THE UNIVERSITY OF CHICAGO

ESSAYS ON ECONOMICS OF IMMIGRATION

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

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY
TUNG LE

CHICAGO, ILLINOIS
DECEMBER 2020

I am from there. I am from here.

I am not there and I am not here.

I have two names, which meet and part,

and I have two languages.

I forget which of them I dream in.

— Mahmoud Darwish

CONTENTS

LIST OF FIGURES v

LIST OF TABLES vi

ACKNOWLEDGMENTS vii

ABSTRACT viii

1 IMMIGRANTS, LANGUAGE BARRIER, AND ASSIMILATION 1
1.1 Introduction 1
1.2 Model 7
1.3 Immigrant assimilation patterns 26
1.4 Concluding remarks 43
1.5 Appendix 45

2 DIVERSITY, GROWTH, AND INEQUALITY 64
2.1 Introduction 64
2.2 Model 67
2.3 Concluding Remarks 80
2.4 Appendix 82

3 REGIONAL IMPACTS OF IMMIGRATION 91
3.1 Introduction 91
3.2 Model 94
3.3 Reduced-form analysis 116
3.4 Concluding remarks 128
3.5 Appendix 129

REFERENCES 133

LIST OF FIGURES

1.1 Fraction of newcomers reside in the ethnic enclave 19
1.2 Fraction of newcomers reside in the ethnic enclave 20
1.3 Marginal impact of immigrant stock on newcomers' value of learning . . . 21
1.4 Ratio of average productivity, English proficiency, labor supply, and earnings between newcomers locate away from the enclave vs. enclave residents of the same characteristics 25
1.5 Location choice, all immigrants 30
1.6 Location choice, immigrants migrated less than 5 years before the survey . 31
1.7 Location choice, all immigrants 32
1.8 Location choice, immigrants migrated less than 5 years before the survey . 33
1.9 Regression residuals 52
1.10 Residuals vs. Fitted values 53
1.11 1990 first quintile by income 61
1.12 1990 third quintile by income 62
1.13 1990 fifth quintile by income 63

2.1 Multiequilibria of the diverse economies 79

3.1 Homothetic preference for locations 103
3.2 Close to home preference 104
3.3 Comparative (dis)advantages in skills 106
3.4 O*NET Survey 117
3.5 Scatterplot of ODI vs. RDI 119
3.6 ODI with immigrants' duration of stay 122
3.7 Immigrants, natives employment by English skills requirements 123
3.8 Immigrants/natives employment by English skills requirement, 10 years after 124
3.9 Industry by English skills requirement 125
3.10 Immigrants, Natives Employment by English skills requirement (industry level) 126
3.11 Immigrants/natives employment by English skills requirement, 10 years after 127

LIST OF TABLES

1.1 Education, English Proficiency, and Spatial Assimilation 34
1.2 Wage growth and English proficiency 41
1.3 English proficiency improvement 42
1.4 Wage growth and English proficiency 50
1.5 English proficiency improvement 51
1.6 Education, English Proficiency, and Spatial Assimilation, Control for Income Level - 1990 55
1.7 Education, English Proficiency, and Spatial Assimilation - 2000 56
1.8 Education, English Proficiency, and Spatial Assimilation, Control for Income Level - 2000 57
1.9 Education, English Proficiency, and Spatial Assimilation - 1980 58
1.10 Education, English Proficiency, and Spatial Assimilation, Control for Income Level - 1980 59

3.1 Determinants of ODI/RDI 121

ACKNOWLEDGMENTS

The path forward toward this dissertation has been intellectually challenging but not without joy. Its completion is thanks in large part to the special people who challenged, suppported, and stayed with me along the way.

I am grateful to my advisors Nancy L. Stokey, Robert E. Lucas Jr., and Steve J. Davis. I am extremely fortunate to be their student and have opportunities to learn and seek their advices from time to time. They, their works are the inspiration for my work. Their continued support and encouragement help me through the most challenged times.

I am also grateful to my old-time professor and mentor Erzo G.J. Luttmer for his encouragement and support. In addition, I would like to thank my colleagues and friends, especially Paphon Kiatsukuldecha, Fang Fu, John van Den Berghe, for their company and support through my study. Last but most important, I am grateful to my parents, my brother and my loving wife for their support, and encouragement over a great part of my life. There is no word that can describe my gratitude to them and I am forever in their debt.

ABSTRACT

This dissertation consists of three essays on the topics of immigration, utilizing economic frameworks to (i) study the rational behaviors and human capital development of immigrants and (ii) study the potential short- and long-run impacts of immigration on wages, growth, and inequality. In particular, these essays develop economics insights into rational choices of immigrants such as their residential and occupational choices; their human capital development such as English language skills; and take these insights to guide empirical work with US data.

How does the language barrier affect the assimilation process of immigrants? Chapter 1, *Immigrants, Language Barrier, and Assimilation,* proposes a model of location choice, time allocation, and learning-from-others to study the three dimensions of immigrant assimilation: economic, linguistic, and spatial. The model shows that the characteristics that govern the newcomers' experience of communication friction and their ability to overcome the language barrier are also determinants of their location choices at arrival – into or away from their ethnic enclave. Subsequently, the composition and quality of agents in the newcomers' residential area influence their incentive to acquire the native language as well as determine their post-migration productivity growth, in the short- and long-run. The model generates a set of predictions about the variation in newcomers' assimilation process that conforms to immigrant assimilation patterns in the US. For US immigrants who migrated from the 1970s to the 1990s, the fraction of a newcomer's residential population are immigrants of the same ethnicity declines monotonically with both his English proficiency and educational attainment. Furthermore, an increase of one level of English proficiency at arrival (on a scale from 0 - cannot speak English to 3 - speak English very well or speak English only) is associated with a real wage gain of more than 7 percent ten years after. This number is even higher among newcomers who either reside away from the ethnic enclave

or are from non-English-speaking countries where the English language is not frequently used.

What are the long-term economic impacts of communication friction in a diverse economy? Chapter 2, *Diversity, Growth, and Inequality,* provides economic insights on the long-run impact on economic growth and income inequality of immigration. This chapter proposes a simple knowledge diffusion model to study the economic impacts of the social friction that arises in meetings between agents from different cultural backgrounds or "groups." In the long-run, social friction leads to a lower aggregate human capital accumulation rate and a constant productivity gap between minority and majority groups. The productivity gap is the source of the higher income inequality, measured by the Gini index. The economic value of a group's cultural identity is scaled to the group's relative population size. When the groups segregate geographically, the impact of social friction on growth rate is smaller, yet, the productivity gap persists in the long run. In the immigration scenario, when communication friction arises because of the lack of native-specific human capital of immigrants at arrival (English language skill for example), immigrants invest less than the social optimal effort into acquiring such human capital as they do not take into account the external gains of natives and other immigrants encountering them. When communication friction arises from the difference in agents' cultural identity and agents can choose to adopt cultural identity (religion for example), multiple equilibria could arise: stable majority-minority equilibria and a more-evenly-distributed population but unstable equilibrium, depending on agents' prior belief about their expected lifetime earnings and/or initial conditions.

Chapter 3, *Regional Impacts of Immigration,* utilizes the spatial economics framework to study immigrants' location choices and economic impacts of immigration across regions and industries. In this model, residential and occupational choices of immigrants are driven by their comparative (dis)advantage compared to natives due to the differences in their

endowed skills at arrival or their tastes for locations. Through price effect, their residential and occupational choices influence the natives' decision to relocate to different regions and switch to occupations where they have comparative advantages over immigrants. In the dynamic setup, the difference between immigrants and natives' occupational choices is smaller with the duration of immigrants' stay as they acquire the native-specific skill that closes the comparative (dis)advantages between them and the natives. With US data, this paper shows the difference in residential and locational choices of immigrants compared to natives, captured by the dissimilarity index, increases with immigrants' age but decreases with their English proficiency, years of education, and having US education. Furthermore, the occupational dissimilarity indices decline with immigrants' duration of stay, especially among younger immigrants.

CHAPTER 1

IMMIGRANTS, LANGUAGE BARRIER, AND ASSIMILATION

1.1 Introduction

After the end of the National Origins Formula in 1965, the size and composition of immigrant inflow has changed significantly. Since then, the foreign-born population has grown almost five-fold and has tripled its population share. The majority of recent newcomers are from non-English speaking countries and are not proficient in English at arrival. In 2017, the foreign-born population accounted for 15 percent of the US population (44.5 million), and approximately 48 percent reported that they could not speak English very well[1]. For these new immigrants, the language barrier is often considered one of the major obstacles to their assimilation process. This paper studies the mechanism how the language barrier affects the assimilation process of immigrants – from their location choices at arrival to their post-migration English proficiency improvement and wage growth.

To study the effect of the language barrier on newcomers' assimilation process, I develop a partial equilibrium model has two components, a host economy and a small group of newcomers who migrate to this economy. In the host economy, natives and pre-existing immigrants accumulate productivity through random encounters (Lucas 2009). The learning opportunity comes less often in a cross-group meeting between a native and a pre-existing immigrant because of the language barrier between them (Le 2019). The newcomers who migrate to this economy differ by their education, initial English proficiency, and initial productivity. On arrival, they choose to reside in the ethnic enclave or away from the enclave where they encounter the pre-existing immigrants more or less frequently, maximizing their lifetime utility (Ahlfeldt et al. 2015). After settling down, they allocate their

[1]2017 Current Population Survey and American Community Survey.

time between working and learning English (Ben-Porath 1967; Lucas and Moll 2014), and acquire productivity through random encounters the pre-existing immigrants and natives. Newcomers with higher education can learn faster, and those with better English have better odds at learning from natives.

The model shows that newcomers with lower education and/or English proficiency at arrival are more likely to locate in an ethnic enclave, where they have a higher present value of earnings. From their perspective, the ex-ante value of random encounters with natives is less valuable at arrival and in the near future than that with the pre-existing immigrants of the same ethnicity, even when the pre-existing immigrants are less productive than their native counterparts. Furthermore, the lower the newcomers' education and/or English proficiency, the lower the ex-ante value of their interactions with natives is. That leads to the monotonic relationship between the newcomers' incentive to reside in the enclave and both their education and English proficiency.

Another factor that influences the location choices of newcomers is the quality of the pre-existing immigrants residing in the enclave. As the pre-existing immigrants become less productive compared to their native counterparts, the more recent newcomers are less likely to reside in the ethnic enclave, regardless of their education or initial English skill. By design, when the pre-existing immigrants' aggregate productivity lowers, the quality of the learning source for enclave residents declines far more than that outside of the enclave and so does the ex-ante present value of living in the enclave compared to that outside. As a result, fewer newcomers are willing to reside in the enclave. The model suggests that if the influx of newcomers from less developed nations is sufficiently large that would lead to the decline in the quality of the learning source at the ethnic enclave, the more recent newcomers find it optimal to settle elsewhere. In the past 30 years, there has been an increasing trend of immigrants in the US settling outside of what would generally be considered the traditional immigrant destinations. In the 2000s, the foreign-born population grew by 24 percent overall but by roughly 49 percent or more in 14 states: South Carolina, Alabama, Tennessee, Delaware, Arkansas, South Dakota, Nevada, Georgia, Kentucky, North Carolina, Wyoming, Idaho, Indiana, and Mississippi – many of which had little prior experience with the foreign-born (Terrazas 2017). This came not too long after the end of National Origins Formula in 1965 and the significant increase of the immigrant inflow into the US.

Location choices then play an important role in newcomers' post-migration human capital development process. Among newcomers with the same education and English proficiency, enclave residents encounter natives less often, and hence, they spend less time learning English, and more time working. In the short-run, these newcomers experience less communication friction and more productivity growth. In the long-run, however, they are less fluent in English and if the pre-existing immigrants are less productive than the natives, the new enclave residents are also less productive. That is because their learning source, the mixture of more frequent encounters with the pre-existing immigrants and less frequent encounters with natives, is less productive than that of those residing away from the enclave. Their earnings might be higher in the short-run but lower in the long-run. On the other hand, among newcomers residing in one location, those with higher English proficiency at arrival and/or pick up English more quickly experience less communication friction in random encounters with natives. As a result, they experience more productivity growth by imitating the more productive natives. Under the assumption that newcomers with higher education can learn faster (either productivity or English skills but not necessarily both), they have a higher incentive to learn English, improve English at a faster pace, and also experience less communication and more wage growth.

As the pre-existing immigrant population grows, it has heterogeneous effects, in sign and magnitude, on newcomers' productivity growth and welfare. When the pre-existing immigrants are less productive than their native counterparts, the effects are more likely to be favorable to newcomers with lower education and/or English proficiency compared to those with higher education and/or English proficiency. From the perspective of the less able newcomers, the pre-existing immigrants are more useful sources of learning when they arrive, and that remains so for a long duration of time afterward. When the pre-existing immigrant population is larger, they encounter the pre-existing immigrants more frequently, can acquire productivity more quickly, and hence, have higher welfare. For newcomers with higher education and/or English proficiency, the language barrier is a small obstacle which they quickly overcome as they become fluent in English. Subsequently, they benefit from more frequent interactions with the more-productive natives. The growing pre-existing immigrant population that crowds out the interactions with natives hence harms their productivity growth and welfare. The bigger pre-existing immigrant population lowers the incentive of all newcomers to learn English, regardless of their characteristics.

Taking the model to the data, at each cross section of the US census 1980-2010, I show the fraction of the population in an immigrant's residential area are of the same ethnicity declines monotonically with his education and English proficiency. This observation holds when restricting sample to those recently migrated within five years before the survey. In the reduced-form analysis with a pseudo-panel created by grouping immigrants by their characteristics, I find that an increase of one level of initial English proficiency at arrival is associated with a real wage gain of more than 8 percent ten years after arrival. For immigrant cohorts who reside in regions with fewer immigrants of the same ethnicity, this number is almost 11 percent. I also find a negative and statistically significant association between the newcomers' English proficiency improvement and the fraction of same ethnic immigrants. As expected from the model's prediction, after controlling for the newcomers' improved English proficiency, there is no statistically significant association between the newcomers' wage growth and the fraction of immigrants of the same ethnicity residing in the newcomers' residential area.

The theoretical intuition and empirical evidence suggest potential problems in recent cross-sectional studies in the immigration literature. In particular, cross-sectional studies that use instrumental variables to correct the measurement error and endogeneity problem arise when estimating the return English proficiency to earnings such as Bleakley and Chin (2004) and Dustmann and Van Soest (2002) might not identify causal relation. In the former paper, the interaction between age-at-arrival and having a non-English speaking country of origin was used as the identifying instrument under the assumption that the non-language effects are the same for newcomers from English- and non-English speaking countries. As shown, the rate at which the newcomers can acquire the English language skills does not only affect their location choices at arrival, their learning source but also determines their experience of communication friction and productivity growth. As a result, this instrument likely to picks up the unobserved component of human capital – says the knowledge of municipal laws and customs that would give an edge to earnings, for example. In the latter paper, the lead-and-lag measurement of the native language skill (German in this case) was used. As argued, the native language skill of newcomers, especially its lag, is a determinant of their current unobserved productivity and hence, does not satisfy the standard exogeneity assumption of an instrument.

Related literature and contributions The major contribution of this paper is the study of the role and importance of newcomers' language skill in their post-migration human capital development and wage growth. This paper views language skill as a "gateway" human capital that enables newcomers to quickly acquire useful knowledge from their new environment. It complements the literature's existing view on English proficiency as an input of production (Chiswick and Miller 1992; Lazear 1999; Dustmann and Van Soest 2002; Bleakley and Chin 2004).

This paper offers a different perspective on immigrants' location choices that is in contrast to Borjas (1998). This paper shows the less able newcomers sort toward the enclave and the ethnic enclave is a transitional point at which newcomers can seamlessly assimilate to the host economy. Furthermore, it provides another economic rationale of immigrants' location choices and extends the literature's findings (LaLonde and Topel 1997; Borjas 1998; Bauer, Epstein, and Gang 2005; Damm 2009; Bleakley and Chin 2010) in the sense that it shows and explains the monotonic relation between the tendency of an immigrant to reside in an enclave and his education and English proficiency.

Finally, this paper offers a different perspective how the pre-existing immigrant population affects the newcomers' location choice at arrival as well as their post-migration human capital development. While agreeing with Lazear 1999; Borjas 2000; Borjas 2015 that the pre-existing immigrant population negatively affects newcomers' incentive to learn English, this paper shows that the growing pre-existing immigrant population could have a mixed effects on skill accumulation and welfare of the newcomers, more likely to be beneficial for less able ones. The theoretical results conform to the findings of Edin, Fredriksson, and Aslund 2003; Beaman 2011; Warman 2007 who found living in the enclave have mixed effects on the newcomers' economic outcomes.

Paper outline Forward, I describe the model and discuss the theoretical results in section 2. Section 3 documents the variation in immigrant assimilation and the determinants of the assimilation process under the guidance of the proposed theory. Section 4 concludes.

1.2 Model

In this section, I set out a partial equilibrium framework that builds on my earlier theoretical work (Le 2019) to study the newcomers' location choice, their incentive to learn English, and their subsequent wage growth. I briefly discuss the relevant materials below.

Setup

The host economy Consider the host economy with heterogeneous agents identified by their type and productivity $(i,z) \in \{a,b\} \times R^{+}$. Type a agents are natives and type b are pre-existing immigrants with the population of π^a and π^b respectively. In this economy, the pre-existing immigrants are the minority group, $\pi^b < \pi^a$. Each agent randomly meets others according to a continuous arrival process at rate α, and upon a meeting, he has an opportunity to imitate the other's productivity[2]. The likelihood that a native and an immigrant can imitate the other's productivity upon a random meeting is less one due to the language barrier between them. The inverse of this likelihood is defined as communication friction.

The host economy's productivity distribution of type a evolves according to the following forward equation

$$\partial_t \log F(z,a,t) = -\alpha\pi^a \{\log\pi^a - \log F(z,a,t)\} - \alpha\pi^b p \left\{\log\pi^b - \log F(z,b,t)\right\} \quad (1.1)$$

that is the change in the scaled productivity distribution is equal to the outflows of agents upon successful meetings with more productive agents. The analogous forward equation can be derived for type b.

[2]That is in a very small Δ fraction of time, an agent will have $\alpha\Delta$ meetings that include $\alpha\Delta\pi^a$ meetings with another type a agents and $\alpha\Delta\pi^b$ meetings with type b agents.

The host economy is assumed to be on a balanced growth path (BGP). That is on the log scale, the productivity distributions of natives and pre-existing immigrants constantly shift to the right at a constant rate of γ while retaining their shape. As shown in Le 2019, the BGP productivity distributions of type a and b are Frechet distributions with scale parameters k^a, k^b and shape parameter θ. For type $i \in \{a,b\}$, the productivity CDF at time t is given by $F(z,i,t) = \pi^i \exp\left[-k^i e^{\frac{\gamma}{\theta}t} z^{-\frac{1}{\theta}}\right]$ and the detrended CDF is $\Phi^i(x) = \pi^i \exp\left[-k^i x^{-\frac{1}{\theta}}\right]$ where $x = ze^{-\gamma t}$ is the detrended productivity. The link between the two CDFs is that for all time t, the fraction of type i agents with productivity $z' < z$ at time t, $F(z,i,t)$, is equal to the fraction of type i agents with productivity $x' < x = ze^{-\gamma t}$ at time 0, $\Phi^i(x)$. Along the BGP, the pre-existing immigrants are less productive than their native counterparts, $k^b < k^a$, because the communication friction has asymmetric effects on pre-existing immigrants and natives, stronger on the former whose population size is smaller.

Natives and pre-existing immigrants reside at two locations that are denoted by $\{a,b\}$. Here, location b represents the ethnic enclave where agents of type b live and location a represents "away from the enclave" where agents of type a live. So $\{a,b\}$ index both locations and agent types. Agents in each location interact with each other more often. Let q_r^j denote the fraction of time an agent in location r randomly encounters an agent of type j, then $q_a^a > \pi^a > q_b^a$ and $q_b^b > \pi^b > q_a^b$.

The newcomers A small group of newcomers, type b, migrate to this economy. Newcomers differ by their initial productivity z, English proficiency E, educational attainment e, and permanent taste for location $\{Z_r\}$. On arrival, they have a one time choice of where to live $r \in \{a,b\}$. Once settled down, they allocate their unit of time between working and learning English to ascend the English proficiency ladder. Newcomers randomly encounter agents from location $r \in \{a,b\}$ and have the opportunity to imitate their productivity if they wish to do so.

Here, z, e and E are state variables. Productivity z is a continuous variable that changes when an agent imitates another agent's productivity. Education e is a time invariant discrete variable. English proficiency E is a discrete variable that changes when an agent improves his English proficiency. To match the US data, I assume there are five levels of educational attainment $e \in \{$1– not graduated high school, 2– graduated high school, 3– some college, 4– graduated college, 5– master's degree and above$\}$ and four levels of English proficiency $E \in \{$0– cannot speak English at all, 1– speaks English but not well, 2– speaks English well, 3– speaks English only or speaks English very well$\}$. To avoid complication, I discuss the taste for location $\{Z\}$ in the location choice problem later.

A newcomer with better English is more likely to be able to understand and imitate a native's productivity upon a meeting with him. Let $p(E)$ denote the probability that an immigrant with English proficiency E can successfully imitate a native's productivity in a random encounter. $p(E)$ then is an increasing function with respect to E. A newcomer with better English also commands a higher wage rate that is denoted by $w(z;E,r)$. For simplicity, the wage rate is given by $w(z;E,r) = z\chi(E,r)$ and is proportional to individual productivity z and a (down) scaling factor $\chi(E,r)$. Here, $\chi(E,r) \leq 1$ is a measure of how effectively an immigrant can utilize his productivity z given his English proficiency E at location r.[3]

Once settled down at location r, a newcomer allocates his unit of time between working and learning English (Ben-Porath 1967; Lucas and Moll 2014). Let s denote the fraction of time that he spends on learning English. Then according to an exponential process of rate $\eta(s;e)$ as a function of individual effort s and education e, he will jump up the English proficiency ladder from E to $E+1$. The jump rate $\eta(s;e)$ is increasing with education e and individual effort s, and has the diminishing return to scale property with respect to

[3]The wage rate is independent of education. I assume that all the productivity that immigrants acquire through their education is absorbed into z and E.

individual effort. For simplicity, $\eta(s;e) = N(e)s^{\sigma}$ with $\sigma < 1$ while $N(e)$ is an increasing function with respect to e.

In this model, the state variable e represents the methodological human capital that newcomers accumulate through their education. A newcomer with higher education has gone through rigorous training and therefore, has better fundamental skills that allow him to learn faster. Therefore, individual effort s fixed, a newcomer with higher education can acquire English more quickly so $N(e)$ increases with respect to education e. A newcomer with higher education also has more opportunities to meet other agents. The random encounter rate for a newcomer with education e is given by $\alpha A(e)$ where α is governed by the characteristics of the host economy while $A(e)$ is an increasing function with respect to education e.[4] In general, any other factor that positively affects the ability of the newcomer to learn can be set in place of education and the later analysis applied.

Time allocation problem

Conditional on locating in r, an immigrant with educational attainment e and English proficiency E takes the evolution of the productivity distributions as given. Then, he chooses the optimal time to learn English $s(\tau)$ to maximize his expected present value earnings with discounted rate ρ:

$$V^{E}(z,t;e,r) = \max_{s(\tau)} E_t \left\{ \int_t^{\infty} e^{-\rho(\tau-t)} \left[1 - s(z(\tau),\tau)\right] w(z(\tau);E(\tau),r)\, d\tau \mid z(t) = z \right\}.$$

Since education level e and location choice r are invariant with time, I suppress the indices r,e in the value function leaving $V^{E}(z,t)$. For an incoming immigrant who cannot

[4]Newcomers with higher education e work in human capital intensive occupations and have more opportunities to learn. Later, I show that this assumption is equivalent to that where newcomers with higher education learn from more capable workers that is a segregated learning environment differs depending on the individual's education e.

speak English very well, $E < 3$, his Bellman equation is

$$\begin{aligned}\rho V^{E}(z,t) &= \max_{s} z\chi(E,r)(1-s) \\ &+ \alpha A(e)\, q_r^a p(E) \int_z^\infty \left[V^E(z',t) - V^E(z,t)\right] D_{z'} \log \frac{F(z',a,t)}{\pi^a} \\ &+ \alpha A(e)\, q_r^b \int_z^\infty \left[V^E(z',t) - V^E(z,t)\right] D_{z'} \log \frac{F^b(z',t)}{\pi^b} \\ &+ \eta(s;e)\left[V^{E+1}(z,t) - V^E(z,t)\right] \\ &+ D_t V^E(z,t).\end{aligned} \tag{1.2}$$

where the components are the earnings flow from working, the benefit flow of random meetings with natives and pre-existing immigrants, the benefit flow from jumping the English proficiency ladder, and the change with respect to time of his present value of earnings.

To solve this system of the Bellman equations, I first use a change of state variable of $x = ze^{-\gamma t}$ where γ is the growth rate of the productivity distributions, to make them time-invariant and hence, detrended the system of Bellman equations. Since the productivity distributions of natives and pre-existing immigrants are Frechet with the same tail shape θ, the Bellman equation is further simplified to

$$\begin{aligned}\rho V^{E}(x) = \max_{s} x\chi(E,r)(1-s) + \alpha \int_x^\infty \left[V^E(x') - V^E(x)\right] D_x \log \Phi(x') \\ + \eta(s;e)\left[V^{E+1}(x) - V^E(x)\right] - D_x V^E(x)\, x\gamma + V^E(x)\,\gamma.\end{aligned} \tag{1.3}$$

where x is the detrended productivity, and $\Phi(x') = \exp\left[-k(E,e;r)\, x^{-\frac{1}{\theta}}\right]$ is the effective productivity distribution with an effective scale parameter

$$k(E,e;r) = A(e)\left[q_r^a p(E)\, k^a + q_r^b k^b\right]. \tag{1.4}$$

Proposition 1 The model is isomorphic to one in which newcomers learn at the same rate but have different learning sources. Those with higher education and/or English proficiency learn from more productive agents.

Proof The Bellman equation (1.2) is equivalent to (1.3) for all newcomers, and $k(E,e;r)$ is an increasing function with respect to education e and English proficiency E. QED.

All else being equal, a newcomer with higher education e and/or English proficiency E also acquires productivity more quickly since he has more opportunities to learn. One way to look at it is that he is more likely to be able to acquire a native's productivity now and in the future because he has more successful random encounters: higher $A(e)$ and $N(e)$. Proposition 1 implies that it can also be interpreted as he has more opportunities to learn because the effective probability that he encounters a more productive agent, $1-\Phi(x)$, is higher because the effective scale parameter $k(E,e;r)$ increases with respect to education e and English proficiency E. On the other hand, as in Le 2019, the rate of productivity acquisition declines with the newcomers' initial productivity x since he is less likely to encounter agents who are more productive than him, $1-\Phi(x)$ decreases with respect to x.

I solve the Bellman equation (1.3) with the finite difference method described in Lucas and Moll 2014 with the appropriate boundary conditions. The boundary conditions, their reasoning, and the computation method are presented in appendix A.

Incentive to learn English

The optimal time that an immigrant spends acquiring English s^* satisfies the first-order condition (FOC)

$$x\chi(E,r) = \frac{d}{ds}\eta(s^*;e) \times \left[V^{E+1}(x) - V^{E}(x)\right] \tag{1.5}$$

where the left-hand side (LHS) is the marginal cost of time a newcomer spends learning English and the right-hand side (RHS) is the marginal benefit of learning English. In this model, the benefit of learning English comes from the gain in the wage rate and the gain from being more likely to be able to learn from natives upon random encounters. The gain in the wage rate comes from the increase in the effectiveness of a newcomer in utilizing his productivity z, from $\chi(E,r)$ to $\chi(E+1,r)$. On the other hand, the gain in learning from natives goes through the second benefits flow in the Bellman equation (1.2) or the effective scale parameter (1.4). Both are components of the term $V^{E+1}(x) - V^{E}(x)$ on the RHS.

First, newcomers' incentive to learn English increases when the return from English skills to the wage rate is higher. The percentage return on the wage rate for an increase of one level of English proficiency is given by the ratio $\frac{\chi(E+1,r)-\chi(E,r)}{\chi(E,r)}$. The numerator of this ratio is inside the term $V^{E+1}(x) - V^{E}(x)$ of the RHS of equation (1.5), and the ratio shows up on the RHS when dividing both sides by $\chi(E,r)$. When this ratio is higher, the marginal benefit of learning English is higher and the newcomer spends more time learning English.

On the other hand, $\chi(E,r)$ can be specified as a function of education e as well such that

$$\frac{\chi\left(E+1,r,e'\right)-\chi\left(E,r,e'\right)}{\chi\left(E,r,e'\right)} < \frac{\chi\left(E+1,r,e\right)-\chi\left(E,r,e\right)}{\chi\left(E,r,e\right)}$$

for $e' < e$. Immigrants with lower education e are more likely to work in occupations that require less communication skills, for example. Therefore, those with lower education will put less effort into learning English because by the nature of their occupations, the from being fluent in English to the wage rate is not significant.

Second, the size of the pre-existing immigrant population has a negative effect on newcomers' incentive to learn English while the aggregate productivity of natives has a positive effect. When the frequency of interactions with pre-existing immigrants q_b increases and those with natives q_a decreases, the marginal gain of being fluent in English is lower since the opportunity to use English (in random encounters with natives) comes less often. From the expression of the effective scale parameter in (1.4), the marginal gain in learning from being proficient in English is a function of $q^a k^a [p(E+1) - p(E)]$ that decreases with q^a. As a result, newcomers who reside in location b or migrate to a host economy with a higher fraction of pre-existing immigrants π^b, which affects q_b^j, put less effort into learning English. Since the term $q^a k^a [p(E+1) - p(E)]$ increases with k^a, newcomers have a higher incentive to learn English when natives are more productive since the ex-ante value of random encounters with natives is more valuable.

Third, newcomers with higher education e have a higher incentive to learn English and therefore, put more effort s^* into doing so. From the FOC equation (1.5), the marginal benefit of learning English increases with respect to education e for two reasons. First, because education and individual effort are complementary in learning English[5], a newcomer with higher education also has a higher marginal return for his time to learn English and therefore, spends more time doing so. Second, since the random meeting rate is increasing with respect to education, a newcomer with a higher education encounters natives more often after settling down and therefore has a higher gain from being fluent in English.

Since the model is isomorphic to a model with a segregated learning environment (proposition 1), the second reason can be interpreted as the newcomer with higher education has a higher gain from being fluent in English as he is more likely to encounter natives who are more productive than him. As a result, he puts more effort into learning English so that he can effectively learn from them.

Finally, newcomers with higher initial productivity x have a lower incentive to learn English. For them, the value of learning from natives is lower, and therefore, the marginal

[5] $\frac{\partial}{\partial e}\left(\frac{\partial \eta}{\partial s}\right) > 0$ under the model's specification.

benefit of learning English converges to the marginal gain in their wage rate. On the other hand, the marginal cost of learning English is higher since their wage rate is higher. As a result, they put less effort into doing so compared to those with lower initial productivity.

Location choice problem

To generalize the newcomers' location choice problem, I use the standard recipe in the spatial economic literature. Similar to Ahlfeldt et al. (2015), a newcomer's utility flow of settling in location r is a product of his permanent taste for location Z_r and his earnings flow $w(z;r)(1-s^*)$. The permanent taste for location captures other unobserved factors that drive newcomers' location choice behavior. His location choice problem is to choose location r where he would have the highest ex-ante value of lifetime utility, which is given by $Z_r V^E(x)$ since Z_r is a constant and can be pulled outside of the Bellman equation in (1.3). The taste for location $\{Z_r\}$ are drawn from independent Frechet distributions with the same tail shape ε but different scale parameters G_r, $Z_r \sim \exp\left\{-G_r Z^{-\frac{1}{\varepsilon}}\right\}$. The fraction of newcomers that reside in location r is then given by[6]

$$\Pi_r(x,E,e) = \frac{G_r V^E(x;e,r)^{\frac{1}{\varepsilon}}}{G_a V^E(x;e,a)^{\frac{1}{\varepsilon}} + G_b V^E(x;e,b)^{\frac{1}{\varepsilon}}} \tag{1.6}$$

where $V^E(x;e,r)$ is the ex-ante present value of earnings of a newcomer with productivity x, education e, and English proficiency E at location r.

The fraction of newcomers residing away from the ethnic enclave is higher when the ratio between the ex-ante present value of earnings $\frac{V^E(x;e,a)}{V^E(x;e,b)}$ is higher. In the extreme case when there is no dispersion in taste preference (ε goes to zero), the location choice of

[6]Follows the derivation of Eaton and Kortum 2002; Ahlfeldt, Redding, Sturm, and Wolf 2015.

newcomers is completely determined by the ex-ante present value of earnings,

$$r = \arg\max_{r\in\{a,b\}} \left\{ V^E(x;e,a), V^E(x;e,b) \right\}.$$

On the other hand, when the tail is very thick (ε goes to infinity), the location choices of newcomers are entirely driven by the taste preference with $\pi_r = G_r/(G_a + G_b)$.

In order to analyze the location choice of incoming immigrants, we need to know how the ex-ante present value of earnings changes with respect to newcomers' characteristics. Since there is no closed-form solution to the Bellman equations system, the (only) viable option is to run a numerical example and deduce economic reasons from the results.

Consider a numerical example with the following specifications. Regarding the random encounter rate and jump rate across the English proficiency ladder $\alpha A(e)$ and $N(e)s^{\sigma}$, assume that $A(e), N(e)$ are concave, increasing functions with respect to education e. $\sigma < 1$ so that the jump rate is concave in individual effort s. The probability of a successful meeting in a random encounter with a native, $p(E)$, is a concave, increasing function with respect to English proficiency E. I set $\chi(E)$ so that wage rate is an increasing function with respect to newcomers' English proficiency but are the same across locations. Regarding the location taste preference, I set $G_a = G_b = 1$ so that on average, newcomers do not prefer one over the other. In the host economy, I set $k^b < k^a$ so that the pre-existing immigrants are less productive than their counterpart natives. The frequency of interactions with pre-existing immigrants in the ethnic enclave is q_b^b is higher than that in location a is q_a^b.[7]

In figure 1.1, I plot the fraction of newcomers who reside in the ethnic enclave. I hold education fixed in each LHS panel and English proficiency fixed in each RHS panel. First, we can see the fraction of newcomers who reside in the ethnic enclave converges to 0.5

[7]The specific numbers in this example are not important for the qualitative analysis. In the next section, I calibrate the model with US census data and discuss the results then.

as newcomers' productivity increases, regardless of their education or English proficiency. This is because for very productive immigrants, the present value of learning from others is very small compared to the present value of working. Hence, the major component that drives their location choices is the present value of the earnings from working, which are identical across locations. And as a result, the fractions of newcomers located in a and b converge to 0.5 as their productivity increases. Furthermore, for newcomers with characteristics (x, E, e), if this fraction is higher than 0.5, they have a higher present value of earnings when residing in the enclave since the utility from location taste for locations a and b are, on average, the same. If it is lower than 0.5, then it is the opposite.

Regardless of their initial productivity, newcomers with higher education and/or English proficiency at arrival are more likely to reside away from the enclave, $\pi_a(x, E+1, e) > \pi_a(x, E, e)$ and $\pi_a(x, E, e+1) > \pi_a(x, E, e)$. In figure 1.2, I plot the fraction of newcomers who reside in the ethnic enclave as a function of education and English proficiency while holding their initial productivity fixed. The component of the present value of earnings that drives newcomers' location choices is the difference in the value of learning from others now and in the future in each location. For newcomers who are not proficient in English and/or are unable to pick up English quickly due to their lack of education, they find it optimal to locate in the ethnic enclave to learn from pre-existing immigrants. This is because the present value of their interactions with natives is actually smaller than that with the pre-existing immigrants (despite the fact that natives are more productive than pre-existing immigrants) after accounting for communication friction, the ability to learn, and the potential loss of time in learning English. The opposite is true for immigrants with higher education and/or English proficiency.[8]

[8]The rationale of the constant fraction of newcomers with $x = 0, E = 3$ (top right corner of figure 1.2) is that the location is driven entirely by the value of learning that is roughly the same regardless of their education under the assumption of the continuous arrival process.

From the location choices of newcomers (data and above results), by revealed preference, the pre-existing immigrant population has heterogeneous effects on newcomers' welfare, depending on their education and initial English proficiency. In this model, the pre-existing immigrant population's effect on the newcomers' welfare is through its effect on the newcomers' productivity growth. The effect on productivity growth can be observed through the effective scale parameter in (1.4), the change in the effective scale parameter with respect to an increase of the frequency of interactions with pre-existing immigrants is:

$$\frac{\partial}{\partial q_r^b} k(E,e;r) = A(e)\left[k^b - p(E)\,k^a\right].$$

This quantity is the marginal effect now. It can be either positive or negative that depends on how productive the pre-existing immigrants are (measured by k^b) compared to their native counterparts from each individual perspective (measured by $p(E)\,k^a$). When the pre-existing immigrants are less productive than their native counterparts, this quantity is more likely to be negative for those who are proficient in English. However, the lifetime effect is the discounted sum of all "period" effects which change over time as newcomers become more proficient in English. Among newcomers who are not English proficient at arrival, those with higher education can acquire English quickly and hence, the effects of pre-existing immigrants are mostly negative through lifetime. For those with lower education, the effects are positive for most of their lifetime. See figure 1.3 for the illustration.

Third, the fraction of newcomers who reside in the ethnic enclave decreases when either the pre-existing immigrant population is less productive or the native population is more productive. That is, all the curves in figures 1.1 and 1.2 shift downward when k^b decreases or k^a increases. As the pre-existing immigrant population becomes less productive, the ex-ante present value of earnings for all newcomers across locations are lower because the value of learning from pre-existing immigrants is lower across locations. How-

Figure 1.1. Fraction of newcomers reside in the ethnic enclave

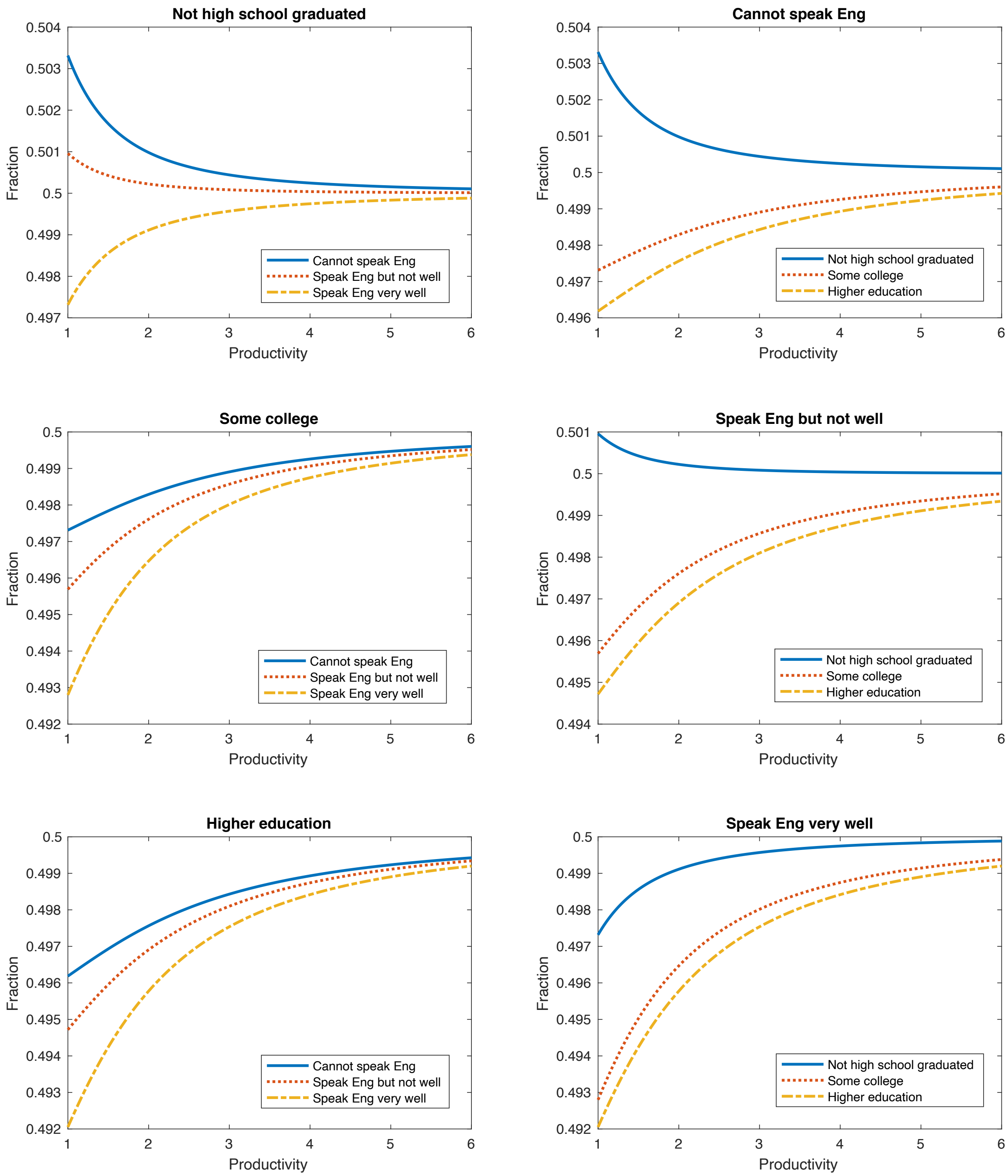


Figure 1.2. Fraction of newcomers reside in the ethnic enclave

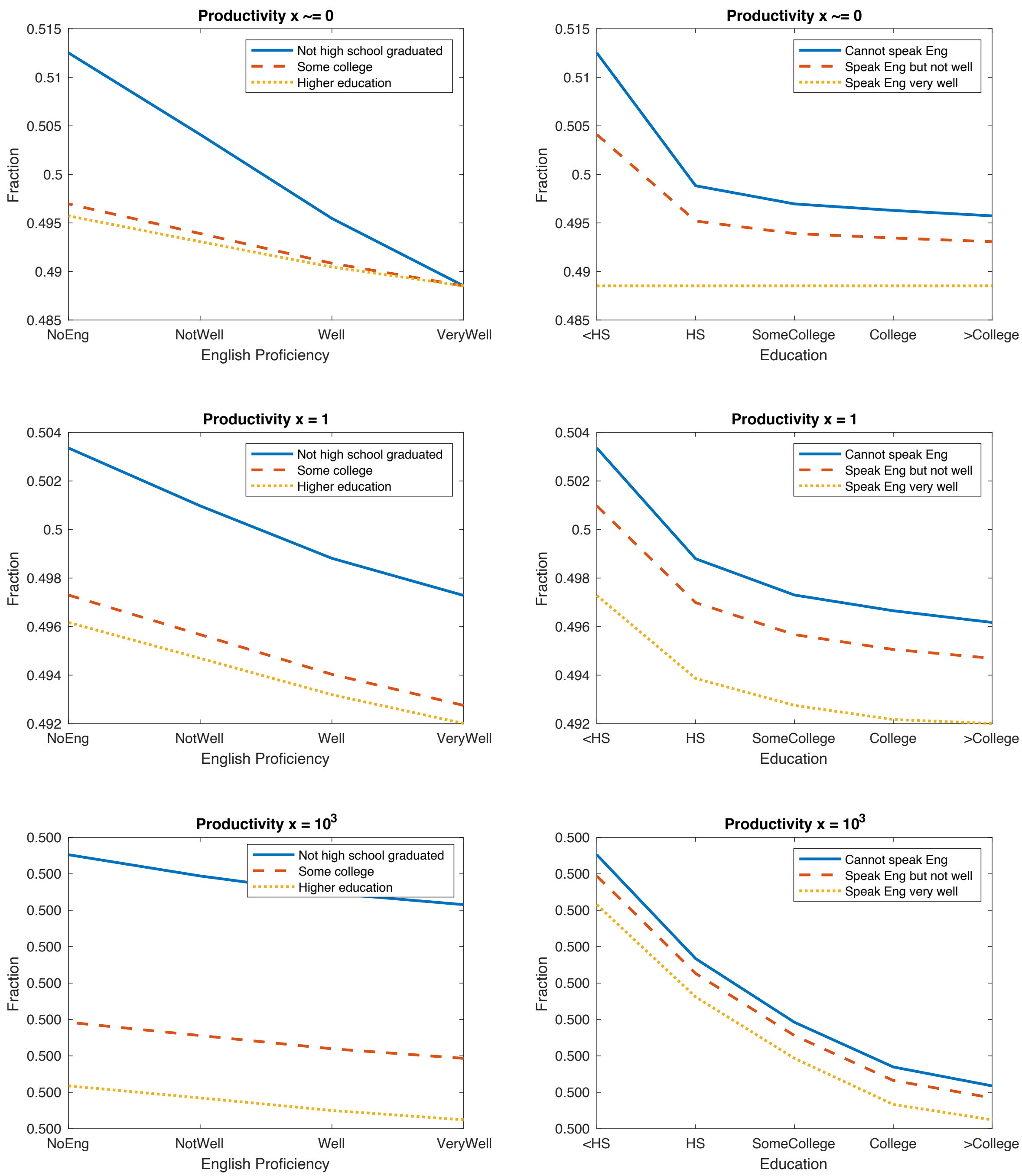
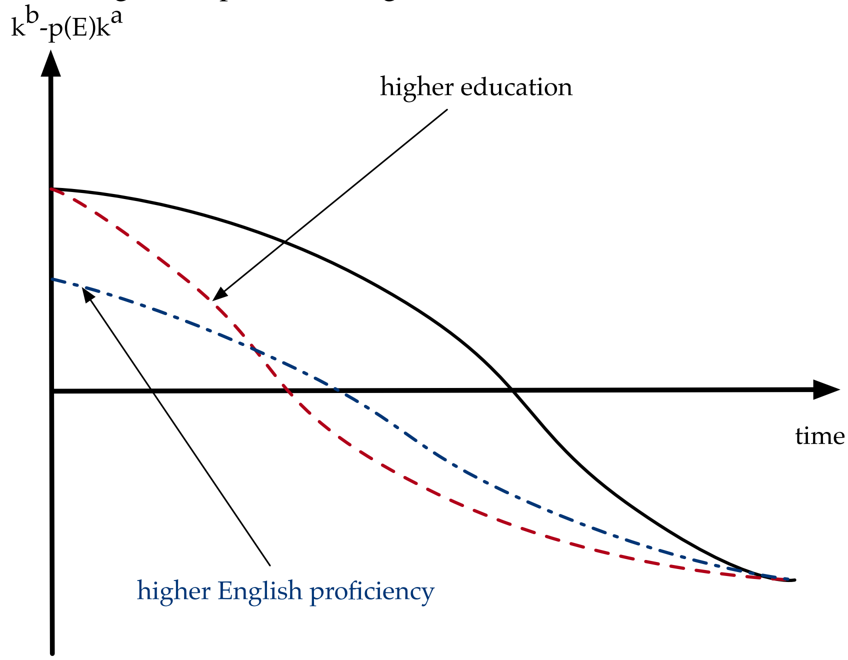


Figure 1.3. Marginal impact of immigrant stock on newcomers' value of learning



ever, the magnitude of this decline is bigger in location b since newcomers interact with pre-existing immigrants more often if they reside in the enclave. As a result, the ratio $V^E(x;e,a)/V^E(x;e,b)$ increases and the fraction of newcomers who reside away from the ethnic enclave increases.

This change can be seen directly by differentiating the effective scale parameter in (1.4) by k^b and see that

$$\frac{\partial}{\partial k^b}k(E,e;a) = \alpha A(e)q_a^b < \alpha A(e)q_b^b = \frac{\partial}{\partial k^b}k(E,e;b)$$

since $q_a^b < q_b^b$. That is the change in quality of the pre-existing immigrant population has a stronger effect on the value of learning from others for those who reside in the ethnic enclave as they interact with pre-existing immigrants more often. The asymmetry of the effects then pushes newcomers toward or away from the ethnic enclave that depends on whether the quality of the pre-existing immigrant population improves or declines. The same thing happens when natives becomes more productive.

Discussion

Education, English proficiency, and the quality of the pre-existing immigrant population all play important roles in newcomers' choices of where to live and their subsequent development of English and productivity. Holding initial productivity fixed, newcomers with higher education and/or English proficiency are more likely to locate away from the ethnic enclave. From their perspective, the language barrier is but a small obstacle compared to the potential gain in productivity that they could have from more interactions with the more productive natives. Hence, they reside away from the enclave so they can interact with natives more often. They allocate more time to learning English and consequently, they pick up English and acquire productivity more quickly. On the other hand, for newcomers with lower education and/or English proficiency, they are better off residing in the enclave, benefiting from encountering and learning from pre-existing immigrants more frequently. From their perspective, the pre-existing immigrants are productive than their native counterparts now and in the distant future. The result conforms to the location choice pattern of immigrants documented later[9].

Furthermore, the result shows that instrumental variables such as the interaction between having a non-English-speaking country and the age at arrival (Bleakley and Chin 2004) and the leads and lags of self-reported language proficiency (Dustmann and Van Soest 2002) might not be valid to identify the causal effect of English proficiency on earn-

[9]Since the fraction of pre-existing immigrants in the residential area of a newcomer with productivity x, education e, and English proficiency E is given by $\Pi_b(x,E,e)$ since in location b, 100% of the population are pre-existing immigrants. This fraction declines with respect to the newcomer's education and English proficiency.

ings. As argued earlier, initial English proficiency and English proficiency at any point in time has an important role in newcomers' productivity growth. The instrumental variable of leads and lags of self-reported language proficiency is then likely to be correlated with the unobserved component of human capital and therefore, is upwardly biased. Furthermore, the rate at which newcomers pick up English does not only affect newcomers' choices of where to live but also, their subsequent development of human capital. Therefore, the instrumental variable of having a non-English speaking country interacts with the age-at-arrival is positively correlated with the difference in unobserved productivity of newcomers from an English speaking country and those from a non-English speaking country.

Labor supply, English proficiency, and productivity growth in the short- and long-run

The results show that the initial labor participation/supply of newcomers who reside away from the ethnic enclave is lower compared to that of enclave residents. Furthermore, the human capital augmented labor supply and earnings of newcomers who reside away from the enclave are lower in the short-run but are higher in the long-run. As shown earlier, newcomers who reside away initially put more time into learning English and less time working. Therefore, their initial labor participation/supply is lower. In the short-run, since newcomers residing away from the enclave encounter natives more frequently, they experience more communication friction, and therefore, they acquire productivity more slowly and are less productive compared to enclave residents. In the long-run, as they become fluent in English and efficiently learn from natives without communication friction, they become more productive compared to enclave residents because their learning source is more productive than that of enclave residents. Consequently, their human capital augmented labor supply and earnings are higher in the long-run. To illustrate this, figure 1.4

shows the ratio of the expected productivity, English proficiency, labor supply, and earnings with respect to time after arrival between a newcomer who resides away from the enclave versus one of the same characteristics who reside in the enclave (in the numerical example).

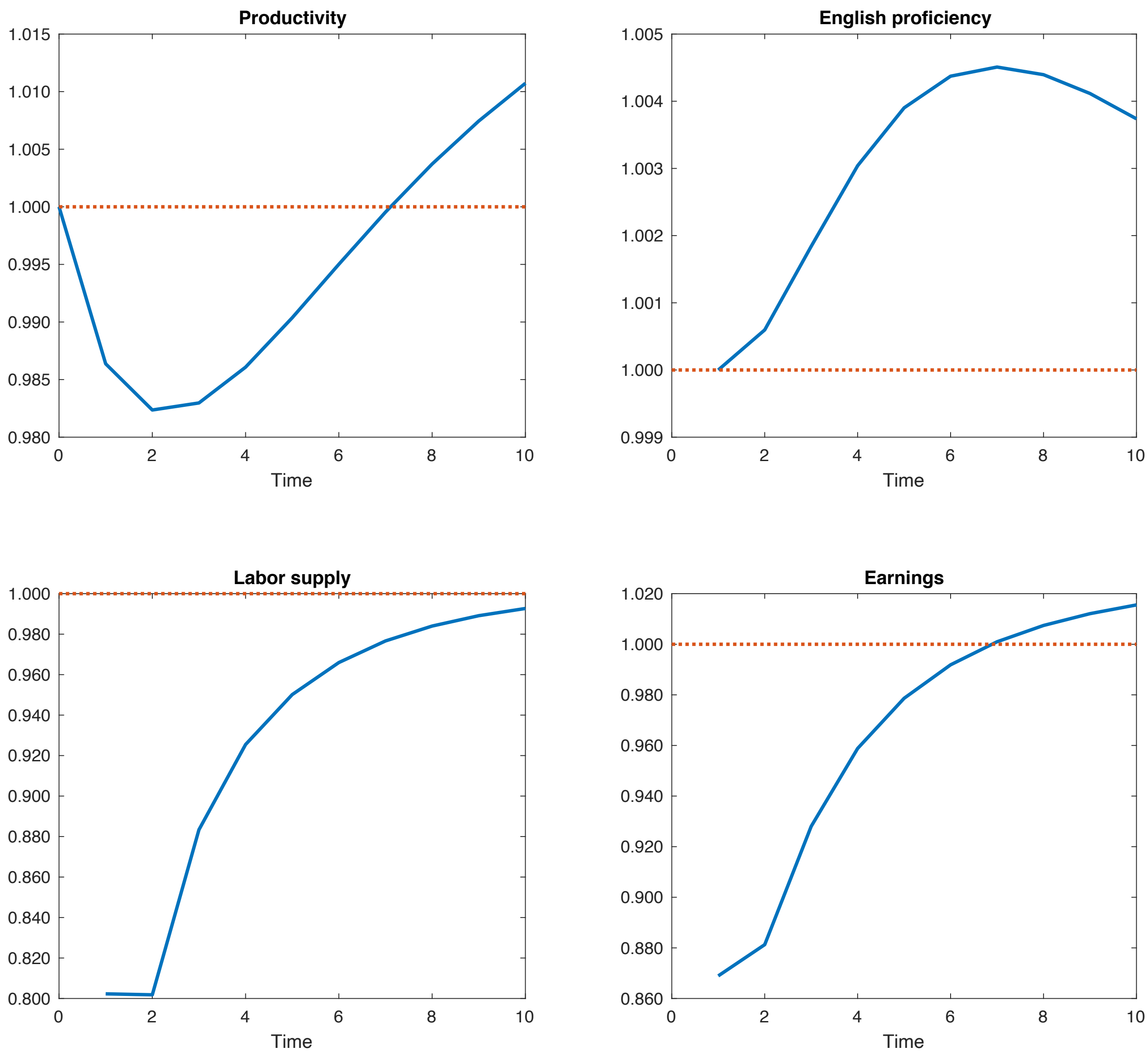
Altruism and the immigrant bargain

The results imply that immigrant households have a higher altruistic value toward their children are more likely to reside away from the ethnic enclave. The location decision of a household with young children can be modeled by combining the value functions of different initial productivity x, education e, English proficiency E, and some altruism factor ζ. The fraction of households H with characteristics $\{x, E, e\}$ whose children of characteristics $\left\{x', E', e'\right\}$ reside in location r is then

$$\Pi_r^H = \frac{G_r V^H(r)^{\frac{1}{\varepsilon}}}{G_a V^H(a)^{\frac{1}{\varepsilon}} + G_b V^H(b)^{\frac{1}{\varepsilon}}}$$

where $V^H(r) = V^E(x; e, r) + \zeta V^{E'}\left(x'; e', r\right)$ is the head of household's present value of utility. When the children's English proficiency and/or education is sufficiently high, they have a higher ex-ante present value of earnings when locating away from the enclave. Therefore, a household head who places a significant weight on his "more able" children or has a high expectation for his children is more likely to locate away from the enclave for the benefit of his children since the value of altruism then $\zeta V^{E'}\left(x'; e', r\right)$ constitutes a sizable part of the household head's utility. This example illustrates a case of the "immigrant bargain" concept: immigrants reside away from the enclave and willingly accept jobs with little prospect of advancement in exchange for a better future for their children.

Figure 1.4. Ratio of average productivity, English proficiency, labor supply, and earnings between newcomers locate away from the enclave vs. enclave residents of the same characteristics



1.3 Immigrant assimilation patterns

In this section, I document the assimilation patterns of immigrants in the US under the guidance of the model.

Data

After the end of the National Origins Formula in 1965, the composition of immigrant inflow changed significantly and the majority of more recent newcomers came from non-English speaking countries. The English-speaking-ability self assessment question was first added to the census in 1980 and hence, in this paper, I use the US census data from 1980 forward. In particular, the data are the US decennial census from 1980-2000 and the pooled 2009-2011 American Community Survey[10] (ACS), which is labeled as the 2010 census. In each cross-section, I select a sample consisting of male immigrants aged 25-64 at the time of the survey, who worked during the survey year, and were not enrolled in school. The data set was obtained from the Integrated Public Use Microdata Series (IPUMS) website in April 2016.

[10]The US census of 2010 data is not available for public use and hence, is substituted by the American Community Survey (ACS). The pooled ACS 2009-2011 is defined to be US census of 2010. The ACS is an ongoing survey by the US census Bureau. It regularly gathers information previously contained only in the long form of the decennial census, such as ancestry, citizenship, educational attainment, income, language proficiency, migration, disability, employment, and housing characteristics.

Location choices

For each immigrant from country k in the sample of 1990, I compute the fraction of the population in their PUMA (public-use micro-data area)[11] that has the same country of origin . I estimate the following regression model

$$y_{okr} = \alpha + \text{EDU}_{okr} \times \text{ENG}_{okr} + \gamma_{okr} + \varepsilon_{okr}$$

for individual o from country k who migrated to the US resided in location r (PUMA r)

- y_{okr} is the fraction of the PUMA r population that has the same country of origin as individual o, that is from country k,
- EDU_{okr} is a vector of dummy variables that indicates whether he is not graduated high school, graduated high school, some college, graduated college, master's degree and above,
- ENG_{okr} is a vector of dummy variables that indicates whether he cannot speak English at all, speaks English but not well, speaks English well, speaks English only or speaks English very well,
- γ_{okr} is a vector of age at time of survey, age-at-arrival, country of birth, and residential state dummy variables.

The regression results are reported in table 1.1. In columns 1 and 2, I report the result when interacting English proficiency with each level of education, $\text{EDU}_{oki} \times \text{ENG}_{oki}$. In columns 3 and 4, I report the result when rearranging the interaction term to $\text{ENG}_{oki} \times \text{EDU}_{oki}$. In

[11] Illinois has 79 PUMAs, California 198, Texas 134, South Dakota 6, and so on.

columns 2 and 4, I restrict the sample to immigrants who recently migrated, from 1985 to 1990. Essentially, the regression captures the average fraction of immigrants of the same ethnicity who reside in the same PUMA of a representative immigrant at each level of education and English proficiency.

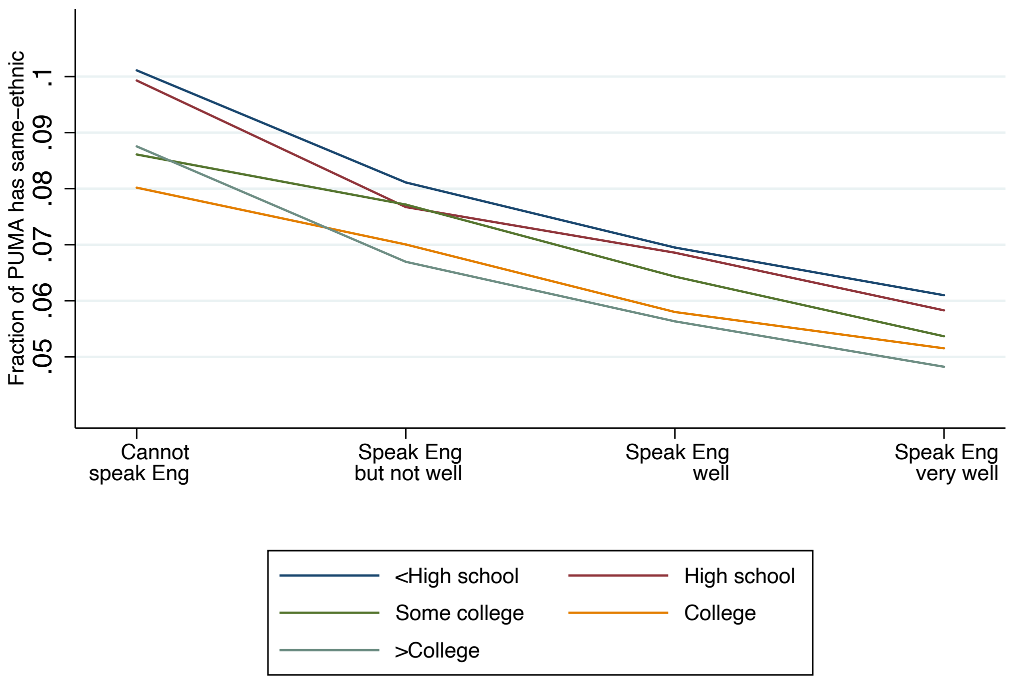
There are five levels of educational attainment {1– not graduated high school, 2– graduated high school, 3– some college, 4– graduated college, 5– master's degree and above} and four levels of English proficiency {0– cannot speak English at all, 1– speaks English but not well, 2– speaks English well, 3– speaks English only or speaks English very well}, resulting in 20 groups of educational attainment and English proficiency. Every four groups in columns 1 & 2 will have the same educational attainment and four different English proficiency levels, in ascending order. Every five groups in columns 3 & 4 will have the same English proficiency level but different educational attainment levels, in ascending order.

I illustrate these regression results in figures 1.5, 1.6, 1.7, and 1.8. For each ethnic group and age at arrival, after conditioning on the same education level, immigrants with better English proficiency reside in PUMA with a lower count of same-ethnic immigrants. Holding English proficiency fixed, those with higher education reside in a PUMA with fewer immigrants of the same ethnicity. By restricting the sample to immigrants who migrated no more than five years before the survey in columns 2 and 4 of table 1.1 (and figure 1.6 and 1.8), the results show that it is not a case of reverse causality – that is immigrants reside randomly across locations and those who do not reside in an enclave attain more education and pick up English more quickly.

In appendix B, I show the results generally hold when I either include fixed effects of income level at each state (interacting dummies of income quintile with dummies of states) or restrict the sample to each income quintile. When using the quintile subsample, the monotonicity breaks down for some cases and the magnitude of the phenomenon is lessen, especially in the first income quintile. These observations are also consistent when using US censuses of 1980, 2000, and 2010 and when excluding Mexican immigrants[12].

[12] The Mexican immigrant population constitutes a large fraction of the total immigrant population and hence, could skew the regression results.

Figure 1.5. Location choice, all immigrants



Wage growth and English proficiency improvement

Motivation for the reduced-form specification Under the assumption of the model that wage of an individual o is given by

$$W(o) = Z(o)\Psi(E(o), r)$$

that is real wage is the product of the unobserved productivity and a function $\Psi(E(o), r)$ that measures of the effectiveness at which he can translate his unobserved productivity to realized productivity in the new environment r given his English proficiency E. Taking log both sides of the expression and averaging across immigrants within the same cohort q

$$w_q = z_q + \sum \chi(E, r)\pi_q(E)$$

Figure 1.6. Location choice, immigrants migrated less than 5 years before the survey

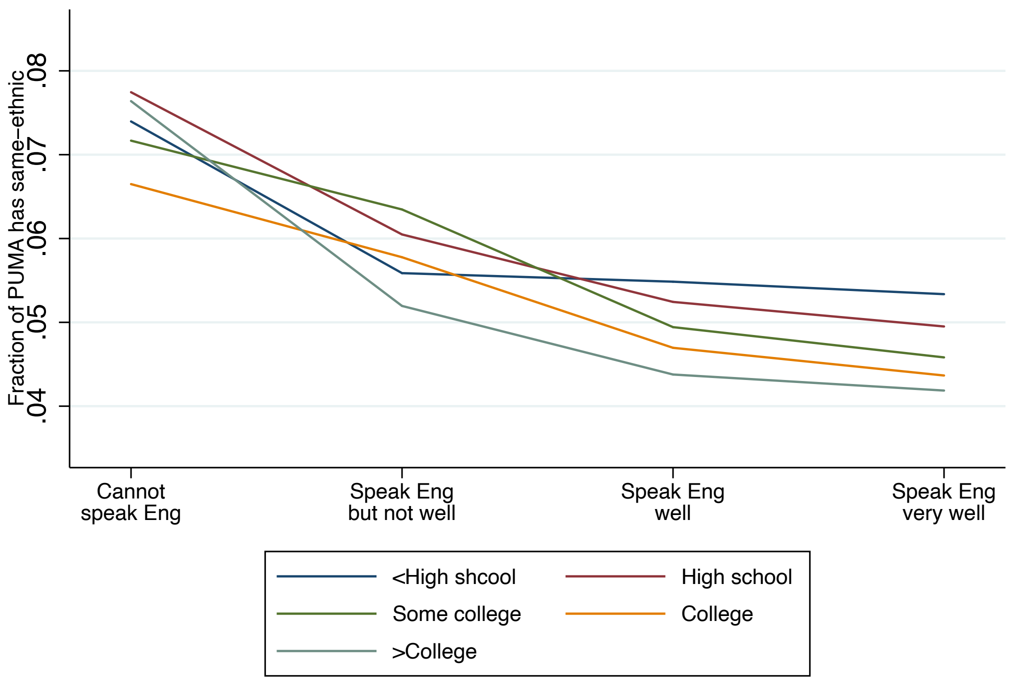


Figure 1.7. Location choice, all immigrants

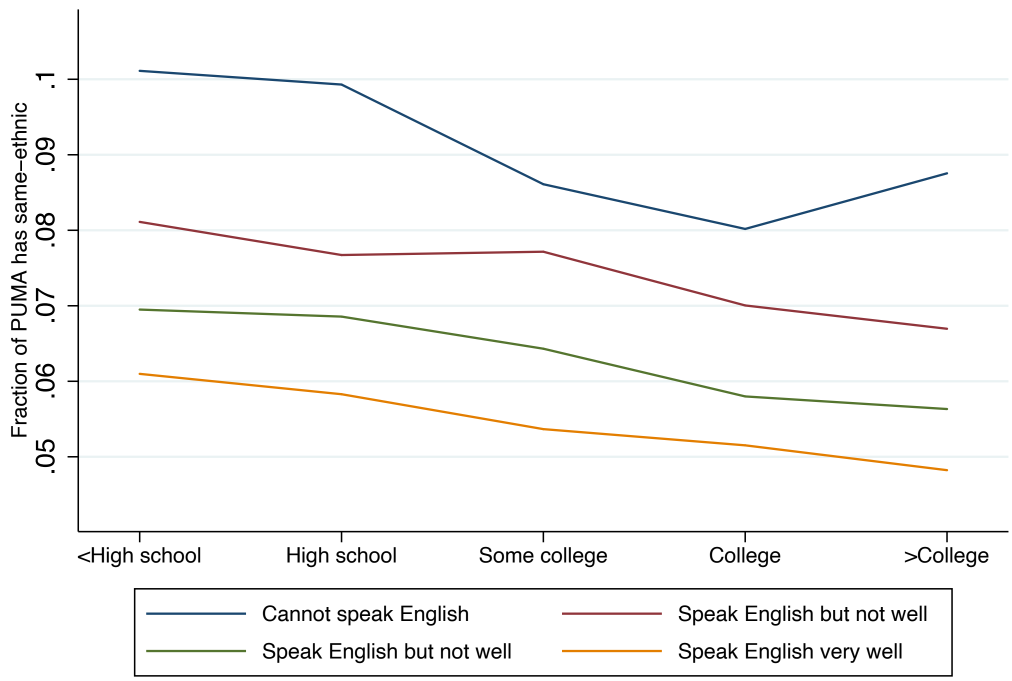


Figure 1.8. Location choice, immigrants migrated less than 5 years before the survey

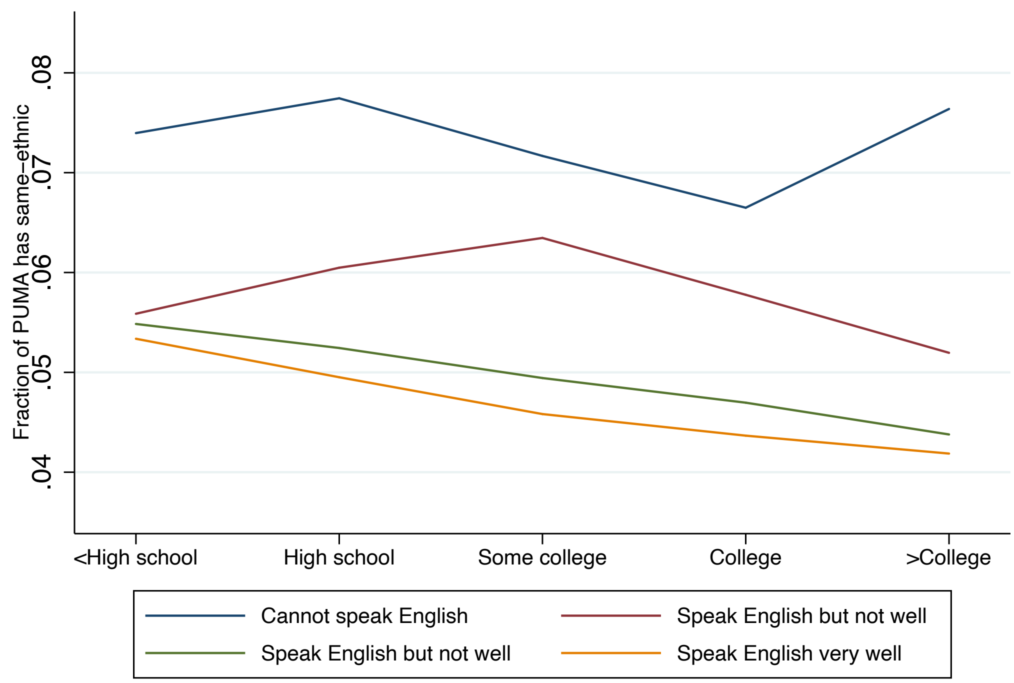


Table 1.1: Education, English Proficiency, and Spatial Assimilation

	(1) EDU × ENG	(2) EDU × ENG, Recent	(3) ENG × EDU	(4) ENG × EDU, Recent
GROUP=1	0	0	0	0
	(.)	(.)	(.)	(.)
GROUP=2	-0.0200***	-0.0181***	-0.00181	0.00349
	(-4.79)	(-5.00)	(-0.38)	(0.74)
GROUP=3	-0.0316***	-0.0191***	-0.0150**	-0.00229
	(-6.32)	(-4.83)	(-2.64)	(-0.36)
GROUP=4	-0.0401***	-0.0206***	-0.0209**	-0.00748
	(-9.07)	(-5.23)	(-3.05)	(-0.85)
GROUP=5	-0.00181	0.00349	-0.0136	0.00242
	(-0.38)	(0.74)	(-1.53)	(0.16)
GROUP=6	-0.0244***	-0.0135***	-0.0200***	-0.0181***
	(-6.36)	(-4.22)	(-4.79)	(-5.00)
GROUP=7	-0.0325***	-0.0215***	-0.0244***	-0.0135***
	(-6.56)	(-5.00)	(-6.36)	(-4.22)
GROUP=8	-0.0428***	-0.0245***	-0.0240***	-0.0105**
	(-8.19)	(-7.50)	(-5.66)	(-2.65)
GROUP=9	-0.0150**	-0.00229	-0.0311***	-0.0162***
	(-2.64)	(-0.36)	(-5.95)	(-3.85)
GROUP=10	-0.0240***	-0.0105**	-0.0342***	-0.0220***
	(-5.66)	(-2.65)	(-6.60)	(-5.09)
GROUP=11	-0.0368***	-0.0245***	-0.0316***	-0.0191***
	(-6.72)	(-6.84)	(-6.32)	(-4.83)
GROUP=12	-0.0475***	-0.0281***	-0.0325***	-0.0215***
	(-8.32)	(-8.11)	(-6.56)	(-5.00)
GROUP=13	-0.0209**	-0.00748	-0.0368***	-0.0245***
	(-3.05)	(-0.85)	(-6.72)	(-6.84)
GROUP=14	-0.0311***	-0.0162***	-0.0431***	-0.0270***
	(-5.95)	(-3.85)	(-7.14)	(-7.37)
GROUP=15	-0.0431***	-0.0270***	-0.0448***	-0.0302***
	(-7.14)	(-7.37)	(-7.87)	(-7.99)
GROUP=16	-0.0496***	-0.0303***	-0.0401***	-0.0206***
	(-7.72)	(-8.05)	(-9.07)	(-5.23)
GROUP=17	-0.0136	0.00242	-0.0428***	-0.0245***
	(-1.53)	(0.16)	(-8.19)	(-7.50)
GROUP=18	-0.0342***	-0.0220***	-0.0475***	-0.0281***
	(-6.60)	(-5.09)	(-8.32)	(-8.11)
GROUP=19	-0.0448***	-0.0302***	-0.0496***	-0.0303***
	(-7.87)	(-7.99)	(-7.72)	(-8.05)
GROUP=20	-0.0529***	-0.0321***	-0.0529***	-0.0321***
	(-8.28)	(-8.94)	(-8.28)	(-8.94)
Constant	0.101***	0.0740***	0.101***	0.0740***
	(12.86)	(13.63)	(12.86)	(13.63)
Observations	252998	44324	252998	44324
Adjusted R^2	0.406	0.417	0.406	0.417

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

where $\pi_q(E)$ is the fraction of immigrants with English proficiency E in cohort q; ω, z denotes the mean of log W, Z; and χ denotes log Ψ.

When the probability that an immigrant in cohort q can speak English very well is used as a measure of cohort q's English proficiency as in Borjas (2015), then there are two levels on the English proficiency ladder $\{0, 1\}$ and the above equation can be rewritten as

$$\begin{aligned} w_q &= \tilde{\chi}(E=0,r) + z_q + \pi_q(E=1) \times \{\chi(E=1,r) - \chi(E=0,r)\} \\ w_q &= \beta_0 + z_q + ENG_q \times \beta_1 \end{aligned} \tag{1.7}$$

where $ENG_q = \pi_q(E=1)$ is the English proficiency of cohort q and

$$\beta_1 = \{\chi(E=1,r) - \chi(E=0,r)\}$$

is the percentage wage incremental when the newcomers improve their English language skills. With the ordinal measure of English proficiency, the above expression holds if the percentage wage incremental is constant for each level of improvement – $\chi(E+1,r) - \chi(E,r) = \text{constant}$.

Hence, without loss of generality, I propose an analogous empirical specification of equation (7) : the average log wage of a cohort can be expressed as a linear combination of cohort q's human capital

$$w_\tau^q = b_0 + b_1 z_\tau^q + b_2 ENG_\tau^q + b_3 EDU^q + u_\tau \tag{1.8}$$

with $z_\tau^q, ENG_\tau^q, EDU^q$ are the log unobserved productivity, English proficiency, and educational attainment at time τ. u_τ is i.i.d. error term normally distributed. On the other hand, The model implies that the growth of log productivity can be expressed in a linear form

$$z_1^q - z_0^q = c_0 + c_1 \left(Z_0 - z_0^q\right) + c_{21} ENG_0^q + c_{22} \left(ENG_1^q - ENG_0^q\right) + c_3 EDU^q + e_0 \quad (1.9)$$

with all the coefficients $c > 0$. That is the productivity growth increases with the initial productivity gap $Z_0 - z_0^q$ between the newcomers and everyone else – with Z_0 is the aggregate level of productivity at time 0, their initial English proficiency, their English proficiency improvement, and their educational attainment.

Subtracting the two periods of equation (1.8) and combine with (1.9), I obtain the following equation

$$\begin{aligned} \Delta w^q = & \, b_1 c_0 + b_1 c_1 Z_0 - b_1 c_1 z_0^q \\ & + b_1 c_{21} ENG_0^q \\ & + (b_1 c_{22} + b_2) \left[ENG_1^q - ENG_0^q\right] \\ & + b_1 c_3 EDU^q \\ & + \varepsilon \end{aligned} \quad (1.10)$$

that describes wage growth as a linear function of multiple determinants: initial productivity gap $Z_0 - z_0^q$, initial English proficiency ENG_0^q, English proficiency improvement $ENG_1^q - ENG_0^q$, and education EDU^q. The initial productivity of the newcomers z_0^q is not observed in the data but can be partially proxied by the log per capita GDP of that countries. Similarly, Z_0 is not observed but approximately the same for newcomers arrived in the same year – a cohort fixed effects could be used to control its variation across time.

Implementation and results Following Borjas (2015), immigrant cohorts can be tracked across censuses using their age, year of immigration, and country of origin to study their economic and linguistic assimilation. A cohort of immigrants age i from country k migrated to the US in year t in census τ will be ten years older in the next census. This allows keeping track of cell of immigrants across panels using their ages. Hereafter, the unit of observation will be an immigrant cell age $i \in \{25\text{-}34, 35\text{-}44, 45\text{-}54\}$, from country k, and migrates to the US in $t \in \{1970\text{-}1974, 1975\text{-}1979, ..., 1995\text{-}1999\}$. The sample includes immigrant cohorts from 80 countries. The sample size is 1422, which is smaller than $1440 = 3 \times 6 \times 80$ since some immigrant cohorts are missing from the data.

From the logic of (1.10), I estimate the following regression model

$$dWAGE^q = \phi + \phi_k lGDP^k + \phi_{E_0} ENG_0^q + \phi_{dE} dENG^q + \phi_{edu} EDU^q + \phi_{frac} frac^q + \varepsilon \tag{1.11}$$

where

- $dWAGE^q$ is the change in log real weekly wage ten years after of immigrants in cell $q = (i,k,t)$ of age i, from country k, and migrated to the US in time t,

- $lGDP^k$ is the per capita GDP of the source country k[13], ENG_0 is average English proficiency at arrival, $dENG$ is English proficiency improvement ten years after arrival, EDU is the average years of education of cohort q, and $frac^q$ is the fraction of working adults who live in cohort q's state of residence (at PUMA level) with the same ethnicity as cell q

- ϕ is the vector of age and cohort fixed effects.

Under the assumption that the individual errors are i.i.d. with variance σ^2, the real wage growth for each cell has a variance of $\frac{\sigma^2}{n_0} + \frac{\sigma^2}{n_1} = \sigma^2 \frac{n_0+n_1}{n_0 n_1}$ where n_0 is the cell size at the

[13] Per capita GDP data are real per capita GDP at the starting year for each entry cohort. The data come from Borjas 2015 who draws from the Penn World Table (Heston, Summers, and Aten 2012).

beginning of the decade and n_1 is the respective cell size ten years later. The regression needs to be weighted by $\frac{n_0 n_1}{n_0+n_1}$, putting more weight on large and balanced cells. The standard errors are reported in parentheses and clustered at the source country level.

The regression results are reported in table 1.2. In column 1 of table 1.2, the regression results show compared to the mean, immigrant cohorts whose English proficiency is one level higher have an average of 7.04 percent increase in wage growth after ten years. In column 2, I interact English proficiency with the indicator of the quartile that divides the observations of each survey year by their fraction of immigrants of the same ethnicity. The magnitude of the coefficient on the initial English proficiency is highest in the first quartile and smaller in other quartiles. In column 3, I include the country of origin fixed effects and the results show that within-group, initial English proficiency has a statistically significant association with the wage growth in the first quartile, but the magnitude of the association still declines with the higher order of the quartile.

One possible explanation for the statistical insignificance in column 3 is that the result is partially driven by immigrants from English speaking countries who were mostly proficient in English at arrival or had exposure to the English language early in their lives. For these newcomers, the initial English proficiency plays a smaller role in their wage growth as they pick up English quickly after arrival (Bleakley and Chin 2004). In columns 4, 5, and 6, I report the results when restricting the sample to immigrants from non-English speaking countries. Non-English speaking countries are defined by whether their average measure of English proficiency is less than 2.5 (column 4), the minimum English proficiency is less than 2 (column 5), and English is listed as an official language in their country of origin (column 6). The results show that within-group, the initial English proficiency is positively associated with the wage growth of immigrants and the association is statistically significant. The magnitude of the association is stronger for those located in regions with a fewer immigrants of the same ethnicity.

The coefficient for the log GDP is negative, which is similar to that documented by Borjas (2015). Although the coefficient on years of education is not statistically significant and even negative, its contribution to wage growth through immigrants' improved English proficiency improvement is netted out when including the English improvement as a regressor. In the regression model, log per capita GDP is a proxy for newcomers' initial productivity. Since immigrants' rate of productivity growth declines with respect to their initial productivity, that explains why the coefficient for log per capita GDP is negative. After netting out the contribution of education to immigrants' English improvement, the regression of wage growth on education is essentially the regression of productivity growth on education, which then suffers a downward bias. Therefore, the coefficient for years of education is downwardly biased and becomes either statistically non-significant and even negative since the log per capita GDP of country of origin cannot sufficiently control the variation in initial productivity of newcomers across different education levels.

In table 1.3, I report the results when estimating an analogous model to (1.10) with the improved English proficiency as the explanatory variables. Compared to column 1, I add the fraction of same-ethnic immigrants as a regressor in column 2 as well as country of origin fixed effects in column 3. The difference between the coefficients for years of education and initial English proficiency in column 1 and 2 is a case of omitted variable bias. Newcomers with higher education and/or English proficiency are less likely to reside in an enclave and therefore, there is a positive correlation between the fraction of immigrants of the same ethnicity and both years of education and initial English proficiency. As a result, when the fraction of immigrants of the same ethnicity is not included, the coefficients for years of education and initial English are upwardly biased. In column 3, the variable years of education has a significant association with newcomers' improved English proficiency after controlling for the cross-group variation that muddles the result with fixed-effects. This cross-group variation includes the variation in quality of education and the age of first exposure to the English language in newcomers' country of origin. Across the results, I include a measure of language score represents the closeness of the immigrants' native language to the English language, first used by Chiswick and Miller 2005, to capture the ease at which the newcomers can learn the English language. It's straightforward to see the improvement of immigrants' English proficiency is positively correlated to this measure.

Table 1.2: Wage growth and English proficiency

	(1)	(2)	(3)	(4)	(5)	(6)
	dWAGE	dWAGE	dWAGE	dWAGE	dWAGE	dWAGE
Years of education (EDUC)	-0.00303	-0.00326	-0.0288*	-0.0307*	-0.0296*	-0.0307*
	(0.00550)	(0.00536)	(0.0129)	(0.0146)	(0.0145)	(0.0133)
English Improvement (dENG)	0.201***	0.200***	0.185**	0.282***	0.275***	0.223***
	(0.0521)	(0.0522)	(0.0652)	(0.0705)	(0.0679)	(0.0640)
Initial English proficiency (ENG0)	0.0704**					
	(0.0237)					
log per capita GDP	-0.0280**	-0.0288**	-0.00301	-0.0189	-0.0174	-0.0147
	(0.00866)	(0.00847)	(0.0507)	(0.0555)	(0.0539)	(0.0515)
Fraction of same-ethnic (frac)	-0.00134	-0.00134	0.0123*	0.0125**	0.0126**	0.0135**
	(0.00165)	(0.00166)	(0.00478)	(0.00452)	(0.00446)	(0.00460)
Quartile frac=1 $\times$ ENG0		0.112***	0.149*	0.249*	0.242*	0.201*
		(0.0317)	(0.0723)	(0.105)	(0.104)	(0.0820)
Quartile frac=2 $\times$ ENG0		0.0724**	0.127	0.240*	0.230*	0.178*
		(0.0255)	(0.0708)	(0.0976)	(0.0968)	(0.0806)
Quartile frac=3 $\times$ ENG0		0.0672**	0.108	0.199*	0.193*	0.157*
		(0.0236)	(0.0685)	(0.0958)	(0.0945)	(0.0785)
Quartile frac=4 $\times$ ENG0		0.0711**	0.0894	0.169	0.163	0.139
		(0.0233)	(0.0687)	(0.0995)	(0.0979)	(0.0789)
Constant	0.342***	0.350***	0.337	0.313	0.297	0.354
	(0.0855)	(0.0836)	(0.462)	(0.538)	(0.524)	(0.485)
Country of origin fixed effects	NO	NO	YES	YES	YES	YES
average ENG0<2.5 (48 countries)				YES		
min ENG0<2 (51 countries)					YES	
ENG is not official language (71)						YES
Observations	1422	1422	1422	823	876	1224
Adjusted R^2	0.361	0.363	0.429	0.447	0.439	0.435

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Real wage is log weekly earnings, adjusted for inflation and deflated by education-age deflator. English proficiency is an ordinal measure that equals to 0 if an immigrant cannot speak English at all, 1 if he speaks English but not well, 2 if he speaks English well, 3 if he speaks English only or speaks English very well; Fraction of pre-existing immigrants is percentage point.

Table 1.3: English proficiency improvement

	(1)	(2)	(3)
	dENG	dENG	dENG
Years of education (EDUC)	0.0147*	0.00867	0.0244***
	(0.00573)	(0.00529)	(0.00458)
Initial English proficiency (ENG0)	-0.198***	-0.211***	-0.490***
	(0.0240)	(0.0251)	(0.0357)
log per capita GDP	0.0144	0.0221**	-0.00170
	(0.00798)	(0.00803)	(0.0179)
Fraction of same-ethnic (frac)		-0.0103***	-0.00699*
		(0.00157)	(0.00267)
Constant	0.264***	0.322***	0.865***
	(0.0750)	(0.0821)	(0.140)
Country of origin fixed effects	NO	NO	YES
Observations	1422	1422	1422
Adjusted R^2	0.572	0.597	0.702

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

English proficiency is an ordinal measure that equals to 0 if an immigrant cannot speak English at all, 1 if he speaks English but not well, 2 if he speaks English well, 3 if he speaks English only or speaks English very well; Fraction of pre-existing immigrants is percentage point.

1.4 Concluding remarks

This paper proposes a parsimonious knowledge diffusion model to study the assimilation of immigrants: how English skills, education, and pre-existing immigrants affect their location choices, their incentive to learn English, and their skill accumulation. English proficiency and education play an important role in the location choice of newcomers and, subsequently, their effort in learning English and their productivity growth. The quality of the pre-existing immigrant population also influences the newcomers' location choices. More newcomers would reside in the enclave if the pre-existing immigrants, who reside in the enclave, are more productive. Furthermore, the model predicts that the pre-existing immigrant population has heterogeneous effects, in sign and magnitude, on newcomers' productivity growth and welfare that depends on its quality and the newcomers' education and English proficiency.

The model's theoretical predictions are consistent with patterns of immigrant assimilation that are observed with US data. With US census data, this paper shows that within the same country of origin group, newcomers with higher education and/or English proficiency reside in regions with fewer immigrants of the same ethnicity. Furthermore, the reduced-form results show that an increase of one level of initial English proficiency is associated with a real wage gain of more than 8 percent, ten years after later. This number is higher for those residing in regions with fewer immigrants of the same ethnicity. The reduced-form regressions show no statistically significant between real wage gain and the fraction of immigrants of the same ethnicity, as predicted by the model when aggregating the heterogeneous effects together.

There are several policy implications and forward research inquiries prompted by the findings of this paper. First, as shown in this paper, location choices of newcomers partially depend on their characteristics, a shift in immigration policy that alters the composition of the immigrant inflow will have differential regional economic impacts across locations. That leads to the following line of questions: What are the potential economic impacts across regions of alternative immigration policies? Which states, cities should pay which upon a change of immigration policy? Second, initial English proficiency has important value in the economic assimilation of the newcomers and its value depends on their country of origin, education, and so on. If one wishes to design a immigration point system that maximizes the newcomers' economic integration (and hence, minimizes their dependence on public welfare system), how to allocate points optimally across newcomers' characteristics – such as their native language skills and education? I believe it is within the literature interest and would be beneficial to policymakers to further research along these lines of inquiry.

1.5 Appendix

A – Bellman equations, boundary conditions and computation methods

Boundary conditions

With the continuous arrival process, I use the boundary condition $\lim_{x \to 0^+} D_x V^E(x) = 0$. The short argument for the boundary conditions is that the value function of type x at time t has three components, expected earnings while remaining at x, the expected earnings after acquiring $x' > x$, and the expected change in value due to a jump in English skill ladder. With the continuous arrival process, the agent with approximately zero productivity will instantly meet and acquire higher productivity which results in the diminished left tail of the Frechet productivity distribution. Hence, for all agents with approximate zero productivity, the value of the first component is zero while the value of the second component is roughly constant with respect to x. Furthermore, it can be shown that the effort in learning English is equal to 1 because of the zero marginal cost of time, that implies $\eta(s;e)$ does not change with respect to x. As a result, the derivatives of all components with respect to x are zero, $\lim_{x \to 0+} D_x v(x) = 0$.

Similar to Luttmer 2015, another set of boundary conditions can be used is the value function of x for large x is approximately equal to the fundamental value, present value of a constant flow of income. The first set of boundary conditions is uniquely applicable for the continuous arrival process while the second set of boundary conditions could be used with any arrival processes.

Computation methods

The computational method is inspired by Lucas and Moll 2014. The steps of the computational method are laid out as follows

Step 1: Compute the BGP Denote the largest eigenvalue of matrix A is Λ_1 and the relevant eigenvector is v_1. The BGP growth rate γ is equal to $\Lambda_1\theta$. The scale parameters of k^{b_0}, k^{b_1}, and k^{b_2} can be computed using k^a the ratio between elements of v_1. At time 0, the productivity distribution of type $i \in I = \{a, b_2, b_1, b_0\}$ is a scaled Frechet

$$\Phi(x,i) = \pi^i \exp\left(-k^i x^{-\frac{1}{\theta}}\right).$$

Step 2: Compute the value function and the optimal allocation of time Here we can solve for the value functions and optimal allocation using the finite difference method and the smooth pasting boundary condition. Denote $x \in \{x_1, ..., x_I\}$ for the finite grid in Matlab and $v_k\left(p_j, \varepsilon\right) = v\left(x_k; p_j, \varepsilon\right)$ for the value of the immigrant with productivity x_k, linguistic skill p_j, and educational attainment ε. Denote ζ_k^i for $D_x \log \frac{\Phi(x_k,a)}{\pi^a}$. For sufficiently large x_I and I, The Bellman equation in (**??**) can be written as

$$\begin{aligned}
\rho v_k\left(p_j,\varepsilon\right) \approx \max_s\, & w_k\left(p_j,\Pi\right)(1-s) \\
& + \alpha\pi^a p_j \sum_k^I \left[v_{k'}\left(p_j,\varepsilon\right) - v_k\left(p_j,\varepsilon\right)\right]\zeta_{k'}^a h_{k'} \\
& + \sum_{h\in\{1,..,N\}} \alpha\pi_j^b p_N \sum_k^I \left[v_{k'}\left(p_j,\varepsilon\right) - v_k\left(p_j,\varepsilon\right)\right]\zeta_{k'}^h h_{k'} \\
& + \eta\left(s;p_j,\Pi\right)\left[v_k\left(p_{j+1},\varepsilon\right) - v_k\left(p_j,\varepsilon\right)\right] \\
& - \frac{v_k\left(p_j,\varepsilon\right) - v_{k-1}\left(p_j,\varepsilon\right)}{h_k} x\gamma + v_k\left(p_j,\varepsilon\right)\gamma
\end{aligned}$$

where $h_k = x_k - x_{k-1}$ since

$$D_x v\left(x;p_j,\varepsilon\right) \approx \frac{v\left(x;p_j,\varepsilon\right) - v\left(x-\Delta;p_j,\varepsilon\right)}{\Delta}.$$

The boundary condition then implies $v_0 = v_1$. From here, we can make a guess of $v_k^0\left(p_j, \varepsilon\right)$, solve for the optimal allocation $s_k^t\left(p_j, \varepsilon\right)$ and $v_k^{t+1}\left(p_j, \varepsilon\right)$ for $t = 0, 1, 2, ...$ as follows.

Step 2.0: Reduce state variable While an immigrant's educational attainment is constant, the immigrant can improve his linguistic skill over time. Hence, it is easier to stack the value functions with different linguistic skill together. That is to write $v^{t+1}(\varepsilon) = \left\{v_{1,2}^{t+1}(\varepsilon), ...v_{I,2}^{t+1}(\varepsilon), ...v_{k,j}^{t+1}(\varepsilon), ...v_{I,0}^{t+1}(\varepsilon)\right\}$. Similar for the optimal allocation of time s^t.

Step 2.1: Solve for optimal allocation of time Given $v_{k,j}^t(\varepsilon)$, search for $s_{k,j}^t(\varepsilon)$ that solves

$$\max_s w_k\left(p_j, \Pi\right)(1-s) + \eta\left(s; p_j, \Pi\right)\left[v_{k,j+1}^t(\varepsilon) - v_{k,j}^t(\varepsilon)\right].$$

Step 2.2: Solve for $v_{k,j}^{t+1}(\varepsilon)$ Given the vector $s^t(\varepsilon)$, we can compute $\eta^t(\varepsilon)$. The Bellman equation implies that

$$\left\{B^t - C^t\right\} v^{t+1}(\varepsilon) = b^t$$

where the right hand side is

$$b^t = w \cdot \left(1 - s^t(\varepsilon)\right)$$

and the matrix C^t is the block matrix

$$C^t = \begin{bmatrix} C_2^t & 0 & 0 \\ 0 & C_1^t & 0 \\ 0 & 0 & C_0^t \end{bmatrix}$$

with each block given by

$$C_j^t =$$

$$\begin{bmatrix} \alpha \sum_i \pi^i p_i^j \zeta_1^i h_1 & \alpha \sum_i \pi^i p_i^j \zeta_2^i h_2 & \cdots & \cdots & \alpha \sum_i \pi^i p_i^j \zeta_I^i h_I \\ 0 & \alpha \sum_i \pi^i p_i^j \zeta_2^i h_2 & & & \vdots \\ \vdots & 0 & \ddots & & \vdots \\ & & 0 & \ddots & \vdots \\ 0 & \cdots & & 0 & \alpha \sum_i \pi^i p_i^j \zeta_I^i h_I \end{bmatrix}$$

for $j \in \{0,1,2\}$, $i \in \{a, b_2, b_1, b_0\}$, and $p_i^j = p_i^{b_j}$. Matrix B^t has two components, block B_j^t and block D_j^t

$$B^t =$$

$$\begin{bmatrix} B_2^t & 0 & 0 \\ D_1^t & B_1^t & 0 \\ 0 & D_0^t & B_0^t \end{bmatrix}$$

where elements of B_j^t are given by

$$\begin{aligned} B_j^t[1,1] &= \rho - \gamma + \alpha \sum_i \pi^i p_i^j \sum_1^I \zeta_{k'}^i h_{k'} + \eta_{1,j}^t \\ B_j^t[k,k] &= \rho - \gamma + \gamma \frac{x_k}{h_k} + \alpha \sum_i \pi^i p_i^j \sum_k^I \zeta_{k'}^i h_{k'} + \eta_{k,j}^t \text{ for } k > 1 \\ B_j^t[k,k-1] &= -\gamma \frac{x_k}{h_k} \text{ for } k > 1 \\ B_j^t[k,l] &= 0 \text{ for } l \neq k \text{ or } l \neq k-1 \end{aligned}$$

where the $[1,1]$ element differs from other diagonal elements because of the boundary condition. The off diagonal element $[k,k-1]$ comes from the linear approximation of $D_x v(\cdot)$. Block B_j^t is as follows

$$B_j^t =$$
$$\begin{bmatrix} \rho-\gamma+\alpha\sum_i \pi^i p_i^j \sum_1^I \zeta_{k'}^i h_{k'} + \eta_{1,j}^t & 0 & \cdots & \cdots & 0 \\ -\gamma\frac{x_2}{h_2} & \rho-\gamma+\gamma\frac{x_2}{h_2}+\alpha\sum_i \pi^i p_i^j \sum_1^I \zeta_{k'}^i h_{k'} + \eta_{2,j}^t & 0 & \cdots & \vdots \\ 0 & \cdots & \ddots & 0 & \vdots \\ \vdots & \cdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & & \end{bmatrix}.$$

Block D_j^t is given by

$$D_j^t =$$
$$\begin{bmatrix} -\eta_{1,j}^t & 0 & \cdots & \cdots & 0 \\ 0 & -\eta_{2,j}^t & 0 & \cdots & \vdots \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & & 0 & -\eta_{I,j}^t \end{bmatrix}.$$

which link the flow from jumping in the ladder of linguistic skill. We, hence, can solve for v^{t+1} and repeat step the process from Step 2.1 until the distance $\mid v^{t+1}-v^t \mid$ is smaller than a certain threshold.

C – Robustness checks

Excluding Mexican immigrants

Table 1.4: Wage growth and English proficiency

	(1) drweekly	(2) drweekly	(3) drweekly	(4) drweekly	(5) drweekly	(6) drweekly
yeduc0	-0.000773	-0.00103	-0.0146	-0.00919	-0.00818	-0.0134
	(0.00576)	(0.00565)	(0.01000)	(0.0113)	(0.0110)	(0.0103)
deng_o	0.233***	0.236***	0.201**	0.260***	0.254***	0.225***
	(0.0448)	(0.0449)	(0.0597)	(0.0672)	(0.0644)	(0.0598)
eng_o0	0.0612*					
	(0.0236)					
lgdp	-0.0299***	-0.0306***	-0.0307	-0.0337	-0.0323	-0.0407
	(0.00809)	(0.00781)	(0.0464)	(0.0527)	(0.0511)	(0.0481)
asep0	-0.0130*	-0.0154*	-0.0350	-0.0323	-0.0276	0.000124
	(0.00556)	(0.00620)	(0.0383)	(0.0433)	(0.0426)	(0.0337)
qasep0=1 × eng_o0		0.0954**	0.0702	0.0993	0.0958	0.105
		(0.0314)	(0.0592)	(0.0767)	(0.0747)	(0.0642)
qasep0=2 × eng_o0		0.0618*	0.0526	0.0927	0.0850	0.0837
		(0.0264)	(0.0583)	(0.0712)	(0.0695)	(0.0631)
qasep0=3 × eng_o0		0.0559*	0.0361	0.0578	0.0548	0.0652
		(0.0243)	(0.0561)	(0.0701)	(0.0680)	(0.0613)
qasep0=4 × eng_o0		0.0642**	0.0221	0.0313	0.0291	0.0512
		(0.0228)	(0.0563)	(0.0730)	(0.0709)	(0.0620)
Constant	0.343***	0.349***	0.621	0.519	0.497	0.594
	(0.0825)	(0.0805)	(0.462)	(0.546)	(0.530)	(0.476)
Observations	1404	1404	1404	805	858	1206
Adjusted R^2	0.394	0.396	0.454	0.462	0.454	0.452

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

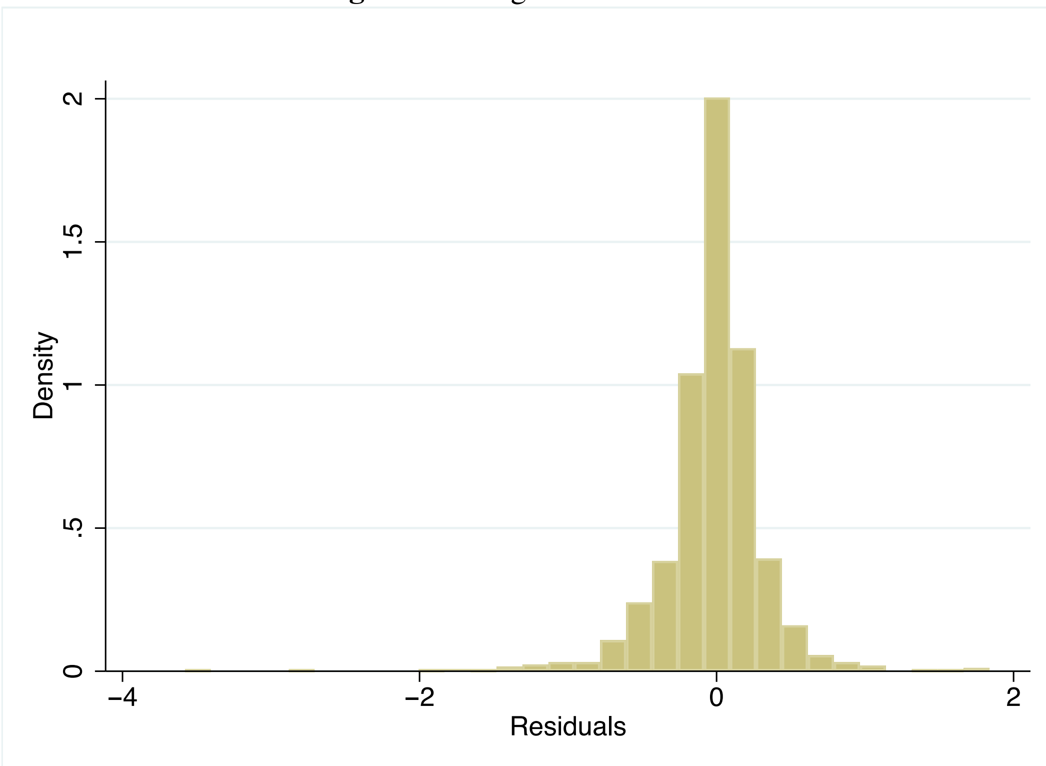
Table 1.5: English proficiency improvement

	(1)	(2)	(3)
	deng_o	deng_o	deng_o
yeduc0	0.00675	0.00666***	0.0185***
	(0.00536)	(0.00193)	(0.00445)
eng_o0	-0.202***	-0.206***	-0.472***
	(0.0244)	(0.00853)	(0.0217)
lgdp	0.0241**	0.0234***	0.0119
	(0.00793)	(0.00308)	(0.0142)
asep0		-0.0150**	-0.0718***
		(0.00478)	(0.0214)
Constant	0.300***	0.325***	0.878***
	(0.0818)	(0.0310)	(0.122)
Observations	1404	1404	1404
Adjusted R^2	0.516	0.519	0.642

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1.9. Regression residuals



Residuals of regression wage growth

Figure 1.10. Residuals vs. Fitted values

Within-group location choices

In this section, I report the robustness checks for the within-group location choices of immigrants.

Table 1.6: Education, English Proficiency, and Spatial Assimilation, Control for Income Level - 1990

	(1) EDU × ENG	(2) EDU × ENG, Recent	(3) ENG × EDU	(4) ENG × EDU, Recent
GROUP=1	0	0	0	0
	(.)	(.)	(.)	(.)
GROUP=2	-0.0202***	-0.0181***	-0.00183	0.00346
	(-4.83)	(-5.04)	(-0.38)	(0.74)
GROUP=3	-0.0319***	-0.0192***	-0.0151**	-0.00247
	(-6.41)	(-4.91)	(-2.71)	(-0.39)
GROUP=4	-0.0403***	-0.0206***	-0.0209**	-0.00793
	(-9.16)	(-5.23)	(-3.03)	(-0.90)
GROUP=5	-0.00183	0.00346	-0.0131	0.00213
	(-0.38)	(0.74)	(-1.47)	(0.14)
GROUP=6	-0.0246***	-0.0136***	-0.0202***	-0.0181***
	(-6.44)	(-4.25)	(-4.83)	(-5.04)
GROUP=7	-0.0329***	-0.0217***	-0.0246***	-0.0136***
	(-6.67)	(-5.12)	(-6.44)	(-4.25)
GROUP=8	-0.0431***	-0.0247***	-0.0240***	-0.0107**
	(-8.33)	(-7.65)	(-5.68)	(-2.71)
GROUP=9	-0.0151**	-0.00247	-0.0310***	-0.0164***
	(-2.71)	(-0.39)	(-5.97)	(-3.93)
GROUP=10	-0.0240***	-0.0107**	-0.0340***	-0.0223***
	(-5.68)	(-2.71)	(-6.72)	(-5.21)
GROUP=11	-0.0370***	-0.0248***	-0.0319***	-0.0192***
	(-6.85)	(-7.00)	(-6.41)	(-4.91)
GROUP=12	-0.0474***	-0.0284***	-0.0329***	-0.0217***
	(-8.49)	(-8.39)	(-6.67)	(-5.12)
GROUP=13	-0.0209**	-0.00793	-0.0370***	-0.0248***
	(-3.03)	(-0.90)	(-6.85)	(-7.00)
GROUP=14	-0.0310***	-0.0164***	-0.0429***	-0.0274***
	(-5.97)	(-3.93)	(-7.27)	(-7.61)
GROUP=15	-0.0429***	-0.0274***	-0.0442***	-0.0305***
	(-7.27)	(-7.61)	(-8.02)	(-8.17)
GROUP=16	-0.0491***	-0.0307***	-0.0403***	-0.0206***
	(-7.87)	(-8.34)	(-9.16)	(-5.23)
GROUP=17	-0.0131	0.00213	-0.0431***	-0.0247***
	(-1.47)	(0.14)	(-8.33)	(-7.65)
GROUP=18	-0.0340***	-0.0223***	-0.0474***	-0.0284***
	(-6.72)	(-5.21)	(-8.49)	(-8.39)
GROUP=19	-0.0442***	-0.0305***	-0.0491***	-0.0307***
	(-8.02)	(-8.17)	(-7.87)	(-8.34)
GROUP=20	-0.0517***	-0.0324***	-0.0517***	-0.0324***
	(-8.46)	(-9.15)	(-8.46)	(-9.15)
Constant	0.101***	0.0742***	0.101***	0.0742***
	(13.02)	(13.84)	(13.02)	(13.84)
Observations	252998	44324	252998	44324
Adjusted R^2	0.406	0.418	0.406	0.418

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.7: Education, English Proficiency, and Spatial Assimilation - 2000

	(1) f_bpld_s	(2) f_bpld_s	(3) f_bpld_s	(4) f_bpld_s
GROUP=1	0	0	0	0
	(.)	(.)	(.)	(.)
GROUP=2	-0.0200***	-0.0181***	-0.00181	0.00349
	(-4.79)	(-5.00)	(-0.38)	(0.74)
GROUP=3	-0.0316***	-0.0191***	-0.0150**	-0.00229
	(-6.32)	(-4.83)	(-2.64)	(-0.36)
GROUP=4	-0.0401***	-0.0206***	-0.0209**	-0.00748
	(-9.07)	(-5.23)	(-3.05)	(-0.85)
GROUP=5	-0.00181	0.00349	-0.0136	0.00242
	(-0.38)	(0.74)	(-1.53)	(0.16)
GROUP=6	-0.0244***	-0.0135***	-0.0200***	-0.0181***
	(-6.36)	(-4.22)	(-4.79)	(-5.00)
GROUP=7	-0.0325***	-0.0215***	-0.0244***	-0.0135***
	(-6.56)	(-5.00)	(-6.36)	(-4.22)
GROUP=8	-0.0428***	-0.0245***	-0.0240***	-0.0105**
	(-8.19)	(-7.50)	(-5.66)	(-2.65)
GROUP=9	-0.0150**	-0.00229	-0.0311***	-0.0162***
	(-2.64)	(-0.36)	(-5.95)	(-3.85)
GROUP=10	-0.0240***	-0.0105**	-0.0342***	-0.0220***
	(-5.66)	(-2.65)	(-6.60)	(-5.09)
GROUP=11	-0.0368***	-0.0245***	-0.0316***	-0.0191***
	(-6.72)	(-6.84)	(-6.32)	(-4.83)
GROUP=12	-0.0475***	-0.0281***	-0.0325***	-0.0215***
	(-8.32)	(-8.11)	(-6.56)	(-5.00)
GROUP=13	-0.0209**	-0.00748	-0.0368***	-0.0245***
	(-3.05)	(-0.85)	(-6.72)	(-6.84)
GROUP=14	-0.0311***	-0.0162***	-0.0431***	-0.0270***
	(-5.95)	(-3.85)	(-7.14)	(-7.37)
GROUP=15	-0.0431***	-0.0270***	-0.0448***	-0.0302***
	(-7.14)	(-7.37)	(-7.87)	(-7.99)
GROUP=16	-0.0496***	-0.0303***	-0.0401***	-0.0206***
	(-7.72)	(-8.05)	(-9.07)	(-5.23)
GROUP=17	-0.0136	0.00242	-0.0428***	-0.0245***
	(-1.53)	(0.16)	(-8.19)	(-7.50)
GROUP=18	-0.0342***	-0.0220***	-0.0475***	-0.0281***
	(-6.60)	(-5.09)	(-8.32)	(-8.11)
GROUP=19	-0.0448***	-0.0302***	-0.0496***	-0.0303***
	(-7.87)	(-7.99)	(-7.72)	(-8.05)
GROUP=20	-0.0529***	-0.0321***	-0.0529***	-0.0321***
	(-8.28)	(-8.94)	(-8.28)	(-8.94)
Constant	0.101***	0.0740***	0.101***	0.0740***
	(12.86)	(13.63)	(12.86)	(13.63)
Observations	252998	44324	252998	44324
Adjusted R^2	0.406	0.417	0.406	0.417

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.8: Education, English Proficiency, and Spatial Assimilation, Control for Income Level - 2000

	(1) f_bpld_s	(2) f_bpld_s	(3) f_bpld_s	(4) f_bpld_s
GROUP=1	0	0	0	0
	(.)	(.)	(.)	(.)
GROUP=2	-0.0202***	-0.0181***	-0.00183	0.00346
	(-4.83)	(-5.04)	(-0.38)	(0.74)
GROUP=3	-0.0319***	-0.0192***	-0.0151**	-0.00247
	(-6.41)	(-4.91)	(-2.71)	(-0.39)
GROUP=4	-0.0403***	-0.0206***	-0.0209**	-0.00793
	(-9.16)	(-5.23)	(-3.03)	(-0.90)
GROUP=5	-0.00183	0.00346	-0.0131	0.00213
	(-0.38)	(0.74)	(-1.47)	(0.14)
GROUP=6	-0.0246***	-0.0136***	-0.0202***	-0.0181***
	(-6.44)	(-4.25)	(-4.83)	(-5.04)
GROUP=7	-0.0329***	-0.0217***	-0.0246***	-0.0136***
	(-6.67)	(-5.12)	(-6.44)	(-4.25)
GROUP=8	-0.0431***	-0.0247***	-0.0240***	-0.0107**
	(-8.33)	(-7.65)	(-5.68)	(-2.71)
GROUP=9	-0.0151**	-0.00247	-0.0310***	-0.0164***
	(-2.71)	(-0.39)	(-5.97)	(-3.93)
GROUP=10	-0.0240***	-0.0107**	-0.0340***	-0.0223***
	(-5.68)	(-2.71)	(-6.72)	(-5.21)
GROUP=11	-0.0370***	-0.0248***	-0.0319***	-0.0192***
	(-6.85)	(-7.00)	(-6.41)	(-4.91)
GROUP=12	-0.0474***	-0.0284***	-0.0329***	-0.0217***
	(-8.49)	(-8.39)	(-6.67)	(-5.12)
GROUP=13	-0.0209**	-0.00793	-0.0370***	-0.0248***
	(-3.03)	(-0.90)	(-6.85)	(-7.00)
GROUP=14	-0.0310***	-0.0164***	-0.0429***	-0.0274***
	(-5.97)	(-3.93)	(-7.27)	(-7.61)
GROUP=15	-0.0429***	-0.0274***	-0.0442***	-0.0305***
	(-7.27)	(-7.61)	(-8.02)	(-8.17)
GROUP=16	-0.0491***	-0.0307***	-0.0403***	-0.0206***
	(-7.87)	(-8.34)	(-9.16)	(-5.23)
GROUP=17	-0.0131	0.00213	-0.0431***	-0.0247***
	(-1.47)	(0.14)	(-8.33)	(-7.65)
GROUP=18	-0.0340***	-0.0223***	-0.0474***	-0.0284***
	(-6.72)	(-5.21)	(-8.49)	(-8.39)
GROUP=19	-0.0442***	-0.0305***	-0.0491***	-0.0307***
	(-8.02)	(-8.17)	(-7.87)	(-8.34)
GROUP=20	-0.0517***	-0.0324***	-0.0517***	-0.0324***
	(-8.46)	(-9.15)	(-8.46)	(-9.15)
Constant	0.101***	0.0742***	0.101***	0.0742***
	(13.02)	(13.84)	(13.02)	(13.84)
Observations	252998	44324	252998	44324
Adjusted R^2	0.406	0.418	0.406	0.418

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.9: Education, English Proficiency, and Spatial Assimilation - 1980

	(1) f_bpld_s	(2) f_bpld_s	(3) f_bpld_s	(4) f_bpld_s
GROUP=1	0	0	0	0
	(.)	(.)	(.)	(.)
GROUP=2	-0.0200***	-0.0181***	-0.00181	0.00349
	(-4.79)	(-5.00)	(-0.38)	(0.74)
GROUP=3	-0.0316***	-0.0191***	-0.0150**	-0.00229
	(-6.32)	(-4.83)	(-2.64)	(-0.36)
GROUP=4	-0.0401***	-0.0206***	-0.0209**	-0.00748
	(-9.07)	(-5.23)	(-3.05)	(-0.85)
GROUP=5	-0.00181	0.00349	-0.0136	0.00242
	(-0.38)	(0.74)	(-1.53)	(0.16)
GROUP=6	-0.0244***	-0.0135***	-0.0200***	-0.0181***
	(-6.36)	(-4.22)	(-4.79)	(-5.00)
GROUP=7	-0.0325***	-0.0215***	-0.0244***	-0.0135***
	(-6.56)	(-5.00)	(-6.36)	(-4.22)
GROUP=8	-0.0428***	-0.0245***	-0.0240***	-0.0105**
	(-8.19)	(-7.50)	(-5.66)	(-2.65)
GROUP=9	-0.0150**	-0.00229	-0.0311***	-0.0162***
	(-2.64)	(-0.36)	(-5.95)	(-3.85)
GROUP=10	-0.0240***	-0.0105**	-0.0342***	-0.0220***
	(-5.66)	(-2.65)	(-6.60)	(-5.09)
GROUP=11	-0.0368***	-0.0245***	-0.0316***	-0.0191***
	(-6.72)	(-6.84)	(-6.32)	(-4.83)
GROUP=12	-0.0475***	-0.0281***	-0.0325***	-0.0215***
	(-8.32)	(-8.11)	(-6.56)	(-5.00)
GROUP=13	-0.0209**	-0.00748	-0.0368***	-0.0245***
	(-3.05)	(-0.85)	(-6.72)	(-6.84)
GROUP=14	-0.0311***	-0.0162***	-0.0431***	-0.0270***
	(-5.95)	(-3.85)	(-7.14)	(-7.37)
GROUP=15	-0.0431***	-0.0270***	-0.0448***	-0.0302***
	(-7.14)	(-7.37)	(-7.87)	(-7.99)
GROUP=16	-0.0496***	-0.0303***	-0.0401***	-0.0206***
	(-7.72)	(-8.05)	(-9.07)	(-5.23)
GROUP=17	-0.0136	0.00242	-0.0428***	-0.0245***
	(-1.53)	(0.16)	(-8.19)	(-7.50)
GROUP=18	-0.0342***	-0.0220***	-0.0475***	-0.0281***
	(-6.60)	(-5.09)	(-8.32)	(-8.11)
GROUP=19	-0.0448***	-0.0302***	-0.0496***	-0.0303***
	(-7.87)	(-7.99)	(-7.72)	(-8.05)
GROUP=20	-0.0529***	-0.0321***	-0.0529***	-0.0321***
	(-8.28)	(-8.94)	(-8.28)	(-8.94)
Constant	0.101***	0.0740***	0.101***	0.0740***
	(12.86)	(13.63)	(12.86)	(13.63)
Observations	252998	44324	252998	44324
Adjusted R^2	0.406	0.417	0.406	0.417

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.10: Education, English Proficiency, and Spatial Assimilation, Control for Income Level - 1980

	(1) f_bpld_s	(2) f_bpld_s	(3) f_bpld_s	(4) f_bpld_s
GROUP=1	0	0	0	0
	(.)	(.)	(.)	(.)
GROUP=2	-0.0202***	-0.0181***	-0.00183	0.00346
	(-4.83)	(-5.04)	(-0.38)	(0.74)
GROUP=3	-0.0319***	-0.0192***	-0.0151**	-0.00247
	(-6.41)	(-4.91)	(-2.71)	(-0.39)
GROUP=4	-0.0403***	-0.0206***	-0.0209**	-0.00793
	(-9.16)	(-5.23)	(-3.03)	(-0.90)
GROUP=5	-0.00183	0.00346	-0.0131	0.00213
	(-0.38)	(0.74)	(-1.47)	(0.14)
GROUP=6	-0.0246***	-0.0136***	-0.0202***	-0.0181***
	(-6.44)	(-4.25)	(-4.83)	(-5.04)
GROUP=7	-0.0329***	-0.0217***	-0.0246***	-0.0136***
	(-6.67)	(-5.12)	(-6.44)	(-4.25)
GROUP=8	-0.0431***	-0.0247***	-0.0240***	-0.0107**
	(-8.33)	(-7.65)	(-5.68)	(-2.71)
GROUP=9	-0.0151**	-0.00247	-0.0310***	-0.0164***
	(-2.71)	(-0.39)	(-5.97)	(-3.93)
GROUP=10	-0.0240***	-0.0107**	-0.0340***	-0.0223***
	(-5.68)	(-2.71)	(-6.72)	(-5.21)
GROUP=11	-0.0370***	-0.0248***	-0.0319***	-0.0192***
	(-6.85)	(-7.00)	(-6.41)	(-4.91)
GROUP=12	-0.0474***	-0.0284***	-0.0329***	-0.0217***
	(-8.49)	(-8.39)	(-6.67)	(-5.12)
GROUP=13	-0.0209**	-0.00793	-0.0370***	-0.0248***
	(-3.03)	(-0.90)	(-6.85)	(-7.00)
GROUP=14	-0.0310***	-0.0164***	-0.0429***	-0.0274***
	(-5.97)	(-3.93)	(-7.27)	(-7.61)
GROUP=15	-0.0429***	-0.0274***	-0.0442***	-0.0305***
	(-7.27)	(-7.61)	(-8.02)	(-8.17)
GROUP=16	-0.0491***	-0.0307***	-0.0403***	-0.0206***
	(-7.87)	(-8.34)	(-9.16)	(-5.23)
GROUP=17	-0.0131	0.00213	-0.0431***	-0.0247***
	(-1.47)	(0.14)	(-8.33)	(-7.65)
GROUP=18	-0.0340***	-0.0223***	-0.0474***	-0.0284***
	(-6.72)	(-5.21)	(-8.49)	(-8.39)
GROUP=19	-0.0442***	-0.0305***	-0.0491***	-0.0307***
	(-8.02)	(-8.17)	(-7.87)	(-8.34)
GROUP=20	-0.0517***	-0.0324***	-0.0517***	-0.0324***
	(-8.46)	(-9.15)	(-8.46)	(-9.15)
Constant	0.101***	0.0742***	0.101***	0.0742***
	(13.02)	(13.84)	(13.02)	(13.84)
Observations	252998	44324	252998	44324
Adjusted R^2	0.406	0.418	0.406	0.418

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Restricting subsample using wage rate quintile

Figure 1.11. 1990 first quintile by income

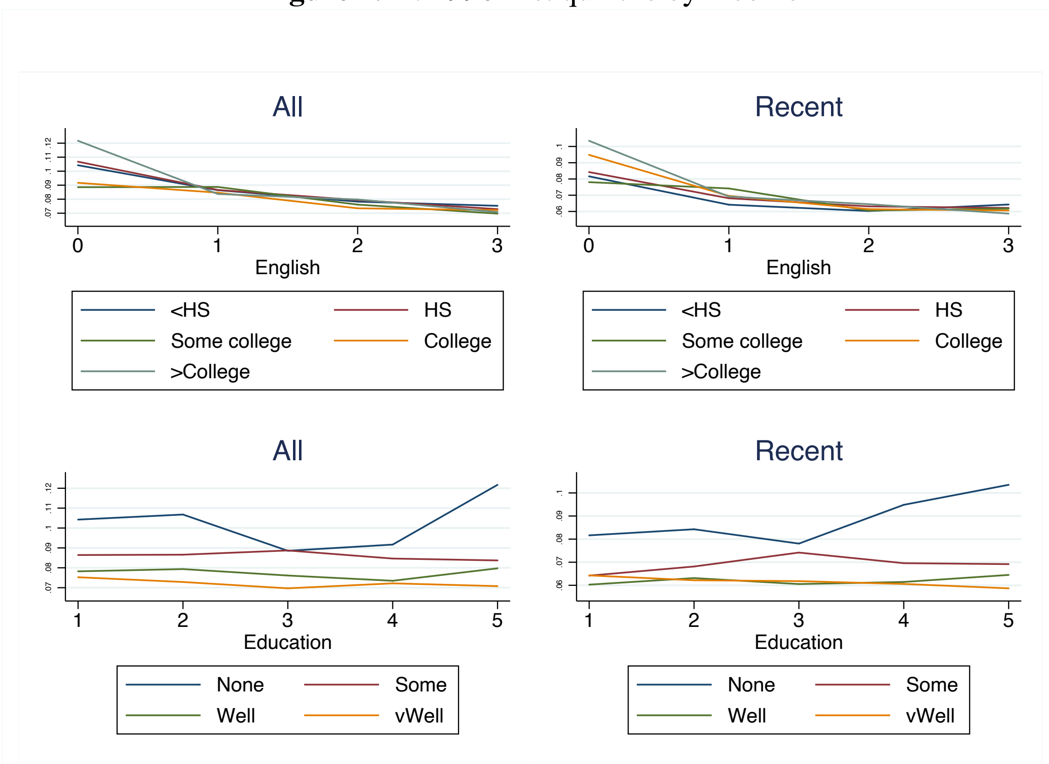


Figure 1.12. 1990 third quintile by income

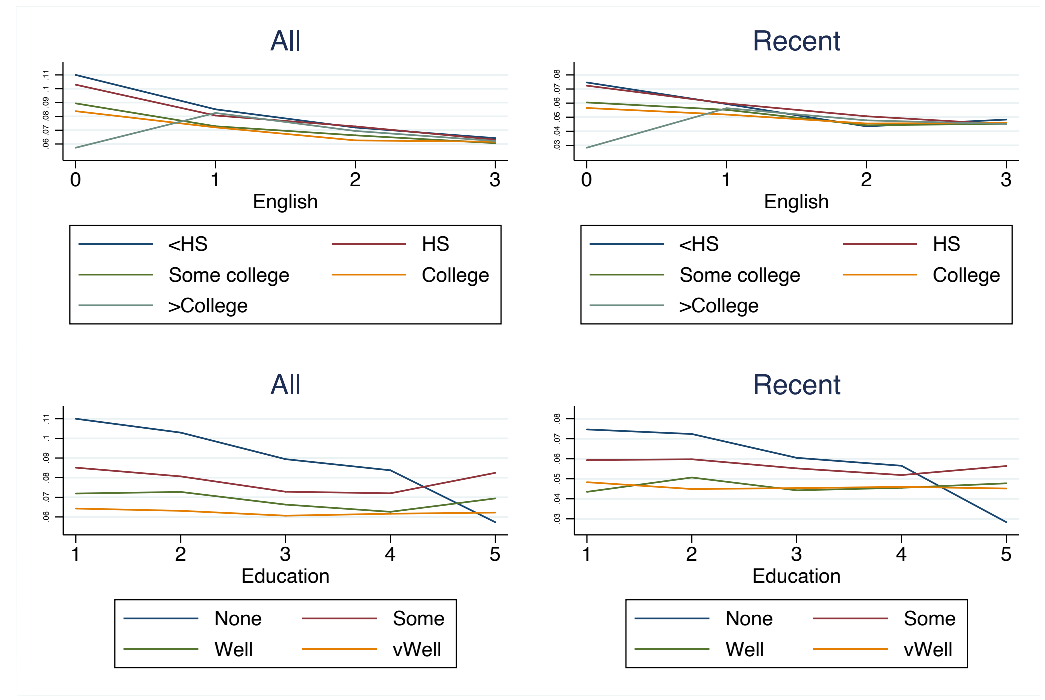
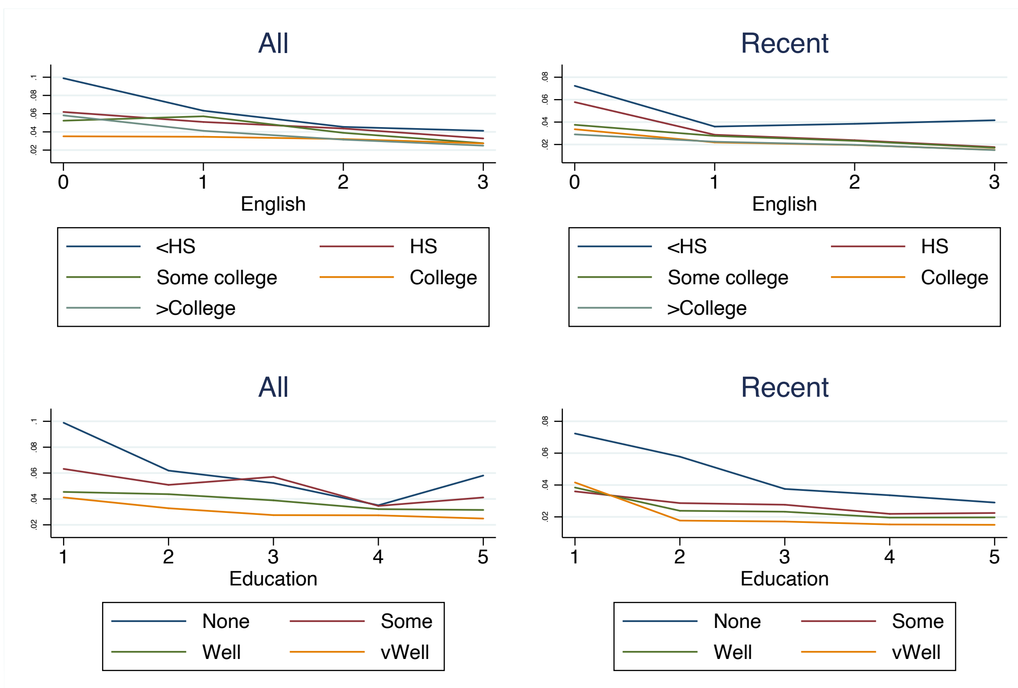


Figure 1.13. 1990 fifth quintile by income



CHAPTER 2

DIVERSITY, GROWTH, AND INEQUALITY

2.1 Introduction

A rich mixture of individual expertise and perspective is indisputably beneficial to the welfare of all. On the other hand, the dissimilarity among individuals could be the source of friction and coordination problems. Social friction, the animosity caused by a clash of ideologies and perspectives between individuals of different backgrounds, presents in daily encounters and interactions between these individuals. Such force limits the degree to which important information could be exchanged and hence, limits the degree to which individuals could learn from each other. Examples of the background differences that could give rise to social friction include the differences in citizenship status, gender, race, political affiliation, and sexual orientation. This paper studies the impact of social friction on the human capital accumulation of individuals in a diverse economy.

This paper extends Lucas (2009), where agents learn from each other through random meetings. A distinct feature of the model is the social friction that limits the extent to which agents from different groups can learn from each other. In particular, the probability that individual skill is successfully revealed in a cross-group meeting, which is a meeting between agents from two different groups, is smaller than that in a same-group meeting.

In the long-run, minority groups lag behind the majority in their aggregate level of human capital due to the asymmetric effects of social friction on the majority and minority groups. The rationale behind the asymmetric effects is that on average, the minority members randomly meet the members from the majority group more often than the members of their group while the opposite is true for the majority members. Therefore, the minority members have fewer opportunities than those in the majority to learn from others because they experience a lower number of successful random meetings. Unsurprisingly, social friction leads to a lower aggregate human capital accumulation rate. The growth rate is lower due to the loss of efficiency in cross-group learning. From the perspective of the majority group, this loss of efficiency is not only due to the failed meetings caused directly by social friction itself but also because of the unproductive meetings with agents from the minority groups who lag behind in the long run. The productivity gap can be closed if the social friction disappears which happens when either there is an exogenous shock that changes the perspective of agents in the economy or members from one group fully integrate into the other.

The existence of the productivity gap suggests that social friction is the potential source of high income-inequality in a diverse economy. When meetings occur continuously, the productivity distribution in the long run for each group is a Frechet distribution with the same tail index. The Gini coefficients for each group are then the same, as a function of the tail index alone. The Gini coefficient for the economy as a whole is higher as it captures additional "between groups" inequality which comes from the productivity gap between the minority and majority groups. This model suggests that the value of group identity, e.g. group language, varies depending on the demographics of the whole economy. A group identity's economic value is scaled by the group's relative population size compared to that of other groups. This implies that members of minority groups could be at a disadvantageous position compared to those from the majority group simply because of their smaller population, all else equal.

Since social friction arises due to differences in cultural values between groups, it would consequently diminish when agents from one group - most likely, agents from the minority groups - assimilate into other groups. For example, immigrants acquire the native language over the time. As a result, agents from the minority groups, who assimilate, would experience less friction, accumulate human capital quicker, become more productive, and close the productivity gap with members from the majority group. The decline of social friction also benefits agents from other groups during meetings with the assimilated agents for two reasons: (i) it is more likely that skills would be be revealed in low friction situations and (ii) the skills of the assimilated agents are likely to be higher. That means in a diverse economy with social friction, the effort that its members take to resolve the friction is below the social optimum as they do not take the benefits others receive into account.

Geographical segregation between agents with different backgrounds is frequently observed in diverse economies. This paper shows that when agents geographically segregate by group, which leads to more frequent meetings with same-ethnic members, the loss of efficiency in the knowledge diffusion process is smaller and hence, the overall human capital accumulation rate is higher in the long-run. Surprisingly, under segregation, the constant productivity gap persists. As a result, the high (income) inequality and related social issues remain even when agents rarely interact with those from a different cultural background. The issue is, now more than ever, relevant.

In an extension of the model where agents can choose to adopt cultural identity, deriving utility from cultural identity and earnings, multiple equilibria could arise. One possible equilibrium is a majority-minority equilibrium where members of the majority group have higher expected utility compared to those of the minority group. Those of the minority group choose to be a minority member if their draw of cultural utility of being a minority is sufficiently high. An equilibrium where the population of each group is more evenly distributed is also possible. However, a distinction between the two equilibrium types is that the majority-minority equilibrium is stable while the more-evenly-distributed equilibrium is not. Starting from the more-evenly-distributed equilibrium, a small disturbance pushes the economy toward a majority-minority equilibrium.

Related Literature

This paper extends Lucas 2009, incorporating friction that has similar flavor to that described in Lucas and Moll 2014. The results of this paper suggests that the decline in native's sentiment toward immigrants (Fouka, Mazumder, and Tabellini 2019) could be a factor that leads to the decline in the economic assimilation of the more recent immigrant cohorts (Borjas 2015, Cassidy 2019). Although the framework is different, this paper is related the work of Durlauf (2006, 2007).

2.2 Model

Consider an economy with heterogeneous agents identified by their type and productivity $(i,z) \in \{a,b\} \times R^{+}$. Denote π^a and π^b for the total mass of type a and b agents. The population is normalized to 1 so that $\pi^a + \pi^b = 1$. For $i \in \{a,b\}$, denote $F(z,i,t)$ for the cumulative distribution at time t. Notice that $\lim_{z\to\infty} F(z,i,t) = \pi_i$. Similar to Lucas 2009, each agent randomly meets others according to a continuous arrival process[1] at rate α. That is in a very small Δ fraction of time, an agent will have $\alpha\Delta$ meetings which include $\alpha\Delta\pi^a$ meetings with another type a agents and $\alpha\Delta\pi^b$ meetings with type b agents. For any $i, j \in \{a,b\}$, a meeting between agent i with another agent j is considered to be a successful meeting if agent i has the opportunity to adopt agent j's productivity. Assume that the probability of a successful meeting between agent i with agent j is p^i_j. Due to the social friction between two groups, assume that $p^i_{j\neq i} = p < p^i_i = 1$. Effectively, in a very small fraction Δ of time, a type a agent has $\alpha\pi^a\Delta$ and $\alpha\pi^b p\Delta$ successful meetings with type a and b agents. In this model, agents wish to maximize their present value of income

[1]The main results throughout the theoretical framework still hold with the Poisson arrival process. The uniqueness of the transitional path cannot be proven since the forward equation system is a complex version of the Lotka-Volterra system.

and thus, upon a successful meeting, agent i adopts agent j's productivity if the latter is more productive.

As shown in Appendix A0, the evolution of the productivity distribution of type a is described by

$$\partial_t \log F(z,a,t) = -\alpha\pi^a \{\log\pi^a - \log F(z,a,t)\} - \alpha\pi^b p \left\{\log\pi^b - \log F(z,b,t)\right\}$$

that is, the change in the scaled productivity distribution is equal to the outflows of agents upon successful meetings with more productive agents. The analogous forward equation can be derived for type b.

Assumption 1 The initial productivity distribution of type a has a Pareto tail with its shape and scale characterized by $0 < \theta < 1$ and $k^a > 0$. That is

$$\lim_{z\to\infty} \frac{\log\pi^a - \log F(z,a,0)}{z^{-1/\theta}} = k^a.$$

Type $b's$ productivity distribution has a thinner tail.

Theorem 1 Under Assumption 1, the economy will eventually converge to a unique non-degenerate balanced growth path (BGP) with growth rate γ, that is $F(z,i,t) = \Phi\left(ze^{-\gamma t}, i\right)$ with $i \in \{a,b\}$. The transition path is unique. The BGP productivity distributions of both groups are Frechet with a tail characterized by θ and the ratio between scale parameters k^a, k^b satisfies

$$\frac{k^a}{k^b} = \frac{\pi^a k^a + \pi^b p k^b}{\pi^a p k^a + \pi^b k^b}. \tag{2.1}$$

The growth rate γ is given by

$$\gamma = \theta\alpha \left[\pi^a + \pi^b p \frac{k^b}{k^a}\right]. \tag{2.2}$$

Proof Shown in Appendix A1.

Productivity gap

Equation (4) can be rewritten in a quadratic form of the ratio of scale parameters. The ratio is the positive root

$$\frac{k^a}{k^b} = \frac{-\left(\pi^b - \pi^a\right) + \sqrt{\left(\pi^b - \pi^a\right)^2 + 4\pi^b\pi^a p^2}}{2\pi^a p}$$

which is a function of π^j and p. Furthermore, given that the BGP productivity distribution for each group is a Frechet, the mean and median of the productivity for each group $j \in \{a,b\}$ are given by

$$\mu^j = \left[k^j\right]^\theta \Gamma(1-\theta)$$
$$m^j = \left[k^j\right]^\theta (\log 2)^{-\theta}.$$

The results imply that the ratio between group income means or medians is a constant equals to $\left(\frac{k^a}{k^b}\right)^\theta$. On the logarithmic scale, the constant productivity gap between two groups is approximately equal to $\theta \log\left(\frac{k^a}{k^b}\right)^2$. In a diverse economy with minority and majority groups, $\pi^b < \pi^a$. Without social friction, $p = 1$, $\frac{k^b}{k^a} = 1$ while with social friction,

[2] Using the linear approximation $E[\log x] \approx \log E[x] - \frac{V[x]}{2E[x^2]}$ as in Y. W. Teh, D. Newman and M. Welling (2006) and that $\frac{V[x]}{2E[x^2]} = \left\{\left(\begin{array}{c} -2\theta \\ -\theta \end{array}\right) - 1\right\}/2\left(\begin{array}{c} -2\theta \\ -\theta \end{array}\right)$ with the derivations from Berg (2008).

$p < 1$, then $\frac{k^b}{k^a} < 1$. Theorem 1 implies that on the BGP, the minority group lags behind the majority in terms of their aggregate productivity level, that is $k^b < k^a$ when $p < 1$. This is because any member of the minority group, on average, experiences more social friction than those from the majority group do, $\alpha\pi^a(1-p)$ as compared to $\alpha\pi^b(1-p)$ failed meetings. When the minority group is less productive compared to the members of the majority group, the minority group has more opportunities to learn from the majority conditional on a successful meeting since on average, members from the majority group are more productive. The opposite is true for the majority group. These two forces equate the growth rates of the two groups in the long-run. The asymmetric effects of social friction could be the source of the unexplained income gap between majority and minority groups in the US.

In scenario of minority and majority groups, the productivity gap cannot be closed unless the social friction completely diminishes, which would likely happen when either there is some shock that helps agents resolve their cultural identity difference or when agents from one group fully integrate into the other group. Could the belief that Asians and Asian Americans are the model minority which was partly propagated to combat racism within the country during World War II (Wu 2013) be a factor the helped the Asian minorities close the wage gap between them and the white majority group?

Theorem 1 also implies that when the present of the minority group in the knowledge diffusion space (workforce for example) increases, the productivity gap between them and the majority group also gets narrower. That is because the social friction's effects become more symmetrical for both minority and majority groups.

However, unsurprisingly, social friction leads to a lower economic growth rate, which is implied by the slower human capital accumulation process. The growth rate is lower due to the loss of efficiency due to social friction. For the majority group, this loss of efficiency is not only due to the failed meetings caused directly by social friction itself $p < 1$ but also because of the unproductive meetings with members of the minority group who lag behind in the long run $\frac{k^b}{k^a} < 1$ - an indirect effect of social friction.

Furthermore, it is simple to show that $\frac{\partial \gamma}{\partial \pi^b} < 0$, that is when the population becomes more even, the aggregate human capital accumulation rate is smaller due to the higher number of failed cross-group meetings.

Inequality

Following Dorfman (1979) and Berg (2008), I compute the productivity Gini coefficient

$$G = \frac{1}{\mu}\int_0^{\infty} F(x)\left[1 - F(x)\right] dx$$

where μ is the mean of $F(x)$. With the Frechet distribution, the Gini index of productivity of each group depends only on the tail parameter θ. In particular, the Gini coefficient for each group is given by $G^j = 2^{\theta} - 1$ which measures the "within group" inequality. On the other hand, the Gini index for the whole economy is given by

$$G = \frac{\left[2^{\theta} - 1\right]\left[(\pi^a)^2 (k^a)^{\theta} + \left(\pi^b\right)^2 \left(k^b\right)^{\theta}\right] + \pi^a \pi^b \left\{2\left(k^a + k^b\right)^{\theta} - (k^a)^{\theta} - \left(k^b\right)^{\theta}\right\}}{\pi^a (k^a)^{\theta} + \pi^b \left(k^b\right)^{\theta}}.$$

It immediately follows that the difference between the two is the measurement of "between groups" inequality,

$$G - G^j = \frac{\pi^a \pi^b \left\{2\left(k^a + k^b\right)^{\theta} - 2^{\theta} (k^a)^{\theta} - 2^{\theta} \left(k^b\right)^{\theta}\right\}}{\pi^a (k^a)^{\theta} + \pi^b \left(k^b\right)^{\theta}} \geq 0$$

since for any $\theta < 1$, $f(x) = x^{\theta}$ is a concave function and therefore, $2\left(\frac{x+y}{2}\right)^{\theta} \geq x^{\theta} + y^{\theta}$ by Jensen's inequality with strict inequality when $k^a \neq k^b$. Without social friction, the normalized productivity distributions of two groups are identical. With the social friction, the minority group is behind the majority group in terms of aggregate productivity and hence, the Gini coefficient is higher as it captures both "within group" inequality G^j and "between groups" inequality $G - G^j$.

Furthermore, the existence of the productivity gap on the BGP suggests that the economic value of a group's cultural identity is scaled to the relative population size of the group. Members of the majority group get ahead compared to those of the minority group because they exchange information more efficiently because they encounter members of their group more frequently. In contrast, members of the minority groups are at disadvantageous position and lag behind simply because of their smaller population size. This does not suggest inequality can be reduced if the population size of the minority group increases since other factors such as education, occupation need to be taken into account.

Geographical Segregation

Geographical segregation between the minority and majority groups is frequently observed across diverse economies. When groups segregate geographically, they randomly meet members of their group more often. Denote q_a^a and q_b^a for the fractions of meetings of type a with another a and b, respectively. Without segregation, the ratio between meetings should be equal to the ratio of the population between the two types $\frac{q_a^a}{q_b^a} = \frac{\pi_a}{\pi_b}$. With segregation, the native members interact with themselves more often, that is $\frac{q_a^a}{q_b^a} > \frac{\pi_a}{\pi_b}$. If the number of meetings that type a has with type b is the same as the number of meetings that type b has with type a, that is $\pi_a q_b^a = \pi_b q_a^b$, it follows that $\frac{q_b^b}{q_a^b} > \frac{\pi^b}{\pi^a}$. When groups segregate geographically, the forward equation becomes

$$\partial_t \log F(z,a,t) = -\alpha q_a^a \{\log \pi^a - \log F(z,a,t)\} - \alpha q_b^a p \left\{\log \pi^b - \log F(z,b,t)\right\}.$$

Proposition With segregation, the growth rate is higher in the long run while the productivity gap persists.

Proof Shown in appendix A3.

With segregation, the economy is different from the perspectives of the minority and majority groups in the knowledge diffusion space. In the perspective of the majority, the economy has a population of q_a^a and q_b^b of type a and b respectively. This implies the number of unproductive meetings with members of the minority group that the majority agents encounter is smaller than that without segregation. Hence, the loss of efficiency in the knowledge diffusion process is smaller for the majority members. Similarly, members of the minority group experience less failed meetings as well. As a result, the BGP growth rate is higher with segregation due to the lower loss of efficiency in the learning technology.

Surprisingly, the long-run persistence of the productivity gap between the minority and majority groups depends only on the demographics and the social friction level as shown in Appendix A2. This implies that income inequality and social issues related to income inequality persist when the two group geographically segregate. In a general equilibrium set-up, agents are likely to have a lower incentive to integrate under segregation as the gain from the smaller number of successful meetings with members from other groups is smaller.

Rate of convergence

Proposition 2 The rate of convergence to the BGP is given by

$$\mid \Lambda_1 - \Lambda_2 \mid = \alpha \sqrt{1 - 4\pi^a \pi^b \left(1 - p^2\right)}. \tag{2.3}$$

In a segregated economy, the rate of convergence is given by

$$|\Lambda_1 - \Lambda_2| = \alpha \frac{q_b^a}{\pi^b} \sqrt{1 - 4\pi^a \pi^b \left(1 - p^2\right)}. \tag{2.4}$$

Proof Shown in appendix A2.

Consider the scenario where the minority group's productivity distribution initially lags behind that of majority group, the model implies that aggregate productivity growth rate of the minority group is lower if social friction is higher, $p \uparrow$. Starting from the same point, the minority group whose cultural value does not significantly differ from that of the majority group experiences less social friction, has more opportunity to learn and quickly catch up with the majority group. The model also implies that the size of minority group reduces the convergence rate. This is due to the meetings with more productive type a is crowded out by random encounters with other type b.

When the two groups segregate geographically or socially, the convergence rate is lower as a result of lower random meetings with more productive type a agents $q_b^a \downarrow$. The convergence rate also declines as the population of the minority group increases $\pi^b \uparrow$ for the same reason as earlier, productive meetings with type a agents and/or type b agents who successfully acquire skill from type a are crowded out. In other words, since the source of ideas q_b^a is scarce while the receivers of ideas π^b are abundant, it takes a longer time for an idea to spread across all receivers, given fixed the idea spreading technology.

Gain from the reduction of friction

Agents in a diverse economy with friction naturally have an incentive to integrate or assimilate into the mainstream culture. Agents who experience the friction would gradually adopt a cultural identity that closely resembles the mainstream culture. For example, immigrants acquire the native language over time, which not only opens up new employment opportunities but also helps them acquire new skills from natives. Generally, it is more likely that members from the minority groups would have a higher incentive to integrate into the mainstream culture if they undergo friction because of their cultural identity.

If friction decreases, the probability that a cross group meeting succeeds is higher and from the BGP growth rate equation (2),

$$\frac{d\gamma}{dp} = \theta\alpha\pi^b \frac{k^b}{k^a} + \theta\alpha\pi^b \frac{d}{dp}\left[\frac{k^b}{k^a}\right] > 0$$

since

$$\frac{d}{dp}\left[\frac{k^b}{k^a}\right] > 0$$

when $\pi^b < \pi^a$. That is, as social friction declines, the majority group's growth rate increases due to (i) the direct benefit from the increasing number of successful cross-group meetings with members of minority groups and (ii) the indirect benefit from meetings with more productive members of the minority groups as the productivity gap becomes smaller.

In a generalized model where agents can put effort into assimilating into mainstream culture, for example type b can put effort into acquire type a language skill, they put in less effort into doing so as compared to social optimum as they don't take into account the external benefits that they would bring to other agents, type a agents who benefit from experiencing less communication friction upon encounter with them and other type b agents who encounter them as they are more productive as a result. The theoretical ground for this argument is Lucas and Moll 2014 in which they show agents, who could invest effort into the learning from others technology so they could learn faster, invest less than social optimum as they don't take the value that others gain by learning from them into account.

One would argue that the externalities would be smaller when the groups segregate. This is not entirely true. Recall that the productivity gap persists under segregation. As long as he does not lose his own group identity entirely, a member from the minority group would become a knowledge bridge to other members in his community as he partially integrates into the mainstream culture and acquires more knowledge from members of the majority group.

Multiple Equilibria of Diverse Economy

Consider the case where agents can choose to one of the cultural identities to adopt at the beginning of their lives, for example religions. Agents' utility from being type a and b comes from two components, cultural identity and earnings. The utility from the cultural identities are permanent taste shocks that are randomly drawn from independent Frechet distributions, $c^i \sim \exp\left(-G^i c^{-1/\varepsilon}\right)$ where G^i is the scale parameter governs average utility of being type i and ε governs the dispersion of the taste shocks. Agents have one unit of time which they inelastically spend working, earning a flow of real income flow equals to their productivity. The lifetime utility of being type i is

$$V(z,t;i) = E_t\left\{\int_0^\infty e^{-\rho(\tau-t)} c^i z(\tau)\, d\tau \mid z(\tau) = z\right\}$$

where ρ is the discount rate. On the BGP, the Bellman equation of type i is

$$\rho V(x;i) = c^i x + \alpha \int_x^\infty \left\{V\left(x';i\right) - V(x;i)\right\} D_{x'} \log \Phi^i\left(x'\right) - V'(x;i)\, x\gamma + V(x;i)\,\gamma$$

where $V(x;i)$ is the detrended present value of utility, x is the detrended productivity, $\Phi^i\left(x'\right)$ is the effective productivity distribution from the perspective of type i with $\Phi^i\left(x'\right) = \exp\left(-k^i x^{-\frac{1}{\theta}}\right)$ and $k^i = \pi^i k^i + \pi^{-i} p k^{-i}$ is the effective scale parameter from the perspective of type i. Since c^i is the permanent taste shock, it can be pulled out of the Bellman equation and hence, $V(x;i) = c^i v(x;i)$ where $v(x;i)$ is the present value of earnings.

A new-born agent in this economy chooses to adopt any cultural identity to maximize his lifetime utility. After adopting a cultural identity i, he instantaneously become a random existing type i agent in the economy. On the BGP, the fraction of type i is given by

$$\pi^i = \frac{G^i \left[E_x v(x;i)\right]^{\frac{1}{\varepsilon}}}{G^a \left[E_x v(x;a)\right]^{\frac{1}{\varepsilon}} + G^b \left[E_x v(x;b)\right]^{\frac{1}{\varepsilon}}}, \tag{2.5}$$

following the derivation of Ahlfeldt et al 2015. Note that when $\varepsilon \to \infty$, the utility draws of cultural identity are extremely dispersed and govern the entire identity choices. Hence, we have a unique equilibrium with the population of type i is given by $\pi^i = G^i / \left(G^a + G^b\right)$.

In the other extreme, when $\varepsilon \to 0$, the present value of earnings completely determine the cultural identity choices. In this case, when there is social/social friction $p > 0$, there are two possible equilibria, a homogeneity equilibrium with only one type and a diverse economy with equal population of two types. The homogeneity equilibrium is stable while the latter is unstable as a small deviation from the equilibrium eventually pushes new-born agents toward choosing the cultural identity of the majority group.

When ε is finite, multiple equilibria can naturally arise in this diverse economy. Consider the case where $G^a = G^b$, the average cultural utility of being type a and type b are the same. (2.5) implies that the ratio of the two type populations equal to the ratio of expected lifetime earnings to the power of $1/\varepsilon$,

$$\frac{\pi^a}{\pi^b} = \frac{\left[E_x v(x;a)\right]^{\frac{1}{\varepsilon}}}{\left[E_x v(x;b)\right]^{\frac{1}{\varepsilon}}}. \tag{2.6}$$

On the other hand, the RHS, the ratio of expected lifetime earnings, is also a function of the ratio of the population since it depends on the economy's growth rate and its productivity gap between type a and b. Consider an example with $\rho = 0.06, \alpha = 0.05, \theta = 0.4$, and

$p = 0.2$. I solve the Bellman equations using finite difference method and compute the RHS of (2.6) as a function of the population ratio.

Figure 2.1 shows the RHS and LHS of (2.6) for a range of value of ε in which where the two lines intersect is an equilibrium. The result shows that when ε goes to infinity, there is only one possible equilibrium which is completely determined by tastes. In this case since the average cultural utilities of being type a and type b are identical, the populations of type a and b are half and half. As ε declines, the cultural identity draws are less dispersed, we start to see the rise of multiple equilibria in this economy even though on average, the cultural utilities of being type a and type b are the same. These equilibria can be either interpreted as a self-fulfilling equilibrium or one that arises from the pre-historic conditions, i.e. exogenous reasons that lead to such equilibrium. Furthermore, note that when $\varepsilon$is small, i.e. $\varepsilon = 1/4; 1/6$, the equal population diverse economy is unstable while the majority-minority equilibria are stable.

Diversity, Clans, and Growth

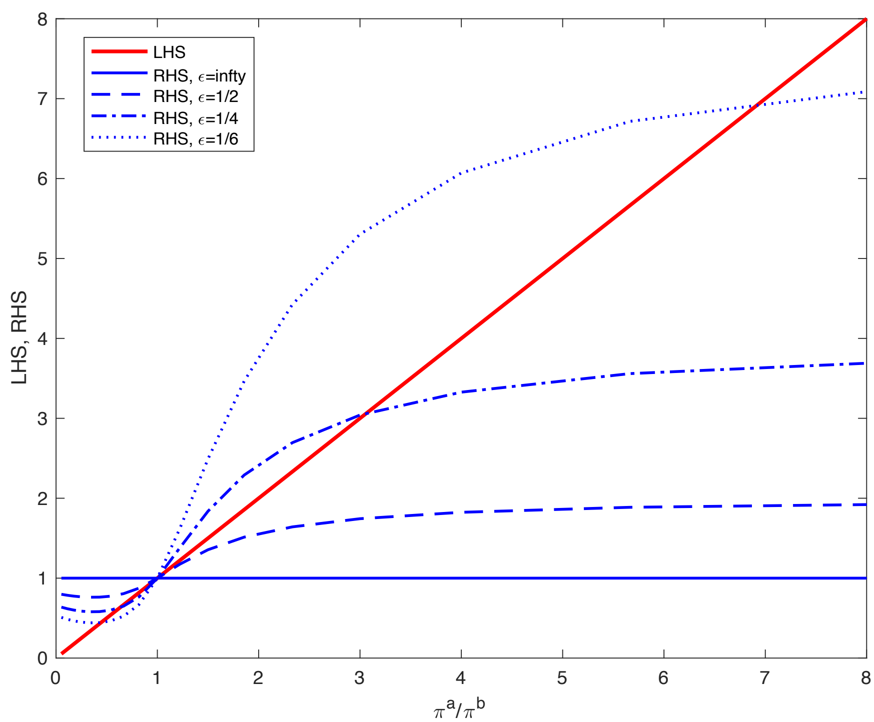
The results can be generalized for more than two groups. If there are N groups indexed by $i = 1,..,N$, the growth rate of this economy along the BGP is given by

$$\gamma = \theta\alpha \sum_{j=1}^{N} \pi^j p^i_j \frac{k^j}{k^i}.$$

where k_j is the scale parameters of group j's productivity distribution and p^i_j is the probability of a successful meetings between agents from group i and j in a cross- group meeting. In the symmetric case, that is $\pi^i = \pi = \frac{1}{N}, p^i_i = 1$, and $p^i_{j\neq i} = p < 1$ for all i, the growth rate is given by

$$\gamma = \theta\alpha \left[\frac{1}{N} + \frac{N-1}{N} p\right].$$

Figure 2.1. Multiequilibria of the diverse economies



The model suggests that if the economy is more diverse, or fractionalized, the aggregate economic growth rate is smaller. The growth rate along the BGP converges to $\theta\alpha p$ as N becomes very large. That is as $N \to \infty$, the economy is asymptotic to an economy in which meetings arrive continuously at rate $\alpha' = \alpha p$. This can be looked at from the opposite direction. Suppose that the formation of groups, or clans, allows members within a group exchange information easier, which leads to a higher human capital accumulation rate. Although with a different mechanism, this result resembles that of Croix, Doepke, and Mokyr (2016) who argue that the formation of clans and guilds helped Western Europe pull ahead compared to other regions. In general, the effect of social friction on growth could be smaller in a more diverse economy due to (i) geographical segregation which, as mentioned earlier, leads to a lower number of cross-group meetings and (ii) agents' higher incentive to integrate into the mainstream culture, e.g. by acquiring the common language (of the majority group - if there is one), when the minority population is smaller.

2.3 Concluding Remarks

This paper proposes a simple framework to study the effect of social friction on the human capital accumulation process. This paper shows that social friction has a negative impact on the productivity growth for both the majority and the minority groups. Interestingly, social friction leads to a constant productivity gap in a logarithmic scale between the majority and minority in the long run. This productivity gap could only be closed when the social friction completely fades away—that is, when either the friction that arises from differences in cultural values is resolved completely or when one group completely integrates into the other groups. When groups segregate geographically, the effect of social friction is mitigated, and the economy grows faster. Surprisingly, the productivity gap persists under segregation which suggests that inequality and other related social issues remain the same under segregation.

This paper shows that in the diverse economy with social friction, income inequality, measured by the Gini coefficient, is higher because of the additional "between groups" inequality that emerges from the productivity gap. This paper suggests that the economic value of a group's cultural identity is scaled by the relative population size of the group. Members from the minority group are at a disadvantageous position purely due to their smaller population size. However, this paper does not suggest that such inequality could be resolved via a high reproduction rate among minority group, as other factors such as the source of knowledge or individual educational attainment have not been considered.

Fouka, Mazumder, and Tabellini 2019 document the change in the attitude of white natives toward black and Hispanic due to the change in the immigrant population over time. This paper shows that in the immigrant scenario, that means higher social friction would lower the human capital assimilation rate of immigrants. Furthermore, higher friction also leads to wider human capital gap between immigrants and the natives, on the balanced growth path (See Theorem 1 and Proposition 2). This could be an explanation for the decline of the economic assimilation of the more recent immigrant cohorts (Borjas 2015, Cassidy 2019).

The broad definition of groups also allows the model to be applied to several other scenarios, giving it the ability to explain potential unexplained wage gap between genders, races, and how that changes with time. Did the increase of workforce participation of women lead to the narrower earning gap between them and their male counterparts? How's about black and white Americans? It is natural to further investigate these questions under the new perspective offered through the lens of this paper.

2.4 Appendix

A0

For type a agents, given the cumulative productivity distribution of type a at time t is $F(z,a,t)$, for a small increment of time h, the cumulative productivity distribution at time $t+h$ is given by

$$F(z,a,t+h) = \Pr\{\text{type } a \text{ with productivity} \leq z \text{ at } t+h\}$$

$$F(z,a,t+h) = \Pr\{\text{type } a \text{ with productivity} \leq z \text{ at } t\ \} \times$$
$$\Pr\left\{\text{all } \alpha_a^a h \text{ and } \alpha_b^a h \text{ draws} \leq z \text{ at } t\right\}$$

$$F(z,a,t+h) = F(z,a,t)\left\{\left(\frac{F(z,a,t)}{\pi^a}\right)^{\alpha_a^a h}\left(\frac{F(z,b,t)}{\pi^b}\right)^{\alpha_b^a h}\right\}.$$

By manipulating the above equation we obtain,

$$\frac{F(z,a,t+h) - F(z,a,t)}{F(z,a,t)\,h} = \frac{\left\{\left(\frac{F(z,a,t)}{\pi^a}\right)^{\alpha_a^a h}\left(\frac{F(z,b,t)}{\pi^b}\right)^{\alpha_b^a h} - 1\right\}}{h}$$

and since

$$\lim_{h\to 0}\frac{f(t+h) - f(t)}{f(t)\,h} = \partial_t \log f(t),$$
$$\lim_{h\to 0}\frac{f(t)^{ah} - 1}{h} = a\log f(t),$$

the evolution of the productivity distribution of type a is described by

$$\partial_t \log F(z,a,t) = \alpha_a^a\{\log F(z,a,t) - \log \pi^a\} + \alpha_b^a\left\{\log F(z,b,t) - \log \pi^b\right\}.$$

A1

This proof is inspired by Lucas' lectures on International Trade & Growth. For simplicity, the proof of Theorem 1 is broken down into three parts. Lemma 1 guarantees the existence of the BGP. Lemma 2 shows that the BGP productivity distributions are Frechet distributions. Combining with Lemma 1 and 2, the last part proves the unique transition path of the productivity distribution.

Assumption 1 For $i \in \{a, b\}$, the initial distribution of productivity for type i has its shape and scale characterized by θ and k_i. That is

$$\lim_{z \to \infty} \frac{\log F(z, i, 0) - \log \pi_i}{z^{-1/\theta}} = -k_i. \tag{2.7}$$

Assumption 2 The scale parameters satisfy

$$\frac{k^a}{k^b} = \frac{\alpha_a^a k^a + \alpha_b^a k^b}{\alpha_a^b k^a + \alpha_b^b k^b}. \tag{2.8}$$

Lemma 1 Under Assumption 1 and 2, there exists a non-degenerate BGP such that two groups grow at the same rate γ, that is $F(z, i, t) = F\left(ze^{-\gamma t}, i, 0\right) = \Phi\left(ze^{-\gamma t}, i\right)$ for $i \in \{a, b\}$. The growth rate γ is given by

$$\gamma = \left[\alpha_a^a + \alpha_b^a \frac{k^b}{k^a}\right] \theta$$

Proof Denote $x = ze^{-\gamma t}$. For type a, along the BGP, $F(z, a, t) = \Phi\left(ze^{-\gamma t}, a\right) = \Phi(x, a)$. The forward equations become

$$\frac{-\gamma \phi(x, a) x}{\Phi(x, a)} = \alpha_a^a \{\log \Phi(x, a) - \log \pi^a\} + \alpha_b^a \left\{\log \Phi(x, b) - \log \pi^b\right\} \tag{2.9}$$

From (8) and the above conjecture, $F(z,a,t) = F\left(ze^{-\gamma t},a,0\right) = \Phi\left(ze^{-\gamma t},a\right)$,

$$\lim_{x\to\infty} \frac{\log\Phi(x,a) - \log\pi^a}{x^{-1/\theta}} = \lim_{z\to\infty} \frac{\log F\left(ze^{-\gamma t},a,0\right) - \log\pi^a}{\left(ze^{-\gamma t}\right)^{-1/\theta}} = -k^a. \tag{2.10}$$

By L'Hospital rules,

$$\lim_{x\to\infty} \frac{\frac{\phi(x,a)}{\Phi(x,a)}}{-\frac{1}{\theta}x^{-1/\theta-1}} = -k^a$$

Plugging in and for large x, the forward equations imply

$$-\frac{\gamma}{\theta}k^a x^{-1/\theta} = -\alpha_a^a k^a x^{-1/\theta} - \alpha_b^a k^b x^{-1/\theta}$$

Simplifying above, we obtain

$$\gamma = \left[\alpha_a^a + \alpha_b^a \frac{k^b}{k^a}\right]\theta.$$

Since this equation hold for both a and b, k^a and k^b must satisfy Assumption 2. That is,

$$\frac{k^a}{k^b} = \frac{\alpha_a^a k^a + \alpha_b^a k^b}{\alpha_a^b k^a + \alpha_b^b k^b}.$$

QED.

Lemma 2 The BGP productivity distributions in Lemma 1 are Frechet with tail characterized by θ and the ratio of scale parameters k^a, k^b satisfy

$$\frac{k^a}{k^b} = \frac{\alpha_a^a k^a + \alpha_b^a k^b}{\alpha_a^b k^a + \alpha_b^b k^b}.$$

Proof The system of differential equations in (10) is a system of first-order linear ordinary differential equations, the solution exists and is unique. Denote

$$X(x) = \begin{bmatrix} \log\Phi(x,a) - \log\pi^a & \log\Phi(x,b) - \log\pi^b \end{bmatrix}'$$

then (10) can be written as

$$xX' = -\frac{1}{\gamma}AX$$

where

$$A = \begin{bmatrix} \alpha_a^a & \alpha_b^a \\ \alpha_a^b & \alpha_b^b \end{bmatrix}.$$

This system has the solution in the form

$$\begin{aligned} X &= x^{-\frac{1}{\gamma}A}C \\ &= e^{-\frac{1}{\gamma}A\log x}C \end{aligned}$$

where $C = \begin{bmatrix} c_1 & c_2 \end{bmatrix}'$ is a vector of constant terms. Denote $\{v_1, v_2\}, \{\Lambda_1, \Lambda_2\}$ for the eigenvectors and eigenvalues of matrix A. V, Λ for the matrix forms. If k is any constant, then the eigenvectors and the eigenvalues of matrix kA is $V, k\Lambda$. Thus, the above solution can be simplified as

$$\begin{aligned} X &= Ve^{-\frac{1}{\gamma}\log x\Lambda}V^{-1}C \\ &= V\begin{bmatrix} x^{-\frac{\Lambda_1}{\gamma}} & 0 \\ 0 & x^{-\frac{\Lambda_2}{\gamma}} \end{bmatrix}V^{-1}C \end{aligned}$$

where C is obtained by using the boundary conditions (8). Denote $C' = V^{-1}C = \begin{bmatrix} c_1' & c_2' \end{bmatrix}'$ for the new constant terms. Then the solution above can be simplified as

$$\log\Phi(x) - \log\pi = c_1' v_1 x^{-\frac{\Lambda_1}{\gamma}} + c_2' v_2 x^{-\frac{\Lambda_2}{\gamma}}$$

As showed later, $\frac{\Lambda_1}{\gamma} = \frac{1}{\theta}$ and $\Lambda_2 < \frac{\gamma}{\theta}$ and $kv_1 = \begin{bmatrix} k^a & k^b \end{bmatrix}'$ for some constant k. By imposing the boundary conditions, the constant term c_2' is equal to zero. The BGP productivity distribution is, therefore,

$$\log\Phi(x) - \log\pi = c_1' v_1 x^{-\frac{1}{\theta}}.$$

QED.

Assumption 3 The initial productivity distribution of type a has its shape and scale characterized by θ and k^a. That is

$$\lim_{z\to\infty} \frac{\log F(z,a,0) - \log\pi^a}{z^{-1/\theta}} = -k^a.$$

The initial productivity distribution of type b either has thinner tail or finite support.

Theorem 1 Under Assumption 3, the economy will eventually converge to a unique non-degenerate BGP with growth rate γ, that is $F(z,i,t) = \Phi\left(ze^{-\gamma t}, i\right)$ with $i \in \{a,b\}$. The BGP productivity distributions of both groups are Frechet with tail characterized by θ and the ratio between scale parameters k^a, k^b satisfies

$$\frac{k^a}{k^b} = \frac{\pi^a k^a + \pi^b p k^b}{\pi^a p k^a + \pi^b k^b}.$$

The growth rate γ is given by

$$\gamma = \theta\alpha\left[\pi^a + \pi^b p \frac{k^b}{k^a}\right].$$

The transition path is unique.

Proof Rewrite the system of forward equations in the matrix form

$$\partial_t X(z,t) = AX(z,t)$$

where

$$X(z,t) = \begin{bmatrix} \log F(z,a,t) - \log\pi^a & \log F(z,b,t) - \log\pi^b \end{bmatrix}'$$
$$A = \begin{bmatrix} \alpha_a^a & \alpha_b^a \\ \alpha_a^b & \alpha_b^b \end{bmatrix}$$

The general solution is given by

$$X(z,t) = e^{At}X(z,0)$$

Denote $\{v_1, v_2\}, \{\Lambda_1, \Lambda_2\}$ for the eigenvectors and eigenvalues of matrix A. V, Λ for the matrix forms. We can write the solution as

$$X(z,t) = Ve^{\Lambda t}V^{-1}X(z,0).$$

Denote $T = \alpha_a^a + \alpha_b^b$ and $D = \alpha_a^a\alpha_b^b - \alpha_b^a\alpha_a^b$. The eigenvalues of the system are given by

$$\Lambda_1 = T/2 + \left(T^2/4 - D\right)^{1/2}$$
$$\Lambda_2 = T/2 - \left(T^2/4 - D\right)^{1/2}$$

where both eigenvalues are positive and $\Lambda_1 > \Lambda_2$. Thus, the cumulative productivity distribution of type a at time $t$along the trend, given $F(z,a,0)$, is

$$\log F\left(ze^{\gamma t},t\right) - \log \pi = Ve^{\Lambda t}V^{-1}\left\{\log F\left(ze^{\gamma t},0\right) - \log \pi\right\}.$$

In the case that group b initially has finite support, for large t, the RHS will have the form $c_1 v_1 e^{\Lambda_1 t} e^{-\frac{\gamma}{\theta}t} + c_2 v_2 e^{\Lambda_2 t} e^{-\frac{\gamma}{\theta}t}$ where c_1, c_2 are some functions of z which are constant with respect to time. Therefore, in order to show that for large t, the cumulative productivity distribution along the trend is stationary, we only need to show that

$$\frac{\gamma}{\theta} = \Lambda_1 > \Lambda_2.$$

From Lemma 1,

$$\frac{\gamma}{\theta} = \left[\alpha_b^b + \alpha_a^b \frac{k^a}{k^b}\right]. \tag{2.11}$$

By solving $\frac{k^a}{k^b}$ that satisfies (9), we obtain

$$\frac{k^a}{k^b} = \frac{-\left(\alpha_b^b - \alpha_a^a\right) + \sqrt{\left(\alpha_b^b - \alpha_a^a\right)^2 + 4\alpha_a^b\alpha_b^a}}{2\alpha_a^b}. \tag{2.12}$$

By substituting (13) into (12), we obtain

$$\frac{\gamma}{\theta} = \frac{\left(\alpha_b^b + \alpha_a^a\right) + \sqrt{\left(\alpha_b^b - \alpha_a^a\right)^2 + 4\alpha_a^b \alpha_b^a}}{2}$$

which is exactly the same as Λ_1! Furthermore, since the system of forward equations is a system of linear differential equations and A is a system of continuous functions, the solution is unique for any given initial condition. The transition path is, therefore, unique. The economy follows a unique transition path and converges the unique BGP in Lemma 1. QED.

A2

From equation (5), the non-stationary component $c_2 v_2 e^{\Lambda_2 t} e^{-\frac{\gamma}{\theta} t}$ of the productivity distribution is diminished at rate $\frac{\gamma}{\theta} - \Lambda_2$ with respect to time. Since I have shown that $\Lambda_1 = \frac{\gamma}{\theta}$, it is straightforward that the convergence rate for the productivity distribution is given by $\Lambda_1 - \Lambda_2$. Recall that the eigenvalues of the system are given by

$$\Lambda_1 = T/2 + \left(T^2/4 - D\right)^{1/2}$$
$$\Lambda_2 = T/2 - \left(T^2/4 - D\right)^{1/2}$$

and $T = \alpha_a^a + \alpha_b^b$ and $D = \alpha_a^a \alpha_b^b - \alpha_b^a \alpha_a^b$. Here $\alpha_a^a = \alpha \pi^a, \alpha_b^a = \alpha \pi^b p$, and so on. By substitution, I obtain the desired results. Similar can be done for the segregated economy.

A3

The income gap between the natives and immigrants, which is a function of $\frac{k^a}{k^b}$, does not depend on the degree of segregation in this model.

$$
\begin{aligned}
\frac{k_a}{k_b} &= \frac{-\left(q_b^b - q_a^a\right) + \sqrt{\left(q_b^b - q_a^a\right)^2 + 4q_a^b q_b^a p^2}}{2q_a^b p} \\
&= \frac{-\left(1 - \frac{\pi^a}{\pi^b}\right)(1 - q_a^a) + \sqrt{\left(1 - \frac{\pi^a}{\pi^b}\right)(1 - q_a^a)^2 + 4\frac{\pi^a}{\pi^b}(1 - q_a^a)^2 p^2}}{2\frac{\pi^a}{\pi^b}(1 - q_a^a)\, p} \\
&= \frac{-\left(\pi^b - \pi^a\right) + \sqrt{\left(\pi^b - \pi^a\right)^2 + 4\pi^a \pi^b p^2}}{2\pi^a p}
\end{aligned}
$$

CHAPTER 3

REGIONAL IMPACTS OF IMMIGRATION

3.1 Introduction

This paper proposes model to study the regional impact of immigration. In particular, this paper seeks to understand the determinants of immigrants' residential and occupational choices, why their choices differ from that of natives. Furthermore, this paper proposes a model to evaluate (i) change in regional demographics, (ii) change in wages across occupations and regions, (iii) change in natives' residential and occupational choices in response to immigration, and (iv) change in immigrants' incentive to acquire English language skill due to a shift in immigration policy that alters the magnitude and composition of immigrant inflow.

This paper proposes a spatial economy model built upon the generalized framework of Eaton and Kortum 2002; Allen and Arkolaskis 2014; Ahlfeldt, Redding, Sturm, and Wolf 2015 and incorporate agent heterogeneity. In particular, agents differ by their types, native or immigrant, and the type of agents govern (i) labor productivity across occupation that reflect their endowed human capital and (ii) their taste for location/occupation that reflect the historical/social connections that influence these choices besides wage.

The model shows that the determinants of immigrants' geographic and occupation concentration are the comparative advantages that immigrants have over natives and other immigrants. The comparative advantages of living in a certain location that a group of immigrants have include but are not limited to utility derived from living close to their relatives (other existing immigrants with the same ethnicity), proximity to their home country, preference for local weather climate, etc. These are unobserved comparative advantages and are modeled with the random taste shocks drawn from Frechet distributions.

Through a simple example of a line economy with one industry, this paper shows the difference in taste shocks causes one group to have a biased preference to live in some regions which lead spatial segregation and labor shortage in other regions. Through the price effect, the biased preference of one group for some regions then gives other groups the comparative advantage to relocate to other regions. Productivity spillover due to agglomeration reduces spatial segregation while a potential negative externality that arises from communication friction leads to further spatial segregation.

The model is extended to incorporate the occupation choices of immigrants. Similar to location, immigrants decide which occupation depending on wage that reflect the price of their human capital augmented labor and their personal preference – taste. Due to their limited English language skill, immigrants are less productive when working in English-language-skill-intensive occupations. Hence, immigrants are sorted into occupations that require less English usage as they have a comparative advantage over natives in these occupations. Because of the tie between an occupation labor productivity and a location, immigrants are also further sorted to locations where less English-intensive occupations are more productive compared to other regions as shown in the second numerical example – the line economy with two industries. The underlying assumption of the connection between occupation labor productivity and location includes the availability of natural resources, weather climate, private and public capital, etc. that favor one occupation over the other.

Furthermore, this paper lays out the dynamic framework to incorporate the inflow of immigrants over time and their decision to invest into the native-specific human capital. In the dynamic model, agents, or immigrants specifically, decide how much time to work and to learn English. As immigrants become proficient in English, the comparative advantage edges of working in less English-intensive occupations become smaller and hence, they might choose to reside in a different location and or work a different job. As a result, immigrant cohorts' residential and occupation choices become more and more alike with that of natives. In this model, the incentive to assimilate of immigrants come from the (i) increment in wage within the same occupation or by switching to a different occupation and (ii) increment in utility derived from the taste of working in a different occupation. It's straightforward that as the newcomers acquire the native language skill, their residential and occupational choices become more comparable to that of natives as the comparative (dis)advantages become smaller.

Taking the model to the data, this paper uses American Community Survey 2001-2016 and Occupational Information Network Survey and finds that:

1. There is a high positive correlation the residential dissimilarity index and occupation dissimilarity index – the measures of the differences in residential and occupational choices of immigrants vs. natives.

2. The difference in residential and occupation choices between immigrants and natives increase with their age but decreases with their English proficiency, years of education, having a US education, and the size of pre-existing immigrant stock.

3. The difference in occupation choices between immigrants and natives decline with their duration of stay in the US, stronger among younger immigrants.

4. Analogous to the previous point, immigrants are highly concentrated in occupations that are less English-intensive and become less concentrated over time.

The model's predictions as seen through numerical examples conform to these observed patterns in the data. The economic insights derived from the model, hence, are potentially relevant to policymakers. These insights include (i) size and composition of immigrant inflow could have differential effects on wage, living cost across regions and industries as well as the incentive to integrate of existing immigrants and (ii) the impact of a immigrant cohort varies with the duration of their stay as they acquire relevant skillsets and the comparative (dis)advantages between them and natives get smaller.

3.2 Model

Set up

In this section, I will describe the spatial economy with heterogeneous agents who differ by their taste (because of their ethnic origin), their education, and their English proficiency. Let S, I, J denote the set of locations, agent types, and goods types. For convenience, let $s \in S, i \in I, j \in J$ index location s, agent type i, and goods type j. Here, the agent type index $i = \{eth, edu, Eng\}$ has three components: ethnic origin, educational attainment, and English proficiency. Agents from different ethnic origins draw taste preference for locations and occupations from different taste distributions. The discrepancy in geographical taste distributions between different types is assumed to come from factors such as the geographical proximity to home country, the distribution of relatives, the similarity between the destination's climate and the home country's, etc. Furthermore, agents with higher educational attainment and/or are more proficient in English have higher labor productivity. Because of the difference between types, agents of one type then have comparative advantages/disadvantages versus those of other types in living in certain locations and/or working in certain industries. Such comparative advantages/disadvantages, hence, lead to the observed residential and occupation concentration patterns of natives and immigrants.

Locations and transportation cost

Location s differs from others by its set of labor productivities $\left\{A^i_{s,j}\right\}_{i\in I, j\in J}$ and its amenity U_s. The set of labor productivities differ across locations because of the difference in local natural resources (commercial land, proximity to water source, weather climate), private capitals (physical and human capitals of firms), and public capitals (infrastructures, taxes, access to political institutions). Locations are connected by a bilateral transportation network that can be used to ship goods subject to the iceberg trade cost. Denote $T^j_{ks} \geq 1$ for the number of units of goods j must be shipped from location k in order to receive one unit of that good in location s. Goods that are produced and consumed within the same region s incur no iceberg trade cost, $T^j_{ss} = 1$. Non-tradable goods has infinite iceberg cost, $T^j_{ks} = \infty$ for $k \neq s$. For simplicity, assume that the iceberg trade costs are symmetric, $T^j_{ks} = T^j_{sk}$.

Firms

Representative firms at location s produce goods type j using constant returns to scale technology with labor as the only input. The labor market is assumed to be perfectly competitive. The total output of goods j at location s $y_{s,j} = \sum_{ij} A^i_{s,j} L^i_{s,j}$ where $A^i_{s,j}$ is the labor productivity of agent type i works in industry j at location s.

Goods

Variety of goods are imperfect substitute and differ by type j and manufactured location s. Variety of goods are combined into a composite bundle C^i_s and consumed by households in location s. The composite consumption bundle is defined as

$$C_s^i = \Omega\left(\left\{q_{ks,j}\right\}\right) = \left[\sum_j \alpha_j \left\{\sum_k \left(q_{ks,j}\right)^{\frac{\sigma-1}{\sigma}}\right\}^{\frac{\sigma}{\sigma-1}\times\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}$$

$$C_s^i = \Omega\left(\left\{q_{ks,j}\right\}\right) = \left[\sum_j \alpha_j q_{s,j}{}^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}$$

where $\Omega(\cdot)$ is an aggregator of variety of goods and $q_{s,j}$ is the composite goods of type j

$$q_{s,j} = \left\{\sum_k \left(q_{ks,j}\right)^{\frac{\sigma-1}{\sigma}}\right\}^{\frac{\sigma}{\sigma-1}}.$$

That is, goods are differentiated by the location they come from k and the industry j and α_j is the share of goods type j. Here, the elasticity of substitution between goods of the same type from different locations is σ and the elasticity of substitution between different types j of composite goods $q_{s,j}$ is ρ.

Price level

Denote P_s for the price index at location s, or the price of the composite goods, in location s

$$P_s = \left[\sum_j \alpha_j^\rho \left\{\sum_k \left(p_{ks,j}\right)^{1-\sigma}\right\}^{\frac{1}{1-\sigma}\times(1-\rho)}\right]^{\frac{1}{1-\rho}}$$

$$P_s = \left[\sum_j \alpha_j^\rho P_{s,j}{}^{1-\rho}\right]^{\frac{1}{1-\rho}}$$

with $p_{s,j}$ is the price index of goods type j at location s and $p_{ks,j} = T_{ks}^j p_k^j$ is price of good type j produced in location k and consumed in location s.

Agents

There is a fixed measure of type L^j agents in this economy. An agent o of type i welfare at location s and working in industry j with a **permanent** taste shock $z^i_{s,j}(o)$ is given by

$$W^i_{s,j}(o) = C^i_s U_s z^i_{s,j}(o)$$

where C^i_s is the composite consumption bundle, U_s is the amenity at location s. He/she chooses where to reside s and which representative firm type j to supply a unit of his time to produce type j good.

Goods demand Regardless of their taste shock, the demand of good j produced in k from any agent with income I in location s is given by

$$\begin{aligned} q_{ks,j} &= [\alpha_j]^\rho \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{-\rho} \frac{I}{P_s} \\ &= [\alpha_j]^\rho \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{1-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{1-\rho} \frac{I}{p_{ks,j}} \end{aligned}$$

which is valued at

$$p_{ks,j}q_{ks,j} = [\alpha_j]^\rho \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{1-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{1-\rho} I$$

and the respective quantity shipped from source location is

$$T_{ks,j}q_{ks,j} = T_{ks,j}[\alpha_j]^\rho \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{1-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{1-\rho} \frac{I}{p_{ks,j}}.$$

Wage Assume that the labor market is competitive, the agent receives a wage $w^i_{s,j} = p_{s,j}A^i_{s,j}$ with $p_{s,j}$ is the price of good j sold at s.

Indirect utility, taste shock, and labor supply The welfare of agent o of type i works in industry j at location s can then be rewritten as

$$W^i_{s,j}(o) = \frac{w^i_{s,j}}{P_s} U_s z^i_{s,j}(o).$$

Assume that each element $z^i_{s,j}$ is drawn from an independent Frechet distribution

$$z^i_{s,j} \sim \Phi^i_{s,j}(z) = \exp\left(-G^i_{s,j} z^{-\varepsilon}\right)$$

$G^i_{s,j} = G^i_s \times G^i_j$ the scale parameter which affects how much an average type i agent prefers to work in industry j or prefers to reside in location s. Given the draw $\left\{z^i_{s,j}\right\}$, this agent chooses where he resides s and the industry j to which he will supply labor that maximizes his welfare $\max_{s,j}\left\{W^i_{s,j}\right\}$.

Since taste shocks are randomly drawn from independent Frechet distribution, the fraction of agent type i lives in s and works in j is given by the conditional industry and location distribution

$$\pi^i_{s,j} = \frac{G^i_{s,j}\left(\frac{w^i_{s,j}}{P_s} U_s\right)^{\varepsilon}}{\sum_{s',j'} G^i_{s',j'}\left(\frac{w^i_{s',j'}}{P_s} U_{s'}\right)^{\varepsilon}}.$$

That implies the labor equilibrium labor supply of type i in industry j at location s is simply $L^i_{s,j} = L^i \pi^i_{s,j}$.

Equilibrium

Definition The competitive equilibrium is the set of prices $\left\{p_{s,j}, w^i_{s,j}\right\}_{i\in I, j\in J, s\in S}$ and the conditional occupation and location distribution $\left\{\pi^i_{s,j}\right\}_{i\in I, j\in J, s\in S}$ that satisfy

1. The goods market clearing conditions

$$\sum_i w_{s,j}^i L_{s,j}^i = (\alpha_j)^\rho \sum_k \left(\frac{p_{sk,j}}{p_{k,j}}\right)^{1-\sigma} \left(\frac{p_{k,j}}{P_k}\right)^{1-\rho} \sum_{i',j'} w_{k,j'}^{i'} L_{k,j'}^{i'}$$

that is, the total income of industry j at location s equals to the value of total export to all locations k.

2. Labor market clearing conditions

$$L_{s,j}^i = \pi_{s,j}^i L^i$$

where $\pi_{s,j}^i$ is the fraction of type i population prefers to locate in s and works in industry j given the prices

$$\pi_{s,j}^i = \frac{G_{sj}^i \left(\frac{w_{s,j}^i}{P_s} U_s\right)^\varepsilon}{\sum_{kj'} G_{kj'}^i \left(\frac{w_{k,j'}^i}{P_s} U_k\right)^\varepsilon}.$$

3. Representative firms' profit maximization condition

$$w_{s,j}^i = A_{s,j}^i p_{s,j}$$

and price normalization condition

$$\sum_{s,j} p_{s,j} = 1.$$

Existence and uniqueness of the equilibrium

To show that an equilibrium exists and unique, I follow the arguments was made in Mas-Colell's Chapter 17.B,C. When labor supply, $L^i_{s,j}$, is exogenous, it is straightforward to show that when locational characteristics, $A^i_{s,j}, U_s$, and individual characteristics, G^i_s, G^i_j, are strictly positive and finite, the exceed demand exhibits gross substitution property if ρ and σ are greater than 1 and hence, there exists a unique equilibrium price vector that clears the goods market. In our model, it's slightly more complicated as when prices change, agents will move from their current location and/or industry to a different location/industry and their net change in demand for any good needs to be taken into account.

Assumption 1 There is an upper-bound to the iceberg cost of tradable goods, $\max T_{ks,j} = \overline{T}$. Furthermore, assume that there are upper- and lower-bound to G, U, A, denoting $\overline{G}, \overline{U}, \overline{A}$ and $\underline{G}, \underline{U}, \underline{A}$. Finally, assume that $\sigma > \rho > 1 + \zeta\varepsilon$ where

$$\zeta = \left(1 - \frac{\sqrt{4a^2 + 4a\frac{b}{c}} - 2a}{2\frac{b}{c}}\right)\left(\frac{\sqrt{4a^2 + 4a\frac{b}{c}} - 2a}{\sqrt{4a^2 + 4a\frac{b}{c}}}\right)$$

and

$$\begin{aligned} a &= \overline{T}^{1-\sigma}\frac{\underline{A}}{\overline{A}} \\ b &= \left(\overline{GU}\right)^{\varepsilon} \\ c &= \left(\frac{\underline{GU}}{\overline{T}}\right)^{\varepsilon}. \end{aligned}$$

Lemma 1 When locational characteristics, $A^i_{s,j}, U_s$, and individual characteristics, G^i_s, G^i_j, are strictly positive and finite, the exceed demands is continuous, homogeneous of degree zero in prices, and satisfy Walras' law. Furthermore, under Assumption 1, the exceed demands also exhibit gross substitute property.

Proof In the Appendix.

Proposition 1 When locational characteristics, $A^i_{s,j}, U_s$, and individual characteristics, G^i_s, G^i_j, are strictly positive and finite, there exists an regular spatial equilibrium and under Assumption 1, this equilibrium is unique.

Proof Directly follows from **Lemma 1** and arguments in Chapter 17.B,C of Mas-Colell, there exists a positive and unique price vector that clears the exceed demand. Given that the taste shock is drawn from a Frechet distribution with no upper-bound, there is always a positive measure of agents in any location and/or occupation and therefore, it is a regular spatial equilibrium.

In the case where there are no trade cost and no variation in G, U, A, the elasticity of substitution across industries ρ has to be greater than $1+\left(\sqrt{2}-1\right)^2\varepsilon \approx 1+0.172\varepsilon$ to guarantee that the exceed demand will have gross substitution property for any price vector and hence, unique equilibrium.

Numerical examples

The following numerical examples demonstrate how the variation in taste and labor productivities across type of agents leads to variation in the geographical and occupation concentration patterns.

The line economy with one industry This example is based on the numerical example given by Allen and Arkolakis (2014). Suppose there are S cities that are uniformly distributed along the line $[-1,1]$ and let a real number $s \in S$ represent each city. In this example, assume that there is only one industry and the iceberg trade cost between any two locations k and s is given by $T_{ks} = T_{sk} = \exp(\tau|k-s|)$. There are two types of agents in this model, $i \in \{a,b\}$. The labor productivity of both types are identical $A_s^i = 1$ for all locations s. The amenity at all locations s are also a constant $U_s = 1$. Total population of type a is exactly the same as that of type b $L^a = L^b = \frac{1}{2}$.

Homothetic location preference In the case where both types have identical tastes for all the locations s so that $G_s^i = \frac{1}{S}$ for all s we have a symmetric equilibrium. The population distribution of both types are identical, concentrating at the center due to the overall lower trade iceberg costs (Figure 3.1). This result is identical to that shown in Allen and Arkolakis (2014).

Close to home country preference Now suppose that home country of type b is to the right of the border of city $s = 1$. Type b agents, hence, have a stronger preference for cities that are closer to their home country, that is $G_s^b = \frac{\exp(-|1-s|)}{\sum_k \exp(-|1-k|)}$. As a result, the model results in a asymmetric equilibrium in which not only the distribution of type b shifts to the right because of their preference to be closer to home but also a shift in type a distribution (Figure 3.2). The type a distribution shifts left due to the price effects on wages of locations closer to 1. In other words, type b agents have a comparative advantage living in cities closer to 1 because of their personal preference and that results in a price effect that give type a a comparative advantage of living off center, to the left.

Discussion In this example, the segregated equilibrium does not exist if the taste distribution of both types are identical. Given the observed variation in immigrants' geographical and occupation concentration, it is reasonable to postulate that the variation is partly driven by the difference in taste preference between different immigrant groups, and

Figure 3.1. Homothetic preference for locations

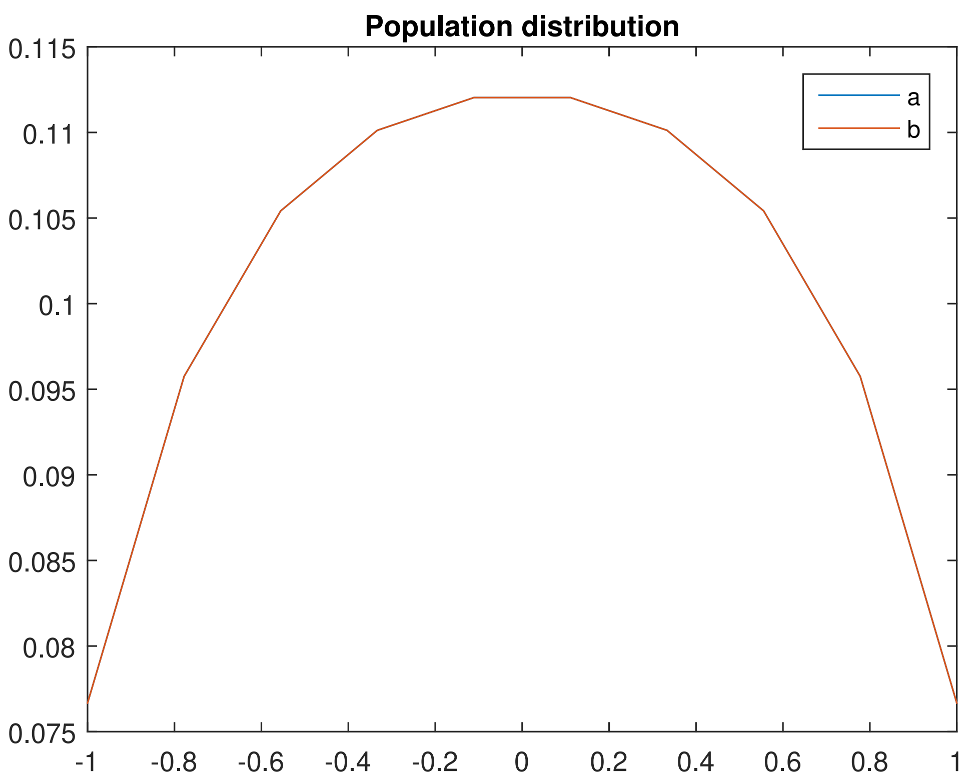
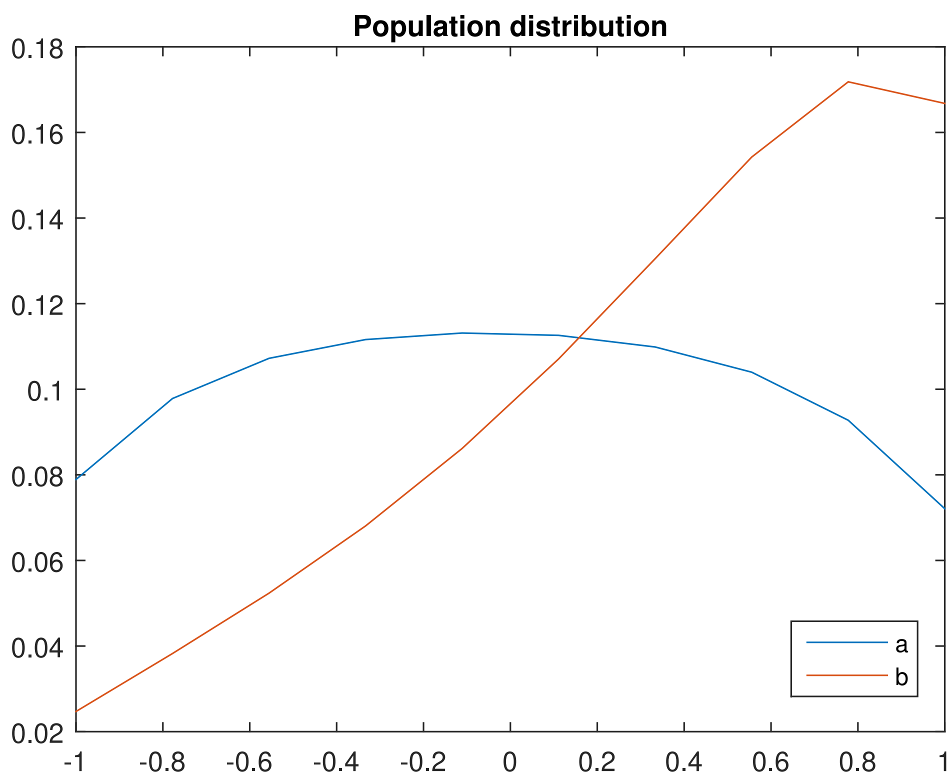


Figure 3.2. Close to home preference



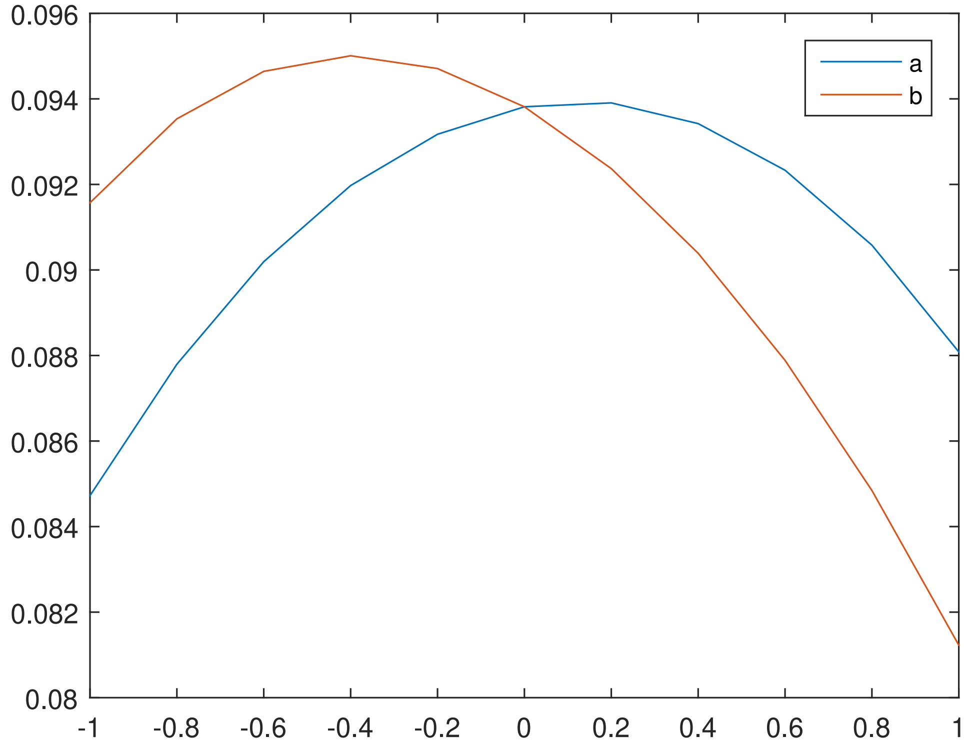
between immigrants and natives. The discrepancy in geographical taste distributions between different types of agents could be due to multiple factors such as the geographical proximity to home country, the distribution of relative members, the similarity between the destination's climate and the home country's, and so on.

The line economy with two industries Immigrants, in general, have a comparative advantage/disadvantage versus natives in some occupations due to their English skill and/or education background. Because of this comparative advantage/disadvantage, immigrants are more likely to sort into occupation/location where they have a comparative advantage.

In this example, I assume that there are two industries, agriculture and finance. Assume that labor productivity in agriculture is increasing as we go west (smaller s) due to its richness in fertile soil and favorable climate, $A^{agr}(s) = \frac{\exp(1-\xi|-1-s|)}{\sum_{s'}\exp(1-\xi|-1-s'|)}$. On the other hand, labor productivity in finance is increasing as we go east due to its proximity to the capital city $s = 1$, $A^{fin}(s) = \frac{\exp(1-\xi|1-s|)}{\sum_{s'}\exp(1-\xi|1-s'|)}$. As before, there are two types of agents, a and b. Because of their limited English skill, the labor productivity of type b in the finance is lower than that of type a by a factor ζ, that is $A_s^{b,fin} = (1-\zeta)A_s^{a,fin}$ for all location s.

As a result, type b has a comparative advantage in agriculture while type a has a comparative advantage in finance. In equilibrium, type b's spatial distribution is more skewed toward the west while the distribution of a is skewed to the east (Figure 3.3). Unsurprisingly, the majority of type a works in the finance while the majority of type b (83.74 percent) works in agriculture.

Figure 3.3. Comparative (dis)advantages in skills



Linguistic assimilation and the dynamics of demographics

For immigrants, having a US degree and/or being fluent in English opens up new opportunities for them. These opportunities range from being able to earn higher wages in their current jobs, to allowing them to pursue different occupations, or even to to reside in a different region that fits their likings. As observed in data, the dissimilarity index of occupation and residential choices are lower for immigrants with higher educational attainment and/or US education and/or more proficient in English. Immigrants' occupation dissimilarity index versus natives also declines with their duration of stay. To capture this idea, I allow agents to make an endogenous choice of time spent acquiring English to become a different type. Agents who become proficient in English then can have a higher wage in their current job and/or switching to a different location/occupation to which they have a stronger preference (a sufficiently high drawn taste shock). Educational attainment can also be implement with identical strategy but will not be pursued in this paper.

Agents

For simplicity, assume that the agent lives for only two periods and the model, effectively, is an overlapping generations model. Young and old agents are indexed by superscript y and o respectively. Recall that agent type index $i = \{eth, edu, Eng\}$ has three components: ethnic origin, educational attainment, and English proficiency. Let 1 and 0 denote agents who are proficient and not proficient in English. Hence, the superscript $\{a,b,0\}$ and $\{a,b,1\}$ denote agent with ethnic a, education b, and are respectively not proficient and proficient in English. Hereafter, the superscript 0 and 1 abbreviate for $\{a,b,0\}$ and $\{a,b,1\}$ indexing agents of the same ethnic and educational attainment but different English proficiency. With or without the parenthesis, the superscripts are now $(y,0)\,,(y,1)\,,(o,0)\,,(o,1)$ indexing agent type. Let t index time and from now on, the subscript sj,t indexes the residential and occupational choices s and j in period t.

In any period, agents spend effort learning English and effectively loosing a fraction h of their utility. **Here, for simplicity, assume that the effort that agents spend acquiring English comes from their leisure time and does not affect their labor supply.** With probability $g(h) = h^{\eta}$ with $\eta > 1$, an agent of type 0 can become fluent in English in the next period, that is to switch from type 0 to type 1. Since old agents will cease to exist in the next period, they will not put effort into acquiring English, $h^{o,0}_{sj,t} = 0$. Furthermore, since natives and immigrant who already are fluent in English have no benefit from this activity, they optimally spend their unit of time working. In this model, agents who are young and not English proficient will spend time acquiring English skill.

For simplicity, assume that the agent lives for only two periods and the model, effectively, is an overlapping generations model. Recall that agent type index $i = \{eth, edu, Eng\}$ has three components: ethnic origin, educational attainment, and English proficiency. Let 1 and 0 denote agents who are proficient and not proficient in English. Hence, the superscript $\{x,y,0\}$ and $\{x,y,1\}$ denote agent with ethnic x, education y, and are respectively not proficient and proficient in English. Given the z vector of taste, denote the original residential and occupational choices $s, j = \arg\max_{s,j}\left\{W^{\{x,y,0\}}_{s,j}(z)\right\}$. The original choices might not be the same as the new residential and occupational choices, denoting

$$s', j' = \arg\max_{s,j}\left\{W^{\{x,y,1\}}_{s,j}(z)\right\},$$

as he/she becomes fluent in English because a shift in the wage matrix $\left\{w^{\{x,y,1\}}_{s,j}\right\}_{s,j}$.

The young agent with a draw z of type 0's problem

$$
\begin{aligned}
& V_t^{y,0}(z) \\
& = \max_{s,j} \max_{h} z_{s,j} \frac{w^0_{sj,t}}{P_s} U_s (1-h) \\
& + \beta \left\{ g(h) \max_{s',j'} z_{s',j'} \frac{w^1_{s'j',t+1}}{P_{s'}} U_{s'} + (1-g(h)) \max_{j',s'} \max_{s',j'} z_{s',j'} \frac{w^0_{s'j',t+1}}{P_{s'}} U_{s'} \right\}.
\end{aligned} \tag{3.1}
$$

That is the young immigrant of type 0 will choose where to live s and industry j to work in the first period. He also chooses how much effort to spent learning English. In the next period, when he becomes old, with the probability $g(h)$ he will become type 1 - fluent in English and receiving the utility of $\max_{s',j'} z_{s',j'} \frac{w^1_{s'j',t+1}}{P_{s'}} U_{s'}$ where s', j' is the optimal locational and occupational choices given the taste shock drawn earlier, conditional on becoming type 1.

Given that there is no cost to moving and switching occupation, the residential and occupational choices are independent of the effort the agent spends learning English. In this paper, I am only interested in the long-term impact of immigration policy and hence, will look at primarily the steady state equilibrium of the model. At the steady state, the young agent of type 0 with a draw z's problem is

$$
V^{y,0}(z) = \max_{s,j} \max_{h} (1-h) W^0_{s,j} + \beta \left\{ g(h) \max_{j',s'} \left\{ W^1_{s',j'} \right\} + (1-g(h)) W^0_{s,j} \right\} \tag{3.2}
$$

where

$$
W_{s,j} = \frac{w_{sj}}{P_s} U_s
$$

not only with the time subscripts gone, the choice j', s' on $W^{0'}_{s',j'}$ become j, s since the prices do not change at steady state.

Furthermore, assume that the transitional probability is given by $g(h) = h^{\eta}$. The optimal effort h^* that agent spent learning English satisfies the FOC

$$0 = -W^0_{s,j} + \beta\eta h^{\eta-1}\left\{\max_{s',j'}\left\{W^1_{s',j'}\right\} - W^0_{s,j}\right\}$$

$$h = \left[\frac{W^0_{s,j}}{\beta\eta\left\{\max_{s',j'}\left\{W^1_{s',j'}\right\} - W^0_{s,j}\right\}}\right]^{\frac{1}{\eta-1}}$$

and the transitional probability is $g^* = g(h^*) = (h^*)^{\eta}$.

Aggregate variables

Denote $k = \{s, j\} = \arg\max\left\{W^0_{s,j}\right\}$ and $k' = \left\{s', j'\right\} = \arg\max\left\{W^1_{s'}\right\}$ respectively for the optimal choices that agent o of type i makes when he was not and when he was proficient in English skill. Let $K = \left\{k'\right\}$ be the set of all the options such that there exists a draw z of taste shock satisfies $W^1_{k'} > W^0_k \geq W^0_{k'}$ (here I drops the superscript i for convenience as this is true for all types i). In other words,

$$\frac{w^1_{k'}U_{k'}}{P_{k'}}z_{k'} > \frac{w^0_k U_k}{P_k}z_k \geq \frac{w^0_{k'}U_{k'}}{P_{k'}}z_{k'}$$

or

$$\frac{w^0_k U_k/P_k}{w^1_{k'}U_{k'}/P_{k'}}z_k < z_{k'} < \frac{w^0_k U_k/P_k}{w^0_{k'}U_{k'}/P_{k'}}z_k$$

$$m_{k'} < z_{k'} < M_{k'}.$$

Furthermore, for k' to be the optimal choices, for any $k" \in K \backslash \{k'\}$,

$$\frac{w_{k'}^{1} U_{k'}}{P_{k'}} z_{k'} > \frac{w_{k"}^{1} U_{k"}}{P_{k"}} z_{k"}$$

and since k was the optimal choice

$$\frac{w_{k}^{0} U_{k}}{P_{k}} z_{k} > \frac{w_{k"}^{0} U_{k"}}{P_{k"}} z_{k"}$$

or

$$z_{k"} < \min\left\{ \frac{w_k^0 U_k / P_k}{w_{k"}^0 U_{k"} / P_{k"}} z_k, \frac{w_{k'}^1 U_{k'} / P_{k'}}{w_{k"}^1 U_{k"} / P_{k"}} z_{k'} \right\}$$

$$z_{k"} < M_{k"}$$

Suppose that from (3), the agent's optimal effort in learning English is $h^*(z_k, z_{k'})$ and his transitional probability is $g^*(z_k, z_{k'})$. The average effort that agents in state k spend learning is

$$E\left[h_k^0\right] = \frac{\sum_{k' \in K} \int_0^\infty \int_{m_{k'}}^{M_{k'}} \overbrace{\int_0^{M_{k"}} \ldots}^{\text{all } k" \in K \backslash \{k'\}} h(z_k, z_{k'})\, d\Theta_t(z_{k"}, \ldots z_{k'}, z_k)}{\sum_{k' \in K} \int_0^\infty \int_{m_{k'}}^{M_{k'}} \overbrace{\int_0^{M_{k"}} \ldots}^{\text{all } k" \in K \backslash \{k'\}} 1\, d\Theta_t(z_{k"}, \ldots z_{k'}, z_k)} \tag{3.3}$$

where $\Theta_t(\cdot)$ denotes the taste distribution at time t. The average probability that type 0 at state k becomes type 1 can be computed similarly.

Population dynamic

Assume that the end of every period t, there is an inflow of $\left\{H_t^i\right\}$ of type i agents into this economy and that all arriving immigrants are at the young age. Let L_t denote the total population at time t. Then, the population of type $\{a,b,0\}$ locates in s and work in j in time $t+1$ includes the old agents of type $\{a,b,0\}$ who were young in the last period and remain not proficient in English and the new young agents of type $\{a,b,1\}$ who arrive to this economy

$$\overbrace{L_{sj,t+1}^{\{a,b,0\}}}^{old+young} = \overbrace{\mathscr{L}_{sj,t}^{\{a,b,0\}}\left(1-\gamma_{sj,t}^{\{a,b,0\}}\right)}^{old} + \overbrace{H_t^{\{a,b,0\}}\pi_{sj,t+1}^{\{a,b,0\}}}^{young}$$

while the population of type $\{a,b,1\}$ locates in s and work in j includes the old agents who were proficient in English when they are young, the old agents who become proficient in English and switch from s', j' to industry s,j, and the new young agents of type $\{a,b,1\}$

$$\overbrace{L_{sj,t+1}^{\{a,b,1\}}}^{old+young} = \overbrace{\mathscr{L}_{sj,t}^{\{a,b,1\}} + \sum_{s'j'}\gamma_{sj|s'j',t}^{\{a,b,0\}}P_{sj|s'j',t}^{i}\mathscr{L}_{s'j',t}^{\{a,b,0\}}}^{old} + \overbrace{H_t^{\{a,b,1\}}\pi_{sj,t+1}^{\{a,b,1\}}}^{young}$$

Here,

$$\begin{aligned}
\mathscr{L}_{sj,t}^{\{a,b,0\}} &= H_{t-1}^{\{a,b,0\}}\pi_{sj,t}^{\{a,b,0\}}\\
\mathscr{L}_{sj,t}^{\{a,b,1\}} &= H_{t-1}^{\{a,b,1\}}\pi_{sj,t}^{\{a,b,1\}},
\end{aligned}$$

is the population of the old who arrived young last period, $\pi_{sj,t}^i$ is the conditional residential and occupational distribution of type i, $\gamma_{sj,t}^{\{a,b,0\}} = \sum_{j',s'}\gamma_{s'j'|sj,t}^{\{a,b,0\}}P_{s'j'|sj,t}^{\{a,b,0\}}$ is the average transitional probability of type 0 becomes type 1 in location s and industry j, $\gamma_{s'j'|sj,t}$ is the average transitional probability of agents who original and new occupational choices are j and j', and $P^{\{a,b,0\}}_{s'j'|sj,t}$ is the fraction of type 0 who new optimal choice is $s'j'$ given the original occupational choice of j.

The aggregate labor supplied in location s and location j by type $\{a,b,0\}$ at time $t+1$ is given by

$$L^{\{a,b,0\}}_{sj,t+1} = \mathscr{L}^{\{a,b,0\}}_{sj,t}\left(1-\gamma^{\{a,b,0\}}_{sj,t}\right) + H^{\{a,b,0\}}_t \pi^{\{a,b,0\}}_{sj,t+1}$$

and labor supplied in location s and industry j by type $\{a,b,1\}$ is given by

$$L^{\{a,b,1\}}_{sj,t+1} = L^{\{a,b,1\}}_{sj,t+1}$$

since they all supply the whole one unit of their time.

Let $Q^{\{a,b,0\}}_{sj,t+1} = \frac{L^{\{a,b,0\}}_{sj,t+1}}{L_{t+1}}$ denote the proportion of type $\{a,b,0\}$ lives in s and works in j at $t+1$ where L_{t+1} is the total population at time $t+1$ which includes the old from last period and the young recently arrives. Denotes $\Lambda^i_t = \frac{H^i_t}{\sum_i H^i_t}$ and $\lambda_t = \frac{\sum_i H^i_t}{\sum_i H^i_{t-1}}$ or $1+\lambda_t = \frac{L_{t+1}}{\sum H^i_{t-1}}$ and $\frac{1+\lambda_t}{\lambda_t} = \frac{L_{t+1}}{\sum H^i_t}$. That is Λ and λ are respectively fraction of the young agents who are of type i and the growth rate of the total population.

From above, the distribution of agents evolves according to

$$\begin{aligned}
Q^{\{a,b,0\}}_{sj,t+1} &= \Lambda^{\{a,b,0\}}_{t-1}\frac{1}{1+\lambda_t}\pi^{\{a,b,0\}}_{sj,t}\left(1-\gamma^{\{a,b,0\}}_{sj,t}\right) + \Lambda^{\{a,b,0\}}_t \frac{\lambda_t}{1+\lambda_t}\pi^{\{a,b,0\}}_{sj,t+1} \\
Q^{\{a,b,1\}}_{sj,t+1} &= \Lambda^{\{a,b,1\}}_{t-1}\frac{1}{1+\lambda_t}\pi^{\{a,b,1\}}_{sj,t+1} + \sum_{j'} \gamma^{\{a,b,0\}}_{sj|s'j',t} P^i_{sj|s'j',t} \Lambda^{\{a,b,0\}}_{t-1}\frac{1}{1+\lambda_t}\pi^{\{a,b,0\}'}_{s'j',t} \\
&\quad + \Lambda^{\{a,b,1\}}_t \frac{\lambda_t}{1+\lambda_t}\pi^{\{a,b,1\}}_{sj,t+1}
\end{aligned}$$

where $\gamma^{\{a,b,0\}}_{sj,t} = \sum_{j',s'} \gamma^{\{a,b,0\}}_{s'j'|sj} P^{\{a,b,0\}}_{s'j'|sj}$.

The steady state version of the population forward equations is given by

$$
\begin{aligned}
Q_{s,j}^{\{a,b,0\}} &= \Lambda^{\{a,b,0\}} \frac{1}{1+\lambda} \pi_{s,j}^{\{a,b,0\}} \left(1-\gamma_{sj}^{\{a,b,0\}}\right) + \Lambda^{\{a,b,0\}} \frac{\lambda}{1+\lambda} \pi_{s,j}^{\{a,b,0\}} \\
Q_{s,j}^{\{a,b,1\}} &= \Lambda^{\{a,b,1\}} \frac{1}{1+\lambda} \pi_{s,j}^{\{a,b,1\}} + \sum_{j''} \gamma_{sj|s'j'}^{\{a,b,0\}} P_{sj|s'j'}^{i} \Lambda^{\{a,b,0\}} \frac{1}{1+\lambda} \pi_{s,j'}^{\{a,b,0\}} \\
&\quad + \Lambda^{\{a,b,1\}} \frac{\lambda}{1+\lambda} \pi_{s,j}^{\{a,b,1\}}
\end{aligned}
$$

The labor supplied by type $\{a,b,0\}$ in j,s in period $t+1$

$$
Q_{s,j}^{\{a,b,0\}} = \Lambda^{\{a,b,0\}} \frac{1}{1+\lambda} \pi_{s,j}^{\{a,b,0\}} \left(1-\gamma_{s,j}^{\{a,b,0\}}\right) + \Lambda^{\{a,b,0\}} \frac{\lambda}{1+\lambda} \pi_{s,j}^{\{a,b,0\}}
$$

while the labor supplied by type $\{a,b,1\}$ is given by

$$
Q_{s,j}^{\{a,b,1\}} = Q_{s,j}^{\{a,b,1\}}.
$$

Equilibrium

Steady state equilibrium When the growth rate for each group λ^i and the ratio between any pair labor productivities remain constant over time, the steady state competitive equilibrium is the set of prices $\left\{p_{s,j}, w_{s,j}^i\right\}_{i\in I, j\in J, s\in S}$, the set of average labor supply $\left\{l_{sj}^i\right\}$, the set of transitional rates and transitional probabilities $\left\{\gamma_{sj}^i, \gamma_{sj'|sj}^i, P_{sj'|sj}^i\right\}$ and the conditional occupation and location distribution $\left\{\pi_{s,j}^i\right\}_{i\in I, j\in J, s\in S}$, and the demographic distribution $\left\{Q_{s,j}^i\right\}_{i\in I, j\in J, s\in S}$ that satisfy

1. The goods market clearing conditions

$$
\sum_i w_{s,j}^i Q_{s,j}^i = (\alpha_j)^\rho \sum_{s'} \left(\frac{p_{ss',j}}{P_{s',j}}\right)^{1-\sigma} \left(\frac{P_{s',j}}{P_{s'}}\right)^{1-\rho} \sum_{i',j'} w_{s',j'}^{i'} Q_{s',j'}^{i'}
$$

that is, the total income of industry j at location s equals to the value of total export to all locations s' where the aggregate labor supply in location s and industry j is given by

$$Q_{s,j}^{\{a,b,0\}} = \Lambda^{\{a,b,0\}} \frac{1}{1+\lambda} \pi_{s,j}^{\{a,b,0\}} \left(1 - \gamma_{s,j}^{\{a,b,0\}}\right) + \Lambda^{\{a,b,0\}} \frac{\lambda}{1+\lambda} \pi_{s,j}^{\{a,b,0\}}$$
$$Q_{s,j}^{\{a,b,1\}} = Q_{s,j}^{\{a,b,1\}}.$$

2. The transitional rates and transitional probabilities come from household's optimal effort in learning English problem (3).

3. Representative firms' profit maximization condition

$$w_{s,j}^{i} = A_{s,j}^{i} p_{s,j}$$

and price normalization condition

$$\sum_{s,j} p_{s,j} = 1.$$

4. Population distribution evolves according to the following forward equations

$$\begin{aligned} Q_{s,j}^{\{a,b,0\}} &= \Lambda^{\{a,b,0\}} \frac{1}{1+\lambda} \pi_{s,j}^{\{a,b,0\}} \left(1 - \gamma_{sj}^{\{a,b,0\}}\right) + \Lambda^{\{a,b,0\}} \frac{\lambda}{1+\lambda} \pi_{s,j}^{\{a,b,0\}} \\ Q_{s,j}^{\{a,b,1\}} &= \Lambda^{\{a,b,1\}} \frac{1}{1+\lambda} \pi_{s,j}^{\{a,b,1\}} + \sum_{j''} \gamma_{sj|s'j'}^{\{a,b,0\}} P_{sj|s'j'}^{i} \Lambda^{\{a,b,0\}} \frac{1}{1+\lambda} \pi_{s,j'}^{\{a,b,0\}} \\ &\quad + \Lambda^{\{a,b,1\}} \frac{\lambda}{1+\lambda} \pi_{s,j}^{\{a,b,1\}}. \end{aligned}$$

3.3 Reduced-form analysis

Data

In this paper, I use data to describe the variation of immigrants' geographical and occupational concentration. I also look at the difference in residential and occupational choices between different immigrant groups, between immigrants and natives and how this difference changes over the time for certain cohorts and across immigrant cohorts. Finally, I use several statistics such as the proportion of an immigrant group that resides to a certain location (a state), the proportion proportion of each immigrant group that works in a certain industry, etc. to calibrate the models that I discuss later on.

The first dataset I use for this paper is the annual American Community Survey (ACS) from 2001 to 2016, which is available to public at IPUMS-USA, University of Minnesota, www.ipums.org. The ACS 2001 - 2004 are respectively 1-in-232, 1-in-261, 1-in-236, 1-in-239 national random sample of the population. ACS 2005-2016 are 1-in-100 national random sample of the population. The data contains micro information which includes

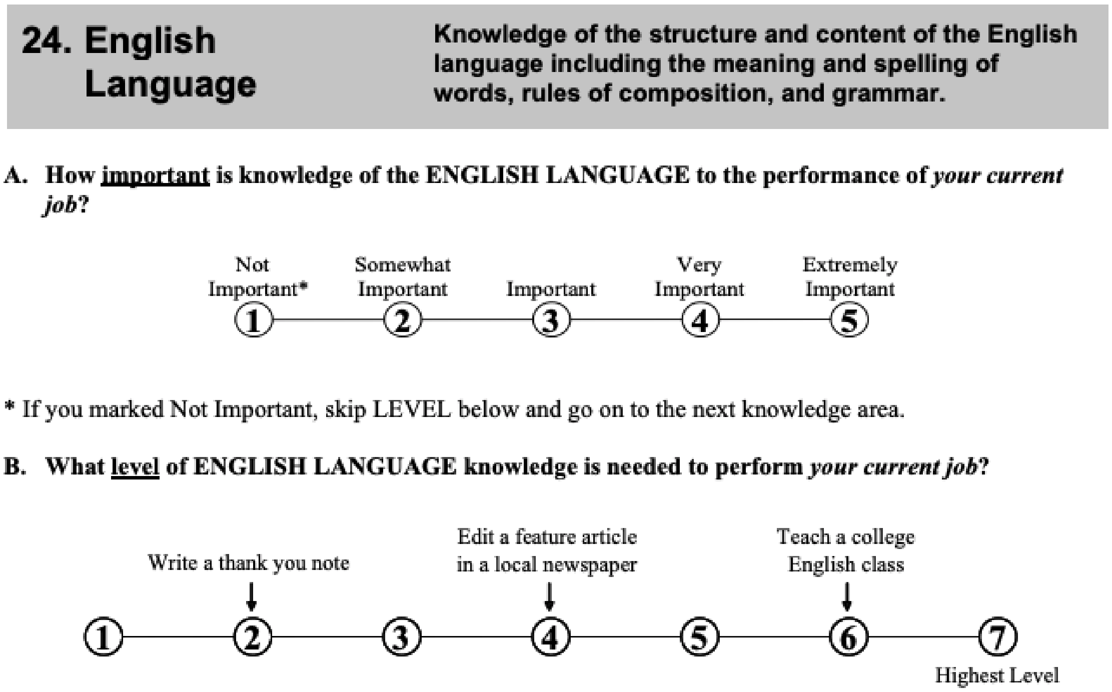
1. individual geographic information – that is their residential location (at the state and at the Public Use Microdata Areas – PUMAs level),

2. individual demographic and race, ethnicity, and nativity information – that is their sex, age, race, birthplace, citizenship status, year of immigration, etc.

3. individual work and income information – that is their employment status, occupation, wage and salary income, etc.

The second dataset I use is the Occupational Information Network, or O*NET, database. This dataset contains information that is needed to classify occupations into categories depending on their English language skill requirement. Specifically, O*NET data is a ques-

Figure 3.4. O*NET Survey



24. English Language — **Knowledge of the structure and content of the English language including the meaning and spelling of words, rules of composition, and grammar.**

A. How important is knowledge of the ENGLISH LANGUAGE to the performance of *your current job*?

Not Important* (1) — Somewhat Important (2) — Important (3) — Very Important (4) — Extremely Important (5)

* If you marked Not Important, skip LEVEL below and go on to the next knowledge area.

B. What level of ENGLISH LANGUAGE knowledge is needed to perform *your current job*?

(1) — (2) Write a thank you note — (3) — (4) Edit a feature article in a local newspaper — (5) — (6) Teach a college English class — (7) Highest Level

tionnaire survey sent out to workers across more than 1000 occupations asking multiple questions, including "*How important is knowledge of the ENGLISH LANGUAGE to the performance of your current job?*" Workers' answers are based on a five-point scale (1) Not important; (2) Somewhat important; (3) Important; (4) Very important; and (5) Extremely Important (Figure 3.4). The data contains the standardized average score

$$S = \frac{O-L}{H-L} * 100$$

where O is the original rating score, L is the lowest rating score for each occupation, and H is the highest rating score for such occupation. The O*NET dataset contains the average score for each occupation.

Dissimilarity indices of residential and occupational choices

To systematically study the geographical assimilation of immigrants, I use the Dissimilarity Index, DI, to measure the difference in how immigrants' residential choices (or occupational choices) compared to natives. The index of dissimilarity of group i at time t is given by

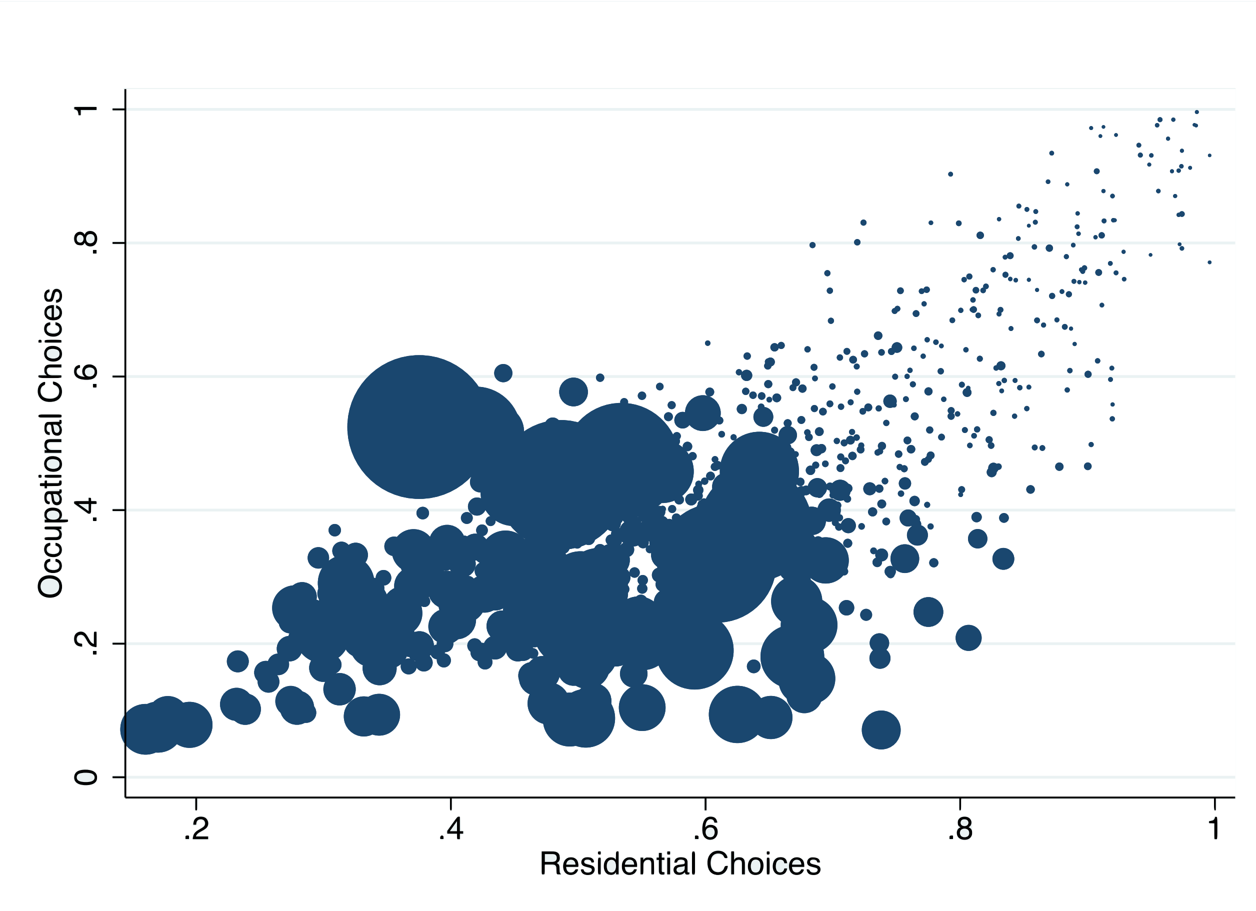
$$DI^i = \frac{1}{2}\sum_s \left|\pi_s^i - \pi_s^n\right|$$

where $\pi_{s,t}^n$ is the fraction of the native population that resides in state s (or works in a category of occupation s defined using the first two digits of the four digit occupation code as on IPUMs USA[1]) at time t. If the share of an immigrant group in any location s is the same as that of natives, the Dissimilarity Index for this group is zero as there is no statistical difference in the likelihood that an immigrant in this group will reside to any location compared to natives. The theoretical range of DI is from 0, when $\pi_s^i = \pi_s^n$ for all s, to 1, when an immigrant group and natives are completely segregated.

With the American Community Survey (ACS) sample 2004-2005-2006, I group immigrants and natives to four age group of 25-34, 35-44, 45-54, 55-65. I further group immigrants by their country of origin and years of migration: before 1985, from 1985-1996, and 1996-2005. I compute the Residential Dissimilarity Index (RDI) and Occupational Dissimilarity Index (ODI) of these immigrant groups against the natives and find a strong positive correlation between RDI and ODI regardless of their group size, as shown in figure 3.5 below. The size of each dot represents the size of each group. The difference index of each group comes from two sources: the difference between the choices of immigrants vs. natives and random selection bias due to the smaller group size. As can be seen in figure 3.5,

[1]https://usa.ipums.org/usa/volii/occ_acs.shtml

Figure 3.5. Scatterplot of ODI vs. RDI



the big dots concentrate in the lower left corner of the scatter plot indicates the presence of random selection bias though the positive correlation remains significant.

To further investigate the determinants of the difference in residential and occupational choices between immigrants and natives, I estimate the following regression model:

$$RDI^i = \phi^i + X^i\beta + \varepsilon$$

$$ODI^i = \phi^i + X^i\beta + \varepsilon$$

where ϕ^i is the vector of age and year of migration fixed effects, X_i is a vector of characteristics including the average years of education, the likelihood of speaking English very well, the likelihood of having US education (an individual is assumed to have USA education if years in USA + years of education + 7 is greater than or equal to age, that is there is at least more than zero year overlap between years in USA and years of education), and same-ethnic immigrant stock. The results are reported in table 3.1. The younger immigrant group tends to have a smaller RDI and ODI. Immigrants came earlier has a lower ODI but higher RDI. Groups with higher educational attainment and or are more proficient in English tend to have a lower ODI and RDI. Groups with a higher likelihood of having a US degree are likely to have a lower ODI. The size of same ethnic immigrant stock has the opposite association with ODI and RDI, positive on the former and negative on the latter. However, when excluding immigrants with Hispanic origin from the sample, the association between same-ethnic immigrant stock and these indices are of the same sign.

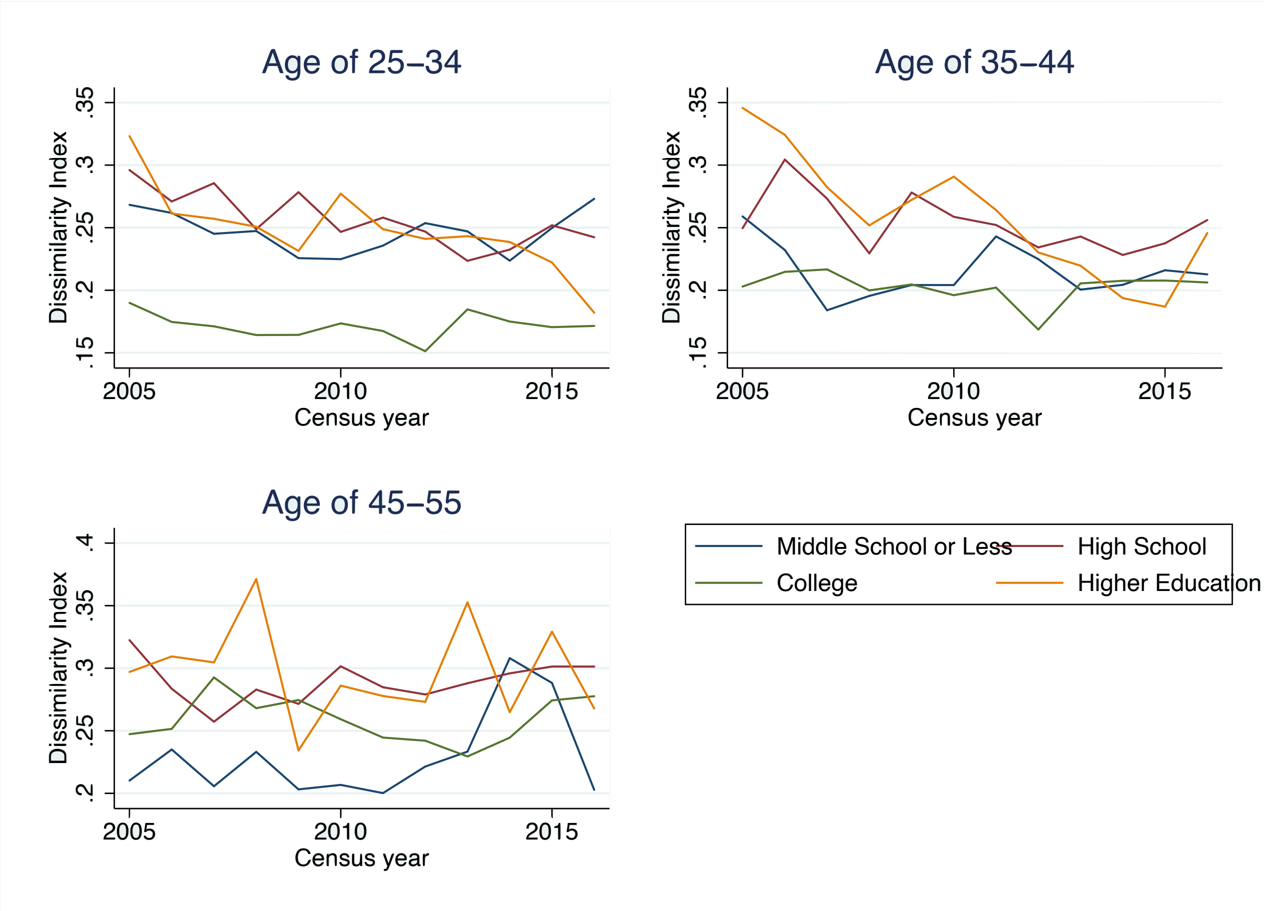
Table 3.1: Determinants of ODI/RDI

	(1) RDI	(2) ODI	(3) RDI(NoHispanic)	(4) ODI(NoHispanic)
Age of 25-34	0	0	0	0
	(.)	(.)	(.)	(.)
Age of 35-44	-0.00133	-0.0177**	-0.0217*	-0.0212**
	(-0.13)	(-2.62)	(-2.02)	(-2.82)
Age of 45-54	-0.00659	-0.0251**	-0.0238	-0.0304**
	(-0.48)	(-2.75)	(-1.72)	(-3.14)
Age of 55-65	-0.0256	-0.0121	-0.0349*	-0.0239*
	(-1.50)	(-1.05)	(-2.04)	(-2.00)
Migrated 96-05	0	0	0	0
	(.)	(.)	(.)	(.)
Migrated 86-95	0.0874***	-0.0205**	0.0721***	-0.00514
	(8.48)	(-2.96)	(6.64)	(-0.68)
Migrated before 85	0.117***	-0.0702***	0.0669***	-0.0460***
	(6.87)	(-6.12)	(3.90)	(-3.84)
English proficiency	-0.000637**	-0.000923***	-0.000848***	-0.00107***
	(-3.27)	(-7.05)	(-4.84)	(-8.73)
Years of education	-0.0158***	-0.0153***	-0.0134***	-0.0123***
	(-6.47)	(-9.30)	(-5.94)	(-7.86)
USA education	-0.000845***	-0.000864***	-0.000899***	-0.000794***
	(-3.72)	(-5.67)	(-4.15)	(-5.24)
Immigrant stock	-0.0255***	0.00938***	-0.253***	-0.196***
	(-7.95)	(4.36)	(-11.80)	(-13.12)
Constant	0.725***	0.633***	0.793***	0.632***
	(23.58)	(30.65)	(28.14)	(32.09)
Observations	1466	1466	1444	1444
Adjusted R^2	0.153	0.608	0.212	0.461

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 3.6. ODI with immigrants' duration of stay

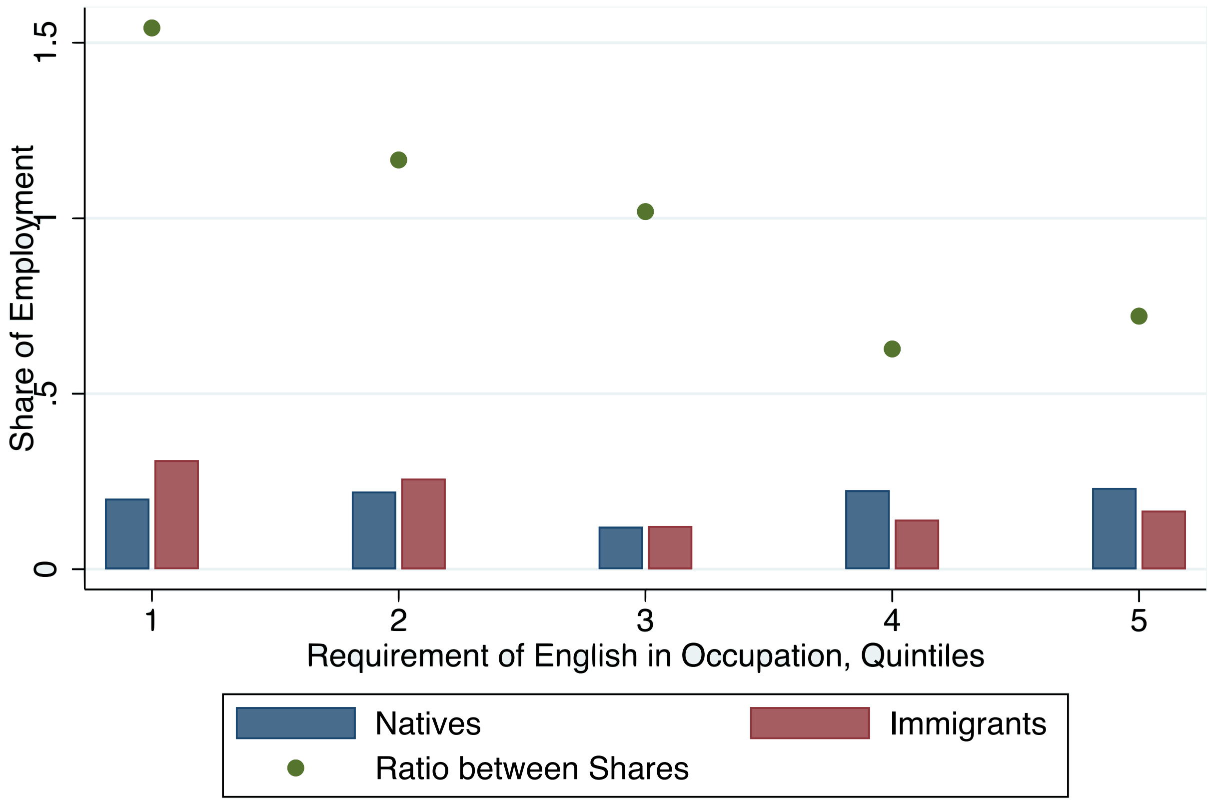


Occupational choices

Dissimilarity index over time

Grouping all immigrants by their age group and education, as shown in figure 3.6, the ODI declines with the duration of stay for younger cohorts, age of 25-44, and across educational attainment. Immigrant cohorts with college educational attainment tend to have the lowest ODI.

Figure 3.7. Immigrants, natives employment by English skills requirements

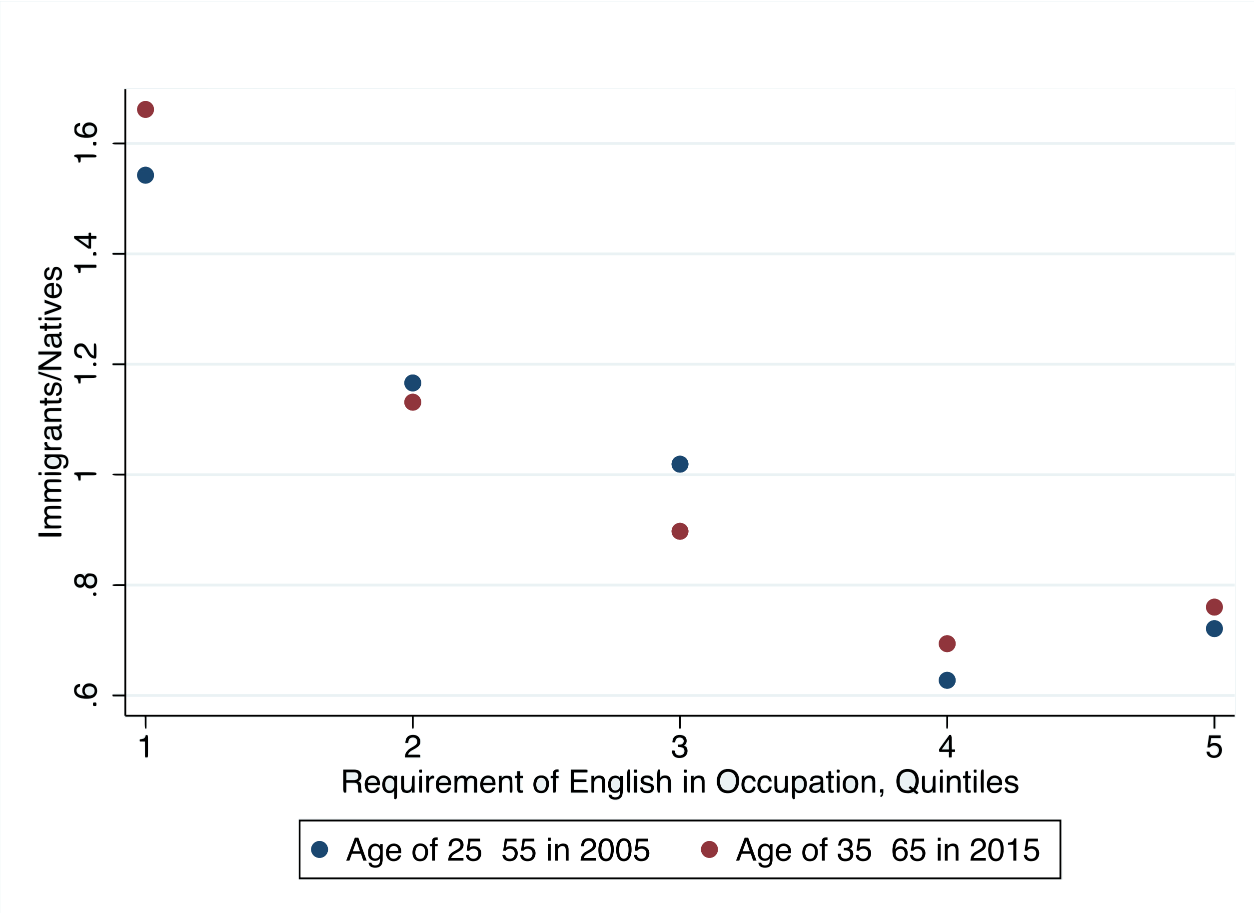


Requirement of English skill in occupation

Using the score ranking the requirement of English provided in O*NET dataset, I divide the sample ACS 04-05-06 into quintiles by English requirement and find the share of natives, immigrant shares in each quintile. As shown in figure 3.7, immigrants are more concentrated in occupations in which English skill is less required compared to natives. The ratio between the shares of immigrants vs natives in the first quintile is 1.54 and is only 0.76 in the fifth quintile where occupation require intensive English language skills.

Similarly, the same exercise on the sample ACS 14-15-16 shows the share of immigrants in the second and third quintiles declines while the share of immigrants in other

Figure 3.8. Immigrants/natives employment by English skills requirement, 10 years after



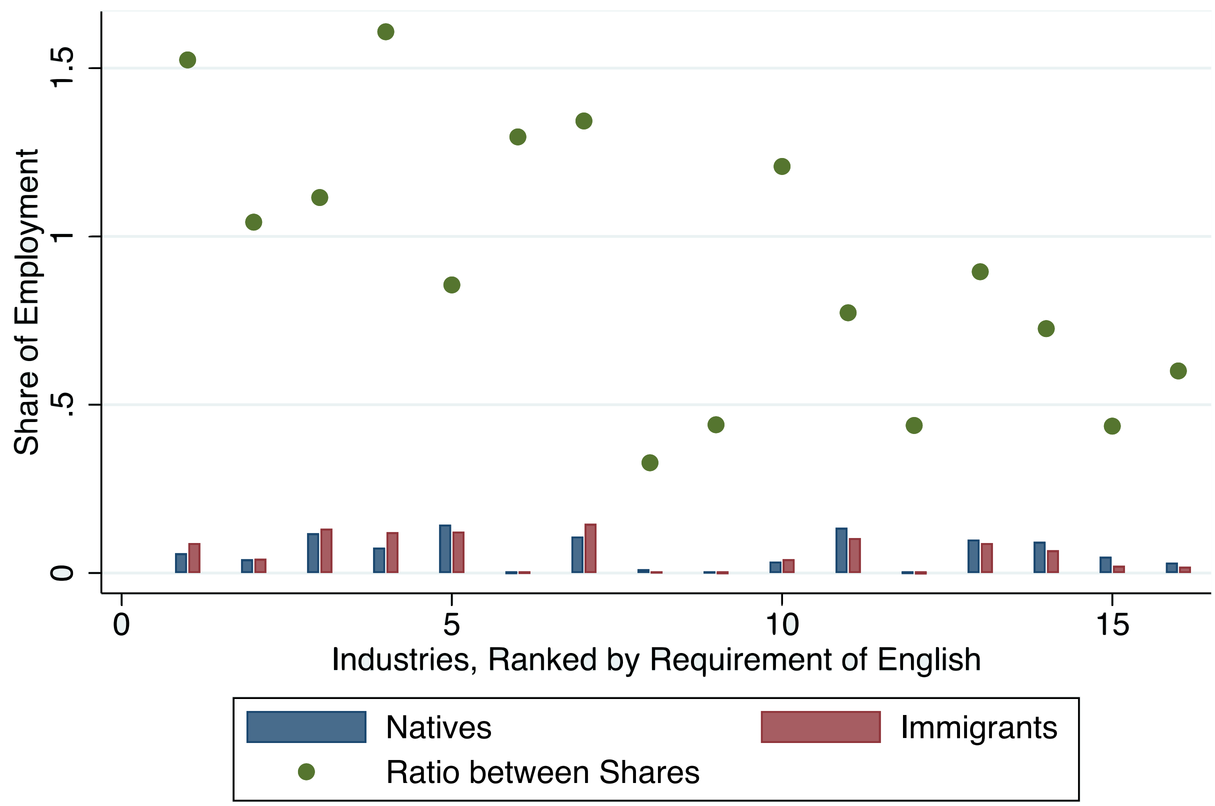
quintiles increase compared to 10 years earlier. That is after 10 years, the share of immigrants sorting into the first, fourth, and fifth quintiles are higher than that of natives (Figure 3.8).

Alternatively, using the 3 digits industry code provided in the ACS sample, I rank industries by the average English requirement for each industry. Industries which listed under *Other Services* require the least English skill include automotive repair and maintenance, car washes, etc. while industries listed under *Information and Communications* have the highest requirement of English skill includes newspaper and book publishers, broadcasting, etc. The ranking is displayed in figure 3.9 below

Figure 3.9. Industry by English skills requirement

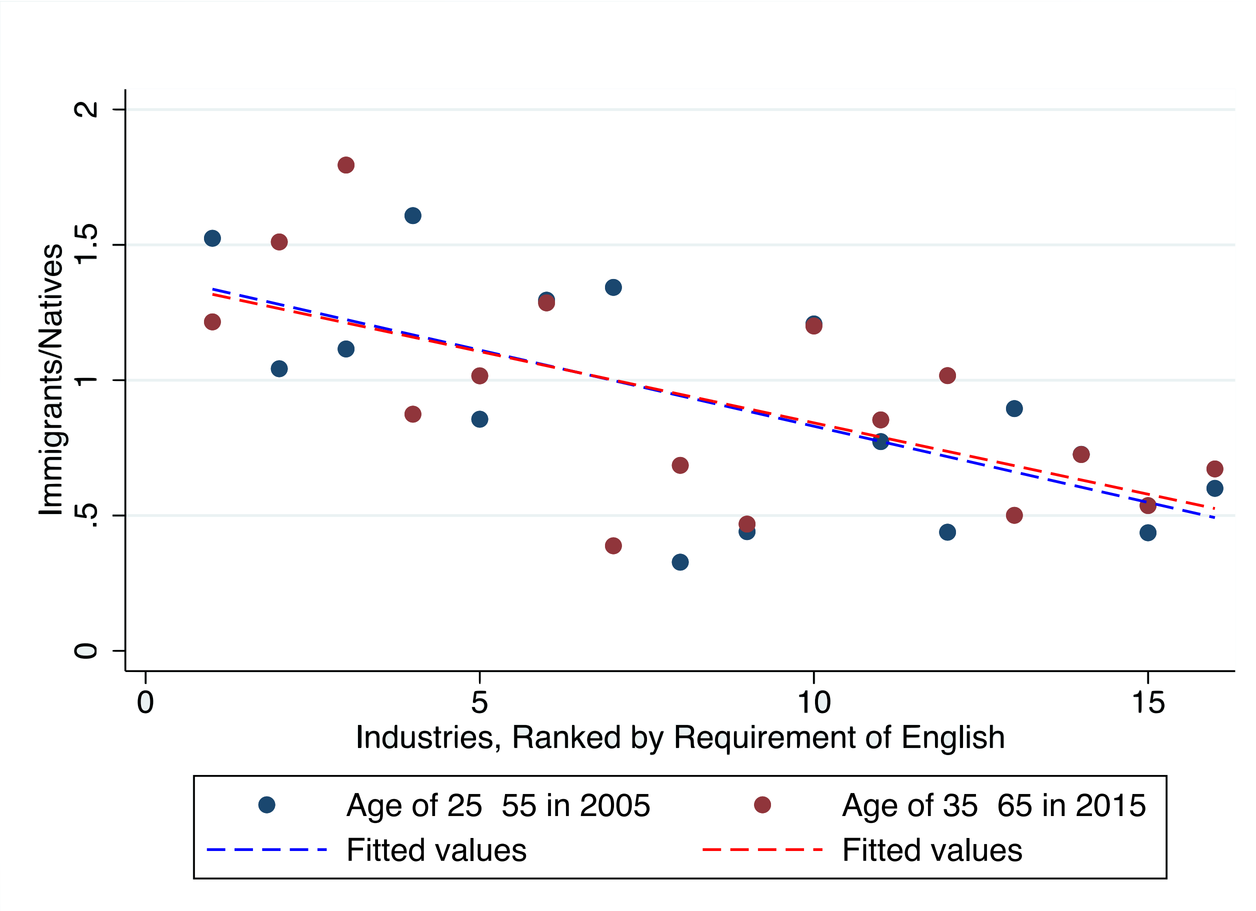
	indcat	ind_engrank
1	OtherSer	1
2	TransportWareHousing	2
3	Manufacturing	3
4	ArtEntFood	4
5	RetailTrade	5
6	AgrFishHunt	6
7	Construction	7
8	Utilities	8
9	Mining	9
10	WholesaleTrade	10
11	EduHealSocial	11
12	Army	12
13	ProfessionalSciMangAdmin	13
14	FinanceInsurRealRentLease	14
15	PublicAdmin	15
16	InformationCommu	16

Figure 3.10. Immigrants, Natives Employment by English skills requirement (industry level)



As before, similar exercises shows that immigrants are concentrated in industries where English skills requirement is lower and after 10 years, immigrants sorting into industries that require more English skills (Figures 3.10 and 3.11).

Figure 3.11. Immigrants/natives employment by English skills requirement, 10 years after



3.4 Concluding remarks

This paper proposes models and develops economic insights into the determinants of immigrants' location/occupational choices, how comparative (dis)advantages in taste and skill compared to natives drive their choices. Through price effect, their residential and occupational choices influences the natives' decision to relocate to different regions and switch to occupations where they have comparative advantages over immigrants. The difference between immigrants and natives' occupational choices is smaller with the duration of immigrants' stay as they acquire native-specific skill that closes their skill gap with them.

The models' predictions are shown to conform with the US data and their economic insights provide valuable lessons to policymakers. These insights include (i) size and composition of immigrant inflow could have differential effects on wage, living cost across regions and industries as well as the incentive to integrate of existing immigrants and (ii) the impact of a immigrant cohort varies with the duration of their stay as they acquire relevant skillsets and the comparative (dis)advantages between them and natives get smaller.

3.5 Appendix

Lemma 1 When locational characteristics, $A^i_{s,j}, U_s$, and individual characteristics, G^i_s, G^i_j, are strictly positive and finite, the exceed demands is continuous, homogeneous of degree zero in prices, and satisfy Walras' law. Furthermore, under Assumption 1, the exceed demands also exhibit gross substitute property.

Proof Demand of good j produced in k from agents in location s

$$d_{ks,j}(p) = \left[\alpha_j\right]^{\rho} \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{-\rho} \frac{\sum_{i',j'} A^{i'}_{s,j'} p_{s,j'} L^{i'}_{s,j'}}{P_s}.$$

Aggregate demand of good j produced in k, including the iceberg trade cost

$$d_{k,j}(p) = \sum_s T_{ks} \left[\alpha_j\right]^{\rho} \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{-\rho} \frac{\sum_{i',j'} A^{i'}_{s,j'} p_{s,j'} L^{i'}_{s,j'}}{P_s}$$

The exceed demand of good type j produced in k is

$$D_{k,j}(p) = \sum_s T_{ks} \left[\alpha_j\right]^{\rho} \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{-\rho} \frac{\sum_{i',j'} A^{i'}_{s,j'} p_{s,j'} L^{i'}_{s,j'}}{P_s} - \sum_i A^i_{k,j} L^i_{k,j}$$

It is immediate that the exceed demand is continuous and homogeneous of degree 0 in p since $L^i_{s,j}$ is homogeneous of degree 0 in p while P_s and $P_{s,j}$ are homogeneous of degree 1 in p. Furthermore, the exceed demand functions also satisfy Walras' law

$$
\begin{aligned}
\sum_{k.j} p_{k,j} D_{k,j}(p) &= \sum_{k,j} p_{k,j} \sum_s T_{ks} [\alpha_j]^{\rho} \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{-\rho} \frac{\sum_{i',j'} A^{i'}_{s,j'} p_{s,j'} L^{i'}_{s,j'}}{P_s} - \sum_{k,j} p_{k,j} \sum_i A^i_{k,j} L^i_{k,j} \\
&= \sum_{k,j} \sum_s [\alpha_j]^{\rho} \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{1-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{1-\rho} \sum_{i',j'} A^{i'}_{s,j'} p_{s,j'} L^{i'}_{s,j'} - \sum_{k,j} \sum_i A^i_{k,j} p_{k,j} L^i_{k,j} \\
&= \sum_s \left\{ \sum_{k,j} [\alpha_j]^{\rho} \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{1-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{1-\rho} \right\} \sum_{i',j'} A^{i'}_{s,j'} p_{s,j'} L^{i'}_{s,j'} - \sum_{k,j} \sum_i A^i_{k,j} p_{k,j} L^i_{k,j} \\
&= \sum_s \sum_{i',j'} A^{i'}_{s,j'} p_{s,j'} L^{i'}_{s,j'} - \sum_{k,j} \sum_i A^i_{k,j} p_{k,j} L^i_{k,j} \\
&= 0.
\end{aligned}
$$

Note that for agent type i in location s working in industry h, the change their demand of goods k,j with respect to a change in price of goods k,j' includes the increase in demand of goods k,j in location s to substitute for goods k,j plus the net change in the demand of goods k,j from agents type i from s working in h moving to k and working in j'

$$
\begin{aligned}
\frac{\partial d^i_{k,j}(p)}{\partial p_{k,j'}} &= T_{ks,j} \left[\alpha_j\right]^{\rho} \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{-\sigma} \left[P_{s,j}\right]^{-\rho} (\rho - 1) P_s^{\rho-2} \left(A^i_{s,h} p_{s,h} L^i_{s,h}\right) \frac{\partial P_s}{\partial p_{k,j'}} \\
&+ T_{ks,j} \left[\alpha_j\right]^{\rho} \left[\frac{p_{ks,j}}{P_{s,j}}\right]^{1-\sigma} \left[\frac{P_{s,j}}{P_s}\right]^{1-\rho} A^i_{s,h} \frac{p_{s,h}}{p_{ks,j}} \frac{\partial L^i_{s,h}}{\partial p_{k,j'}} \\
&- T_{kk,j} \left[\alpha_j\right]^{\rho} \left[\frac{p_{kk,j}}{P_{k,j}}\right]^{1-\sigma} \left[\frac{P_{k,j}}{P_k}\right]^{1-\rho} A^i_{k,j'} \frac{p_{k,j'}}{p_{kk,j}} \frac{\partial L^i_{s,h}}{\partial p_{k,j'}} \\
&= A \{ (\rho - 1) \left[\alpha_{j'}\right]^{\rho} \left[\frac{P_{s,j'}}{P_s}\right]^{1-\rho} \left[\frac{p_{ks,j'}}{P_{s,j'}}\right]^{1-\sigma} \\
&+ \left[T^{\sigma-1}_{ks,j} \left[\frac{P_{k,j}}{P_{s,j}}\right]^{\sigma-\rho} \left[\frac{P_k}{P_s}\right]^{\rho-1} \frac{A^i_{k,j'}}{A^i_{k,h}} \frac{p_{k,j'}}{p_{s,h}} - 1 \right] \pi^i_{k,j'} \varepsilon \} \\
&> A \left\{ (\rho - 1) - (ax - 1) \frac{bx}{bx + c} \varepsilon \right\}
\end{aligned}
$$

where

$$A = T_{ks,j}T_{ks,j'}\left[\alpha_j\right]^{\rho}\left[\frac{p_{ks,j}}{P_{s,j}}\right]^{1-\sigma}\left[\frac{P_{s,j}}{P_s}\right]^{1-\rho}\frac{L^i\pi^i_{s,h}A^i_{s,h}p_{s,h}}{p_{ks,j}p_{ks,j'}}$$

$$a = \overline{T}^{1-\sigma}\frac{\overline{A}}{\underline{A}}$$

$$b = \left(\overline{GU}\right)^{\varepsilon}$$

$$c = \left(\frac{\underline{GU}}{\overline{\overline{T}}}\right)^{\varepsilon}.$$

$\frac{\partial d^i_{k,j}(p)}{\partial p_{k,j'}}$ is positive when $\rho > 1+\zeta\varepsilon$ for any price vector with

$$\zeta = \left(1 - \frac{\sqrt{4a^2+4a\frac{b}{c}}-2a}{2\frac{b}{c}}\right)\left(\frac{\sqrt{4a^2+4a\frac{b}{c}}-2a}{\sqrt{4a^2+4a\frac{b}{c}}}\right).$$

For goods from k', j',

$$\frac{\partial d^i_{k,j}(p)}{\partial p_{k',j'}} = A\{(\rho-1)\left[\alpha_{j'}\right]^{\rho}\left[\frac{P_{s,j'}}{P_s}\right]^{1-\rho}\left[\frac{p_{k's,j'}}{P_{s,j'}}\right]^{1-\sigma} + \left[\frac{T^{1-\sigma}_{kk',j}}{T^{1-\sigma}_{ks,j}}\left[\frac{P_{k',j}}{P_{s,j}}\right]^{\sigma-\rho}\left[\frac{P_{k'}}{P_s}\right]^{\rho-1}\frac{A^i_{k,j'}p_{k,j'}}{A_{s,h}p_{s,h}}-1\right]\pi^i_{k',j'}\varepsilon\}$$

where

$$T_{ks,j}T_{k's,j'}\left[\alpha_j\right]^{\rho}\left[\frac{p_{ks,j}}{P_{s,j}}\right]^{1-\sigma}\left[\frac{P_{s,j}}{P_s}\right]^{1-\rho}\frac{A^i_{s,h}p_{s,h}L^i\pi^i_{s,h}}{p_{k's,j'}p_{ks,j}}$$

and is positive if the same condition above holds.

For goods from k', j,

$$\begin{aligned}
d^{i}_{ks,j}(p) &= T_{ks,j}\left[\alpha_j\right]^{\rho}\left[p_{ks,j}\right]^{-\sigma}\left[P_{s,j}\right]^{\sigma-\rho}P_s^{\rho-1}I^i \\
\frac{\partial d^{i}_{ks,j}}{\partial p_{k',j}} &= T_{ks,j}\left[\alpha_j\right]^{\rho}\left[p_{ks,j}\right]^{-\sigma}\left[P_{s,j}\right]^{\sigma-\rho}P_s^{\rho-1}\frac{\partial I^i}{\partial p_{k',j}} \\
&\quad + T_{ks,j}\left[\alpha_j\right]^{\rho}\left[p_{ks,j}\right]^{-\sigma}\left[P_{s,j}\right]^{\sigma-\rho}I^i(\rho-1)P_s^{\rho-1}\frac{\partial P_s}{\partial p_{k',j}} \\
&\quad + T_{ks,j}\left[\alpha_j\right]^{\rho}\left[p_{ks,j}\right]^{-\sigma}P_s^{\rho-1}I^i(\sigma-\rho)\left[P_{s,j}\right]^{\sigma-\rho-1}\frac{\partial P_s}{\partial p_{k',j}}
\end{aligned}$$

which is positive when $\sigma > \rho$ and $\rho > 1+\zeta\varepsilon$.

Since the exceed demand is homogeneity of degree zero in prices, it implies

$$\nabla d(p)\,p = 0$$

and therefore, $\frac{\partial D_{k,j}(p)}{\partial p_{k,j}} < 0$.

BIBLIOGRAPHY

Ahlfeldt, Gabriel M et al. (2015). “The economics of density: Evidence from the Berlin Wall”. In: *Econometrica* 83.6, pp. 2127–2189.

Alba, Richard and Nancy Foner (2015). *Strangers no more: Immigration and the challenges of integration in North America and Western Europe*. Princeton University Press.

Bauer, Thomas, Gil S Epstein, and Ira N Gang (2005). “Enclaves, language, and the location choice of migrants”. In: *Journal of Population Economics* 18.4, pp. 649–662.

Beaman, Lori A (2011). “Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the US”. In: *The Review of Economic Studies* 79.1, pp. 128–161.

Ben-Porath, Yoram (1967). “The production of human capital and the life cycle of earnings”. In: *Journal of political economy* 75.4, Part 1, pp. 352–365.

Berg, Peter ter et al. (2009). “Unification of the Frechet and Weibull Distribution”. In: *De Nederlandsche Bank*.

Bernstein, Shai et al. (2018). *The contribution of high-skilled immigrants to innovation in the United States*. Tech. rep.

Bleakley, Hoyt and Aimee Chin (2004). “Language skills and earnings: Evidence from childhood immigrants”. In: *Review of Economics and statistics* 86.2, pp. 481–496.

— (2010). “Age at arrival, English proficiency, and social assimilation among US immigrants”. In: *American Economic Journal: Applied Economics* 2.1, pp. 165–92.

Borjas, George J (1998). “To ghetto or not to ghetto: Ethnicity and residential segregation”. In: *Journal of Urban Economics* 44.2, pp. 228–253.

— (2000). “Ethnic enclaves and assimilation”. In: *Swedish Economic policy review* 7.2, pp. 89–122.

— (2015). “The slowdown in the economic assimilation of immigrants: Aging and cohort effects revisited again”. In: *Journal of Human Capital* 9.4, pp. 483–517.

Cassidy, Hugh (2019). “Occupational Attainment of Natives and Immigrants: A Cross-Cohort Analysis”. In: *Journal of Human Capital* 13.3, pp. 375–409.

Chiswick, Barry R and Paul W Miller (1992). “Language in the immigrant labor market”. In: *Immigration, language and ethnicity: Canada and the United States*, pp. 229–296.

— (1995). “The endogeneity between language and earnings: International analyses”. In: *Journal of labor economics* 13.2, pp. 246–288.

— (2005). “Linguistic distance: A quantitative measure of the distance between English and other languages”. In: *Journal of Multilingual and Multicultural Development* 26.1, pp. 1–11.

Croix, David de la, Matthias Doepke, and Joel Mokyr (2016). “Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Pre-Industrial Economy”. In: *Working Paper*.

Damm, Anna Piil (2009). “Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence”. In: *Journal of Labor Economics* 27.2, pp. 281–314.

Dorfman, Robert (1979). “A formula for the Gini coefficient”. In: *The review of economics and statistics*, pp. 146–149.

Dustmann, Christian and Arthur Van Soest (2002). “Language and the earnings of immigrants”. In: *ILR Review* 55.3, pp. 473–492.

Edin, Per-Anders, Peter Fredriksson, and Olof Aslund (2003). “Ethnic enclaves and the economic success of immigrants—Evidence from a natural experiment”. In: *The quarterly journal of economics* 118.1, pp. 329–357.

Florax, Raymond JGM, Thomas de Graaff, and Brigitte S Waldorf (2005). “A spatial economic perspective on language acquisition: segregation, networking, and assimilation of immigrants”. In: *Environment and Planning A* 37.10, pp. 1877–1897.

Gabaix, Xavier (1999a). “Zipf’s Law and the Growth of Cities”. In: *American Economic Review* 89.2, pp. 129–132.

— (1999b). “Zipf’s law for cities: an explanation”. In: *The Quarterly journal of economics* 114.3, pp. 739–767.

Hendricks, Lutz and Todd Schoellman (2017). “Human capital and development accounting: New evidence from wage gains at migration”. In: *The Quarterly Journal of Economics* 133.2, pp. 665–700.

LaLonde, Robert J and Robert H Topel (1992). "The assimilation of immigrants in the US labor market". In: *Immigration and the workforce: Economic consequences for the United States and source areas*. University of Chicago Press, pp. 67–92.

— (1997). "Economic impact of international migration and the economic performance of migrants". In: *Handbook of population and family economics* 1, pp. 799–850.

Lazear, Edward P (1999). "Culture and language". In: *Journal of political Economy* 107.S6, S95–S126.

Le, Tung (2019). "Diversity, Communication Friction, Growth, and Inequality". In: *Working Paper*.

Lucas, Robert E (2009). "Ideas and growth". In: *Economica* 76.301, pp. 1–19.

Lucas, Robert E and Benjamin Moll (2014). "Knowledge growth and the allocation of time". In: *Journal of Political Economy* 122.1, pp. 1–51.

Luttmer, Erzo GJ (2007). "Selection, growth, and the size distribution of firms". In: *The Quarterly Journal of Economics* 122.3, pp. 1103–1144.

— (2012). "Technology diffusion and growth". In: *Journal of Economic Theory* 147.2, pp. 602–622.

— (2015). "An Assignment Model of Knowledge Diffusion and Income Inequality". In: *Working Paper*.

Massey, Douglas S (2008). *New faces in new places: The changing geography of American immigration*. Russell Sage Foundation.

Ottaviano, Gianmarco IP and Giovanni Peri (2006). "The economic value of cultural diversity: evidence from US cities". In: *Journal of Economic geography* 6.1, pp. 9–44.

Schelling, Thomas C (1971). "Dynamic models of segregation". In: *Journal of mathematical sociology* 1.2, pp. 143–186.

Schoeni, Robert F et al. (1996). *The Mixed Economic Progress of Immigrants.* ERIC.

Terrazas, Aaron Terrazas Aaron (Mar. 2017). *Immigrants in New-Destination States*. URL: `https://www.migrationpolicy.org/article/immigrants-new-destination-states`.

Warman, Casey (2007). “Ethnic enclaves and immigrant earnings growth”. In: *Canadian Journal of Economics/Revue canadienne d’économique* 40.2, pp. 401–422.

Young, Malcolm B et al. (1995). “National Evaluation of Adult Education Programs. Executive Summary.” In: