immigration literature has focused primarily on unskilled workers, even though skilled labor immigration is becoming ever more important. I fill that gap by developing a two-country heterogeneous agent general equilibrium model with three novel features: the division of labor into skilled and unskilled labor, the possibility of return migration, and the skill heterogeneity within skilled workers. I use the model to counterfactually analyze the quantitative impact on equilibrium prices, output, and welfare once the H-1B visa quota is permanently doubled. I found that the US experiences a 0.79% gain in output per capita but this aggregate gain hides wide heterogeneity in the effects on wages and welfare both across and within skill groups and age cohorts. I also show that not accounting for return migration results in severe exaggeration of the policy impact. I conclude the paper by drawing attention to two potentially important issues that this paper has abstracted from.

First, the model assumed that the inflow of skilled foreigners do not impact TFP. However, there is an extensive empirical literature on the impact of skilled immigration on innovation. In this paper, it was shown that doubling of the quota leads to an immediate doubling of the inflow of skilled foreigners and that a big portion of the new inflow was comprised of less talented individuals who did not find it optimal to apply for the visa at the old steady state. It would be interesting to quantify how much of the gain in innovation by increased inflow of skilled foreigners is offset by the fact that most of this increase are due to less talented individuals.

Second, this paper takes the proportion of skilled to unskilled workers entering the economy to be exogenous, but it would be interesting to have that be an endogenous decision and examine how a change in immigration policy affects the worker's decision to become a skilled versus unskilled worker. If the H-1B visa quota is doubled, the incentive to become a skilled worker among the foreign populace would increase, which may end up having an even smaller impact on increasing the probability of obtaining the visa. Consideration of these factors will inevitably lead to a better assessment of the welfare effects associated with the reduction in international labor mobility barriers. For now, I leave these issues up for future research.

Bibliography

Aiyagari, S. Rao. (1994). "Uninsured Idiosyncratic Risk, and Aggregate Saving." *Quarterly Journal of Economics*, 109(3): 659-84.

Borjas, George. (1994). "Economics of immigration." *Journal of Economic Literature* 32, 1667–717.

Borjas, George. (2003). "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics*, 118(4): 1335–74.

Borjas, George, and Bernt Bratsberg. (1996). "Who Leaves? The Outmigration of the Foreign-Born." *The Review of Economics and Statistics*, 78(1): 165-76.

Borjas, George, Richard Freeman, and Lawrence Katz. (1997). "How Much Do Immigration and Trade Affect Labor Market Outcomes?" *Brookings Papers on Economic Activity* 1: 1-90.

Bound, John, Breno Braga, Joseph Golden, and Guarav Khanna. (2015). "Recruitment of Foreigners in the Market for Computer Scientists in the United States." *Journal of Labor Economics* 33(S1): S187–S223.

Card, David. (1990). "The impact of the Mariel boatlift on the Miami labor market." *Industrial and Labor Relations Review* 43(2): 245–57.

Card, David. (2001). "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration." *Journal of Labor Economics* 19, 22–64.

Docquier, Frédéric, and Abdeslam Marfouk. (2006). "Measuring international migration by educational attainment, 1990–2000." In: *International Migration, Remittances and the Brain Drain*, 151–199. Eds. C. Özden and M. Schiff. New York: McMillan and Palgrave.

Doran, Kirk, Alexander Gelber, and Adam Isen. (2015). "The Effects of High-Skilled Immigration Policy on Firms: Evidence from H-1B Visa Lotteries." Working Paper.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. (2015). "The Next Generation of the Penn World Table" forthcoming *American Economic Review*, available for download at www.ggdc.net/pwt.

Finn, Michael G. (2014). "Stay Rates of Foreign Doctorate Recipients from U.S. Universities, 2011." Oak Ridge, TN: Oak Ridge Institute for Science and Education.

Friedberg, R., and J. Hunt (1995). "The impact of immigrants on host country wages, employment and growth." *Journal of Economic Perspectives* 9, 23–44.

Goldin, Claudia and Lawrence F. Katz, (2007). "The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890 to 2005." *NBER Working Papers* 12984, National Bureau of Economic Research, Inc.

Grogger, Jeffrey and Gordon H. Hanson. (2015). "Attracting Talent: Location Choices of Foreign-Born PhDs in the United States." *Journal of Labor Economics* 33(S1): S5–S38.

Heckman, James, Lance Lochner, and Christopher Taber. (1998). "Explaining Rising Wage Inequality: Explanations With A Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents." *Review of Economic Dynamics*, 1(1): 1-58.

Huggett, Mark, Gustavo Ventura, and Amir Yaron. (2011). "Sources of Lifetime Inequality." *American Economic Review*, 101: 2923-54.

Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. (2010). "How Much Does Immigration Boost Innovation." *American Economic Journal: Macroeconomics* 2(2): 31-56.

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

Kennan, John. (2013). "Open Borders." *Review of Economic Dynamics*, 16(2): L1-L13.

Kerr, Sari Pekkala, and William R. Kerr. (2011). "Economic Impacts of Immigration: A Survey." *Finnish Economic Papers* 24(1): 1-32.

Kerr, Sari Pekkala, William R. Kerr, and William F. Lincoln. (2015). "Skilled Immigration and the Employment Structures of US Firms." *Journal of Labor Economics* 33(S1): S147–S186.

Kerr, William R., and William F. Lincoln. (2010). "The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention." *Journal of Labor Economics* 28(3): 473-508.

Klein, Paul, and Gustavo Ventura. (2009). "Productivity differences and the dynamic effects of labor movements." *Journal of Monetary Economics*, 56(8): 1059–73.

Krusell, Per, Lee E. Ohanian, Jose-Victor Ríos-Rull, and Giovanni L. Violante. (2000). "Capital-skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica*, 68: 1029–53.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. (2015). "Life-Cycle Wage Growth Across Countries." Working Paper.

LaLonde, Robert, and Robert Topel. (1992). "The assimilation of immigrants in the U.S. labor market." In *Immigration and the Work Force: Economic Consequences for the United States and the Source Areas*, 67–92. Eds. G. Borjas and R. Freeman. Chicago: University of Chicago Press.

Lowell, B. Lindsay. (2000). "H-1B temporary workers: Estimating the population." Working Paper no. 12, University of California, San Diego.

Lubostky, D. (2007). "Chutes or ladders? A longitudinal study of immigrant earnings." *Journal of Political Economy* 115, 820–67.

OECD. (2008). *International Migration Outlook 2008*, OECD Publishing.

Okkerse, L. (2008). "How to measure labour market effects of immigration: A review." *Journal of Economic Surveys,* 22: 1–30.

Ottaviano, Gianmarco I. P., and Giovanni Peri. (2012). "Rethinking the Effects of Immigration on Wages." *Journal of the European Economic Association*, 10: 152–97.

Peri, Giovanni, Kevin Shih, and Chad Sparber. (2015). "STEM Workers, H-1B Visas, and Productivity in US Cities." *Journal of Labor Economics* 33(S1): S225–S255.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. (2015). *Integrated Public Use Microdata Series: Version 6.0* [Machine-readable database]. Minneapolis: University of Minnesota.

Stokey, Nancy L. (1996). "Free Trade, Factor Returns, and Factor Accumulation." *Journal of Economic Growth*, 1(4): 421-47.

Tauchen, George. (1986). "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions." *Economic Letters*, 20: 177–81.

Appendix A

Computation Algorithm

A.1 Recursive Stationary Equilibrium

The recursive stationary equilibrium of the two-country heterogeneous agent model described in Chapter 3 cannot be solved by pen and paper. Hence, I develop a two-country version of the Aiyagari (1994) heterogeneous agent model computation algorithm. The algorithm is as follows:

1. Guess the aggregate demand values for the total world capital $K^* = K_N + K_S$, total human capital in each of the two countries H_N and H_S, and the probability p of visa approval. There is no need to make a guess on L_N and L_S since unskilled workers inelastically supply labor and cannot relocate to a foreign country.

2. Using the assumption that capital moves costlessly across two countries, calculate K_N and K_S by solving the following equation

$$r = A_N F_K(K_N, H_N, L_N) - \delta = A_S F_K(K^* - K_N, H_S, L_S) - \delta$$

3. Prices $\left(r, w_h^N, w_h^S, w_l^N, w_l^S\right)$ are calculated according to the marginal conditions of the aggregate production.

4. Given the terminal value functions $V_{65}^{xx}(k,h) = V_{65}^{xy}(k,h) = W_{65}^{x}(k) = \psi \cdot k$ and prices calculated in the previous step, calculate the value functions and optimal decision rules of all the agents in the model by backward induction. That is, calculate $V_a^{xx}(k,h)$, $V_a^{xy}(k,h)$, and $W_a^{x}(k)$ starting with $a = 64$, all the way until $a = 25$.

5. Compute the stationary distribution using the calculated optimal decision rules.

6. Use the stationary distribution and the optimal decision rules to compute the aggregate supply values for K^*, H_N, and H_S. Aggregate the mass of workers who decide to enter the visa lottery at the stationary distribution to calculate p.

7. Compare these values to the initial guesses on K^*, H_N, H_S, and p. If they are close, then the computation is finished; if not, update the guess and repeat the algorithm.

For maximization in the value function calculation, I use the subplex algorithm, which is a subspace-searching simplex algorithm for unconstrained optimization of general multivariate functions (general version of Nelder-Mead). The z shocks to human capital are approximated by a five-state Markov chain using the Tauchen (1986) procedure. I use a step function to approximate the CDF of the stationary distribution.

Value functions and optimal decision rules are interpolated with piecewise-linear functions over the grid points. I use 300 grid points for capital and 50 grid points for human capital. Because skilled workers have five states—capital, human capital, country of birth, country of current residence, and age—there are a total of $300 \cdot 50 \cdot 2 \cdot 2 \cdot 40 = 2,400,000$ of grid points to evaluate. To speed up the computation process, I code my program in Fortran 90 and use Open MPI to parallelize the program. One iteration of the algorithm takes 3 minutes for completion using 150 processors on Acropolis, which is a clustered Linux system maintained by the Social Science Computing Services at the University of Chicago.

A.2 Transition Path

Let $t \in \{0, 1, \ldots\}$ denote time along the transition path, in which $t = 0$ represents the old steady state, $t = 1$ represents the year in which the H-1B visa quota is permanently doubled, $t = 2$ represents the year after the policy change, etc. Also, let $\hat{V}_a^{xx}(k, h; t)$, $\hat{V}_a^{xy}(k, h; t)$, and $\hat{W}_a^{x}(k; t)$ denote respectively the value function of a skilled worker in home country, a skilled worker in foreign country, and an unskilled worker at time t of the transition path. $\hat{\Omega}_a^{xx}(k, h; t)$, $\hat{\Omega}_a^{xy}(k, h; t)$, and $\hat{\Omega}_a^{x}(k; t)$ represent respectively the distribution of skilled workers in home country, skilled workers in foreign country, and unskilled workers at time t. For ease of notation, I denote $\mathbb{V}_t = \left\{\hat{V}_a^{xx}(k, h; t), \hat{V}_a^{xy}(k, h; t), \hat{W}_a^{x}(k; t)\right\}$ as the set containing various value functions of all of the heterogeneous agents present in the economy at time t. Likewise, $\mathbb{F}_t = \left\{\hat{\Omega}_a^{xx}(k, h; t), \hat{\Omega}_a^{xy}(k, h; t), \hat{\Omega}_a^{x}(k; t)\right\}$ represents the set containing all the relevant distribution of the agents at time t.

I use the following algorithm for computing the transition path:

1. Compute the stationary distribution before and after the policy change using the algorithm described in Appendix A.1, and respectively call these $\mathbb{F}_0$ and $\mathbb{F}_\infty$. Let $\mathbb{V}_\infty$ be the set of terminal value functions at the new steady state.

2. Assume it takes an arbitrary number of periods $\mathcal{T}$ to transition from $\mathbb{F}_0$ to $\mathbb{F}_\infty$.

3. Guess a sequence of aggregate demands for capital and human capital, and the probability p of visa approval. In other words, guess $(K_{N,t} + K_{S,t}, H_{N,t}, H_{S,t}, p_t)_{t=1}^{\mathcal{T}-1}$. Use those values to calculate the sequence of prices $\left(r_t, w_{h,t}^N, w_{h,t}^S, w_{l,t}^N, w_{l,t}^S\right)_{t=1}^{\mathcal{T}-1}$.

4. Starting from $\mathbb{V}_\mathcal{T} = \mathbb{V}_\infty$, compute the time sequence of value functions $\{\mathbb{V}_{\mathcal{T}-1}, \mathbb{V}_{\mathcal{T}-2}, \ldots, \mathbb{V}_1\}$ and all the associated optimal decision rules along the transition path by backward induction, taking as given the sequence of prices $\left(r_t, w_{h,t}^N, w_{h,t}^S, w_{l,t}^N, w_{l,t}^S\right)_{t=1}^{\mathcal{T}-1}$ calculated in the previous step.

5. Use the calculated optimal decision rules to compute the time sequence of distributions $\{\mathbb{F}_1, \ldots, \mathbb{F}_{\mathcal{T}}\}$ and the aggregate supply values for $(K_{N,t} + K_{S,t}, H_{N,t}, H_{S,t}, p_t)_{t=1}^{\mathcal{T}-1}$.

6. Check if the sequence of calculated aggregate supply values for $(K_{N,t} + K_{S,t}, H_{N,t}, H_{S,t}, p_t)_{t=1}^{\mathcal{T}-1}$ is close to the guessed sequence of aggregate demand values of these variables. If not, update $(K_{N,t} + K_{S,t}, H_{N,t}, H_{S,t}, p_t)_{t=1}^{\mathcal{T}-1}$ and repeat from Step 3 on.

7. Check if $\mathbb{F}_{\mathcal{T}} \approx \mathbb{F}_{\infty}$. If not, increase $\mathcal{T}$ and repeat from Step 2 on.

Appendix B

Figures

Figure B.1: Initial Distribution of Human Capital in the Simulation

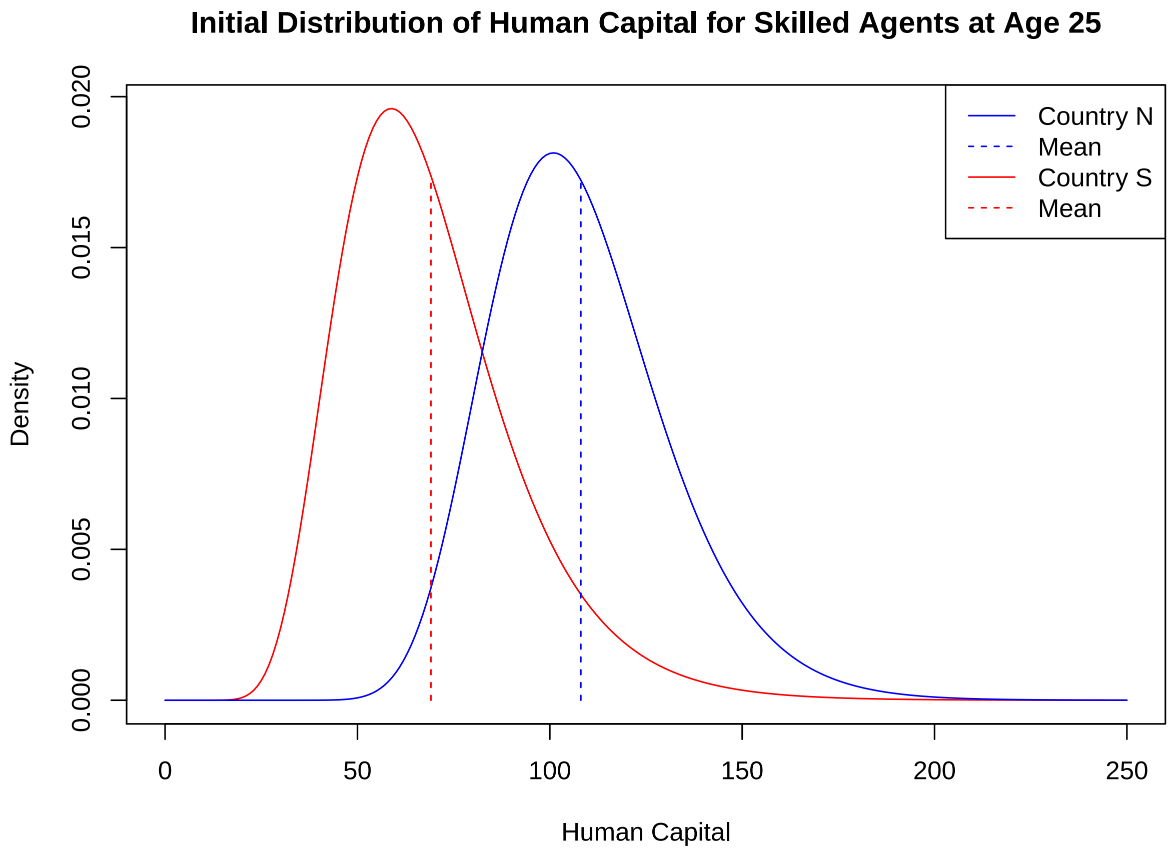


Figure B.2: US Average Wealth Profile: Data vs Model

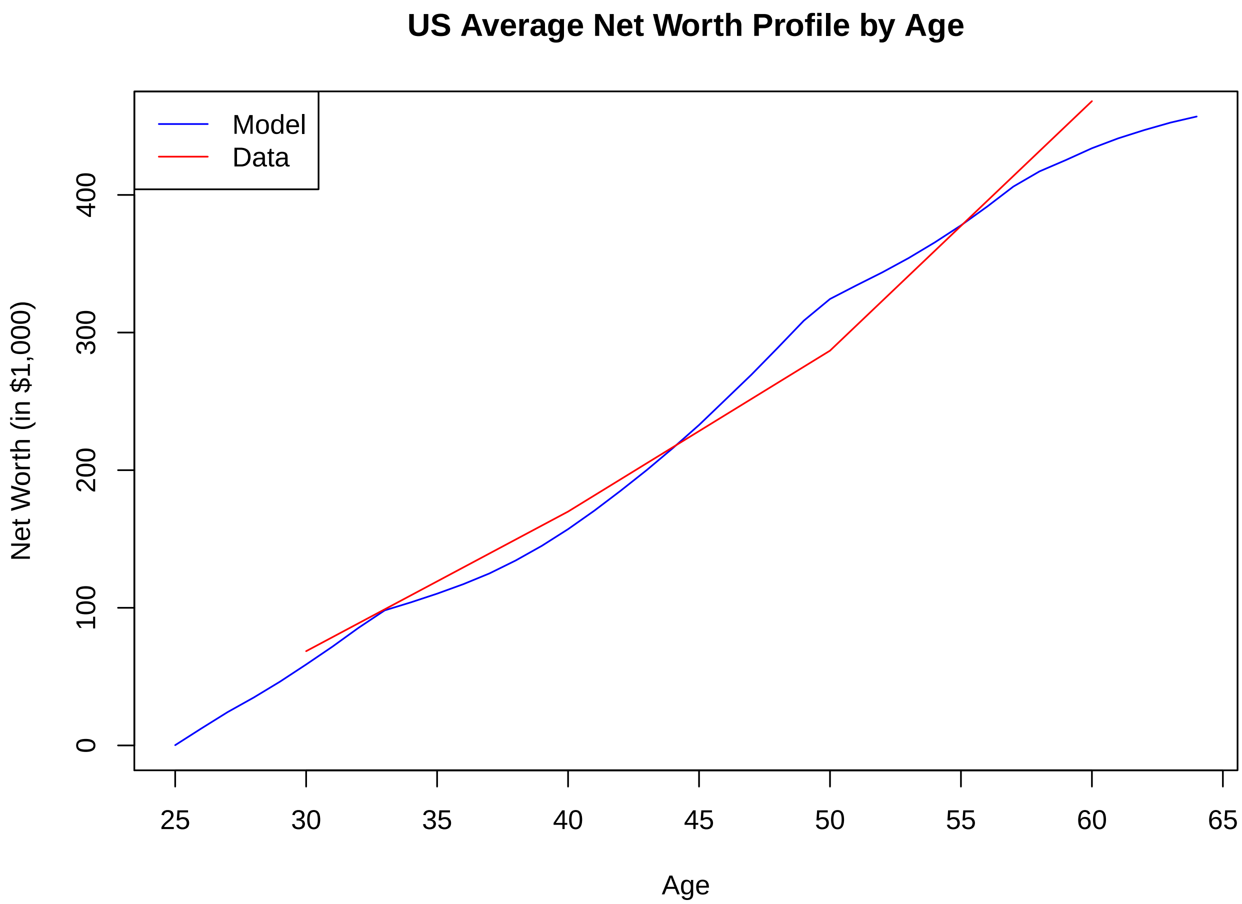


Figure B.3: Profiles over the Life-cycle

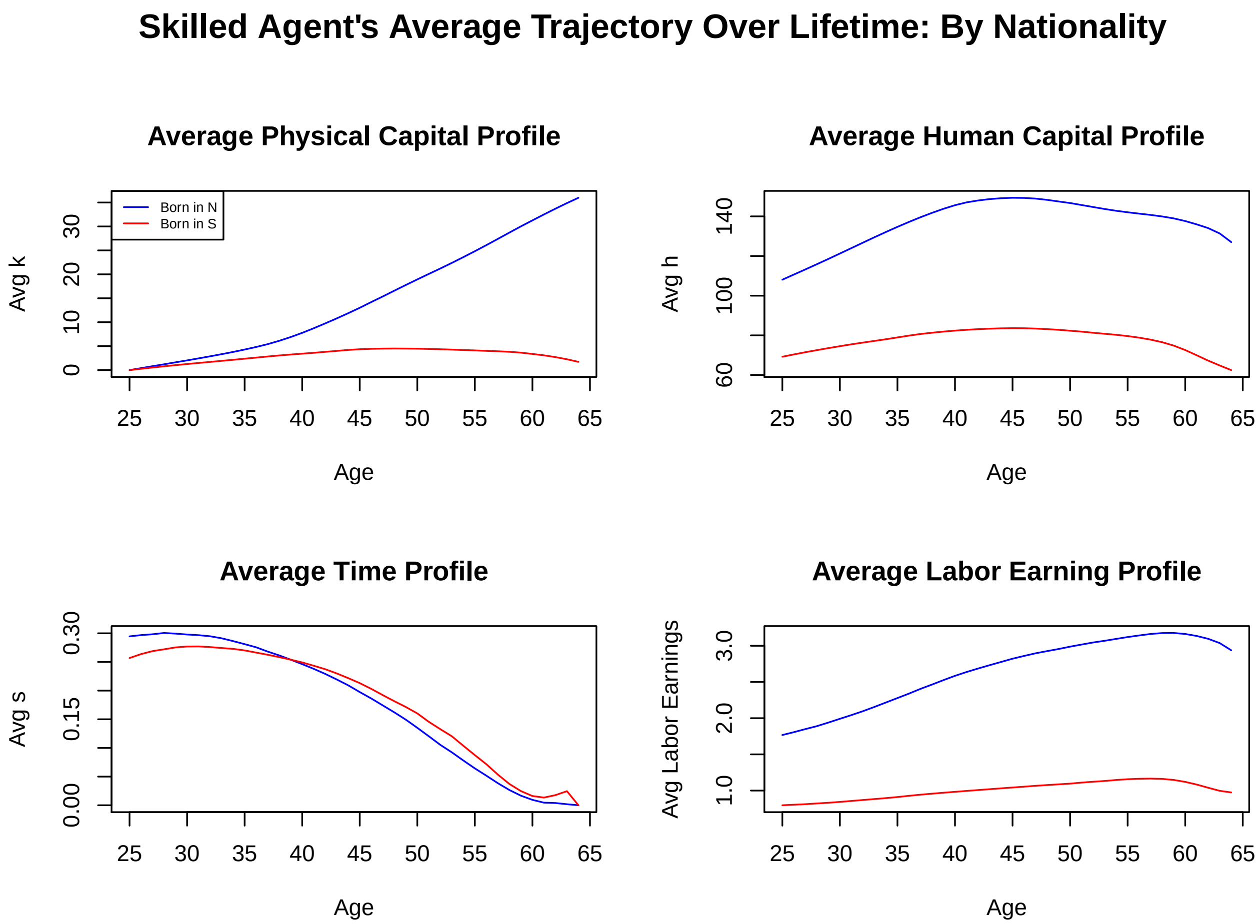
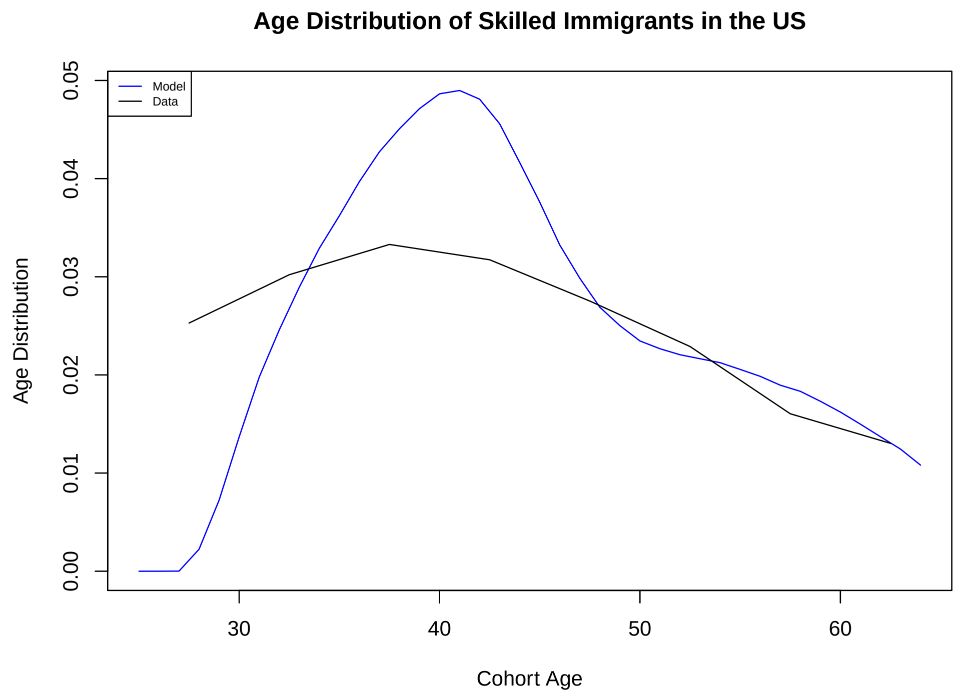


Figure B.4: Immigration and Return Migration



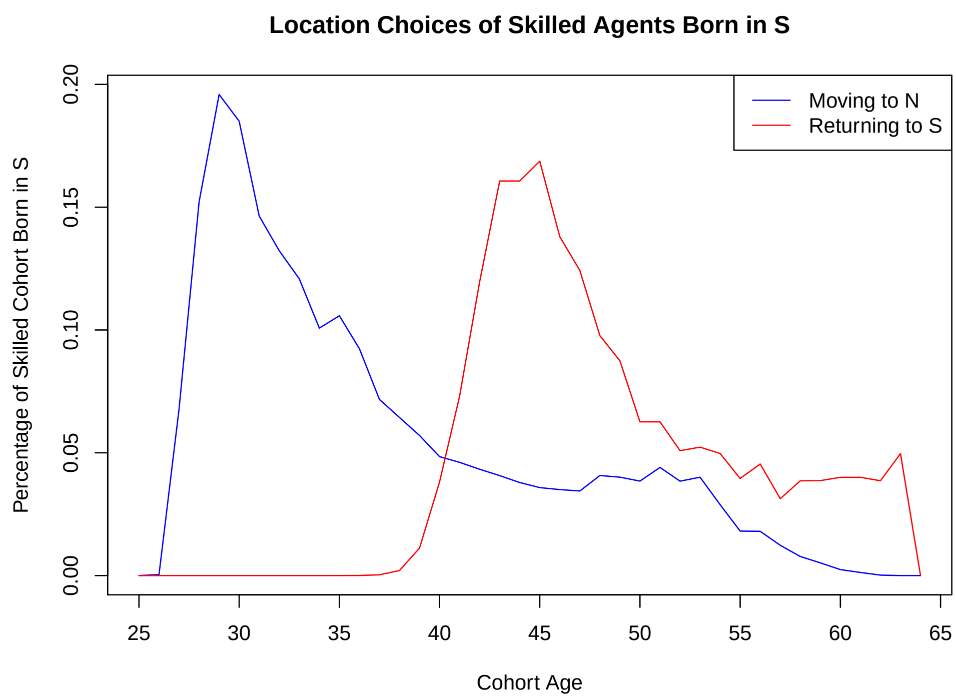
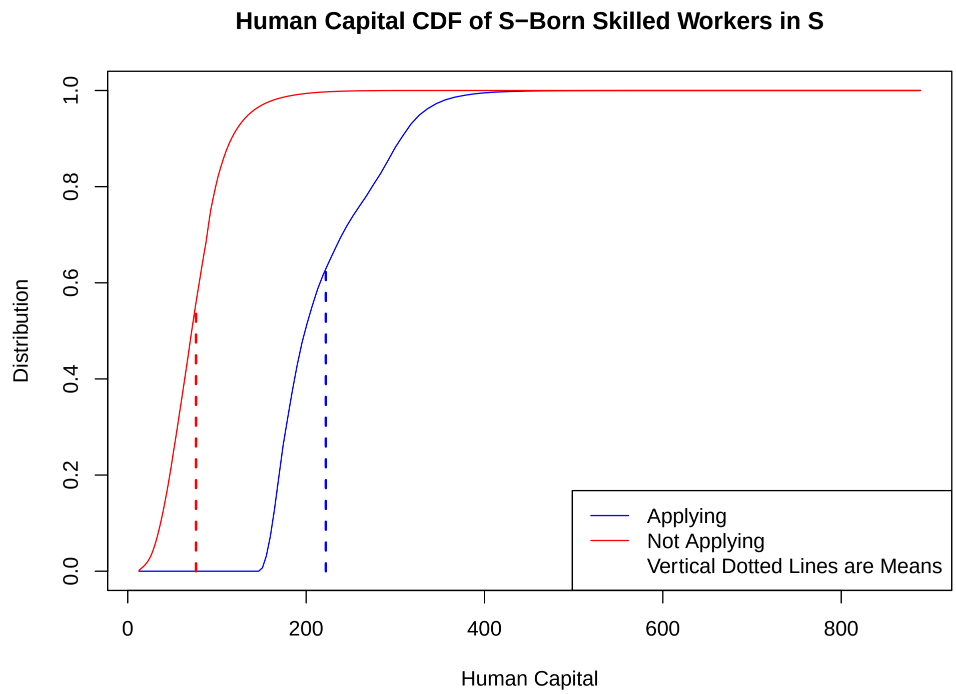


Figure B.5: Comparing Human Capital Distribution



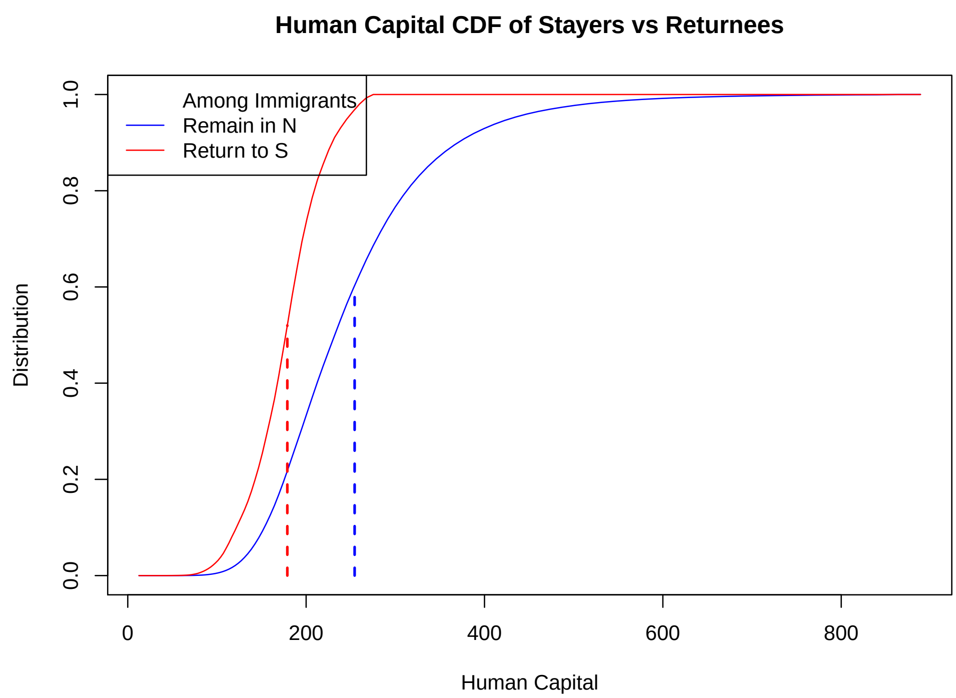


Figure B.6: Policy Experiment: Human Capital Distribution of Immigrants

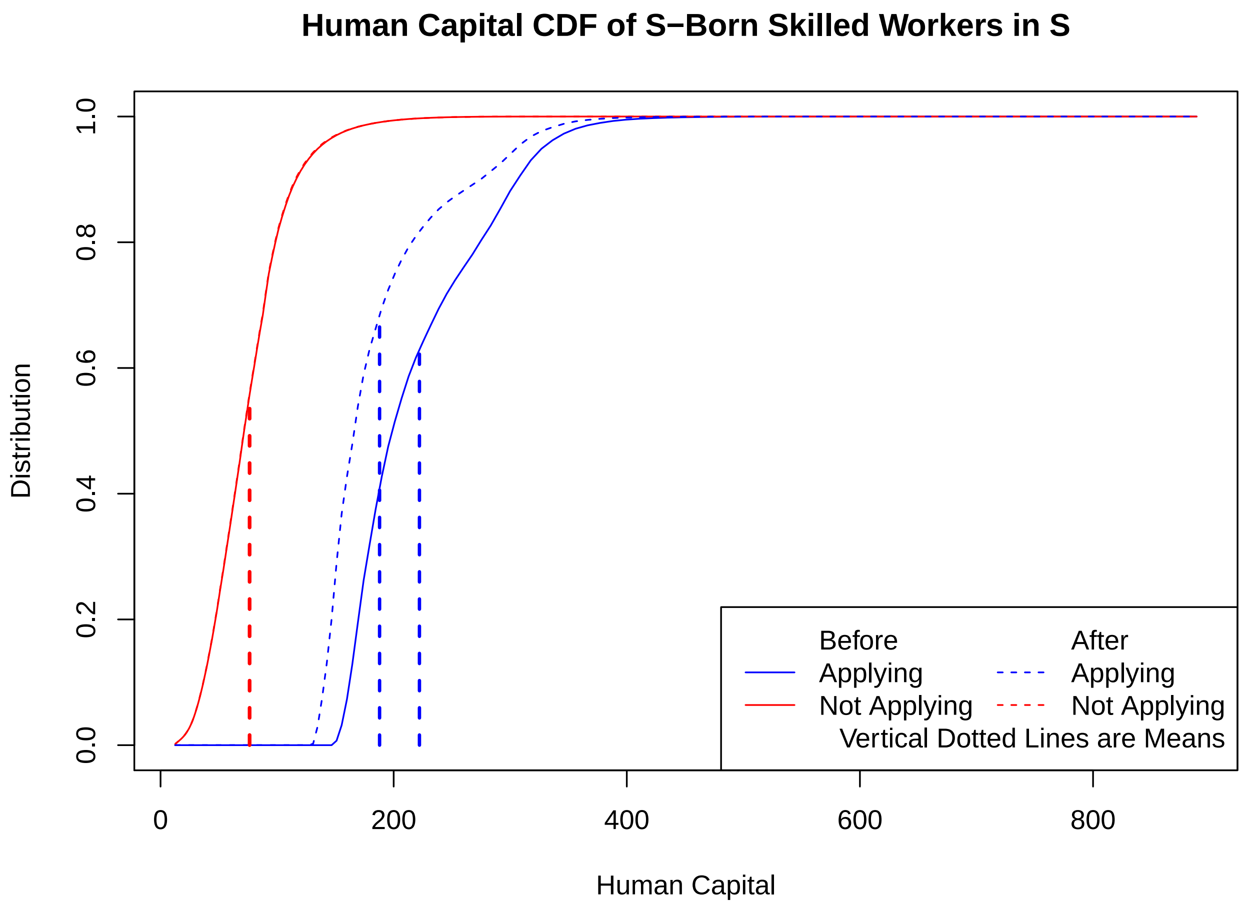


Figure B.7: Policy Experiment: Mass of Immigrants over Human Capital

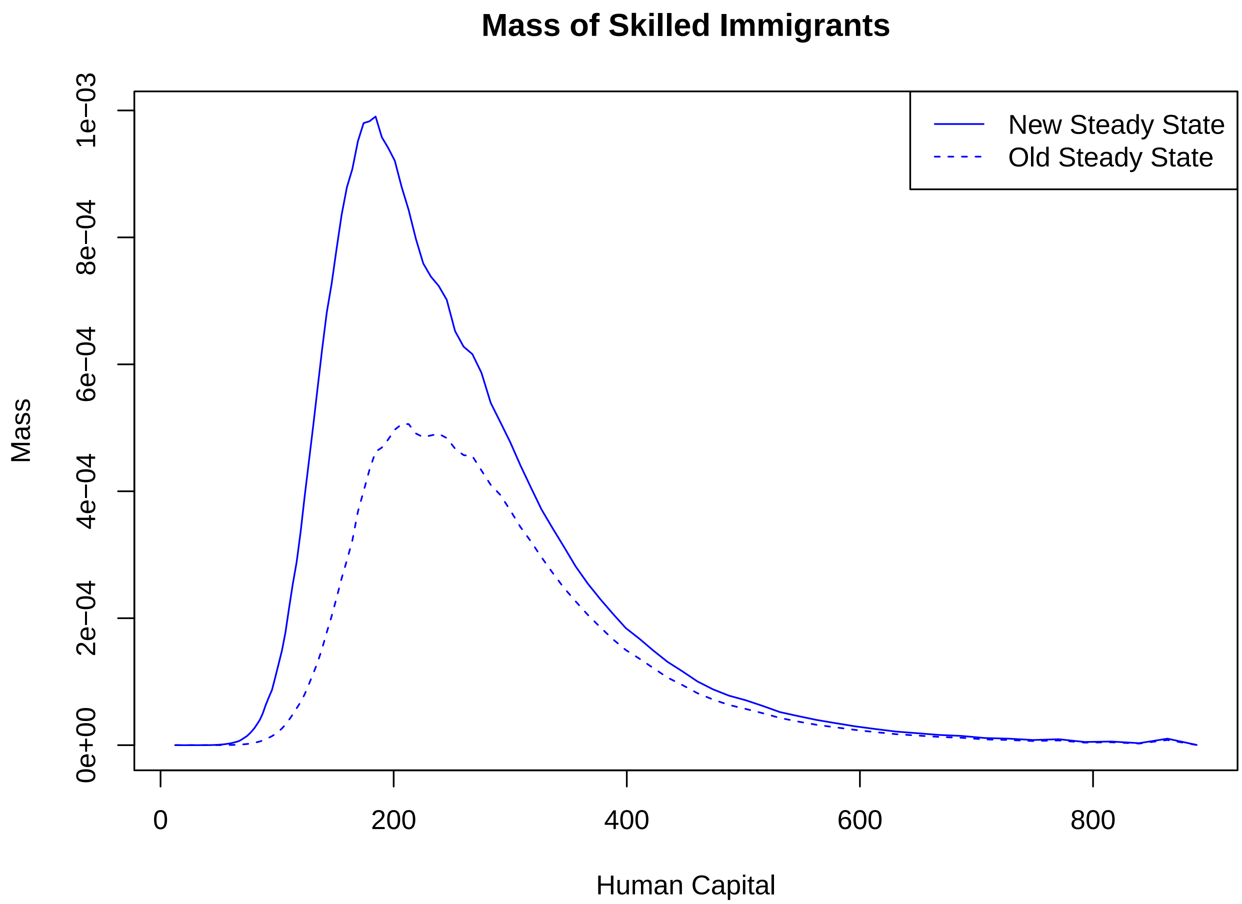


Figure B.8: Policy Experiment: Change in Stock of Immigrants

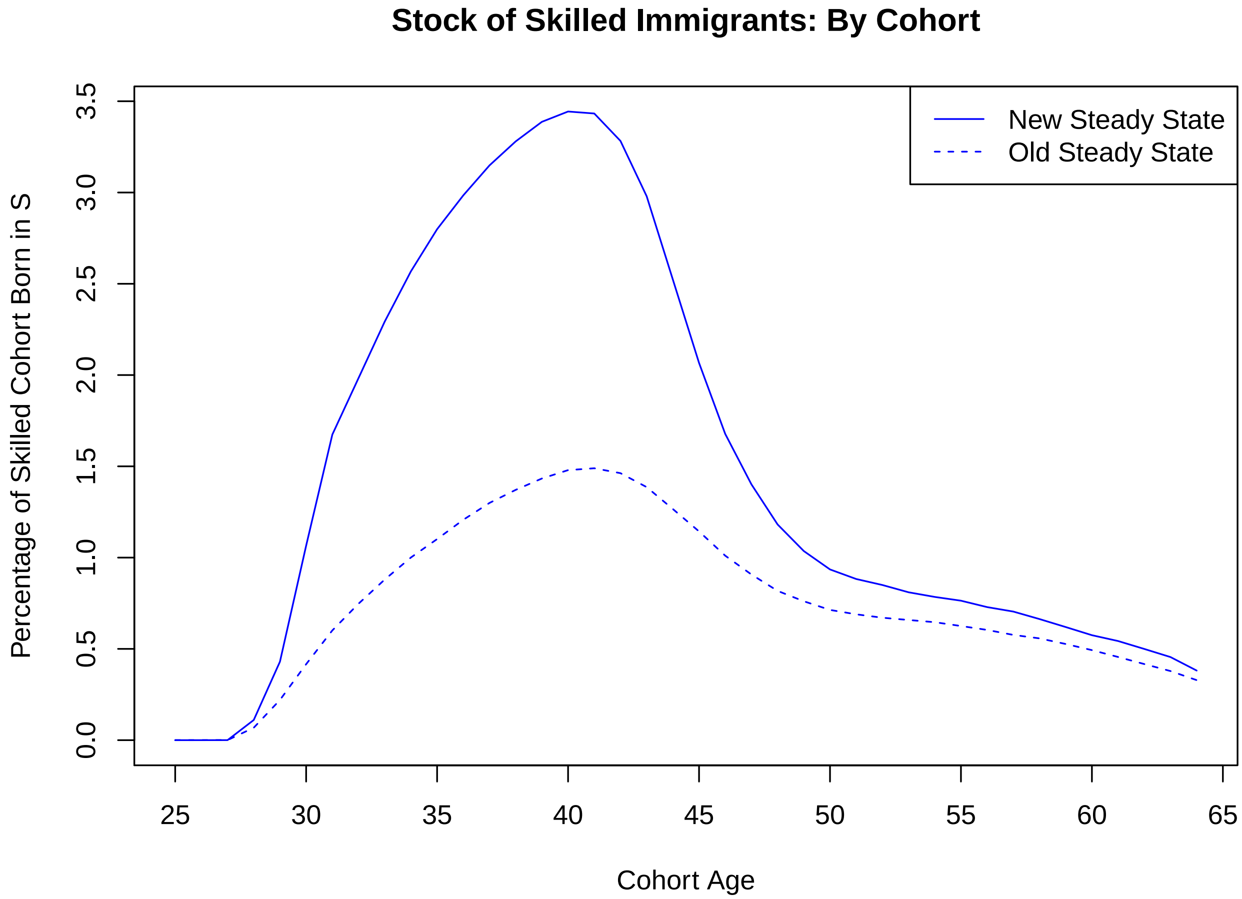


Figure B.9: Transition Path: Flow and Stock of Immigration and Return Migration

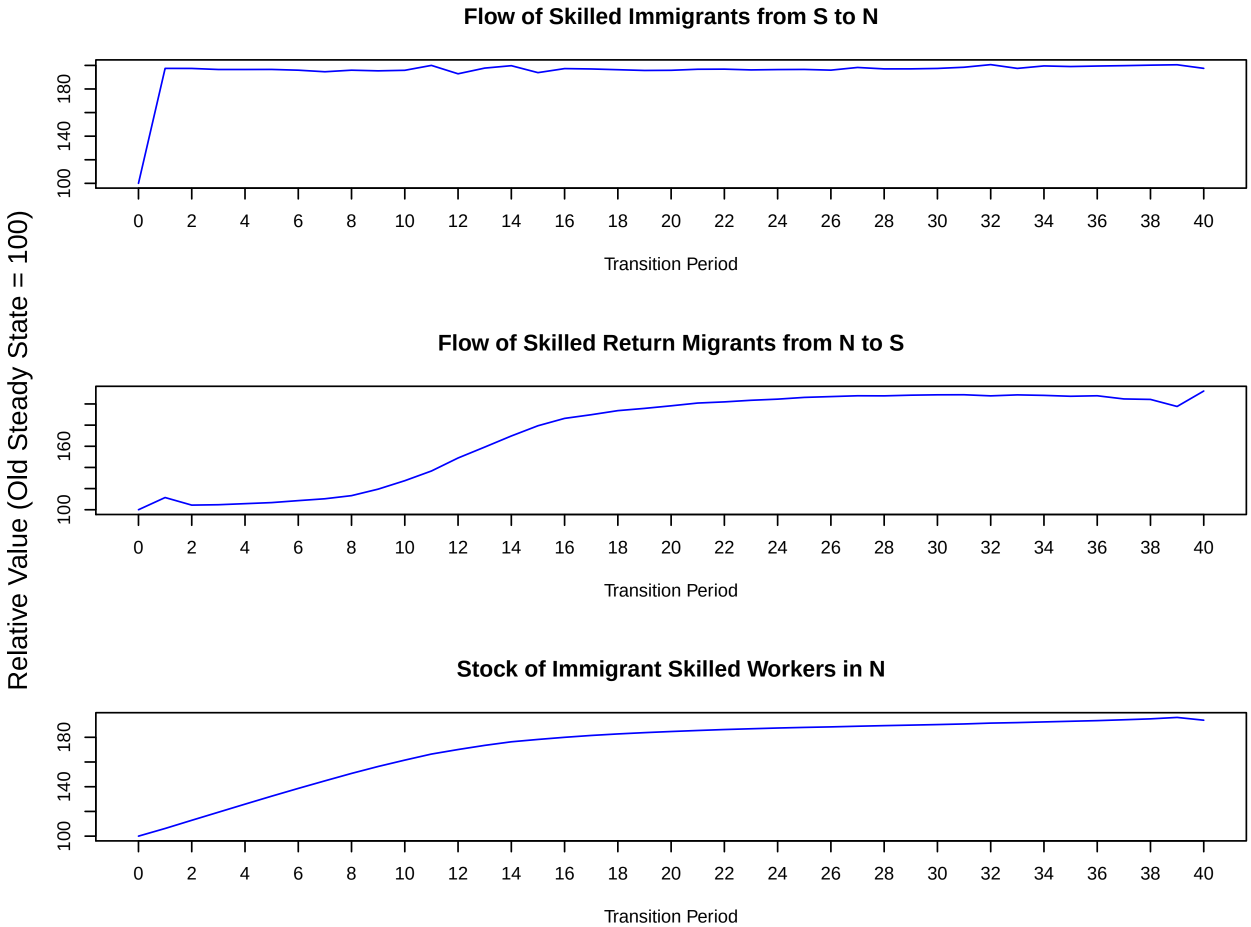


Figure B.10: Transition Path

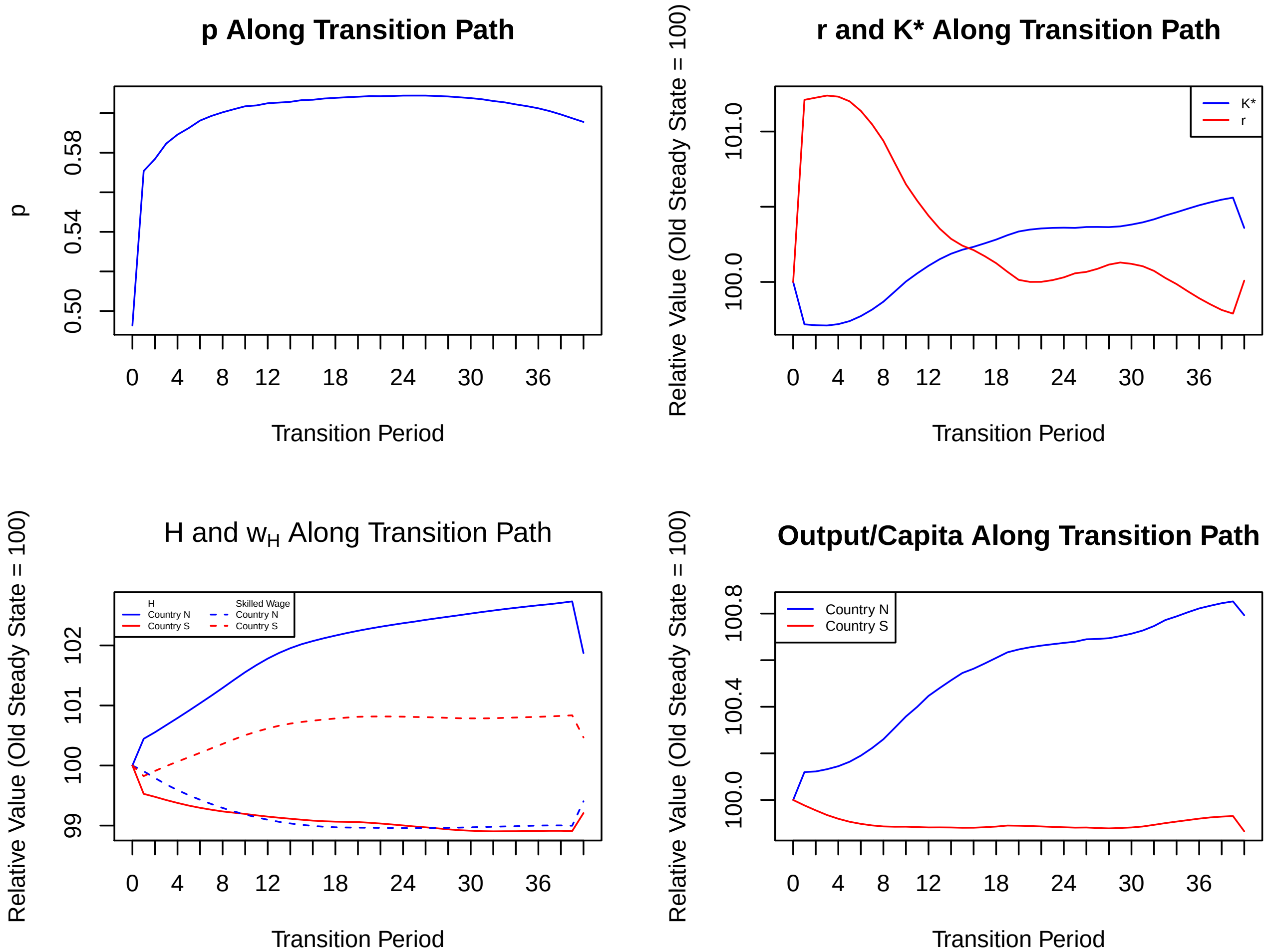


Figure B.11: Welfare Changes of N-born Skilled Workers : By Cohort Age

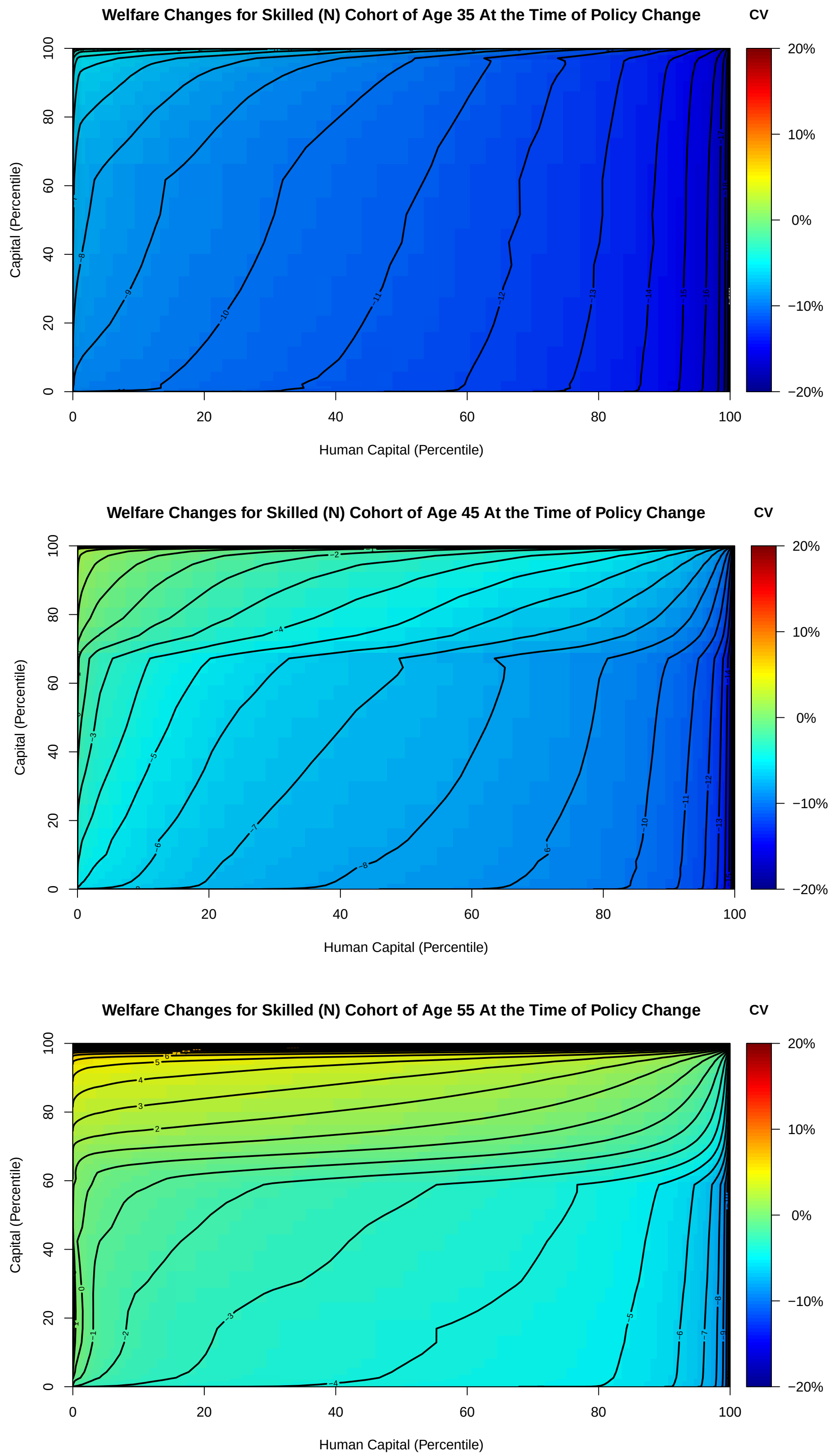


Figure B.12: Welfare Changes of S-born Skilled Workers : By Country of Residence

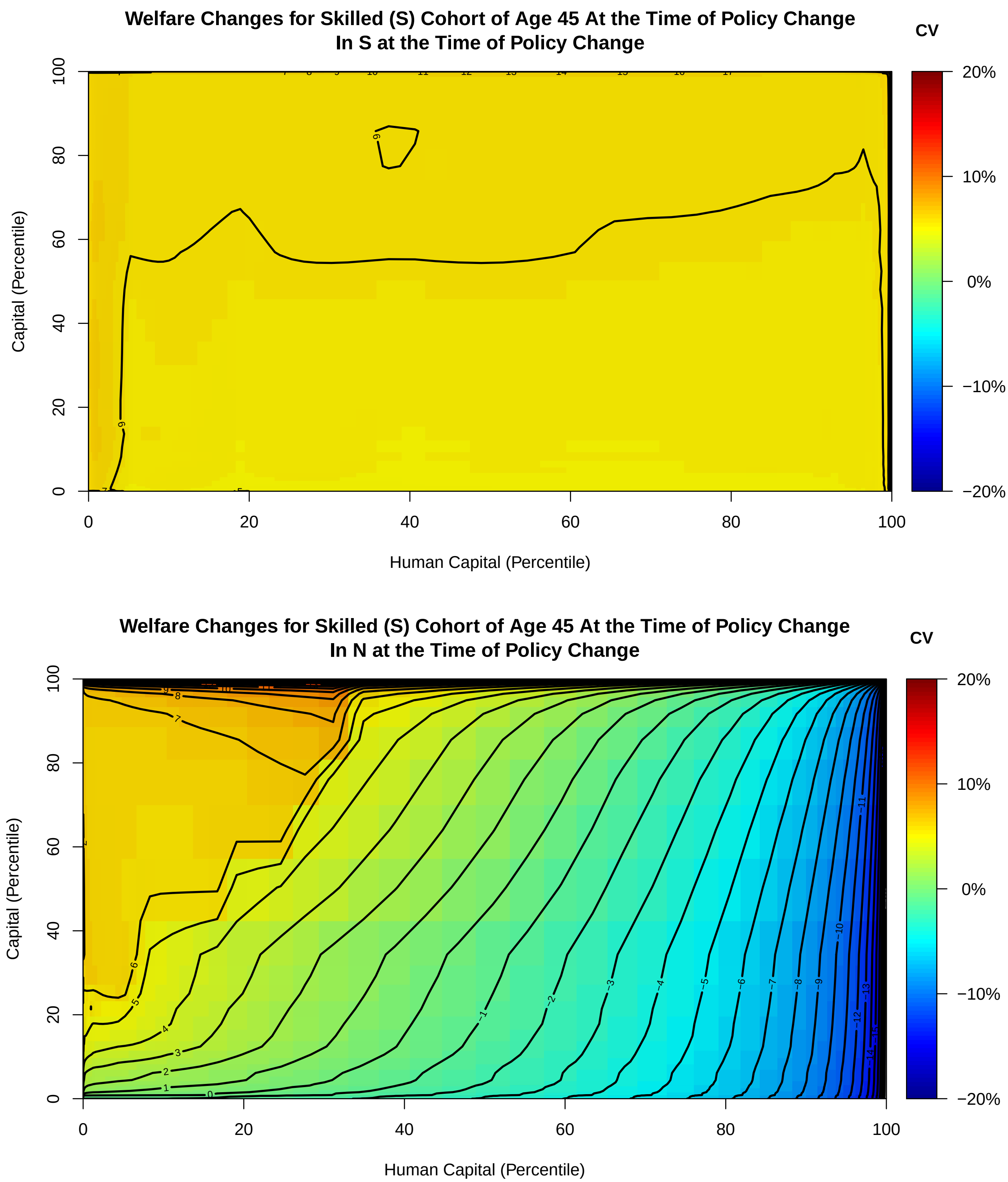


Figure B.13: Welfare Changes of Skilled Workers Aged 45

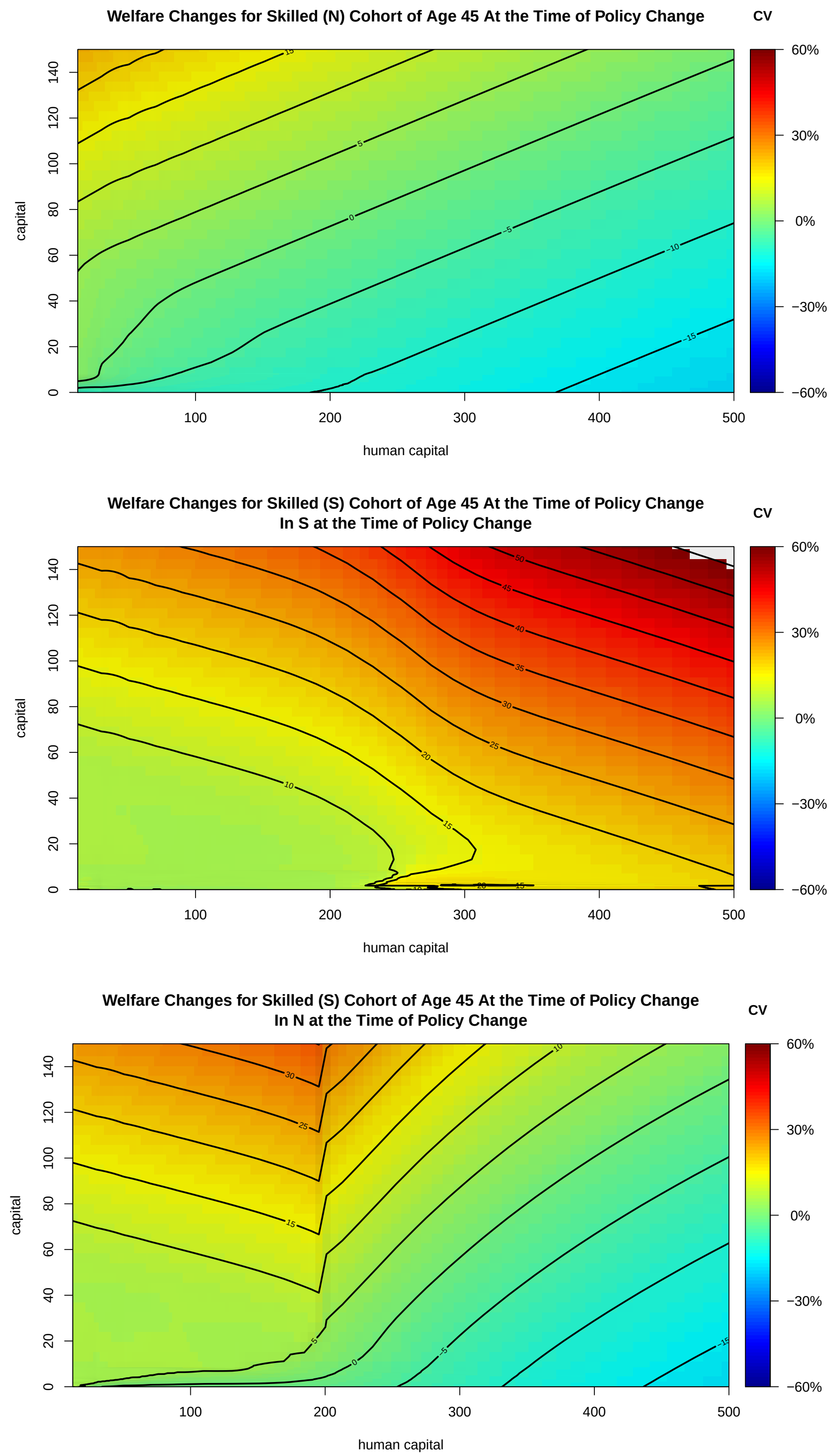
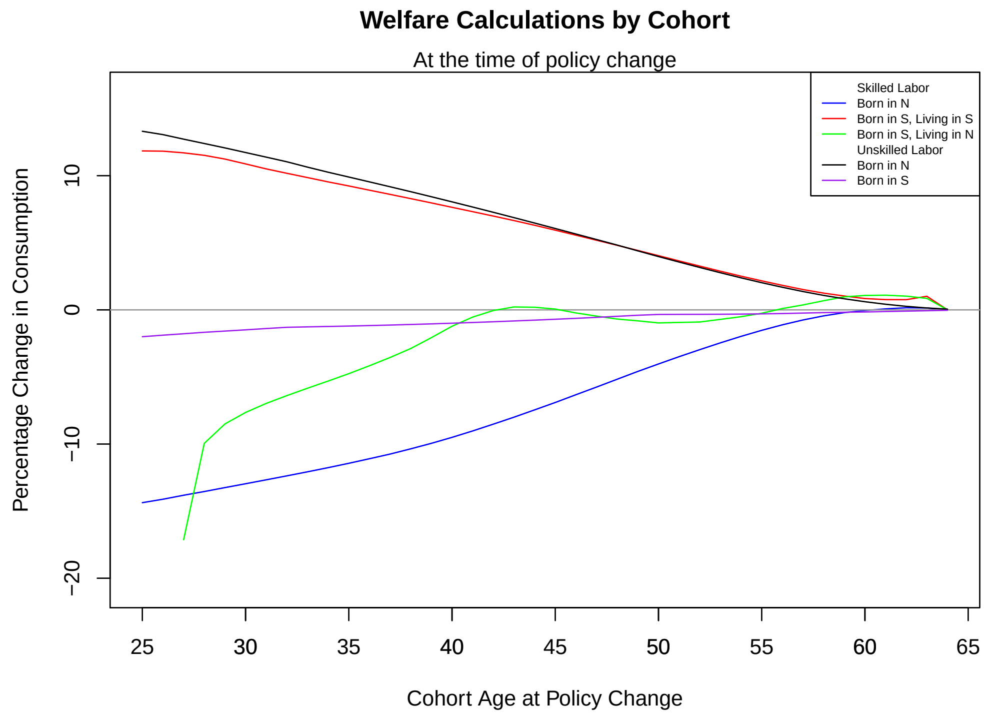


Figure B.14: Welfare Changes



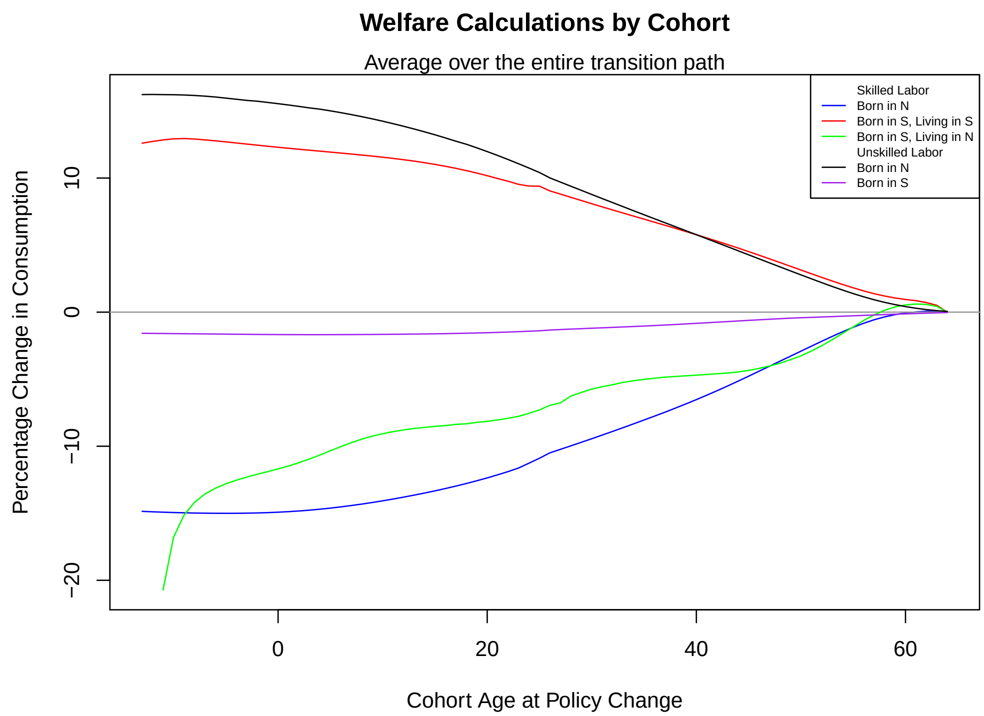


Figure B.15: Baseline vs Alternate: Probability of Visa Approval

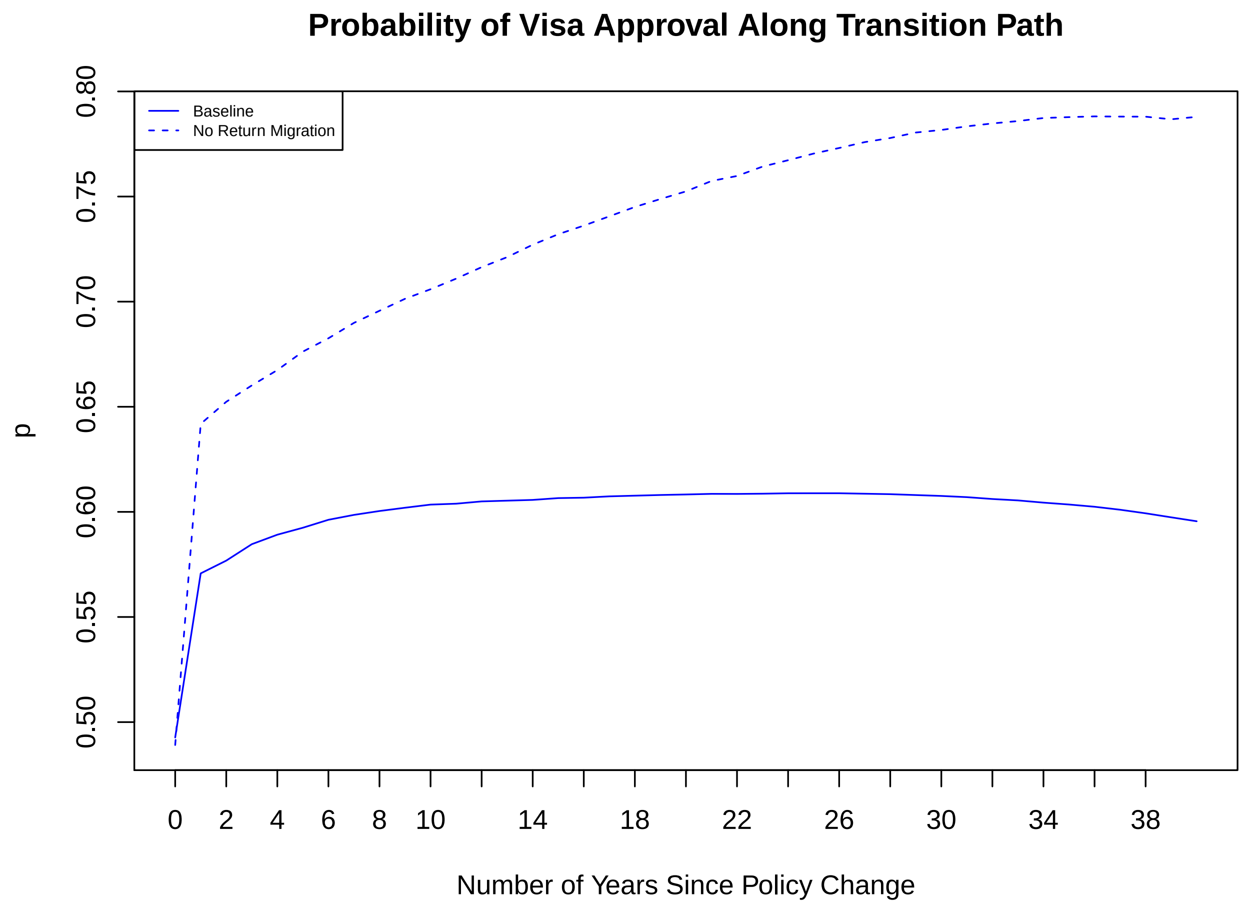
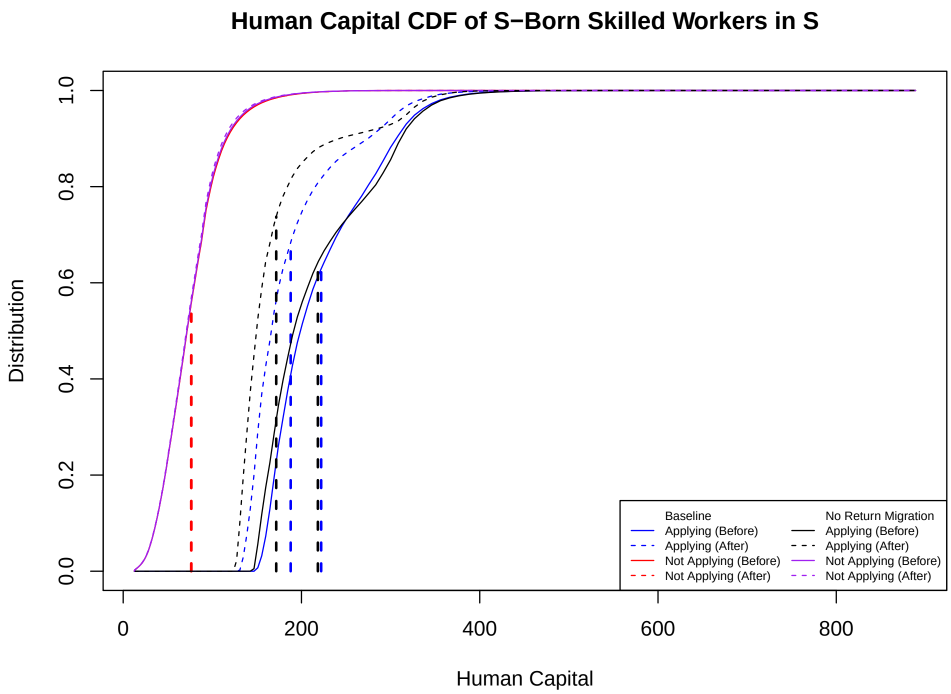


Figure B.16: Baseline vs Alternate: Human Capital Distribution



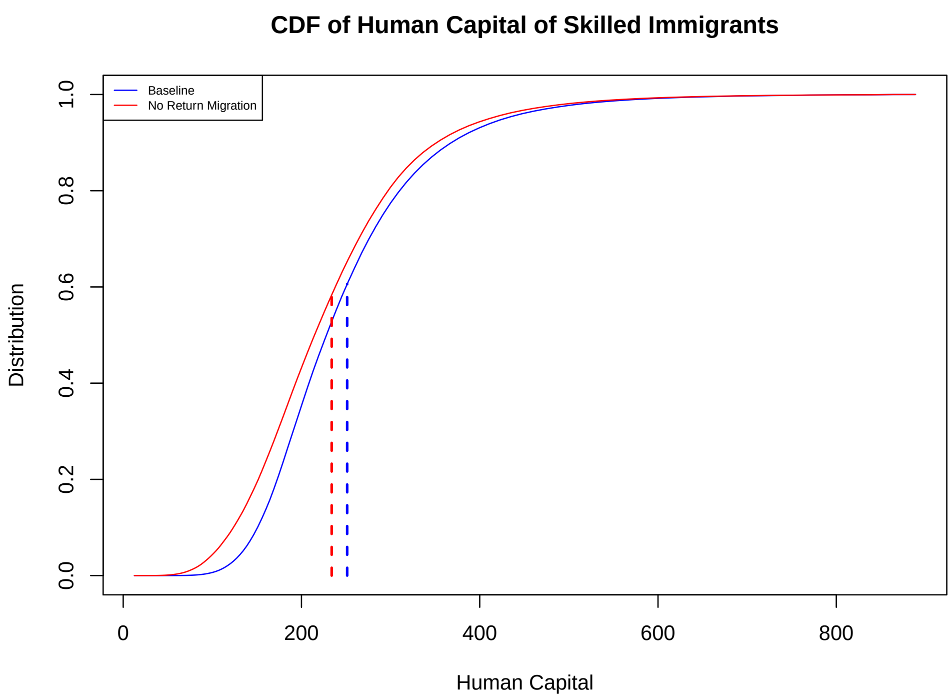


Figure B.17: Baseline vs Alternate: Flow and Stock of Immigration

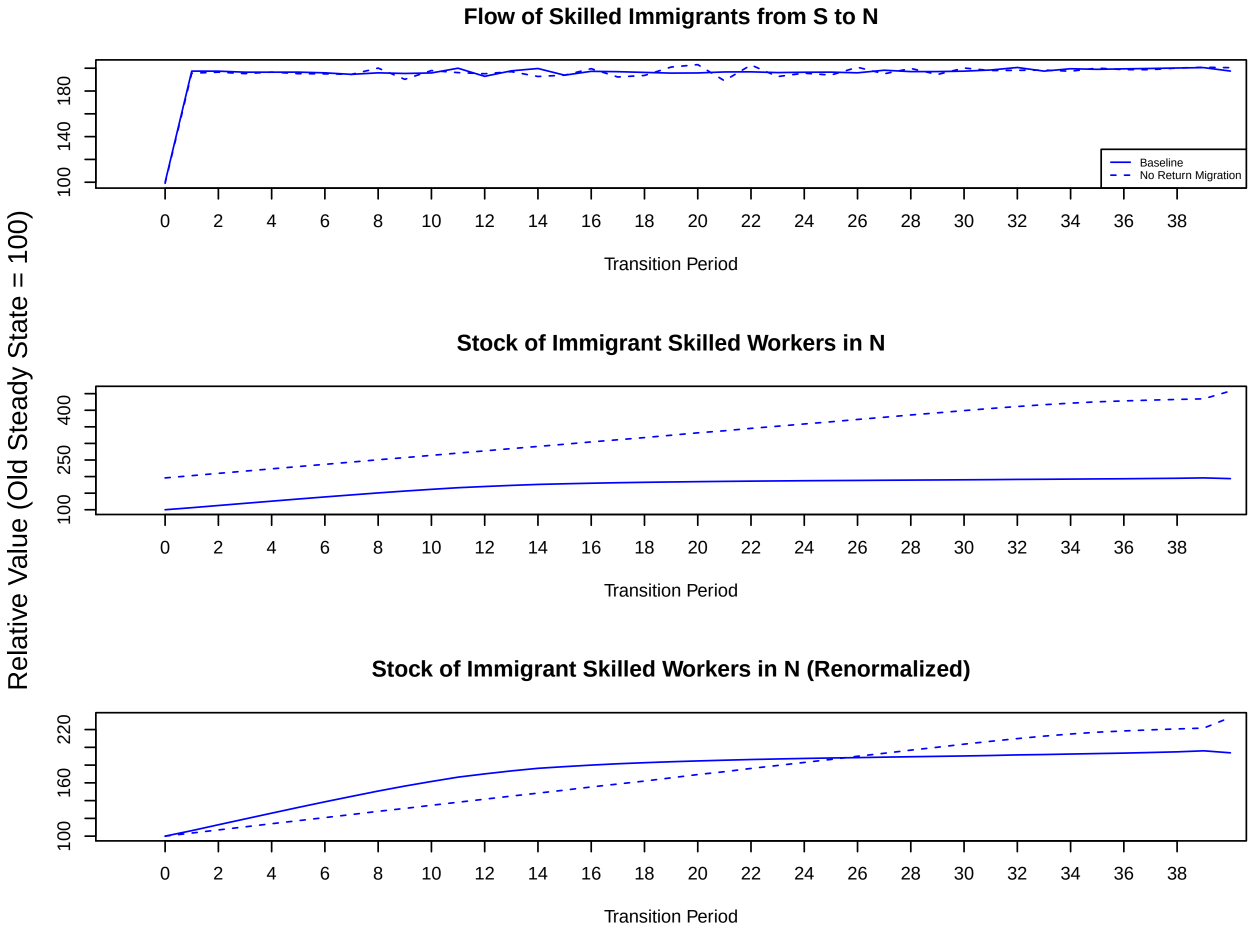


Figure B.18: Baseline vs Alternate: H, w_H, w_L, and Output/capita

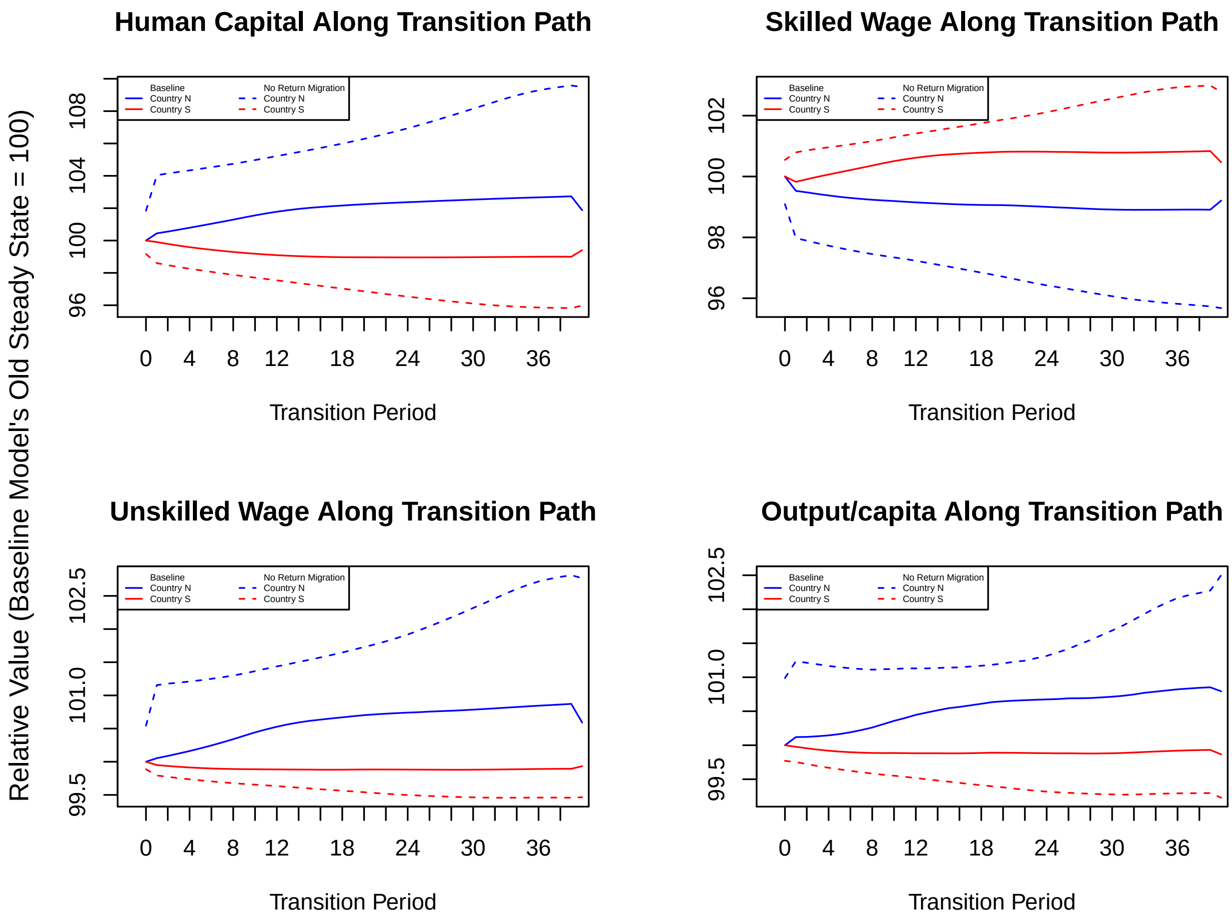


Figure B.19: Baseline vs Alternate: Welfare Changes