ESSAYS ON INTERGENERATIONAL CULTURAL ASSIMILATION

A Dissertation
Presented to
The Faculty of the Department
of Economics
University of Houston

In Partial Fulfillment
Of the Requirements for the Degree of
Doctor of Philosophy

By
Saumya Rana
May 2019

ESSAYS ON INTERGENERATIONAL CULTURAL ASSIMILATION

Saumya Rana

APPROVED:

Vikram Maheshri, Ph.D.
Committee Chair

Chinhui Juhn, Ph.D.

Elaine Liu, Ph.D.

Srishti Nayak, Ph.D.
Princeton University

Antonio D. Tillis, Ph.D.
Dean, College of Liberal Arts and Social Sciences
Department of Hispanic Studies

ESSAYS ON INTERGENERATIONAL CULTURAL ASSIMILATION

An Abstract of a Dissertation
Presented to
The Faculty of the Department
of Economics
University of Houston

In Partial Fulfillment
Of the Requirements for the Degree of
Doctor of Philosophy

By
Saumya Rana
May, 2019

Abstract

This dissertation is comprised of two studies on intergenerational cultural assimilation. The first chapter looks at the effect of preferences on assimilation in terms of residential outcomes. Parent's desire to transmit their culture to their children can lead them to reside with neighbors of the same ethnicity. Ethnic segregation may be an outcome of such sorting. I develop a theoretical model of intergenerational cultural transmission that incorporates moving and neighbor choices. I prove how these choices can generate a dynamic equilibrium in which ethnic segregation will persist in the long run at the disaggregate neighborhood level with diversity in the aggregate. I examine whether model predictions regarding these segregation-inducing parental moves are supported by data. I use U.S Census Microdata (1990, 2000) and U.K Fourth National Survey of Ethnic Minorities (1993-1994). The empirical strategy exploits variation in moves by the presence and age of children, ethnicity shares and cultural motivations. I find that a one standard deviation increase in share of own ethnicity in a location reduces the probability of leaving that location by .8 percentage points (7.6%) for parents compared to non-parents. Parents are also 1.76 percentage points (3.6%) more likely to go from a location with a lower share of own ethnicity to one with a higher share of own ethnicity. These effects are stronger for parents with young children. The findings suggest that young children may be disproportionately more exposed to ethnic enclaves. Cultural transmission appears to be an important causal determinant of the differential sorting patterns and a possible mechanism to help explain persistence in enclaves.

The second chapter examines the effect of linguistic constraints on outcomes. In particular, I study the causal impact of English proficiency among immigrant parents on language skills, educational outcomes and attitudes of second-generation immigrants in the U.K. To address the endogeneity in parent's English proficiency, I take advantage of the phenomenon that younger children learn languages more easily than older children. I employ a difference in difference strategy based on the instrument proposed by Bleakely and Chin (2004)- age at arrival of immigrant parents who came as children from English and non-English speaking countries to the U.K. I find that parent's English proficiency has a significant and positive impact on the English proficiency of their U.K born children. I also find a positive impact on their years of schooling, probability of pursuing higher education and job satisfaction. The impact of parental English proficiency on social attitudes related to risk taking, trustworthiness and political interest is negative. Additionally, individuals to English proficient parents are less likely to perceive religion to be of significance. They are also less likely to consider retention of own identity as crucial. I find parent's educational qualification, labor force status and residential location to be important channels in explaining these language effects.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Prof. Vikram Maheshri for his guidance throughout my Ph.D. studies and research. His knowledge and mentorship enabled me to think independently and stay persistent. Without his support and encouragement, this dissertation would not have been possible. I am immensely grateful to Prof. Chinhui Juhn for her inputs, for stimulating conversations and for the opportunity to work as her research assistant for the last couple of years. I have learnt from her every step of the way and my experience as a student and Economist would not have been the same otherwise. I would also like to sincerely thank Prof. Elaine Liu whose knowledge, attention to detail, advise and prompt help whenever required was indispensable to the success of my dissertation.

A special and sincere thanks to my external reviewer, Prof. Srishti Nayak for her invaluable inputs and fresh perspectives.

I sincerely thank Prof. Aimee Chin, Prof. Kei Mu Yi, Prof. Willa Friedman, Prof. Fan Wang, Prof. Nathan Canen, Prof. Kohlhase and Dr. Andrew Zupann for all the insightful and helpful discussions. I would also like to extend my gratitude to Prof. German Cubas, Prof. Dietrich Vollrath, Prof. Bent Sorensen and Prof. David Papell for their encouragement and guidance. I also thank Ms. Amber Pozo for all the administrative help.

I would also like to thank my colleagues for being part of this amazing journey-Yewande Olapade, Elizabeth Luh, Aritri Banerjee, Vinh Pham, Arpita Chakraborty, Gautham Udupa, Daniel Jacobs, Lalita Thienprasiddhi, Xuejing Zuo and Vasundhara Tanwar. Heesuk Jung has been an incredible friend, colleague and fellow Economist. I applaud her commitment to our daily coffee sessions at Cougar and thank her for adding a little Macro in my every day life.

Swati Sood, Poorna Bhattacharjee, Sonali punhani, Rukmini Bahl, Lily Virguez, Sharanya Chandran and Maroya Faied , all deserve a special mention for being extraordinary women and for being the support system that they always have been. I sincerely appreciate the respect they have shown this journey, for the effort in understanding the value of this process and for not once asking when this would come to an end. I also thank my Houston friends for making the end of the week something to look forward to and Ivy and Akshay for making the weekday evenings something to look forward to.

I would also like to thank my mother and sister in law for their patience, help and some amazing food. My mausi, mama and incredible grandparents have always been encouraging and appreciative of my pursuits and the support has given me both the confidence and the motivation to do well.

I would like to thank my immediate family. My parents have been role models in every possible way. I thank them for teaching me to be fearless and for believing in me. Their selflessness and spirit is exemplary. I thank them for putting me first and owe this success to their unconditional love and support. My brother deserves all the appreciation for being the person that he is - compassionate, persevering and humorous. His qualities have been inspirational to me in my own journey.

Finally, I thank the love of my life - Ishan Rao. He has been there for me at every step, patiently and selflessly. I thank him for listening to me talk endlessly about Economics, for coming every week to see me and for being an incredibly positive influence in my life. He made this pursuit a memorable one.

to taiji and tauji

Contents

List of Figures

List of Tables

Chapter 1

Cultural Transmission and Ethnic Segregation

1.1 Introduction

Parental investments are important determinants of children's economic success (Mulligan, C. B. (1997)). Some of these parental investments may be influenced by cultural preferences that they wish to transmit to the next generation[1]. Bisin, Topa and Verdier (2004) show that marriage homogamy along religious lines among Protestants, Catholics and Jews in the U.S can be explained by the desire to transmit own religious beliefs to children. Dixit (2009) and Gradstein and Justman (2002, 2005) explore the role of collective socialization mechanisms such as the role of education in the transmission of a common culture and social norms. Historical institutions have also been shown to explain intergenerational transmission and persistence in outcomes. Guise, Sapienza, and Zingales (2016) provide evidence of higher civic capital in Italian cities that experienced the formation of free city-states in medieval Italy. Acemoglu et al. (2001) estimate the effects of colonial institutions on the persistence in economic performance.

[1] I will use the term cultural transmission and socialization interchangeably throughout this paper to refer to the transfer of parent's cultural trait to children.

Cultural transmission may result from investments at home, for instance through the time parents devote to teaching their children about their cultural values. On the other hand, it can be an outcome of imitation and learning from peers and neighbors who share those values. If within-family investments are costly or challenging due to reasons such as time constraints, parents may choose to reside with others of their own type. Doing so will allow for cultural transmission through social interactions. Such sorting could lead to segregation.

In this paper, I focus on the choice of neighbors as a mechanism for cultural transmission along ethnic lines[2]. I further investigate whether these choices can explain the rise and persistence of ethnic enclaves.

First, I develop a theory to formalize the above intuition and provide the link between cultural transmission, neighbors and segregation. I generalize the seminal Bisin and Verdier (2001) model of cultural transmission to incorporate moving and neighbor choices. In the model, parents derive utility if the child is successfully socialized to their trait but incur socialization costs. I provide conditions under which moving to a more ethnically similar neighborhood will lower these socialization costs and make it optimal for parents to self-segregate. The model generates empirically testable predictions regarding the optimal moving patterns of parents with a desire for socialization. Specifically, I show that parents are less likely to migrate out if the share of their own ethnicity in the current residence increases. Conditional on moving, they are also more likely to move from ethnically less similar to ethnically more similar neighborhoods.

I then demonstrate how these moving patterns can generate a novel dynamic equilibrium in which segregated ethnic enclaves at a disaggregate neighborhood level will persist with aggregate diversity. According to Bisin and Verdier (2001), if family and society act as substitutes (complements) in the socialization process or in other words if within-family socialization declines (increases) as the fraction of preferred trait in the population increases and vice versa, the

[2] For simplicity, from here on I will use the term ethnic transmission/socialization for transmission of ethnic traits.

long run distribution of cultural traits in the population will be completely heterogenous (homogenous). I theoretically contribute to their paper by showing that if moving is permitted, homogeneity is sustainable at a disaggregate level with heterogeneity in the aggregate under similar socialization assumptions.

Intuitively, one possible way that persistence can arise is if ethnic transmission generates similar socialization preferences among future generations. The empirical existence and persistence of such enclaves is well documented (Portes and Manning (1986)). This paper provides a mechanism to explain such long run hysteresis[3].

Next, I test whether the model predictions regarding these segregation-inducing parental moves are supported by data. However, there is no single data source that contains socialization preferences, residential histories and demographics disaggregated by parental status and geography. I address these obstacles by using multiple and complementary datasets- the *U.S Integrated Public Use Census Microdata (IPUMS; 1990, 2000)* and the *U.K Fourth National Survey of Ethnic Minorities (U.K FNSEM; 1993-1994)*. U.S Census data allows me to track migration and the U.K data has the advantage of a highly disaggregate level of geography and stated cultural preferences.

I employ a difference in difference design to evaluate how migration varies across parents and non-parents and across $PUMAs$ with varying own ethnicity shares[4] using $IPUMS$ data. I find that a one standard deviation increase in share of own ethnicity in a $PUMA$ reduces the probability of leaving that $PUMA$ by .808 percentage points (∼7.6%) for parents compared to non-parents. Parents are also 1.76 percentage points (∼3.6%) more likely to go from a neighborhood with a lower share of own ethnicity to one with a higher share of own ethnicity.

[3]Bosch et al. (2015) focus on ethnic enclaves in Spain and show that differential treatment in the rental market can explain their persistence.

[4]The most disaggregate level of geography that is publicly available in the U.S Census Microdata is $PUMA$. $PUMA$ is a Census designated statistical area with a population of at least 100,000.

To establish a causal link between these findings and ethnic transmission is surely challenging. For example, parents and non-parents differ on dimensions other than preferences for socialization and ethnic composition of locations is endogenous. To address these issues, I include extensive micro-level controls that may be correlated with selection into parenthood and particular residential choices. I also compare the migration pattern of parents and non-parents based on age of child. I find that parents are more likely to sort with ethnically similar neighbors than non-parents but this probability declines with increase in age of eldest child. This systematic variation with age is consistent with the literature that finds early formative years to be important for ethnic and religious transmission[5]. The estimates are also qualitatively consistent when I compare parents with young children to non-parents and to those whose children are above 19 years of age and possibly past the age of socialization. They are also robust to the inclusion of individuals who could have moved in anticipation of having children. One issue is that own ethnicity shares could be correlated with amenities that parents and non-parents value differently. For example, schools in ethnically similar neighborhoods might offer a particular curriculum that parents in that ethnic community value more. I also compare parents with eldest child in age 3-5, 6-8 and 9-11 to parents whose children at the time of moving were between 12-18. To the extent that parents with kids between 6-8 years will care as much about the curriculum as those with kids between say 9-11 years, the bias of this nature will get differenced out. As an additional check, I also use an instrumental variable strategy that exploits plausibly exogenous variation in own ethnic composition at city level. I find the results to be qualitatively similar to before.

To get at cultural motivations, in addition to looking at variation across parental status and own ethnic shares, I also use variation by immigrant arrival

[5]For evidence that ethnic attributes are acquired usually in the early formative years of childhood, see Clark and Worthington (1987); Cornwall (1988); Erickson (1992); Hayes and Pittelkow (1993).

age. There is evidence in the literature that suggests that those with fewer years since arrival in the host country have stronger cultural preferences (Aslund et al. (2015); Cutler et al. (2008); Edin et al. (2003); Chin et al. (2004, 2008, 2010)). Finally, I also use the U.K Fourth National Survey of Ethnic Minorities to evaluate stated socialization, moving and ethnic neighbor preferences at a highly disaggregate level of geography. Estimates are robust to these various specifications and suggest that parents with young children and those who are possibly less culturally assimilated are most likely to sort with ethnically similar neighbors. Consistent with the model, cultural transmission appears to be an important mechanism to explain these patterns.

To the best of my knowledge, this is the first paper to suggest that young children may be disproportionately more exposed to ethnic enclaves than others. Parents face various tradeoffs when making their locational choices. To the extent that cultural amenities (e.g. heritage Sunday schools) and neighbors are substitutes in the socialization process, centrally located and easily accessible cultural amenities may reduce the incentive to find neighbors of the same type. Additionally, if non-cultural moving costs are high (e.g. due to high real estate prices), then despite the incentive, it may not be optimal for parents to reside in proximity with ethnically similar neighbors. According to recent evidence, neighborhood context and years of childhood exposure has a significant impact on intergenerational economic mobility (Chetty et al. (2014), Agostinelli (2018)). It is thus important to consider these tradeoffs that influence the social context of the environment that children grow up in. However, there is little consensus on the nature of this long run impact and it is shown to ultimately depend on the average attributes in that enclave[6]. To the extent that enclaves have more social capital and stronger networks than non-enclaves (Patacchini and Zenou (2012); Cutler et al. (2008); Waldfogel (2003)), the enclave effect on

[6]Studies find that the impact of ethnic segregation on human capital accumulation depends on average attributes of the group (Borjas (1998); Edin, Fredriksson and Aslund (2003); Cutler, Glaeser and Vigdor (2005a)). See Agostinelli (2018); Chetty and Hendren (2016a,b) for years of childhood exposure and impact of peers and neighborhoods.

intergenerational economic mobility could be positive (Chetty et al. (2014)). On the other other hand and as pointed out by Borjas (1992), if the average skill of the parent's ethnic enclave is low, human capital would be lower among future generations as well. For instance, he finds that average schooling of the second generation is lowered by .2 years to a one year decrease in average schooling among parent generation's ethnic group.

Either way, from a policy perspective, investments in enclaves, particularly those that young children benefit from (e.g. schooling investments) are worthwhile goals and may yield positive effects that are larger and more long run than previously expected.

The rest of the paper is organized as follows. Section 1.2 discusses related literature. The model is included in section 1.3. In section 1.4, I discuss the data, empirical strategy and estimates based on regressions using U.S Census Microdata. Section 1.5 provides a discussion of U.K FNSEM data and related results. In section 1.6, I discuss several robustness checks. Section 1.7 concludes.

1.2 Relevant Literature

There is extensive theoretical literature that studies dynamics of cultures in the long run. One branch of this literature is based on evolutionary selection mechanisms (Kockesen (2000); Bester and Guth (1998); Eshel et al. (1998); Fershtman and Weiss (1998); Robson (1996)). In these studies, preferences are intergenerationally transmitted either through genetic transmission or else through imitation. Models of cultural transmission were first introduced in 'Evolutionary Anthropology' by Cavalli-Sforza and Feldman (1973) and Boyd and Richerson (1985). In their models, a child is assumed to inherit parental preferences due to direct socialization effort of the parents. If the child is not directly socialized, he/she is then assumed to pick up the trait indirectly though imitation/learning from the environment. A key feature of anthropological models of cultural transmission such as those of Sforza, Richerson and others such as Walter (1965) and Baumrind (1967) is the assumption of exogenous direct socialization effort[7].

Bisin and Verdier (2001) is the first theoretical paper to endogenise socialization effort and develop an 'Economic Model of Cultural Transmission' with 2 cultural traits. In their framework, direct socialization effort is endogenous in the sense that parents choose their optimum effort based on the relative fraction of the preferred trait in the population. Montgomery (2008) is an extension of the Bisin and Verdier model to n traits. Saez-Marti and Sjögren (2008) look at how culture evolves in a model where parents socialize children to attributes that maximize the child's lifetime utility. I contribute to this theoretical literature by extending the Bisin and Verdier (2001) model to allow for endogenous moving and neighbor choices.

Few papers look at how neighborhoods (exogenously) affect the process of cultural transmission. Patacchini and Zenou (2007) using U.K National Child Development Study (NCDS) provide evidence that parents invest more in affecting children's educational attainment when living in high quality neighborhoods. Other papers such as Bisin, Patacchini, Verdier, and Zenou (2010) using U.K Fourth National Survey of Ethnic Minorities focus on the impact of neighborhood ethnic density on identity formation[8]. Algan et al. (2013) using the French Labor Force Survey look at how neighborhoods influence naming choice and the probability of the child being given an Arabic name. Relative to this literature, I look at the relationship between culture and neighborhoods by endogenizing neighborhoods and allowing for cultural transmission to also affect neighborhood choice.

Other set of studies such as those on school segregation shed light on parental preferences for their children's social interactions. For instance, Re-

[7]See Bisin and Verdier (2001) for a review.

[8]See Benhabib, J., Bisin, A., & Jackson, M. O. (Eds.), (2010), Handbook of social economics for a complete review of theoretical and empirical literature on the economics of cultural transmission.

ber (2005) and Baum-Snow and Lutz (2011) explore the extent to which white enrollment reduced in school districts that were court ordered to be desegregated in the 1960s and 1970s. Cascio and lewis (2012) provide evidence that native flight from the average California school district between 1970-2000 in response to increase in low english hispanic enrollment is attributable to lower preferences for hispanic peers among white parents[9]. However, the distaste for diversity in children's peers suggested by this literature is not specifically related to cultural transmission.

Motivations for various parental investments in children and the relation to cultural assimilation has also been explored in the literature. For instance, Abramitzky et al. (2016) provide evidence of name based assimilation. They find that historical immigrants to the U.S chose less foreign names for their children the longer they were in the country, eventually eliminating almost half the gap with natives. Jia and Persson (2017) show how parents in ethnically mixed marriages choose the ethnicity of their children in China. The key finding is that social motives have a significant influence on the impact of individual incentives on parental choice. My paper differs from this literature in two regards- (1) I look at a different type of cultural investment for children, ethnic neighbors and a different motivation to explain the same, ethnic transmission (2) I demonstrate that cultural investments of parents are higher than those of non-parents.

Other papers on culture and segregation demonstrate how individual's own cultural preferences are reflected in their location choices (Aslund et al. (2015)) or in the higher premium they are willing to pay to reside with ethnically similar neighbors (Maisy Wong (2008)). Card, Mas and Rothstein (2008) focus on the preferences of the white population. They show how an increase in neighborhood minority population beyond a certain threshold leads to white outflows and in turn to extreme residential segregation. Cutler at al. (2008)

[9]For the impact of racial segregation within schools on outcomes, see Echenique, Fryer and Kaufman (2006)

study the determinants of immigrant segregation. They find that much of the resurgence in immigrant segregation since the 1980s in the U.S can be explained by increased suburbanization and immigrant occupation of neighborhoods previously occupied by natives. I contribute to this literature by evaluating the role of cultural transmission as a determinant of long run segregation. Additionally, I focus on children and demonstrate a differential in the desire for enclave residence between parents and non-parents.

There are two closely related papers that are worth mentioning. Bezin and Moizeu (2018) is a theoretical paper that examines cultural preservation and segregation. While their model or theoretical framework is intuitively similar, this paper contributes by also providing empirical evidence on the interlinkages between cultural transmission differential sorting patterns of parents relative to others.

The second paper is Ionnides and Zanella (2007, working paper). This study uses PSID data and a discrete choice framework to show that households with children relative to those without children sort into neighborhoods with various attributes that are conducive to cultural transmission and human capital accumulation. However, the paper does not demonstrate any particular outcome of such location choices. Relative to their paper, I specifically focus on ethnic neighbor choices due to cultural transmission and prove segregation to be an outcome of such choices. I also demonstrate a difference in the moving patterns of households with children of various ages relative to those without children and those with older children. Both these theoretical and empirical findings contribute to Ionnides and Zanella (2007).

1.3 Model

1.3.1 Set Up

Suppose there are two types of cultural traits in the population, $\{r, s\}$ and two neighborhoods, $\{n_1, n_2\}$. At the beginning of each time period t, individuals belong to either of the two neighborhoods. At time t, fraction of the total population in neighborhoods n_1 and n_2 is $x_1(t)$ and $x_2(t) = 1 - x_1(t)$ respectively. Fraction of population with trait i where $i\epsilon\{r, s\}$, in neighborhood k where $k\epsilon\{n_1, n_2\}$ is $q_k^i(t)$.

I follow the set up proposed by Bisin et al. (2001, 2008) and Montgomery (2008). Suppose period length is dt and fraction dt of the population dies between time t and $t + dt$. Fertility and socialization takes place in individual's last period of life. Reproduction is asexual and families in the population consist of a single parent and a single child. After the child is born, the parent decides whether to continue in the current residence or migrate. Children are socialized after that in the neighborhood that parents choose[10].

Next, I examine how cultural traits evolve over time in the two neighborhoods. Given that individuals with traits r and s are symmetric, all the information required to examine the evolution of the two traits can be inferred by deriving the following[11]:

$$\{\frac{\partial q_1^i(q_1^i, q_{2,}^i x_1)}{\partial t}, \frac{\partial q_2^i(q_1^i, q_{2,}^i x_1)}{\partial t}, \frac{\partial x_1(q_1^i, q_{2,}^i x_1)}{\partial t}\} \quad [3.1]$$

Children are born without any cultural traits but get socialized to the parent $i's$ cultural trait with a certain probability. The probabilities of socialization depend on both the direct effort of the parent, τ_k^i and fraction of preferred trait in the population of the neighborhood that the child is socialized in. Since child is socialized after migration, socialization probabilities depend on fraction of *net in-migrants* in neighborhood k with the preferred trait. Probability that a child who is born to parent i and socialized in neighborhood k, acquires the cultural trait of the parent is $p_k^{ii}(q_1^i, q_2^i, x_1)$. Given 2 cultural traits, probability of socialization to trait j $(i \neq j)$ is $p_k^{ij}(q_1^i, q_2^i, x_1) = 1 - p_k^{ii}(q_1^i, q_2^i, x_1)$.

Let the probability of migration of parent i from $n_1\,(n_2)$ to $n_2\,(n_1)$ be

[10] Given this set up, note that only fraction dt will make migration and socialization decisions between time period t and $t + dt$.

[11] I ignore the time notation for simplicity.

$\lambda^i_{12}\,(\lambda^i_{21})$[12]. The net fraction of parents with trait i who move to k, $m^i_k(q^i_1, q^i_2, x^i_1)$ can be written as:

$$m^i_k(q^i_1, q^i_2, x_1) = q^i_k x_k - \lambda^i_{12} x_k + \lambda^i_{21}(1 - x_k) \quad [3.2]$$

Then, the dynamic equations specified in 3.1 can be written as:

$$\frac{\partial q^i_k(q^i_1, q^i_2, x_1)}{\partial t} = (\frac{1}{x_k})[\,(1 - q^i_k)\underbrace{(m^i_k p^{ii}_k + m^j_k p^{ji}_k)}_{(a)} - (q^i_k)\underbrace{(m^i_k p^{ij}_k + m^j_k p^{jj}_k)}_{(b)}\,] \;;\; k \epsilon (n_1, n_2) \quad [3.3]$$

- (a) denotes the fraction of net parent movers into neighborhood k whose children get socialized to trait i between t and $t + dt$ and
- (b) denotes the fraction of net parent movers into neighborhood k whose children get socialized to trait j between t and $t + dt$.

$$\frac{\partial x_1(q^i_1, q^i_2, x_1)}{\partial t} = (1 - x_1)(m^i_1 + m^j_1) - (x_1)(m^i_2 + m^j_2) \;;\; i, j \epsilon (r, s) \; and \; i \neq j \quad [3.4]$$

The derivation for equations 3.3 and 3.4 is included in appendix section 1.12. Endogenous migration probabilities $\lambda' s$ and the socialization probabilities are discussed in the next section[13].

[12] λs are endogenous and will be discussed in greater detail in the next section.

[13] Appendix section 13.1 includes the derivation of the general version shown in 3.3 and 3.4 and of the dynamic equations under specific migration probability and socialization technology.

1.3.2 Preferences, Endogenous Cultural Transmission and Migration

Given the previously mentioned set up, fraction dt parents will optimize between t and $t + dt$. They optimize location choice by comparing the utility they would get in the two neighborhoods (based on time t characteristics) and choose the location where this utility is higher[14].

Parents i derive utility from (1) Consumption/Income, $u^i(w^i)$ and (2) Cultural Transmission, $p_k^{ii}(f(\tau_k^i), q_k^i(t)).v^{ii} + p_k^{ij}(f(\tau_k^i), q_k^i(t)).v^{ij}$[15]. v^{ii} and v^{ij} are the utilities that parent i gets if the child acquires trait i and j respectively. Given that parents are altruistic, $v^{ij} \neq 0$. However, parents are assumed to have a cultural bias towards own trait and $v^{ii} > v^{ij}$ or $\Delta v^i = v^{ii} - v^{ij} \geq 0$. $c(\tau_k^i)$ represents the cost of socialization. Δv^i is exogenous and and let $\Delta v^i = \Delta v^j = \Delta v$.

Thus, for a parent i in neighborhood k, maximizing over time period dt, optimum utility based on time t shares is given by:

$$U_k^i(w*, \tau*) = max_\tau \left[u^i(w*) - c(\tau_k^i) + (p_k^{ii}(\tau_k^i, q_k^i(t)).v^{ii} + p_k^{ij}(\tau_k^i, q_k^i(t)).v^{ij})\right] \quad [3.5]$$

where $w^* = argmax_{w \epsilon W}\, u^i(w^i)$[16]

Let,

$$c(\tau_k^i) = \frac{1}{2}(\tau_k^i)^2 \quad [3.6]$$

[14]Note that parents optimize in every period. However, optimization in itself is myopic and does not account for any future change in cultural traits or socialization costs that their offsprings might incur.

[15]Note that when optimizing their location decision, individuals use the information on time t share of trait i and not the share of trait i after migration. That is why $p_k^{ii} = f(q_k^i, \tau_k^i)$. However, since socialization takes place after migration, actual probabilities depend on the share after migration. Even if individuals optimized location choice based on the (anticipated) shares after migration, given that all i parents are symmetric, optimum location decision will not change.

[16]Similar to Bisin and Verdier (2001), the crucial simplification here is that gain from socio-economic activities and socialization is additively separable. Additionally, parent's utility from consumption does not vary with neighborhoods.

and

$$p_k^{ii} = f(\tau_k^i) + (1 - f(\tau_k^i))(q_k^i(t)) \quad [3.7]$$

According to equation 3.7, probability of socialization to trait i of a child born to parent i depends on the direct effort of the parents, $f(\tau_k^i)$. If the child is not directly socialized, the probability of which is $1 - f(\tau_k^i)$, he/she picks the trait from the society with the probability q_k^i. Thus, even if direct socialization fails, the child could still acquire trait i depending on the relative fraction of i in the neighborhood that the child is socialized in. For simplicity, let $f(\tau_k^i) = \tau_k^i$. Then,

$$p_k^{ii} = \tau_k^i + (1 - \tau_k^i)(q_k^i(t) \quad [3.8]$$

Next, based on where $U_k^i(w*, \tau*)$ is higher, parents will choose that location.

1.3.3 Empirically Testable Moving Patterns

Without Moving costs:

If there are no moving costs, parents will choose neighborhood 2 if $U_2^i(w*, \tau*) > U_1^i(w*, \tau*)$ and vice versa. Under the assumptions of a simple socialization technology such as the one in 3.8, optimum socialization effort is

$$\tau_k^{*i}(q_i^k(t); \Delta v) = (1 - q_k^i(t))\Delta v$$

Optimum utility that parent i expects to get from choosing neighborhood k:

$$U_k^i(w*, \tau*) = u^i(w^*) - \frac{1}{2}((1 - q_k^i(t))\Delta v)^2 + (1 - q_k^i(t) + (q_k^i(t))^2)\Delta v + v^{ij}$$

$$\frac{\partial U_k^i(w*, \tau*)}{\partial q_k^i(t)} = (1 - q_k^i(t))(\Delta v)^2 - \Delta v^i + 2q_i^k(t)\Delta v$$

$0 < \Delta v^i \leq 1$ and thus[17]

$$\frac{\partial U_k^i(w*, \tau*)}{\partial q_k^i(t)} > 0 \quad [3.9]$$

Based on equation 3.9, moving to a location with a higher share of preferred trait relative to the origin (in other words, an increase in q_k^i) would lead to a gain in expected optimum utility and incentivize migration. In this case, *parents with cultural transmission preferences will move from a neighborhood with lower fraction of own type to a neighborhood with higher fraction of own type.*

With Moving costs:

Now suppose there are costs to moving denoted by σ and $0 \leq \sigma \leq 1$. Suppose $q_2^i(t) > q_1^i(t)$. Then, by the same reasoning as above, moving to n_2 will increase optimum utility. However, parents will move only if this gain exceeds moving costs or $\Delta U_{12}^i[\,(q_1^i(t), q_2^i(t)\,] = U_2^i[\,w*, \tau*(q_2^i(t))\,] - U_1^i[\,w, \tau^*(q_1^i(t))\,] > \sigma$.

If however the initial gain in utility does not exceed moving costs, then at no point in the future would there be any migration between the two locations. The reason is that under substitution, for any fraction of trait i such that $q_2^i \geq q_1^i$,

$$\frac{\partial q_2^i(t)}{\partial t} - \frac{\partial q_1^i(t)}{\partial t} \leq 0$$

The difference in the growth rate across the two locations will therefore never be high enough for the utility gain from migration to offset the fixed moving costs.

Thus, even under the moving cost case, either parents will not m*igrate, but if they do, it will be from a neighborhood with a lower fraction of own type to a neighborhood with higher fraction of own type.*

Corollary

Parent i in n_1 will relocate to n_2 if $\Delta U_{12}^i = U_2^i(w*, \tau*) - U_1^i(w*, \tau*) > 0$ and

[17] τ_k^i is also the probability of direct socialization and $\in [0, 1]$. This constraints Δv^i to (0,1]

vice versa. Given [3.5], $\frac{\partial U_1^i(w*,\tau*)}{\partial q_1^i(t)} > 0 => \frac{\Delta U_{12}^i}{\partial q_1^i(t)} < 0$. Similarly, $\frac{\partial U_2^i(w*,\tau*)}{\partial q_2^i(t)} > 0$ $=> \frac{\Delta U_{12}^i}{\partial q_2^i(t)} > 0$. Then, all else equal, if the share of preferred trait in the current neighborhood increases, the utility gain from migration will decline and vice versa. Likelihood of out-migration will drop from 1 to 0 if the increase in share in current neighborhood leads ΔU_{12}^i to go from > 0 to ≤ 0. In the appendix, I also derive continuous migration probabilities as shares of own trait in the two neighborhood. The general intuition still holds. *Parents with cultural transmission preferences are less (more) likely to migrate out of (into) current (destination) neighborhood, the higher the share of their own type in that neighborhood*

1.3.4 Dynamic Equilibrium: Disaggregate Enclaves and Aggregate Diversity

In this section, I show how the above described moving patterns can generate a dynamic equilibrium in which ethnic segregation will coexist in the long run at the disaggregate neighborhood level with diversity in the aggregate.

Equilibrium is defined by points q_1^i, q_2^i and x_1 for which all the three dynamic equations,

$$\frac{\partial q_1^i(q_1^i, q_{2,}^i x_1)}{\partial t}, \frac{\partial q_2^i(q_1^i, q_{2,}^i x_1)}{\partial t}, \frac{\partial x_1(q_1^i, q_{2,}^i x_1)}{\partial t}$$

are stationary and stable[18].

Without Moving Costs

Suppose the initial distribution of cultural traits at time t across the two neighborhoods is such that $q_1^i(t) > q_2^i(t)$ and there are no moving costs. Since optimum utility is increasing in fraction of own type, fraction dt parents with trait $i(j)$ migrate to $n_1(n_2)$ between t and $t + dt$ and fraction $1 - dt$ $i's(j's)$ will remain in $n_2(n_1)$. This is intuitive given that in a location with higher

[18]stability is formally proved in the appendix based on whether the eigenvalues of the jacobian matrix for the dynamic system is negative.

share of own type, the probability that a child picks up the trait through social interactions (e.g. from peers, neighbors) is higher. Parental effort is lower in these neighborhoods. This intuition is embodied in $(\tau_k^i)^*$ that is linearly decreasing in the fraction of preferred trait in society.

Then, at time $t + dt$, fraction of trait i will be even higher in n_1. Similarly, fraction of trait j will be higher in n_2 . The same migration pattern will repeat itself and $\lambda_{12}^i = 1$ and $\lambda_{21}^j = 1$ in every period. Eventually, n_1 would end up with all $i's$ and n_2 with all $j's$. At that point, no individuals will have an incentive to relocate. With a small disturbance, as long as $q_1^i > q_2^i$, we would end up with segregation again and points $(q_1^i, q_2^i = 1, 0)$ will be stable and stationary. Formal proof is included in appendix section 1.12.

Figure 1.1 illustrates the dynamics with migration and without migration. In neighborhood k that experiences in-migration (out migration) of individuals with trait i, $\frac{\partial q_k^i}{\partial t}$ will always be non negative (positive). It is also worth noting the distinct features compared to the dynamics without migration as in Bisin and Verdier(2001)- (1) Migration each period shifts the dynamics and equilibrium is reached faster than without migration (2) Homogeneity, albeit at a local level, is a stable equilibrium outcome.

With Moving Costs

As before, suppose there are costs to moving denoted by σ and $0 \leq \sigma \leq 1$. Equilibrium now depends on the initial distribution of trait i across the two neighborhoods relative to σ. There are two possibilities and I discuss them in detail below.

Case 1: Migration in each period→ segregation

Suppose

$$q_2^i(t) > q_1^i(t)$$

and

$$\Delta U^i_{12}[\,(q^i_1(t), q^i_2(t)\,] = U^i_2[\,w*, \tau*(q^i_2(t))\,] - U^i_1[\,w, \tau^*(q^i_1(t))\,] > \sigma$$

I ignore time notation for simplicity. If the cultural gains from moving to n_2 exceed the moving costs for parents i, they will migrate to n_2 and parents j to n_1 between t and $t + dt$. As was discussed in the previous section, if there is migration, $\frac{\partial q^i_2}{\partial t} \geq 0$ and $\frac{\partial q^i_1}{\partial t} \leq 0$. For that reason, at time $t + dt$, q^i_2 will be even higher and q^i_1 will be lower than before. Given that $\frac{\partial \Delta U^i_{21}}{\partial q^i_2} > 0$ and $\frac{\partial \Delta U^i_{21}}{\partial q^i_1} < 0$, overtime the difference in optimum utility will only increase and exceed σ by an even greater amount. The dynamics would work in the same way as the case with no migration friction and we would end up with homogenous neighborhoods in the long run. In other words, segregation will be a stable and stationary equilibria as long as $\Delta U^i_{21}[(q^i_1(t), q^i_2(t)] > \sigma$. At lower costs, this condition can be satisfied at lower values for ΔU^i_{21} or lower shares of trait i in the two locations, $q^i_2 - q^i_1$. More formally, $\frac{\partial (q^i_2 - q^i_1)}{\partial \sigma} \geq 0$. The dynamics is the same as in figure 1.1.

Case 2: No migration in any period → diversity

Suppose

$$q^i_2(t) > q^i_1(t)$$

and

$$\Delta U^i_{12}[\,(q^i_1(t), q^i_2(t)\,] = U^i_2[\,w*, \tau*(q^i_2(t))\,] - U^i_1[\,w, \tau^*(q^i_1(t))\,] < \sigma$$

Under these conditions, there would be no initial migration. Furthermore, given substitution, the growth rate of trait i in n_2 will be less than in n_1. The difference in growth rates of trait i between the second and the first neighborhood would only decline.Thus, at no point in the future will q^i_2 be so high relative to q^i_1 so as to make it optimal for parents to migrate to n_2 in the future. The two neighborhoods will evolve independently. The continuous time dynamics of the fraction of trait i in the two neighborhoods (separately) is given by[19]:

$$\frac{\partial q_k^i}{\partial t} = (q_k^i)(1 - q_k^i)(1 - 2q_k^i) \ where \ k\epsilon(n_1, n_2)$$

Figure 1.2 illustrates the dynamics that leads to heterogeneity in both neighborhoods in equilibrium. This is similar to the Bisin and Verdier (2001) world without migration.

The model shows how static optimization decisions of parents with socialization probabilities generate dynamics in the aggregate and impact the geographic distribution of cultural traits in the long run. The equilibrium analysis also suggests that (i) when moving costs are low or absent, a small increase in share of own type in (potential) destination neighborhood relative to the current location can incentivize migration and lead to segregation in the long run (ii) Diversity is sustainable under high moving costs. Note that here, even though migration probabilities in each period are assumed to be 0 and 1, equilibrium is not immediately reached. This is because of the assumption that only fraction dt individuals in each neighborhood are faced with a moving decision between t and $t + dt$. In appendix section 1.12, I relax the assumption of a $0, 1$ migration probability and that of a linear socialization technology. I derive continuous migration probabilities as shares of the preferred trait in the two neighborhoods. As before, I demonstrate how shares in current and (potential) destination neighborhoods incentivize parents to move and how these moves lead to segregation in equilibrium at neighborhood level and coexistence/heterogeneity of traits in the population as a whole.

Next, I examine if there is evidence in the data to support these socialization driven and segregation inducing sorting patterns.

[19]This is the dynamics in Bisin and Verdier (2001), but derived under the assumption that $f(\tau_k^i) = \tau_k^i$.

1.4 Data and Empirical Strategy, U.S Census Microdata

I empirically test whether parents with socialization preferences for their children sort into areas with higher share of their own ethnicity. Using individuals without children as a control group, I first examine how share of own ethnicity in a location impacts the moving behavior of parents relative to individuals without kids (non-parents).

To do so, I use Integrated Public Use Microdata, 2000. Individuals in 2000 were asked to report their address and that of the house they lived in 5 years ago. Their current and 1995 'Public Use Microdata Area ($PUMA$)' of residence is imputed accordingly. Based on this retrospective information, I know if the individual is a mover and if he moved across $PUMA$ (either across or within state) or within PUMA.

The raw sample data has a total of 16,900,110 individuals including adults and children. To the extent that members within each household make joint moving decisions, I keep a unique household representative. In most cases, I keep the individual who reports to be the household head. However, there are cases when the household head is not reported. For those households, I keep the adult male unless the female is the sole bread winner or if it is a single mother household. I drop households where individuals report to be living in institutions and households where reported migration status varies across members of the household. I also drop households with mixed ethnicities and few possibly erroneous cases where no member reports to be above 18. With these filters and keeping adults between 19-68 (as of 1995), my sample is restricted to about 18% of the original data.

I define share of ethnicity in a $PUMA$ based on the ancestry questions asked of both natives and immigrants. There are 172 different ancestries in total. However, several of these such as 'Hispanics' or 'African-Americans' for example can be interpreted as race. Other ancestries such as 'South American' for instance can be associated with various different countries of origin. Given the possibly vast heterogeneity within some of these ancestry groups, I keep only those that can be closely linked with a clear country of origin. I keep a total of 138 ancestries and in doing so lose an additional 5.5% of the original sample. The list of ancestries is included in the appendix section 1.10.1. I then calculate for each individual, the total population in his/her ancestry in 1995 as a share of the total population in 1995 in the $PUMA$ of residence that he/she either currently still lives in, moved out of or moved into[20]. I examine moving patterns based on shares as of 1995 to avoid simultaneity issues.

I keep both natives and immigrants who moved either across $PUMA$ or did not move at all. Note that shares are calculated at ancestry-$PUMA$ level. To avoid bias, I drop within $PUMA$ movers who may have moved to a census tract or block within the $PUMA$. For example, suppose a large fraction of non-parents moved from neighborhoods with low share to a neighborhood with higher share of own ethnicity within a $PUMA$. Then treating this group as non-movers will over estimate the probability of parents sorting into $PUMAs$ with higher share of their own ethnicity relative to non-parents.

I define parents based on age of eldest child. I first keep parents whose eldest child is 19 years or younger in 2000, individuals without kids and dropping those who have 0 weights in the data. I end up with a final sample of 1,385,007[21]. Summary statistics based on this sample and by mover status is provided in table 1.1. I estimate the following-

[20] Note that the ethnicity shares are calculated using all individuals who report their ethnicity and not just households.

[21] The raw sample is 1385203 but 196 observations have 0 weights. From here on, I will refer to parents as those with eldest child 19 years and younger because individuals move primarily between 1995 and 1999 and the children at the time of moving would be 18 years and younger. But estimates are quantitatively not very different and qualitatively robust to keeping 18 years as the age cut off, as of 2000.

$$y_{i,e,p} = \alpha + \sum_{g=1}^{6} \beta_{1g} Parent_i \text{x} S_{e,p} + \sum_{g=1}^{6} \beta_{2g} Parent_{i,g} + \beta_3 S_{e,p}$$
$$+ \text{X}'_{i,e,p} + K_e + Z_p + \epsilon_{i,e,p} \quad [4.1]$$

where $y_{i,e,p}$ is a dummy variable equal to 1 if individual i with ancestry/ethnicity e moved out of $PUMA$, p between 1995 and 2000. I define 6 groups that I denote as g based on the age of eldest child. $g = 1$ if age of eldest child is between $0 - 2$, $g = 2$ if age of eldest child is between 3-5 up to $g = 6$ for age of eldest between 15-17. $\text{S}_{e,p}$ is the fraction of individual's own ethnicity in $PUMA$ of residence in 1995. $Parent_{i,g}$ is a dummy that takes the value of 1 if individual reports to have kids in 2000 and if the age of the eldest child is in the age category defined by group g. Since individuals could have moved in anticipation, parents also include those individuals who did not have kids in 1995 but did so in 2000. As a robustness check, I also estimate the above regression after dropping those who may have moved in anticipation and keeping parents whose eldest child was between 5-19 years of age as of 2000. If parents have cultural transmission preferences and in other ways are similar to non-parents, in $PUMAs$ with higher (lower) share of own ethnicity, parents should be much less (more) likely to move out than non-parents. We should expect β_{1g} to be negative.

To incorporate destination choice, I keep movers only and estimate whether parents are more likely than non-parents to move from a $PUMA$ with lower share of own ethnicity to a $PUMA$ with a higher share of own ethnicity using the following:

$$\text{y}_{\text{i,e,p}} = \alpha + \sum_{\text{g}=1}^{6} \beta_{\text{g}} \text{Parent}_{\text{i,g}} + \text{X}'_{\text{i,e,p}} + \text{K}_{\text{e}} + \text{Z}_{\text{p}} + \epsilon_{\text{i,e,p.}} \quad [4.2]$$

I calculate the difference in own ethnicity shares (in 1995) between destination and origin $PUMA's$. $y_{i,e,p}$ is a dummy that equals 1 if the difference is positive and 0 otherwise. If the effects are explained by cultural transmission, I would expect β_g to be positive and for parents to show a higher likelihood of sorting into location with ethnically similar neighbors relative to non-parents.

Note that in the sample of natives and immigrants, those reporting German ancestry constitute almost 20% of the sample and Mexicans constitute about 5%. Among immigrants, those with Mexican ancestry constitute approximately 20%. As a robustness check, I estimate equations 4.1 and 4.2 with natives and immigrants but after dropping those with German and Mexican ancestry. I also look at estimates with immigrants only, both with and without Mexicans. Additionally, I compare parents not just with non-parents but also with parents whose kids are above 19 years of age and possibly past the age of socialization.

One concern is that parents and non-parents are dissimilar in various ways and the difference in self-segregating choices between the two groups could be due to reasons other than cultural transmission for kids. This could be an issue in particular for individuals with older kids. These parents might have already chosen their residence based on kids in 1995 and a lot could have changed from the time the child was born and to the time of migration. To address this issue, I check if the estimates in equations 4.1 and 4.2 hold for parents who moved in anticipation and those with very young children who would have moved immediately after children. I further check whether the estimates systematically change with age of eldest child.

A related threat to identification based on endogeneity of children stems from unobservables that are correlated with fertility and change of residence. For instance, positive income shocks may incentivize parents to have kids and also choose areas with better school districts or suburban homes with higher acreage. I include various individual level controls such as age, sex, education, household income, individual wage, employment, distance and time to work, english proficiency as well as rent paid, home ownership and mortgage status and a full set of ethnicity dummies (K_e). The vector of individual controls is given by $X'_{i,e,p}$. Also, given the interaction with own ethnic shares, this does not pose a challenge unless other home and $PUMA$ amenities are systematically correlated with shares. To address the potential endogeneity in shares, I include $PUMA$ fixed effects, Z_p. Using non-parents as a control group will difference out unobservable amenities that parents and non-parents value equally.

However, there might be certain unobservables that are correlated with share of own ethnicity that parents and non-parents value differently. For example, schools in ethnically similar neighborhoods might offer a particular curriculum that parents in that ethnic community value more. I then also compare parents with eldest child in age 3-5, 6-8 and 9-11 to parents whose children at the time of moving were between 12-18. To the extent that parents with kids between 6-8 years will care as much about the curriculum as those with kids between say 9-11 years, the bias of this nature will get differenced out. As an additional check, I also use an instrumental variable strategy that exploits plausibly exogenous variation in own ethnic composition at city level. I find the results to be qualitatively similar to before. The strategy is discussed in greater detail in section 1.5.

Finally, I also use other measures such as individual's stated preferences provided in the U.K data to show that neighbor preferences of own type among parents is attributable to cultural socialization.

1.4.1 Migration and Own Ethnicity Share, $PUMA$ Level

Likelihood of Leaving a $PUMA$, Comparing Parents and Non-Parents

Figure 1.3, left panel shows the difference between the average number of parents who leave a particular $PUMA$ and the average number of non-parents who leave that $PUMA$ between 1995-2000. I plot this difference on the y-axis and the standardized share of own ancestry in that $PUMA$ on the x-axis. The figure shows that for lower shares of own ethnicity, average number of parents who out-migrate from that $PUMA$ exceeds the average number of non-parents. This difference declines with increase in share of own ethnicity. While the plot is based on raw data, the general differences hold true in a regression framework after controlling for various individual and $PUMA$ level controls and fixed effects that were described in the previous section.

The estimation results are included in Table 1.2. Column 1 is based on parents whose eldest child is 19 years of age or lower. I find that a one standard deviation increase in share of own ethnicity reduces the probability of leaving that $PUMA$ by .808 percentage points (7.6%) for parents compared to non-parents.

I also estimate this equation by age of eldest child. The estimates are included in columns 2-7 of table 2 and shown in figure 1.4. I find a systematic decline in the effects of own ethnicity with the age of eldest child. For parents whose eldest child is between 3-5 years, a one standard deviation increase in share of own ethnicity reduces their probability of leaving the $PUMA$ by 2.28 percentage points compared to non-parents. The estimates decline to 1.58 percentage points for eldest child between 6-8, to 1.09 percentage point for 9-11 and less than 1 percentage point for ages 12 and over.

I then compare within parents and find that parents with younger kids are also much less likely to migrate out of ethnically similar $PUMA's$ compared to parents whose kids at the time of moving are between $12 - 18$ and presumably past the age of socialization. The effects as before are strongest for children below age 5. Relative to parents with older kids, a one standard deviation increase in share of own ethnicity reduces the probability of leaving that $PUMA$ by 1.87 percentage points for parents with eldest child between 3-5, to 1.21 percentage points if eldest child is between ages 6-8 and to .847 percentage points for the 9-11 age group. These estimates are shown in Panel A of table 1.4.

Table 1.5 shows how a one standard deviation increase in share of own ancestry changes the out migration probability of parents with eldest child below a certain age relative to households with kids above 18 at the time of migration. As before, the effect is negative and the out migration probability declines from about 2.35 percentage points for eldest child between 3-5 to 1.65 percentage points for ages 6-8, 1.116 percentage points for ages 9-11 to less than 1 percentage points for age of eldest child over 12.

Likelihood of Moving into (destination) $PUMA$ with Higher Own Ethnicity Share relative to Origin, Comparing Parents and Non-Parents

So far, I have only looked at the probability of leaving a $PUMA$ between 1995 and 2000. However, it might be the case that parents moving out of $PUMAs$ with lower share of own type are moving into even more diverse places relative to non-parents. To see if that is the case, I incorporate destination choice. Conditional on moving, I look at whether parents are more likely to move from a $PUMA$ of lower share of own ethnicity into a $PUMA$ of higher share compared to non-parents. A plot of these probabilities for parents and non-parents is shown in the right panel of figure 1.3.

The raw data analysis suggests that on average among movers, parents are about 2.7 percentage points (5.5%) more likely than non-parents to move into more ethnically similar areas relative to the $PUMAs$ they move from. The difference in these probabilities is statistically significant at 95% confidence level with a t-value of 12.13 approximately.

Equation 4.2 estimates these differences and the results are included in table 1.3. Estimates suggest that the probability of choosing more ethnically similar $PUMAs$ is higher for parents than non-parents by a statistically significant margin. The probability for parents is 1.76 percentage points (3.6%) higher than for non-parents. The estimates are stronger for children of younger ages. For age of eldest child in the range of 3-5 years or 6-8 years, parents are almost 2.7 percentage points (5.6%) more likely to choose destinations with higher share of own type relative to non-parents. For ages 9-11, the effect reduces to about 1.95 percentage points (4.05%) and to about 1.5 percentage points (3.6%) for ages 15-17 years. The results are robust to including parents with kids older than 18 or 19 years in the non-parent sample as well as to keeping peak parenting age adults between 21-45 years.

In so far as neighbors are viewed as socialization agents, the decline in estimates with age are consistent with (previously provided) evidence in the literature that finds early formative years to be important for ethnic and religious transmission. It is also consistent with a general investment story that parents who invest in their children by making costly decisions such as moving will do so earlier in the child's lifecycle to maximize returns.

The results however are not significant when I compare parents with younger kids to parents whose kids at the time of moving are between $12 - 18$. These results are shown in Panel B of table 1.4. This may be driven by the significant reduction in sample size, due to which several ancestry groups end up with very few individuals. But as before, the estimates are robust when I compare parents with eldest child below a certain age to households with kids above 18 at the time of migration. The estimates are show in table 1.6 and suggest that parents with eldest child in the age group 3-5 and 6-8 are about 2.9 percentage points and 2.3 percentage points (respectively) more likely to move from ethnically less to ethnically more similar $PUMAs$. As before, the estimates are not significant for older ages.

1.4.2 Immigrant Migration and Own Ethnicity Share, MSA Level

One endogeneity concern with the previous analysis is that $PUMA$ shares may be correlated with other amenities that parents and particularly parents with young kids value differently from non-parents and from parents with older kids. As an additional check, I exploit plausibly exogenous variation in ethnic composition at the level of 'Metropolitan Statistical Areas (MSA)'[22]. In the

[22] $MSAs$ are Census designated statistical areas with a core urban population of at least

next set of sections, I analyze migration flows of parents and non-parents across $MSAs$ using both OLS and an Instrumental Variable (I.V) strategy.

Likelihood of Leaving an MSA, Comparing Parents and Non-Parents

I use U.S Census 1990 to evaluate how the moving behavior of immigrants who moved before 1980 responds to new immigrant inflows from their own country of birth. The number of immigrant movers are more substantial in the 1990 Census relative to the 2000 Census and for that reason, I switch to 1990 for this part of the analysis. Specifically, I first provide OLS estimates of the probability that a pre-1985 immigrant parent (relative to non-parent) left his MSA between 1985-1990, based on new immigrant inflows into that MSA from his own country of birth. I define new immigrants as those who moved after 1985. I use the following specification:

$$y_{i,c,m} = \alpha + \sum_{g=1}^{g=6} \beta_{1g} Parent_{i,g} \text{x} S_{c,m} + \sum_{g=1}^{g=6} \beta_{2g} Parent_i + \beta_3 S_{c,m} + \text{X}'_{i,c,m} + K_c + Z_m + \epsilon_{i,c,m}. \quad [4.3]$$

$y_{i,c,m}$ is a dummy = 1 if the pre-1985 immigrant from country c who reports to be living in MSA, m moved out of his/her MSA of residence between 1985-1990. $S_{c,m}$ is the share of immigrants from c who moved into m between 1985-90. K_c and Z_m are country of birth and MSA dummies. $X'_{i,c,m}$is a vector of individual controls including age, sex, education, income, employment, distance and time to work, english proficiency as well as rent paid, home ownership and mortgage status, age of immigrant entry and years since arrival.

OLS estimates from 4.3 are included in table 1.8. Similar to $PUMA$ level regressions, I find that relative to non-parents, a 1 standard deviation increase in share of immigrants from own country of birth in MSA reduces the likelihood of leaving that MSA by 1.43 percentage points for parents with eldest child in the age group 0-2. Similar estimates for parents with kids between ages

50,000. MSA typically includes a city and greater-city area.

3-5 and 6-8 are .740 percentage points and 1.06 percentage points respectively. For parents with older kids, above age 8 that is, β_{1g} actually becomes positive suggesting that for later ages, parental investments matter less.

But as was previously discussed, $S_{c,m}$ could be endogenous. I use an instrumental variable (I.V) strategy based on the widely used shift share instrument (Altonji and Card (1991); Card 2001))[23] to proxy for the actual share of immigrants from individual's own country of birth with predicted share. The instrument exploits the idea that new immigrants tend to cluster in cities with previous immigrants from their own country of birth. The validity of this instrument hinges on the assumption that predicted shares are independent of other area unobservables. However, and this is a common criticism of the instrument[24], if some local conditions due to which previous immigrants settled in a certain city persist over time, then new immigrant inflows into that city may not be exogenous to those conditions. In such a case, the exclusion restriction would be invalid and instrumenting for $S_{c,m}$ in 4.3 will not address the discussed endogeneity.

To get around this issue, instead of regressing the probability of leaving an MSA on immigrant shares, I estimate the *difference* in moves based on immigrant shares between parents and non-parents. In particular, I look at difference in net-inflows between pre-1985 immigrant parents and non-parents. The identifying assumption is valid in so far as those time-persistent conditions don't vary with parent status. Next, I discuss OLS and I.V estimates from the regression of net-inflows on immigrant shares.

Net-Inflows of Existing Immigrant and Natives, Comparing Parents and Non-Parents

Let $(\Delta Par_{c,m})_{85-90}$ be the change in the total population of existing immigrant parents from c in m between time period 1985 and 1990. This change is

[23] Cascio and Lewis (2012) use the instrument to study the impact of increased hispanic enrollment on the residential outflows of natives.

[24] see Jaeger et al. (2018)

driven by either the net inflows of parents or of existing non-parents who have children and become parents.

$$(\Delta Par_{c,m})_{85-90} = \underbrace{(\Delta M_{c,m})_{85-90}}_{net\, migrant\, inflows\, of parents} + \underbrace{(\Delta N_{c,m})_{85-90}}_{recently\, become\, parents}$$

Corresponding change in the population of non-parents can be written as:

$$(\Delta NonPar_{c,m})_{85-90} = \underbrace{(\Delta M^{'}_{c,m})_{85-90}}_{net\, migrant\, inflows\, of\ non-parents} - \underbrace{(\Delta N^{'}_{c,m})_{85-90}}_{recently\, become\, parents}$$

Then the difference in the growth rate of the parent population versus the non-parent population that is driven by migration can be written as:

$$\frac{(\Delta M_{c,m})_{85-90}}{(M_{c,m}+N_{c,m})_{85}} - \frac{(\Delta M^{'}{}_{c,m})_{85-90}}{(M^{'}_{c,m}+N^{'}_{c,m})_{85}}$$

where the two denominators refer to the total parent and non-parent population of existing immigrants from c in m.

I first use the following OLS specification to analyze the net inflows of pre 1985 immigrants to new immigrant inflows:

$$\frac{(\Delta M_{c,m})_{85-90}}{(M_{c,m}+N_{c,m})_{85}} - \frac{(\Delta M^{'}{}_{c,m})_{85-90}}{(M^{'}_{c,m}+N^{'}_{c,m})_{85}} = \alpha + \beta_1.(S_{c,m}) + X^{\prime}{}_{c,m}\beta + Z_m + K_c + \varepsilon_{c,m} \quad [4.4]$$

β_1 shows how the percentage change in net inflows of parents versus non-parents responds to a one standard deviation increase in the share of immigrants from own country of birth at MSA level. OLS estimates from this regression are shown in table 1.9. I find that a 1 standard deviation increase in the share of new immigrants from own country of birth leads to a 3.7% increase in net inflow of parents with age of eldest child between 0-2 in that MSA relative to non-parents. For ages 3-5, the estimate is not significant but for ages 6-8 and

9-11, the increase is 2.5% and 2.9% respectively. For higher ages, the estimates are not statistically significant.

Instrumental Variable Strategy

I use an I.V strategy to instrument for $S_{c,m}$ in 4.4. Let $\gamma_{c,us}$ be immigrants from country c who moved into the U.S before 1985 and $\gamma_{c,m}$ be the total number of those pre-1985 immigrants who moved from country c to MSA, m. Suppose $\zeta_{c,u.s}$ denotes new immigrant inflows in to the U.S from c between 1985-1990 and $P_{m,85}$ is the total MSA population as of 1985. Then, share of predicted inflow of immigrants from country c and MSA m between 1985-1990, λ_{cm} is given by :

$$\lambda_{cm} = \frac{1}{P_{m,85}} \frac{\gamma_{c,m}}{\gamma_{c,us}} \zeta_{c,u.s}$$

The general intuition is that if x% of earlier immigrants move to a particular location, x% of future immigrants from the same country can be predicted to move to that location. I use $\lambda_{c,m}$ to instrument for $S_{c,m}$ in 4.4 and estimate the following equations:

$$\frac{(\Delta M_{c,m})_{85-90}}{(M_{c,m}+N_{c,m})_{85}} - \frac{(\Delta M'_{c,m})_{85-90}}{(M'_{c,m}+N'_{c,m})_{85}} = \alpha + \beta_1.(S_{c,m}) + X'_{c,m}\beta + Z_m + K_c + \varepsilon_{c,m}$$

$$\mathrm{S_{c,m}} = \alpha_{\mathrm{fs}} + \beta_{\mathrm{fs}}\lambda_{\mathrm{c,m}} + +\gamma_{\mathrm{fs}}\mathrm{X}'_{\mathrm{i,c,m}} + \mathrm{K_c} + \mathrm{Z_m} + \eta_{\mathrm{i,c,m.}} \quad [4.5]$$

Consistent with the model intuition, I find that $MSAs$ with higher predicted immigrant shares also witness a higher net inflow of immigrant parents from the same country relative to non-parents. The estimates are statistically significant in line with previous predictions, only for parents with younger children. Estimates for parents with age of eldest child between 0-5 are shown in panel A table 1.10. Growth rate of the parent population driven by migration in m increases by 3.66% more than that of non-parents to a 1 standard increase in predicted share of own country immigrants. These estimates are driven primarily by parents with eldest child between 3-5 years of age. For these parents, the effect is 4.6% larger compared to non-parents.

Finally, I also look at native response to new immigrant inflows. Specifically, I keep U.S born individuals who report ancestry of a country c and examine their likelihood of (net) migration into $MSAs$ with higher immigrant inflows from c. For example, if a U.S born individual reports Mexican ancestry, I study her migration response to new immigrants inflows from Mexico.

Given that these U.S born individuals could be presumably fully assimilated and have no cultural socialization preferences, it is unclear at the outset what the migration response might be. Interestingly, even for this group I find a positive and statistically significant effect. Once again these effects only hold for younger ages. I keep parents with eldest child age between 0-5 as before and find an estimate of 1.22%. The result is shown in panel B of table 1.10.

1.5 U.K Fourth National Survey of Ethnic Minorities (1993-1994)

1.5.1 Data

In this section, I analyze the predictions of the model using the U.K Fourth National Survey of Ethnic Minorities (U.K FNSEM), 1993-1994. U.K FNSEM is a 1 year survey data of individuals 16 and over, covering England, Wales and Scotland. The survey dataset has two independent samples. One is a sample of U.K born individuals of British origin. Second is the ethnic sample with 6 main ethnicities- (1) Indian (2) Caribbean (3) African Asian (4) Pakistani (6) Bangladeshi (7) Chinese. Different questions are asked of the two samples and it is not possible to study both jointly. I focus on the ethnic sample.

This data set has two important advantages. First is that the data is available at a highly disaggregate level of geography, '*wards*'. *Wards* are equivalent to census tracts in the U.S with a population of about 3000-4000. Fraction of individuals in the *ward* with a particular ethnicity is categorized into the following - (1) $\leq$ 2% (2) $(2\%, 5\%]$ (3)$(5\%, 10\%]$ (4) $(10\%, 15\%]$ (5) $(15\%, 25\%]$ (6)$(25\%, 33\%]$ (7) 33% $\geq$. Similar to Bisin et al. (2016), I define share of own ethnicity based on the average value of the interval. For values $\geq$ 33% and $\leq$ 2% , the share is equal to 33% and 2% respectively.

The second advantage is that individuals are surveyed on various questions that can be used to proxy for cultural preferences in general and for stated cultural socialization preferences, specifically. The raw data after keeping adults between 19-68 has 3035 observations. However, there are several missing values and the sample reduces significantly based on the particular outcome of interest in the regression. Summary statistics of the sample is included in table 1.11.

1.5.2 Ward Social Composition and Socialization Preferences

Are parents with stronger socialization preferences more likely to settle in *wards* with higher ethnic density of their own type ? Two questions in the survey can be used to gauge socialization (a) When choosing a school for your 11 year old child, would your choice be influenced by the number of people from your ethnic origin ? If yes, then how important would the influence be ? The options to this question are very important, fairly important, not very important, no influence and cant say. (b) If schools were similar in other ways, would you prefer that your 11 year old child go to a school with fewer than half, half or more than half from your ethnic origin? Other options to this question include no preference and cant say.

I drop those who report cant say on either of the two questions and specify socialization preference as a dummy in 2 ways- (a) Dummy that takes the value

1 when (own) ethnic composition of the school is a fairly or very important factor in the choice of school for the 11 year old child. The dummy is 0 if the choice is not very, not at all important or if the individual has no preference. (*b*) Dummy equal to 1 if individuals would prefer to send their child to a school with at least half of the peers from their own ethnic origin.

Regression adjusted correlations between socialization preferences defined in the 2 ways and (mean) *ward* ethnic density of own type is shows in figure 1.6.[25]. The probability that individual i states ethnic socialization preference is positively correlated with with *ward* share of own ethnicity.

1.5.3 Likelihood of reporting a Preference to Leave Current Ward, Comparing Traditional Parents and Non-Parents.

Using U.K FNSEM (1993-1994), I also analyze the willingness to move of parents based on the ethnic density of the *ward* they currently reside in. Individuals are asked if given a choice they would like to move or continue to reside in the current neighborhood. Figure 1.7 shows the mean difference in willingness to move by *ward* ethnic density between non-parents and parents. Parents are defined by whether they have kids between 0-4, 5-11 or 12-15. The figure suggests that *ward* satisfaction increases with increase in ward ethnic density of own type and more so for parents. The pattern is robust to keeping children of different ages.

Next, I estimate the extent to which this pattern is attributable to social-

[25]I regress the socialization dummy on *ward* share of own ethnicity and include various controls for age, sex, income, employment, education, english proficiency, years since arrival, area of residence and homeownership status, if individual has been discriminated and ward level controls such as ward unemployment and ward social housing density. The complete list of controls is included in the summary statistics table 8.

ization preferences using the following specification:

$$\begin{aligned} y_{i,w} = \alpha + \beta_1(Parent_i \text{x} Traditional_i \text{x} S_w) + \beta_2(Parent_i \text{x} Traditional_i) \\ + \beta_3(Traditional_i \text{x} S_w) + \beta_4(Parent_i \text{x} S_w) + \sigma_1 Parent_i \\ + \sigma_2 Traditional_i + \sigma_3 S_w + X'_{i,w}\eta + \mu_{i,a,w} \quad [5.1] \end{aligned}$$

$y_{i,w}$ is a dummy that takes the value 1 if individual i in ward w expresses a willingness to move out of *ward w*. I define traditional as the group with higher average number of traditional responses to specific survey questions. Accordingly, the three groups that I call traditional are Muslims, those born in country dominant in their reported ethnicity (e.g. 1 if born in China if reported ethnicity is Chinese and 0 if born in either U.K or in any country other than China) and those among married couples who share the same ethnicity. The questions and the responses that I use to define traditional are included in appendix section 1.11. The identifying assumption is that traditional parents are likely to have higher cultural socialization preferences than less traditional or more assimilated immigrant parents and that any difference in the willingness to move between parents and non-parents that is driven by non cultural reasons should be the same across traditional and non traditional groups.

As shown in table 1.12, a one standard deviation increase in *ward* share of own ethnicity reduces the preference to leave the current *ward* more for traditional parents compared to traditional non-parents and this effect falls further for the non traditional group. The results are robust to Muslims and married couples with the same ethnicity. For those born in country dominant in their reported ethnicity, the estimates are significant only when I keep peak parenting age individuals between 21-45.

The fact that individuals are willing to move to *wards* different from the ones they currently reside is consistent with the intuition that there could be costs to moving that constrain individuals in less culturally optimum neighborhoods.

1.6 Robustness Checks

$PUMA$ Level Regressions

Recall that the $PUMA$ estimates shown in tables 1.2 and 1.3 are based on a sample of natives and immigrants. Also, the parent sample in these estimates is based on those who did not have children in 1995 but did so in 2000 and therefore could have moved in anticipation.

As a robustness check, I estimate equations 4.1 and 4.2 after dropping those who moved in anticipation and only keeping individuals who were already parents in 1995. Table 1.13, column 1 provides estimates for the probability of leaving a $PUMA$, comparing parents with kids between 5 and 19 as of 2000 and non-parents. The estimates are qualitatively similar and close in magnitude to estimates for parents with kids between 0-19. A one standard deviation increase in own ethnicity share reduces the likelihood of leaving that $PUMA$ by .7 percentage points for parents with kids between 5-19. In column 2, I compare parents with both non-parents and parents with kids above 19. The estimates are similar in both magnitude and sign to previous estimates of .8.

I also look at the probability of moving from a $PUMA$ of lower share of own ethnicity to a $PUMA$ with higher share of own ethnicity for parents with kids between 5-19. As shown by column 1 of table 1.14, these parents are 1.55 percentage points more likely to sort into destinations with similar ethnicity shares relative to non-parents and 1.53 percentage points more likely relative to both non-parents and parents with kids above age 19. The estimate was 1.76 percentage points when parents with kids between 0-19 and non-parents were compared.

Next, I keep immigrants and evaluate the likelihood of leaving a $PUMA$ (equation 4.1) for parents relative to non-parents. These estimates are shown in table 1.15. As before, an increase in share of own ethnicity of origin $PUMA$ reduces the likelihood of leaving that $PUMA$ more for parents than non-parents.

I also find a systematic decline in the effects with increase in age of eldest child. The decline is shown in figure 1.9. The re-estimation of equation 4.2 for the immigrant sample, however, yields statistically insignificant results. In large part, this stems from the decline in the sample size. Note that equation 4.2 that looks at probability of moving from a less to a more ethnically similar $PUMA$ is based on movers only. When I keep immigrants movers and drop within $PUMA$ movers, I end up with a sample of 15,897 individuals. The sample drops even more when I add age and ancestry restrictions similar to the ones on which 4.2 is estimated. The same size and associated (insignificant) estimates from this regression is shown in table 1.16. In the immigrant sample, Mexicans constitute around 23% of the whole sample. I find the results for out-migration to be robust to keeping immigrants and dropping Mexicans. The estimates for the sample without Mexicans is shown in table 1.17. But when I keep movers only and re-estimate equation 4.2 for immigrants without Mexicans, the sample size drops to less than 10,000 individuals and the estimates are not significant.

Note that in the sample with both natives and immigrants, Germans constitute 20% of the sample and Mexican share is around 5%. I find the results to be robust when I re-estimate equations 4.1 and 4.2 for both natives and immigrants but drop those who report German and Mexican ancestry. The estimates are included in table 1.18 and 1.19. Moving patterns are consistent with previous estimates and suggest that even after dropping Germans and Mexicans, parents are more likely to sort with ethnically similar neighbors relative to non-parents.

I do a final robustness check for $PUMA$ level regressions. Figure 1.8, panel A shows the density of $PUMA$ level, share of own ethnicity in the sample. The distribution is disproportionately skewed to the left. I drop extreme values for share of own ethnicity and restrict the share to be between .1 and .4. The distribution after dropping these outliers is shown in Panel B of figure 1.8.

I then estimate the probability of out migration and of moving from a lower to a higher own ethnicity share $PUMA$ for parents and non-parents. The results are robust and included in tables 1.20, panel A and B.

Likelihood of Leaving an MSA and Age at Immigration, Comparing Parents and Non-Parents

Here I exploit age at arrival of immigrants but estimate the impact of predicted shares of individuals from own country of birth on the moving behavior of existing immigrant individuals based on both their age at arrival to the U.S and parental status. Using U.S Census 1990, I estimate the following:

$$
\begin{aligned}
\mathrm{y_{i,c,m}} = \alpha + \beta_1(\mathrm{Parent_i x \lambda_{c,m} x D_a}) + \beta_2(\lambda_{c,m} \mathrm{x} D_a) + \beta_3(Parent_i \mathrm{x} D_a) \\
+ \beta_4(\lambda_{c,m} \mathrm{x} Parent_i) + \sigma_1 A_{c,m} + \sigma_2 Parent_i \\
+ \sigma_3 D_a + \mathrm{X}'_{i,c,m} + K_c + Z_m + \epsilon_{i,c,m.} \quad [6.1]
\end{aligned}
$$

where D_a refers to age at arrival. D_a is a dummy variable equal to 1 if individuals arrive at age a or after and 0 otherwise. The underlying assumption is that non-cultural differences in migration probabilities of parents and non-parents based on predicted shares should be the same among those who arrive before age a and after.

Similar to previous literature (Aslund (2015); Chin et al. (2003, 2007)), I keep existing immigrants who moved to the U.S when they were children. For childhood immigrants, timing and selection into immigration is exogenous (assuming they had no choice as children in this decision).

There is no defined cut off for D_a and typically childhood immigrants are defined as those who moved before age 18. In my specification however, I find the estimates to be significant when I keep the cut off of 9 or 10 and when I keep even younger childhood immigrants, those who moved before age 16.

Table 1.21 shows estimates for arrival age cut off 9 but results are robust to using cut off 10 and to the inclusion of parents with eldest child over 19 years. β_1 is negative and statistically significant. Estimates suggests that a one standard deviation increase in predicted share of immigrants from own country of birth reduces the likelihood of leaving that MSA for parents relative to non-parents by 1.56 percentage points more for late arrivers relative to early arrivers. This finding is consistent with the intuition that if socialization preferences are driving moving behaviors and neighbor preferences for parents, the effects will be stronger for those with presumably stronger home country attachments.

Ethnic Neighbor Preferences and Age at Immigration, Comparing Parents and Non-Parents

I also estimate the impact of arrival age on stated ethnic neighbor preferences for parents and non-parents using U.K FNSEM (1993-1994). Individuals are asked if they would prefer to live in a neighborhood with fewer than half of the neighbors from own ethnic origin, half, more than half or if they have no preference.

I keep individuals who moved when they were 16 or younger and presumably had no choice in the timing of migration. I then estimate the following:

$$y_{i,w} = \alpha + \beta_1(Parent_i \text{x} ArrAge_i) + \beta_2 Parent_i + \beta_3 ArrAge_i + X'_{i,w}\eta + \mu_{i,w} \quad [6.2]$$

$y_{i,w}$ is a dummy that takes the value 1 if individual i in $ward$ w who immigrated to the U.K at age a, states that she would prefer to live with approximately half or more than half of the neighbors from her own ethnic origin. $ArrAge_i$ is the standardized value of arrival age to the U.K. One concern is that other omitted variables related to age at arrival may affect selection into parenthood as well as neighbor preferences. For instance, late arrivers may be more english deficient and also more likely to have children. If english deficiency also affects neighbor preferences, estimates of β_1 will be attributable to language constraints as opposed to cultural transmission preferences. To account for that, I also include various individual and contextual controls such as age, sex, income, employment, education, english proficiency, years since arrival, homeownership status, if individual has been discriminated and ward level controls such as ward unemployment, ward social housing density and ward share of all immigrants. These controls are summarized by the vector $X'_{i,w}$ and included in the summary statistics table 1.11.

Estimates from 6.2 are shown in table 1.22 and suggest that a 1 standard deviation increase in age at arrival increases preference of parents with children ≤ 4 for own type neighbors by 11.9 percentage points compared to non-parents. Estimate for parents with kids between 5-11 is close to 13.1 percentage points. Similar to all the previous findings, the effects are strongest for younger kids. Estimates are not significant for children in age group 12-15[26]. When I compare parents to both non-parents and all other parents, estimates are significant at 95% confidence level, but for ages 5-11 only . For this group, 1 standard deviation increase in age at arrival increases similar neighbor preferences relative to non-parents and other parents by 10.6 percentage points.

Interestingly, the effect of a 1 standard deviation increase in arrival age for non-parents is negative. A plot of ethnic neighbor preferences by arrival age for parents and non-parents separately suggests that the positive β_1 is driven more by the decline in age at arrival effects on neighbor preferences for non-parents rather than an increase in those effects for parents.

1.7 Conclusion

This paper provides a micro-foundational framework to understand the relationship between intergenerational ethnic preferences and ethnic segregation. The underlying intuition is that driven by these preferences, parents may make self-segregating choices. Empirical results suggest that children tend to intensify ethnic neighbor preferences and increase the likelihood of parents sorting

[26] In the data, I only have information on whether the individual is a parent and the age of the child is defined in the 3 categories : 0-4; 5-11; 12-15.

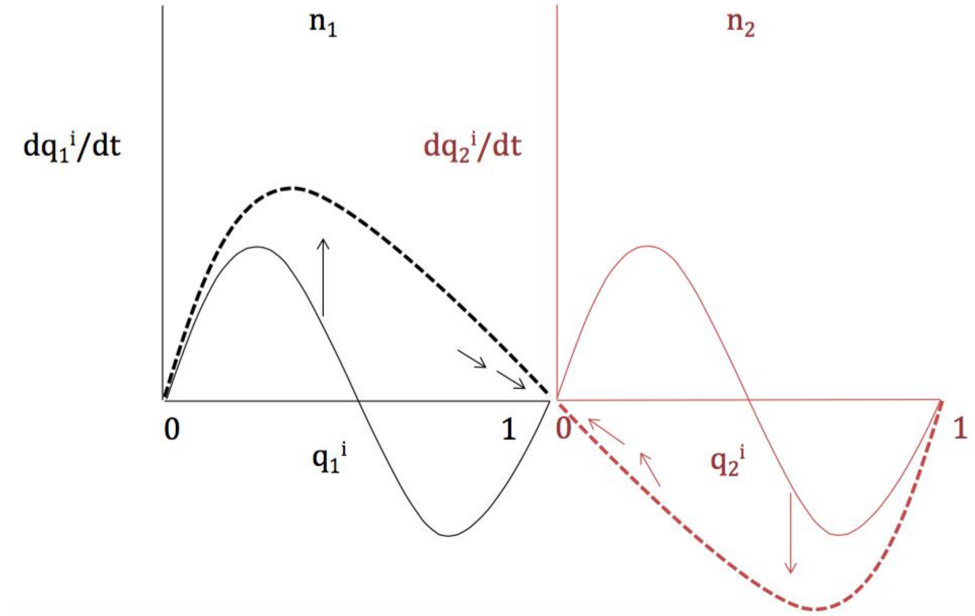
with others of own ethnicity. The novel finding is that this likelihood is higher among parents relative to those without children and even more so among parents with young children and less assimilated immigrant parents. Preference for cultural transmission appears to be an important causal mechanism to explain these results.

The importance of these findings stems from the fact that it can lead to the spread and persistence in ethnic segregation. Ethnic enclaves can certainly form for various reasons. However, if their formation is driven by socialization incentives, it will disproportionately affect families with young children and can create intergenerational dependencies in outcomes. For instance, large number of immigrant children living in segregated communities of Washington D.C are linguistically isolated and consequently faced with economic and academic challenges. This points to intergenerational transmission of language as one of the potential channels through which exposure of young children to ethnic enclaves can matter.

This paper suggests that if moving costs are high, for instance due to high commute costs or high real estate prices on the one hand or if parents have easy access to cultural amenities that they value for their children, incentives to self-segregate with similar neighbors would be lower and doing so would be costlier. Policies that affect such trade-offs can have a large impact on the demographic distribution of the population in the long run and given the social multiplier effect, impact the economic mobility of both minority and non-minority children.

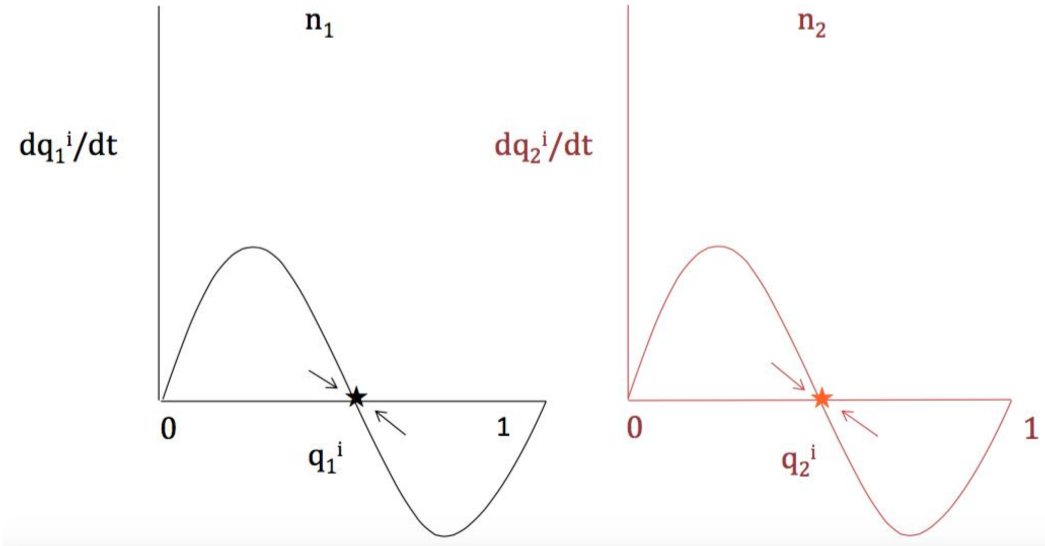
1.8 Figures

Figure 1.1: Evolution of Cultural Traits in Neighborhoods 1 and 2 with Migration



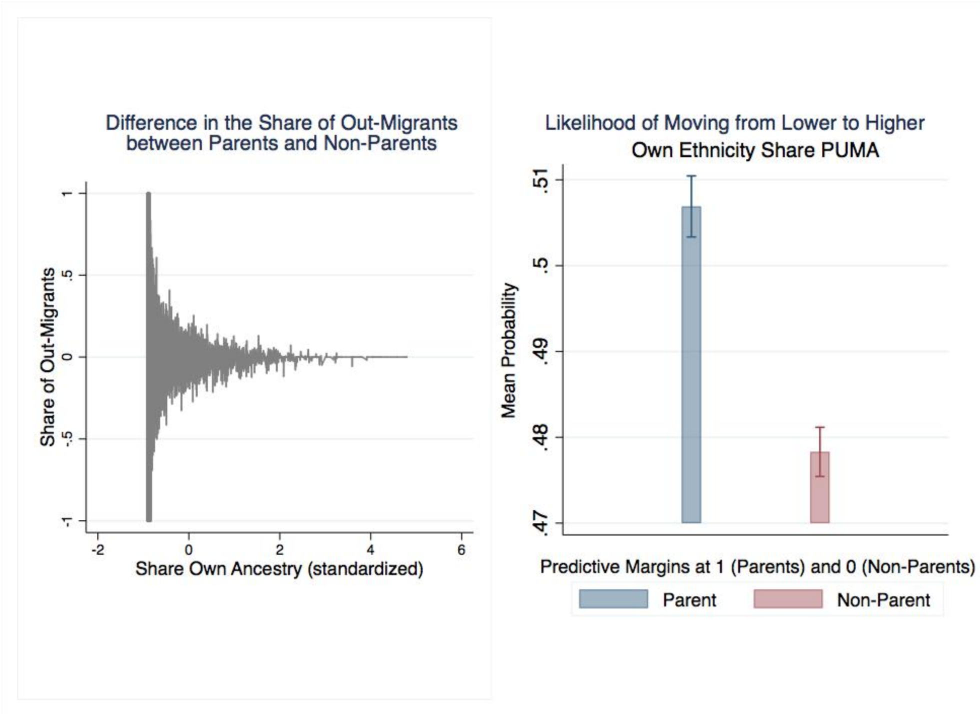
Notes: The solid line indicates the dynamic evolution of cultural trait i in the two neighborhoods, n_1 and n_2, without migration. The dashed line indicates the evolution of cultural trait i under migration from n_1 to n_2. Under low or moving costs, migration leads to segregation of cultural traits across the two neighborhoods.

Figure 1.2: Evolution of Cultural Traits in Neighborhoods 1 and 2 without Migration



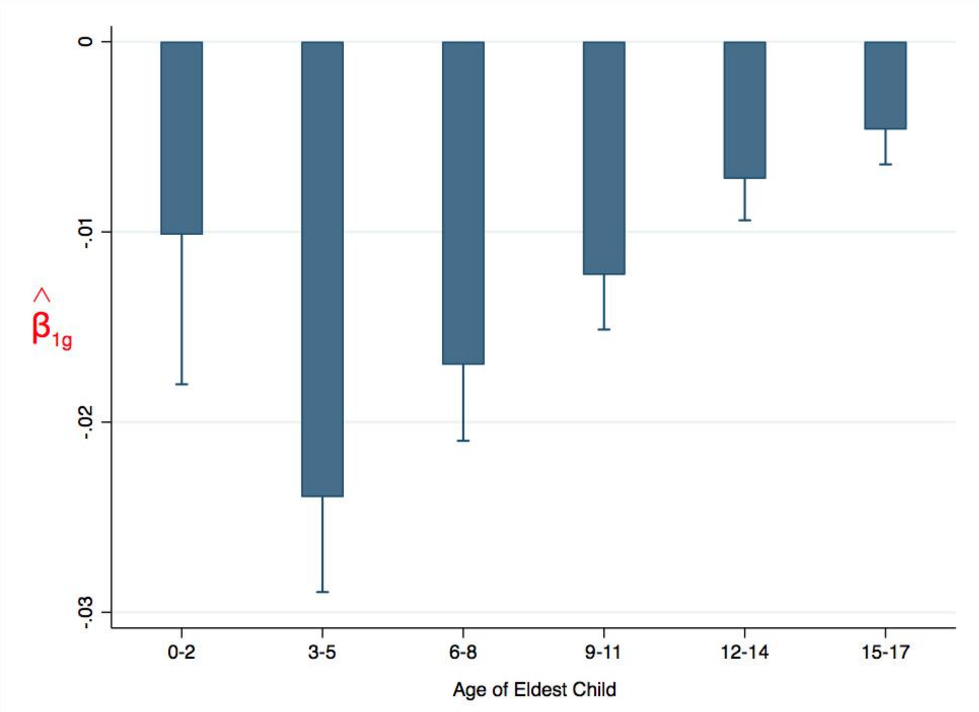
Notes: The left panel indicates the dynamic evolution of cultural trait i in the first neighborhood, n_1. The right panel indicates the dynamic evolution of cultural trait i in the second neighborhood, n_2. Under high moving costs, the two neighborhoods evolve independently. The long run distribution of cultural traits in the two neighborhoods is heterogeneous.

Figure 1.3: Migration Across PUMA, Comparing Parents and Non-Parents



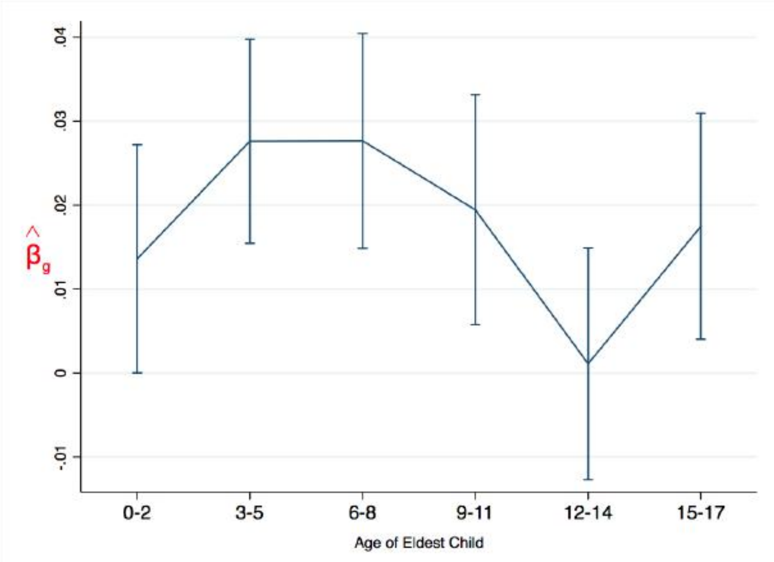
Notes: The left panel shows the difference in the share of out-migrants between parents and non-parents between 1995 and 2000 for parents and non-parents. The right panel shows the average probability of moving from a $PUMA$ of lower share of own ethnicity to a $PUMA$ with a higher share of own ethnicity for parents and non-parents. The probabilities are not regression adjusted. The figure in the right panel is based on a sample of across-$PUMA$ movers only.

Figure 1.4: Likelihood of leaving a PUMA, Comparing Parents and Non-parents by Age of Eldest Child



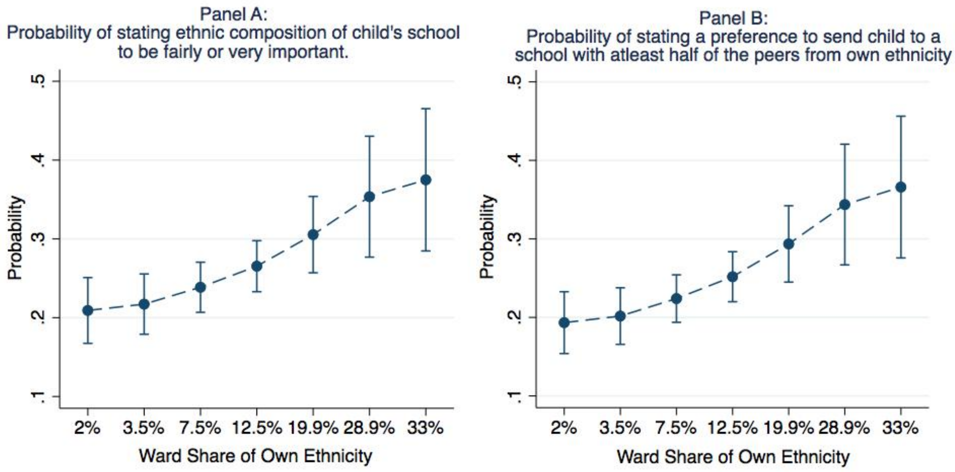
Notes: Y axis shows results from the regression of the dependent variable (dummy = 1 if individual left his/her $PUMA$ of residence) on the interactions between share of own ethnicity in origin $PUMA$ with a dummy that takes the value of 1 if individual is a parent with age of eldest child in a certain age group. There are 6 parent groups (that I denote by g) with age of eldest child between 0-2, 3-5, 6-8 up to 15-17. Each point shows the effect of share own ethnicity on the difference in migration probabilities of parents in group g relative to non-parents. The definition of ethnicity is based on ancestry.

Figure 1.5: Likelihood of Moving from Lower to Higher Own Ethnicity Share PUMA, Comparing Parents and Non-parents by Age of Eldest Child



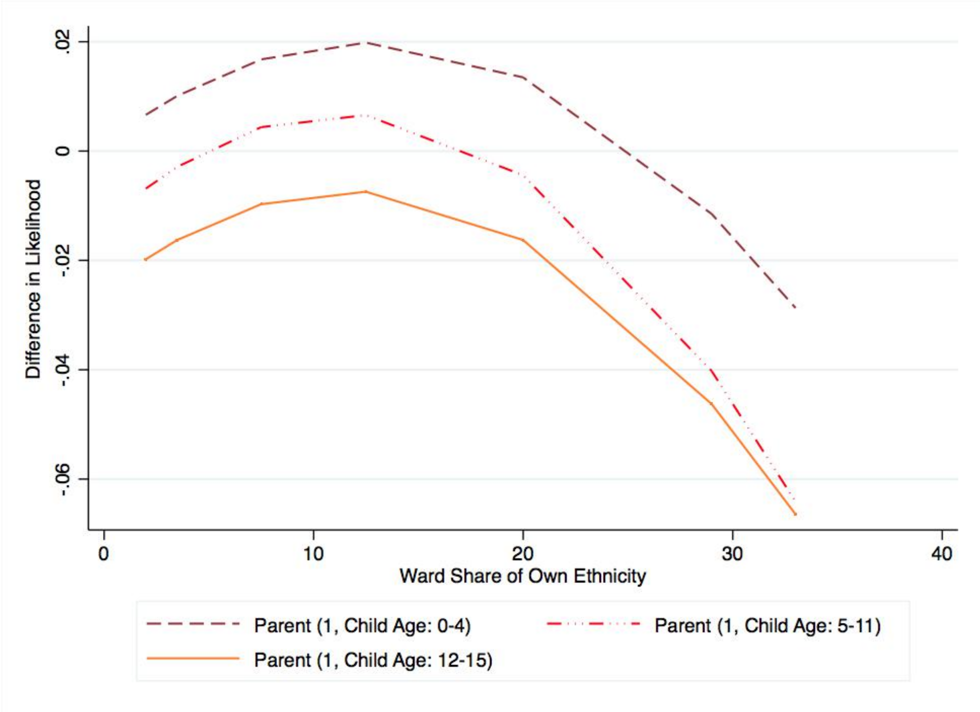
Notes: Y axis shows results from the regression of the dependent variable (dummy = 1 if individual left his/her $PUMA$ of residence) on the interactions between share of own ethnicity in origin $PUMA$ with a dummy that takes the value of 1 if individual is a parent with age of eldest child in a certain age group. There are 6 parent groups (that I denote by g) with age of eldest child between 0-2, 3-5, 6-8 up to 15-17. Each point shows the effect of share own ethnicity on the difference in migration probabilities of parents in group g relative to non-parents. The definition of ethnicity is based on ancestry.

Figure 1.6: Socialization Preferences and Ward Ethnic Density



Notes: Y axis of Panel A and Panel B reports probability estimates from the regression of the dependent variable (described in the legend) on *ward* share of own ethnicity.

Figure 1.7: Willingness to Move and Ward Ethnic Density, Comparing Parents and Non-Parents



Notes: Y axis reports the difference in the probability of stating a preference to leave current *ward* between parents and non-parents. The probability estimates are based on the regression of the dependent variable (dummy=1 if individual prefers to leave current *ward*) on *ward* share of own ethnicity.

Figure 1.8: Distribution of Own Ethnicity Share Across PUMA

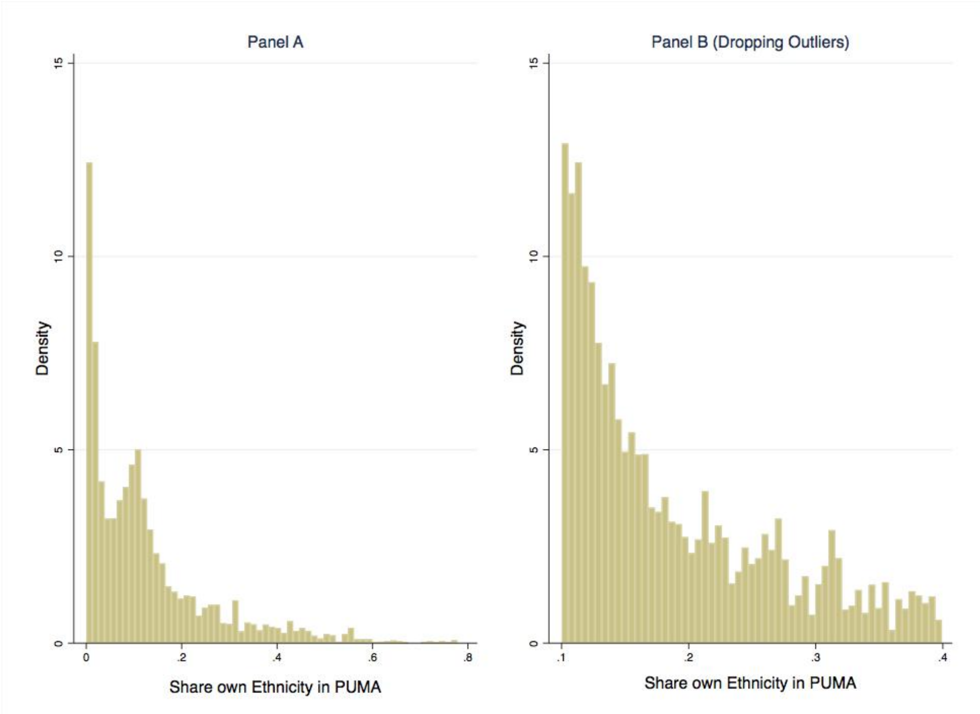
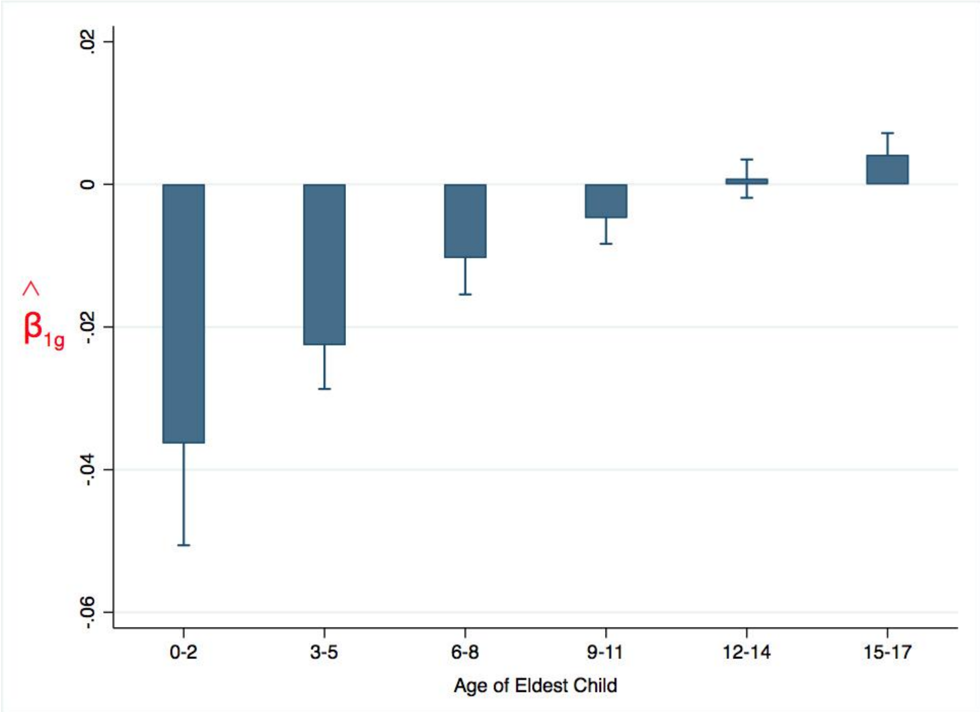


Figure 1.9: Likelihood of leaving a PUMA, Comparing Parents and Non-parents by Age of Eldest Child (Immigrant Sample)



Notes: Y axis shows results from the regression of the dependent variable (dummy = 1 if individual left his/her $PUMA$ of residence) on the interactions between share of own ethnicity in origin $PUMA$ with a dummy that takes the value of 1 if individual is a parent with age of eldest child in a certain age group. There are 6 parent groups (that I denote by g) with age of eldest child between 0-2, 3-5, 6-8 up to 15-17. Each point shows the effect of share own ethnicity on the difference in migration probabilities of parents in group g relative to non-parents. The sample is based on immigrants only. The definition of ethnicity is based on ancestry.

1.9 Tables

Table 1.1: Summary Statistics, U.S Census Microdata (2000)

	(Movers (across PUMA))		(Non-Movers)	
	Mean	S.D	Mean	S.D.
Age	39.4	12.6	50.4	12.1
Male	0.67	0.47	0.72	0.45
Parent (1, eldest kid age 0-2)	0.058	0.23	0.016	0.12
Parent (1, eldest kid age 3-5)	0.062	0.24	0.031	0.17
Parent (1, eldest kid age 6-8)	0.059	0.24	0.045	0.21
Parent (1, eldest kid age 9-11)	0.055	0.23	0.061	0.24
Parent (1, eldest kid age 12-14)	0.053	0.22	0.074	0.26
Parent (1, eldest kid age 15-17)	0.053	0.22	0.10	0.30
Parent (1, eldest kid age 0-19)	0.36	0.48	0.38	0.49
Indiv w/o Kids	0.64	0.48	0.62	0.49
College Educated	0.45	0.50	0.33	0.47
Wage (hourly)	17.5	19.0	20.4	25.2
Annual HH Income (in 1000s)	59.3	56.2	68.3	67.7
Monthly Rent	289.2	390.7	84.1	235.0
Homeowner	0.50	0.50	0.83	0.37
Travel Time to Work	19.7	23.6	18.2	24.0
Have Mortgage	1.34	1.44	1.98	1.23
Married, Spouse Present	0.52	0.50	0.65	0.48
Immigrants	0.080	0.27	0.14	0.35
Move to PUMA with higher share of own ethnicity than origin	0.49	0.50	.	.
Move to PUMA with lower share of own ethnicity than origin	0.51	0.50	.	.
n	141060		1243947	

Note: Based on natives and immigrants between 19-68. Total of 138 ancestries.

Table 1.2: Likelihood of Leaving a PUMA,
Comparing Parents and Non-Parents

Dep. Var: 1 if indiv. moved out of PUMA b/w 1995-2000	Age of Eldest Child						
	All	0-2	3-5	6-8	9-11	12-14	15-17
Dep. Mean:	(.109)	(.117)	(.115)	(.112)	(.109)	(.107)	(.103)
(Share Own Ethnicity in Origin) X (Parent)	-0.00808***	-0.00904**	-0.0228***	-0.0158***	-0.0109***	-0.00584***	-0.00323***
	(0.000625)	(0.00350)	(0.00203)	(0.00149)	(0.00116)	(0.000907)	(0.000854)
Share Own Ethnicity	-0.0101***	-0.0103***	-0.0103***	-0.0103***	-0.0103***	-0.0103***	-0.0103***
	(0.000846)	(0.000861)	(0.000861)	(0.000861)	(0.000861)	(0.000861)	(0.000861)
Parent	-0.0410***	0.0219***	-0.0401***	-0.0511***	-0.0465***	-0.0350***	-0.0326***
	(0.00161)	(0.00305)	(0.00216)	(0.00214)	(0.00201)	(0.00178)	(0.00164)
n	1385007	1385007	1385007	1385007	1385007	1385007	1385007

Note: Standard errors are in parentheses and clustered at PUMA level. Sample is based on parents and individuals without kids between 19-68 (as of 1995). Each unit of observation is a household head. I drop individuals who report to be living in institutions, households where reported migration status or ethnicity varies across members of the household. Sample is based on those who don't move or move across PUMA. I keep ethnicities that are closely related to country of birth (138 total ethnicities). Shares are standardized. * (p<.10), ** (p<.05), *** (p<.01).

Table 1.3: Likelihood of Moving from a lower to a Higher Own Ethnicity Share PUMA, Comparing Parents and Non-Parents

Dep Var: 1 if moved from lower to higher own ethnicity share PUMA b/w 1995-2000	Age of Eldest Child						
	All	0-2	3-5	6-8	9-11	12-14	15-17
Dep. Mean:	(.490)	(.480)	(.481)	(.482)	(.481)	(.479)	(.481)
Parent	0.0176***	0.0136**	0.0276***	0.0277***	0.0195**	0.00109	0.0175**
	(0.00365)	(0.00692)	(0.00620)	(0.00652)	(0.00698)	(0.00702)	(0.00686)
n	141060	141060	141060	141060	141060	141060	141060

Note: Standard errors are in parentheses and clustered at PUMA level. Sample is based on movers only. Refer to table 1.2 notes for other sample restrictions. * (p<.10), ** (p<.05), *** (p<.01).

Table 1.4: Migration Across PUMA, Comparing among Parents

Panel A: Out-Migration	(Age of Eldest Child)			
	0-2	3-5	6-8	9-11
Share Own Ethnicity in Origin) X (Parent)	-0.00463	-0.0187***	-0.0121***	-0.00847***
	(0.00425)	(0.00267)	(0.00216)	(0.00187)
Share Own Ethnicity in Origin	-0.0110***	-0.0110***	-0.0110***	-0.0110***
Parent	0.116***	0.0484***	0.0291***	0.0200***
	(0.00397)	(0.00269)	(0.00236)	(0.00212)
n	268418	268418	268418	268418

Panel B: Migration from less to more ethnically similar PUMA	(Age of Eldest Child)			
	0-2	3-5	6-8	9-11
Parent	0.0000179	0.0179	0.0192	0.0107
	(0.0160)	(0.0161)	(0.0150)	(0.0146)
n	35900	35900	35900	35900

Note: Standard errors are in parentheses and clustered at PUMA level. In panel A, I compare out-migration probabilities of parents with age of eldest child between 0-2 (column 1), 3-5 (column 2), 6-8 (column 3) to parents with all kids between 12-18 at the time of migration. In panel B, I do the same but look at migration from less ethnically similar to more ethnically similar PUMA for movers only. * (p<.10), ** (p<.05), *** (p<.01).

Table 1.5: Likelihood of Leaving a PUMA,
Comparing among Parents

Dep. Var: 1 if indiv. moved out of PUMA b/w 1995-2000	Age of Eldest Child						
	All	0-2	3-5	6-8	9-11	12-14	15-17
(Share Own Ethnicity in Origin) X (Parent)	-0.00873***	-0.00962**	-0.0235***	-0.0165***	-0.0116***	-0.00648***	-0.00380***
	(0.000614)	(0.00357)	(0.00209)	(0.00154)	(0.00118)	(0.000887)	(0.000798)
Share Own Ethnicity	-0.00951***	-0.00973***	-0.00973***	-0.00973***	-0.00973***	-0.00973***	-0.00973***
	(0.000848)	(0.000860)	(0.000860)	(0.000860)	(0.000860)	(0.000860)	(0.000860)
Parent	-0.0415***	0.0215***	-0.0411***	-0.0522***	-0.0475***	-0.0357***	-0.0330***
	(0.00158)	(0.00309)	(0.00216)	(0.00213)	(0.00199)	(0.00176)	(0.00162)
n	1461689	1461689	1461689	1461689	1461689	1461689	1461689

Note: Standard errors are in parentheses and clustered at PUMA level. Each unit of observation is a household head. I drop individuals who report to be living in institutions, households where reported migration status or ethnicity varies across members of the household. Sample is based on those who don't move or move across PUMAs. Here I compare parents with age of eldest child in a certain age to those with all kids above 18. I keep ethnicities that are closely related to country of birth (138 total ethnicities). Shares are standardized. * (p<.10), ** (p<.05), *** (p<.01).

Table 1.6: Likelihood of Moving from a lower to a Higher Own Ethnicity Share PUMA, Comparing among Parents

Dep Var: 1 if moved from lower to higher own ethnicity share PUMA b/w 1995-2000	Age of Eldest Child						
	All	0-2	3-5	6-8	9-11	12-14	15-17
Parent	0.0117***	0.0166**	0.0293***	0.0231***	0.00627	0.00229	0.00306
	(0.00273)	(0.00589)	(0.00510)	(0.00510)	(0.00523)	(0.00491)	(0.00458)
n	265110	265110	265110	265110	265110	265110	265110

Note: Standard errors are in parentheses and clustered at PUMA level. Sample is based on movers only. Here I compare parents with age of eldest child in a certain age to those with all kids above 18. Refer to table 1.5 notes for other sample restrictions. * (p<.10), ** (p<.05), *** (p<.01).

Table 1.7: Summary Statistics, U.S Census Microdata (1990)

	Movers		Non-Movers	
	Mean	S.D	Mean	S.D
Age	40.4	12.1	46.2	12.9
Male	0.75	0.43	0.72	0.45
College Educated	0.39	0.49	0.22	0.41
Wage(hourly)	13.6	13.2	12.6	13.7
Arrival Age in the U.S	22.6	11.7	25.7	11.4
Years since Arrival	17.7	9.86	20.5	10.5
English Proficiency; Scale (1-4); 1: Poor, 4: Proficient)	3.36	0.83	3.10	0.93
Linguistically Isolated	0.20	0.40	0.24	0.43
Annual HH Income (in 1000s)	42.3	37.9	42.3	38.7
Parent (1, eldest kid age: 0-2)	0.067	0.25	0.039	0.19
Parent (1, eldest kid age: 3-5)	0.084	0.28	0.055	0.23
Parent (1, eldest kid age: 6-8)	0.081	0.27	0.063	0.24
Parent (1, eldest kid age: 9-11)	0.076	0.26	0.070	0.26
Parent (1, eldest kid age: 12-14)	0.069	0.25	0.076	0.26
Parent (1, eldest kid age: 15-17)	0.068	0.25	0.089	0.28
Parent (1, eldest kid age: 0-19)	0.48	0.50	0.45	0.50
Indiv w/o Kids	0.46	0.50	0.37	0.48
Monthly Rent	515.9	240.9	447.0	223.6
Homeowner	0.47	0.50	0.58	0.49
Travel time to work	21.2	21.3	19.7	19.9
Have Mortgage	1.25	1.45	1.34	1.43
Married	0.62	0.49	0.64	0.48
n	31145		260239	

Note: Based on existing/pre-1985 immigrants between 19-68.

Table 1.8: Likelihood of Leaving an MSA, Comparing Parents and Non-Parents (OLS)

Dep. Var: 1 if indiv. moved out of MSA b/w 1985-1990.		Age of Eldest Child					
Dep. Mean:	All (.121)	0-2 (.133)	3-5 (.133)	6-8 (.130)	9-11 (.127)	12-14 (.124)	15-17 (.121)
Actual Immig. ShareXParent	0.00226	-0.0143**	-0.00740**	-0.0106**	0.00422	0.00338	0.0115**
	(0.00292)	(0.00514)	(0.00334)	(0.00395)	(0.00399)	(0.00392)	(0.00414)
Actual Immig. Share	-0.0211***	-0.0199**	-0.0199**	-0.0199**	-0.0199**	-0.0199**	-0.0199**
	(0.00599)	(0.00652)	(0.00652)	(0.00652)	(0.00652)	(0.00652)	(0.00652)
Parent	-0.0365***	-0.0269***	-0.0219***	-0.0344***	-0.0320***	-0.0298***	-0.0283***
	(0.00416)	(0.00627)	(0.00513)	(0.00654)	(0.00459)	(0.00557)	(0.00389)
n	240963	240963	240963	240963	240963	240963	240963

Note: Standard errors in parentheses and clustered at MSA level. Immigrant shares are standardized. The estimates show how the existing (pre-1985) immigrants respond to new immigrant inflows between 1985-1990. The sample is based on parents and those without kids and adults between 19-68 years (as of 1985). * ($p<.10$), ** ($p<.05$), *** ($p<.01$)

Table 1.9: Net-Inflows across MSAs of Existing Immigrant Parents and Non-Parents (OLS)

Dep. Var: Difference in netinflows of parents and non-parents.	Age of Eldest Child						
	All	0-2	3-5	6-8	9-11	12-14	15-17
Actual Immig. Share	0.00298	0.0371**	0.0186	0.0252**	0.0299*	0.0220	0.0199
	(0.00952)	(0.0148)	(0.0158)	(0.0120)	(0.0157)	(0.0153)	(0.0124)
n	3371	1579	1799	1794	1803	1785	1889

Note: Standard errors are in parentheses and are clustered at the MSA level. For other sample restrictions, refer to table 1.8. (p<.10), ** (p<.05), *** (p<.001) (p<.01)

Table 1.10: Net-Inflows across MSAs of Existing Immigrant Parents, Native Parents and Non-Parents (I.V)

Dep. Var: Difference in netinflows of parents and non-parents.			
Panel A: Domestic Net Inflows of Existing Immigrants	First Stage	OLS	IV
Pred. Immig. Share from Own Country of Birth	7.246***		
	(0.318)		
Actual. Immig. Share from Own Country of Birth		0.0274**	0.0366**
		(0.0126)	(0.0178)
n	2985	2210	2210
Panel B: Domestic Net Inflows of Natives	First Stage	OLS	IV
Pred. Immig. Share from Own Country of Birth	7.050***		
	(0.185)		
Actual. Immig. Share from Own Country of Birth		0.0149**	0.0122*
		(0.00634)	(0.00707)
n	3334	3025	3025

Note: Standard errors in parentheses and clustered at MSA level. The sample includes individuals without kids and parents with age of eldest child <6. Estimates are not significant for children of older ages. Panel A shows how domestic net inflows of existing immigrants respond to new immigrant inflows from the same country. Panel B shows how domestic net inflows of natives respond to new immigrant inflows from the country of similar ethnicity. * (p<.10), ** (p<.05), *** (p<.01)

Table 1.11: Summary Statistics, U.K Fourth National Survey of Ethnic Minorities (1993-1994)

	Mean	S.D
Age	37.5	12.1
Female	1.45	0.50
Employed	0.71	0.45
Employee	0.85	0.36
Manager	0.045	0.21
Weekly Housheold Income (before tax)	6.74	4.26
houseowner	0.74	0.44
British Qualification	0.57	0.49
English Proficiency (Scale 1-4, 1: fluent)	1.49	1.25
Born in U.K	0.28	0.45
Age at arrival in U.K	18.2	9.27
Years since arrival in U.K	23.5	9.00
Parent, 1(Kid age$<$16)	0.49	0.50
Parent, 1(Kid Age: 0-4)	0.25	0.43
Parent, 1(Kid Age: 5-11)	0.30	0.46
Parent, 1(Kid Age: 12-15)	0.19	0.39
Married	0.69	0.46
1(if had arranged marriage)	0.094	0.29
Homogeous Ethnic Partnership	0.80	0.40
Muslim	0.17	0.38
1(if ever discriminated)	0.13	0.33
1(if prefer to live with $\geq 1/2$ of own ethnic neighbors)	0.35	0.48
1 (if prefer to leave current Ward)	0.34	0.48
South	0.58	0.49
Ward Unemployment	12.6	5.20
Ethnic Share of Ward, (all ethnicities)	25.4	19.8
Social Housing Density of Ward	27.9	17.7
South	0.58	0.49

Note: Based on adults b/w 19-68. The raw data has 3035 observations.

Table 1.12: Likelihood of reporting a Preference to Leave Current Ward, Comparing Parents and Non-Parents

	Religion is Muslim	Couples share Ethnicity	Born in Country of Ethnicity (All Adults)	Born in Country of Ethnicity (Peak Parenting Age)
Share own Ethnicity in Origin	-0.129*	-0.867***	-0.153	-0.401**
XTrad.XParent	(0.0751)	(0.299)	(0.110)	(0.179)
Trad.XParent	-0.0302	-0.591**	-0.0481	-0.270**
	(0.0894)	(0.257)	(0.105)	(0.132)
Share own Ethnicity in OriginXTrad.	0.120*	0.759***	0.149*	0.358**
	(0.0636)	(0.233)	(0.0776)	(0.158)
Share own Ethnicity in OriginXParent	-0.0127	0.780***	0.101	0.382**
	(0.0445)	(0.295)	(0.104)	(0.174)
Share own Ethnicity in Origin	-0.0261	-0.734***	-0.160**	-0.402***
	(0.0360)	(0.231)	(0.0719)	(0.155)
Parent	0.0199	0.604**	0.0498	0.229*
	(0.0521)	(0.246)	(0.0997)	(0.123)
Trad., 1(if traditional)	0.00196	0.527***	0.195***	0.411***
	(0.0740)	(0.199)	(0.0730)	(0.103)
n	871	661	860	599

Notes: Standard errors are robust and in paranthesis. Traditional is defined based on if Muslim or not, born in country of reported ethnicity or not, and if parents share same ethnicity. Column 2 is based on married only, column 3 is based on all adults between 19-68 and column 4 is based on peak parenting age adults between 21-45. See text for controls. * (p<.10), ** (p<.05), *** (p<.01)

1.10 Appendix, Census Microdata

1.10.1 Ancestry list

Alsatian	Rom	Spanish
Austrian	Hungarian	Bahamian
Basque	Latvian	Barbadian
Belgian	Lithuanian	Belizean
Flemish	Macedonian	Jamaican
British Isles	Polish	Dutch West Indies
Danish	Romanian	Trinidadian
Dutch	Russian	British Virgin Islander
English	Serbian	Anguilla Islander
Finnish	Slovak	Grenadian
French	Slovene	St Lucia Islander
German	Ukrainian	West Indian
Prussian	Yugoslavian	Haitian
Italian	Nuevo Mexicano	Moroccan
Sicilian	Costa Rican	Iranian
Luxemburger	Guatemalan	Iraqi
Norwegian	Nicaraguan	Jordanian
Portuguese	Panamanian	Lebanese
Scotch Irish	Salvadoran	Syrian
Scottish	Argentinean	Armenian
Swiss	Chilean	Yemeni
Welsh	Colombian	Palestinian
Albanian	Ecuadorian	Assyrian

Nigerian	Fijian	Bulgarian
Sierra Leonean	Pacific Islander	Croatian
Somalian	Estonian	Uruguayan
South African	Dominican	Ethiopian
Sudanese	Liberian	Hawaiian
Afghan	Chamorro	Eritrean
Nepali	Bohemian	Samoan
Asian Indian	Cuban	Venezuelan
Pakistani	Kenyan	Belourussian
Sri Lankan	Guamanian	Peruvian
Burmese	Czech	Cape Verdean
Cambodian	Puerto Rican	Polynesian
Chinese	Ghanian	Taiwanese
Indonesian	Chicano/Chicana	Tongan
Japanese	Egyptian	Cantonese
Okinawan	Filipino	Brazilian
Laotian	Maltese	Slav
Hmong	Honduran	Greek
Malaysian	Israeli	Irish
Thai	Korean	Hong Kong
Vietnamese	Swedish	Guyanese
Australian	Bolivian	Mexican
New Zealander	Turkish	Icelander

Table 1.13: Likelihood of Leaving a PUMA

	Age of Eldest Child: 5-19	
	Parents vs Non-Parents	Parents vs Non-Parents and Parents with Kids > 19 years
Share Own Ethnicity in OriginXParent	-0.00747***	-0.00804***
	(0.000653)	(0.000622)
Share Own Ethnicity in Origin	-0.00973***	-0.00925***
	(0.000821)	(0.000819)
Parent	-0.0453***	-0.0458***
	(0.00177)	(0.00175)
n	1329106	1397330

Notes: Standard errors are in parentheses and clustered at PUMA level. Column 1 is based on parents and non-parents. Column 2 is based on parents, non-parents and parents with kids above 19. Based on individuals between age 19-68 (as of 1995). Each unit of observation is a household head. I drop individuals who report to be living in institutions, households where reported migration status or ethnicity varies across members of the household. Sample is based on those who don't move or move across PUMA. I keep ethnicities that are closely related to country of birth. Shares are standardized. * ($p<.10$), ** ($p<.05$), *** ($p<.01$).

Table 1.14: Likelihood of Moving from a lower to a Higher Own Ethnicity Share PUMA

	Age of Eldest Child: 5-19	
	Parents vs Non-Parents	Parents vs Non-Parents and Parents with Kids > 19 years
Parent	0.0155***	0.0153***
	(0.00408)	(0.00408)
n	126912	128498

Note: Standard errors are in parentheses and clustered at PUMA level. Sample is based on movers only. Refer to table 1.13 notes for other sample restrictions. * ($p<.10$), ** ($p<.05$), *** ($p<.01$).

Table 1.15: Likelihood of Leaving a PUMA,
Comparing Parents with Non-Parents and Other Parents (Immigrants only)

	Age of Eldest Child						
	All	0-2	3-5	6-8	9-11	12-14	15-17
(Share Own Ethnicity in Origin)	-0.00155	-0.0363***	-0.0225***	-0.0103***	-0.00472**	0.000800	0.00414**
X Parent	(0.00118)	(0.00729)	(0.00313)	(0.00260)	(0.00184)	(0.00137)	(0.00155)
Share Own Ethnicity	-0.00790***	-0.00664**	-0.00664**	-0.00664**	-0.00664**	-0.00664**	-0.00664**
	(0.00199)	(0.00202)	(0.00202)	(0.00202)	(0.00202)	(0.00202)	(0.00202)
Parent	-0.0233***	0.0208***	-0.0199***	-0.0321***	-0.0282***	-0.0241***	-0.0186***
	(0.00217)	(0.00552)	(0.00361)	(0.00328)	(0.00288)	(0.00271)	(0.00226)
n	176447	176447	176447	176447	176447	176447	176447

Note: Standard errors in parentheses and clustered at PUMA level. Sample is based on parents, non-parents and parents with kids above 19 years. Based on individuals between 19-68 (as of 1995). Each unit of observation is a household head. I drop individuals who report to be living in institutions, households where reported migration status or ethnicity varies across members of the household. Sample is based on those who don't move or move across PUMAs. I keep immigrants and ethnicities that are closely related to country of birth. Shares are standardized. ** (p<.05), *** (p<.01)

Table 1.16: Likelihood of Moving from a lower to a Higher Own Ethnicity Share PUMA, Comparing Parents and Non-Parents (Immigrants Only)

	Age of Eldest Child						
	All	0-2	3-5	6-8	9-11	12-14	15-17
Parent	-0.00581	-0.00178	0.00663	-0.0138	0.0128	-0.0328	-0.0144
	(0.0137)	(0.0242)	(0.0200)	(0.0220)	(0.0225)	(0.0226)	(0.0255)
n	10714	10714	10714	10714	10714	10714	10714

Note: Standard errors in parentheses and clustered at PUMA level. Sample is based on movers only. Refer to table 1.16 notes for other sample restrictions. * (p<.10), ** (p<.05), *** (p<.01)

Table 1.17: Likelihood of Leaving a PUMA,
Comparing Parents and Non-Parents (Immigrants, without Mexicans)

	Age of Eldest Child						
	All	0-2	3-5	6-8	9-11	12-14	15-17
(Share Own Ethnicity in Origin) X Parent	-0.0208***	-0.0304***	-0.0128***	-0.00806**	-0.00713**	-0.00107	0.00218
	(0.00130)	(0.00587)	(0.00347)	(0.00281)	(0.00232)	(0.00177)	(0.00159)
Share Own Ethnicity	-0.00838***	-0.00790***	-0.00790***	-0.00790***	-0.00790***	-0.00790***	-0.00790***
	(0.00163)	(0.00160)	(0.00160)	(0.00160)	(0.00160)	(0.00160)	(0.00160)
Parent	-0.0208***	0.0238***	-0.0173***	-0.0270***	-0.0261***	-0.0224***	-0.0168***
	(0.00230)	(0.00582)	(0.00403)	(0.00359)	(0.00321)	(0.00290)	(0.00250)
n	139671	139671	139671	139671	139671	139671	139671

Note: Standard errors in parentheses and clustered at PUMA level. Sample is based on parents and individuals without kids between 19-68 (as of 1995). Each unit of observation is a household head. I drop individuals who report to be living in institutions, households where reported migration status or ethnicity varies across members of the household. Sample is based on those who don't move or move across PUMA. I keep ethnicities that are closely related to country of birth. Sample is based on immigrants but without Mexicans. Shares are standardized. * (p<.10), ** (p<.05), *** (p<.01).

Table 1.18: Likelihood of Leaving a PUMA,
Comparing Parents and Non-Parents (Without Mexicans and Germans)

	Age of Eldest Child						
	All	0-2	3-5	6-8	9-11	12-14	15-17
(Share Own Ethnicity in Origin)	-0.00668***	-0.00532	-0.00949***	-0.00930***	-0.00880***	-0.00563***	-0.00427***
x(Parent)	(0.000825)	(0.00411)	(0.00239)	(0.00150)	(0.00139)	(0.00117)	(0.00122)
Share Own Ethnicity	-0.00784***	-0.00825***	-0.00825***	-0.00825***	-0.00825***	-0.00825***	-0.00825***
	(0.000886)	(0.000900)	(0.000900)	(0.000900)	(0.000900)	(0.000900)	(0.000900)
Parent	-0.0393***	0.0239***	-0.0315***	-0.0470***	-0.0432***	-0.0333***	-0.0321***
	(0.00180)	(0.00363)	(0.00252)	(0.00239)	(0.00229)	(0.00203)	(0.00187)
n	998796	998796	998796	998796	998796	998796	998796

Note: Standard errors in parentheses and clustered at PUMA level. Sample is based on parents and individuals without kids between 19-68 (as of 1995). Each unit of observation is a household head. I drop individuals who report to be living in institutions, households where reported migration status or ethnicity varies across members of the household. Sample is based on those who don't move or move across PUMAs. I keep ethnicities that are closely related to country of birth. Sample is based on immigrants and natives but without Mexicans and Germans. Shares are standardized. * ($p<.10$), ** ($p<.05$), *** ($p<.01$).

Table 1.19: Likelihood of Moving from a lower to a Higher Own Ethnicity Share PUMA, Comparing Parents and Non-Parents (Without Mexicans and Germans)

	Age of Eldest Child						
	All	0-2	3-5	6-8	9-11	12-14	15-17
Parent	0.00861**	0.00969	0.0191**	0.0214**	0.00973	-0.0137	0.00406
	(0.00421)	(0.00823)	(0.00763)	(0.00784)	(0.00824)	(0.00841)	(0.00814)
n	100528	100528	100528	100528	100528	100528	100528

Note: Standard errors in parentheses and clustered at PUMA level. Sample is based on movers only. Refer to table 1.19 notes for other sample restrictions. * (p<.10), ** (p<.05), *** (p<.01)

Table 1.20: Migration Across PUMA,
Comparing Parents and Non-Parents (Dropping Extreme Ancestry Shares)

Panel A	
	Age of Eldest Child: 0-19
Share Own EthnicityXParent	-0.00656***
	(0.00161)
Share Own Ethnicity	-0.00380**
	(0.00174)
Parent	-0.0432***
	(0.00215)
n	558600
Panel B	
	Age of Eldest Child: 0-19
Parent	0.0300***
	(0.00552)
n	57736

Note: Standard errors are in parentheses and clustered at PUMA level. I drop share of own ancestry below .1 and above .4. Panel A reports shows the probability out migration. Panel B reports the probability of moving from ethnically less similar to more similar PUMA. Panel B is based on movers only. See table 1.2 for notes. * (p<.10), ** (p<.05), *** (p<.01).

Table 1.21: Likelihood of Leaving an MSA and Age at Immigration, Comparing Parents and Non-Parents

ParentXPred. Immig. ShareXAge at Arrival Dummy	-0.0156*
	(0.00866)
ParentXAge at Arrival Dummy	-0.0162*
	(0.00879)
Pred. Immig. ShareXAge at Arrival Dummy	0.0167**
	(0.00718)
ParentXPred. Immig. Share	0.0211**
	(0.00866)
Pred. Immig. Share in Origin from Own Country of Birth	-0.0445**
	(0.0143)
Age at Arrival Dummy 1 if >9	0.0112
	(0.0116)
Parent	-0.0344***
	(0.00743)
n	47959

Note: Standard errors are in parentheses and clustered at MSA level. I keep childhood immigrants who moved to the U.S before age 16. Sample is based on parents and non-parents only. Estimates are robust to age at arrival dummy=1 if >10. Estimates are driven mainly by parents with eldest child in age group 6-11. * ($p<.10$), ** ($p<.05$), *** ($p<.01$)

1.11 U.K FNSEM (1993-1994) Appendix

I define traditional based on the following survey questions: (1) Would you personally mind if a close relative were to marry a white person ? 1 if yes and 0 if no. (2) How important is religion in the way you live your life ? 1 if fairly/very important. 0 if not very important or not at all important. (3) Would you prefer to send your child to a school with fewer than half, half or more than half of the same ethnic origin or do you have no preference ? 1 if at least 1/2. 0 if fewer than 1/2, no preference or cant say. (4) In many ways do you think of yourself as British ? 1 if strongly agree. 0 if agree, neither agree or disagree, disagree, strongly disagree or cant say. The difference in mean responses are shown below. Based on the difference, I define traditional in 3 ways- Muslims[27]; those born in country dominant in their reported ethnicity (e.g. 1 if born in China if reported ethnicity is Chinese and 0 if born in either U.K or in any country other than China); those among married couples who share the same ethnicity[28].

Groups	Difference in Mean Traditional Responses
Muslim-Non-Muslim	.4637333**
Born in Country of Ethnicity- Not Born in Country of Ethnicity	.1919438 **
Couple with Same Ethnicity- Couples with Different Ethnicity	.5114247 **
(note **: p<.05)	

[27]Bisin et al. (2010) using U.K Fourth National Survey of Ethnic Minorities provide evidence that Muslims have stronger religious identities and assimilate slower than non-Muslims.

[28]For a review of literature on the impact of homogamous marriage on socialization and religious identities among children, refer to the Handbook of Social Economics (2010).

Table 1.22: Ethnic Neighbor Preferences and Age at Immigration, Comparing Parents and Non-Parents

Dep Var.: Dummy=1 if prefer to live with ≥ half of the neighbors from own ethnic origin	Age of Eldest Child		
	0-4	5-11	12-15
ParentxAge at Arrival in U.K	0.119*	0.131**	0.0530
	(0.0692)	(0.0618)	(0.0811)
Age at Arrival in U.K	-0.145**	-0.140**	-0.152**
	(0.0506)	(0.0518)	(0.0522)
Parent	0.0382	0.00311	0.0285
	(0.0734)	(0.0713)	(0.0881)
Individual Controls	Yes	Yes	Yes
Contextual Controls	Yes	Yes	Yes
n	263	293	209

Notes: Standard errors are robust and in paranthesis. Age at arrival is standardized. Individual controls include age, gender, income, education, employment, english proficiency, dummy if individual has been discriminated. Contextual controls include ward unemployment, fraction immigrants, dummy if region is south. Based on childhood immigrants, parents and non-parents. Based on those who moved to U.K before age 16. Based on adults b/w 19-68. * (p<.10), ** (p<.05), *** (p<.01).

1.12 Model Appendix

Dynamic Equations and Equilibrium

Following Bisin et al.(2001, 2008) and Montgomery(2008), suppose period length is dt and fraction dt of the population dies between time t and $t + dt$. Fertility and socialization takes place in individual's last period of life. Ignoring time notation for simplicity, let total number of individuals with trait i at time t be I_1 where $I_1(t) = q_1^i x N$. Similarly, $I_2(t) = q_2^i (1 - x)N$.

Each individual decides whether to continue in the current residence or migrate to the alternate neighborhood between time t and $t + dt$. Migration is instantaneous. If parents move (stay), children are socialized in the destination (origin) neighborhood. Then,

$$I_1(t+dt) = \underbrace{(1-dt)(q_1^i x_1 N)}_{non\,movers\ b/w\ t\ \&\ t+dt} + \underbrace{dt.N[\underbrace{q_1^i x_1 - \lambda_{12}^i q_1^i x_1 + \lambda_{21}^i q_2^i (1-x_1)}_{m_1^i}]P_1^{ii}}_{net\ parent\ i\ migrants\ in\ n_1} +$$

$$\underbrace{dt.N[\underbrace{q_1^j x_1 - \lambda_{12}^j q_1^j x_1 + \lambda_{21}^i q_2^j (1-x_1)}_{m_j^i}]\ P_1^{ji}}_{net\ parent\ j\ migrants\ in\ n_2}$$

$$J_1(t+dt) = (1-dt)(q_1^j xN) + dt[q_1^i xN - \lambda_{12}^i q_1^i xN + \lambda_{21}^i q_2^i (1-x)N]P_1^{ij} +$$

$$dt[q_1^j xN - \lambda_{12}^j q_1^j xN + \lambda_{21}^i q_2^j (1-x)N]P_1^{jj}$$

$$q_1^i(t+dt) = \frac{I_1(t+dt)}{I_1(t+dt) + J_1(t+dt)}$$

$$\frac{\partial q_1^i(q_1^i, q_2^i, x)}{\partial t} = \underset{dt \to 0}{lt} \frac{1}{dt} \{ q_1^i(t+dt) - q_1^i(t) \}$$

$$= (\frac{1}{x_1})[(1-q_1^i)(m_1^i p_1^{ii} + m_1^j p_1^{ji}) - (q_1^i)(m_1^i p_1^{ij} + m_1^j p_1^{jj})]$$

With no migration frictions and suppose initial $q_2^i > q_1^i$, at each period t, $\lambda_{12}^i = 1$, $\lambda_{21}^i = 0$, $\lambda_{12}^j = 1$ and $\lambda_{21}^j = 1$. Then,

$$\frac{\partial q_1^i(q_1^i, q_2^i, x_1)}{\partial t} =$$

$$(\frac{1}{x_1})[(1-q_1^i)(x_1) + (1-q_2^i)(1-x_1)][(1-q_1^i)P_{ji}^i - q_1^i(P_{jj}^1)] \quad [A1]$$

$I_2(t+dt)$ and $J_2(t+dt)$ can be derived in a similar way. Then ,

$$q_2^i(t+dt) = \frac{I_2(t+dt)}{I_2(t+dt) + J_2(t+dt)}$$

$$\frac{\partial q_2^i(q_1^i, q_2^i, x)}{\partial t} = \underset{dt \to 0}{lt} \frac{1}{dt} \{ q_2^i(t+dt) - q_2^i(t) \}$$

$$\frac{\partial q_2^i(q_1^i, q_2^i, x_1)}{\partial t} =$$

$$(\frac{1}{1-x_1})[(q_1^i)(x_1) + (1-q_2^i)(1-x_1)][(1-q_2^i)P_{ii}^2 - q_2^i(P_{ij}^2)] \quad [A2]$$

$$x(t+dt) = \frac{I_1(t+dt) + J_1(t+dt)}{I_1(t+dt) + J_1(t+dt) + I_2(t+dt) + J_2(t+dt)}$$

$$\frac{\partial x_1(q_1^i, q_2^i, x)}{\partial t} = \underset{dt \to 0}{lt} \frac{1}{dt} \{ x_1(t+dt) - x_1(t) \}$$

$$\frac{\partial x_1(q_1^i, q_2^i, x_1)}{\partial t} =$$

$$(1 - x_1)(1 - q_2^i) - (x_1)(q_1^i) \quad [A3]$$

Note that since socialization happens after migration, actual optimum effort and socialization probabilities are functions of net migrants and not initial q^i's and q^j's. Based on the socialization technology described in the main paper, this yields

,

$$p_k^{ii} = \tau_k^i + (1 - \tau_k^i)(\gamma_k^i(t))$$

where γ_k^i is the share of trait i in neighborhood k *after* migration.

and

$$\tau_k^i = (1 - \gamma_k^i(t))\Delta v$$

For illustrative simplicity, let $\Delta v = 1$(this is exogenous and specifying any other value will not change the equilibrium results.) Then, $p_{ii}^k = (1-\gamma_k^i)+(\gamma_k^i)^2$; $p_{ii}^k = (1-\gamma_k^j)+(\gamma_k^j)^2$; $p_{ij}^k = 1-p_{ii}^k$; $p_{ji}^k = 1-p_{jj}^k$. Values of γ are straightforward to derive. For instance,

$$\gamma_2^i = \frac{(q_2^i(1-x_1) + q_1^i x_1 dt)N \ \to \text{net migrants with trait i in n}_2}{((1-x_1)(1-dt) + q_2^i(1-x_1)dt + q_1^i(x_1)dt)N \ \to \text{total net migrants in n}_2}$$

After substituting these values, all the dynamic equations can be derived as functions of q_1^i, q_2^i and x_1 and the dynamics can be characterized by the following system:

$$\frac{\partial q_1^i(q_1^i, q_{2,}^i x_1)}{\partial t};\ \frac{\partial q_1^i(q_1^i, q_{2,}^i x_1)}{\partial t};\ \frac{\partial q_1^i(q_1^i, q_{2,}^i x_1)}{\partial t}$$

The following jacobian matrix takes the following form:

$$\begin{matrix} \frac{(\partial q_1^i(q_1^i,q_{2,}^i x)/\partial t)}{\partial q_1^i} & \frac{(\partial q_1^i(q_1^i,q_{2,}^i x)/\partial t)}{\partial q_2^i} & \frac{(\partial q_1^i(q_1^i,q_{2,}^i x)/\partial t)}{\partial x} \\ \frac{(\partial q_2^i(q_1^i,q_{2,}^i x)/\partial t)}{\partial q_1^i} & \frac{\partial q_2^i(q_1^i,q_{2,}^i x)/\partial t)}{\partial q_2^i} & \frac{\partial q_2^i(q_1^i,q_{2,}^i x)/\partial t)}{\partial x} \\ \frac{\partial x(q_1^i,q_{2,}^i x)/\partial t)}{\partial q_1^i} & \frac{\partial x(q_1^i,q_{2,}^i x)/\partial t)}{\partial q_2^i} & \frac{\partial x(q_1^i,q_{2,}^i x)/\partial t)}{\partial x} \end{matrix}$$

The model supports segregation in equilibrium. At points $(q_1^i, q_2^i) = (0, 1)$, all the 3 dynamic equations are 0 and hence stationary. Furthermore, the eigenvalues of the Jacobian are negative, supporting stability.

Dynamics and Equilibrium under Non Linear Socialization Mechanisms and Migration Probabilities

I allow for non-linearities in the socialization mechanisms by allowing the probability of direct socialization to be the following[29]

$$d_k^i = \tau_k^i.q_k^i \ where\, i\epsilon\{r, s\}\, and\, k\epsilon\{n_1, n_2\}$$

Probability of socialization to trait i, for child born to trait i, can then be defined (after migration) as :

$$p_k^{ii} = \tau_k^i.\gamma_k^i + (1 - \tau_k^i.\gamma_k^i)(\gamma_k^i)$$

and

$$p_k^{ij} = 1 - p_k^{ii}$$

At the beginning of each period t, parents i in say n_1 migrate with proba-

[29]This specification is borrowed from Bisin and Verdier(2001).

bility λ_{12}^i to n_2. They then choose socialization effort in order to maximize the following -

$$max_{\tau,x^i}\ u^i(x^i) - c(\tau_2^i) + (p_2^{ii}v^{ii} + p_2^{ij}v^{ij})$$

This yields $\tau_2^{i*} = \gamma_2^i(1-\gamma_2^i)\triangle v$. For simplicity , let $\triangle v = 1$. Given that $\triangle v$ is exogenous, specifying any other value will not change the equilibrium results.

We then end up with -

$$\tau_2^{i*} = \gamma_2^i(1-\gamma_2^i)$$

After substituting τ^* in the socialization probabilities , we end up with p_k^{ii} and p_k^{ij} as functions of q_k^i's and x_1.

MIGRATION/MARKOV PROBABILITIES

Let the probability of going from neighborhood 1 to neighborhood 2, between t and $t+dt$ for parent i be:

$$\lambda_{12}^i = Pr.(X_{t+dt} = 2|X_t = 1)....pr.\,of\,moving\,from\,n_1\,to\,n_2$$

Similarly ,

$$\lambda_{11}^i = Pr.(X_{t+dt} = 1|X_t = 1)....pr.\,of\,not\,moving\,from\,n_1$$

$$\lambda_{21}^i = Pr.(X_{t+dt} = 1|X_t = 2)....pr.\,of\,moving\,from\,n_2\,to\,n_1$$

$$\lambda_{22}^i = Pr.(X_{t+dt} = 2|X_t = 2)....pr.\,of\,not\,moving\,from\,n_2$$

Each λ_{xy}^i represents the $(x,y)^{th}$ element of a transient matrix $M(t)$:

$$\begin{bmatrix} \lambda_{11}^i & \lambda_{12}^i \\ \lambda_{21}^i & \lambda_{22}^i \end{bmatrix}$$

$$M^{'}(t) = (1/dt).M(t).Z$$

Differentiation w.r.t t yields the following matrix exponential -

$$M^{'}(t) = e^{(t/dt)(Z)}$$

$$z = \begin{bmatrix} -q_2^i & q_2^i \\ -q_1^i & q_1^i \end{bmatrix}$$

Then, $M(t)$=

$$\begin{bmatrix} \frac{q_1^i}{q_2^i+q_1^i} + \frac{q_2^i}{q_2^i+q_1^i} e^{-(q_1^i+q_2^i)\frac{t}{dt}} & \frac{q_2^i}{q_2^i+q_1^i} + \frac{q_2^i}{q_2^i+q_1^i} e^{-(q_1^i+q_2^i)\frac{t}{dt}} \\ \frac{q_1^i}{q_2^i+q_1^i} - \frac{q_1^i}{q_2^i+q_1^i} e^{-(q_1^i+q_2^i)\frac{t}{dt}} & \frac{q_2^i}{q_2^i+q_1^i} + \frac{q_2^i}{q_2^i+q_1^i} e^{-(q_1^i+q_2^i)\frac{t}{dt}} \end{bmatrix}$$

As $dt \to 0$ or $t \to \infty$, the probability distribution tends to:

$$\begin{bmatrix} \frac{q_1^i}{q_2^i+q_1^i} & \frac{q_2^i}{q_2^i+q_1^i} \\ \frac{q_1^i}{q_2^i+q_1^i} & \frac{q_2^i}{q_2^i+q_1^i} \end{bmatrix}$$

Probability of migration of parent i from n_1 to n_2 is now $\lambda_{12}^i(q_1^i, q_2^i) = \frac{q_2^i}{q_1^i+q_2^i}$ and from $n2$ to $n1$ as $\frac{q_1^i}{q_2^i+q_1^i}$. As shown above, these transition probabilities have a Markov representation and the following characteristics are worth noting-

- $\frac{\partial \lambda_{12}^i(q_1^i,q_2^i)}{\partial q_1^i} < 0\, and\, \frac{\partial \lambda_{12}^i(q_1^i,q_2^i)}{\partial q_2^i} > 0$ if $q_1^i, q_2^i \epsilon \{0, 1\}$ and $\lambda_{12}^i(q_1^i, q_2^i) = 0$ if $q_2^i = 0$. Given that marginal benefit is increasing in the share of own cultural trait in the population, this property is intuitive.

- let $z = [\frac{q_2^i(t+dt)}{q_1^i(t+dt)}] - [\frac{q_2^i(t)}{q_1^i(t)}]$. Then $\frac{\partial \lambda_{12}^i}{\partial t} \geq 0$ if $z \geq 0$ and < 0, otherwise. In other words, if the ratio of shares i in the two neighborhoods, $\frac{q_2^i}{q_1^i}$, do not change from one period to the next, the probability of migration stays stays constant over time.

- Owing to migration frictions, $\lambda_{12}^i(q_1^i, q_2^i)$ could be less than 1 even when $q_2^i > q_1^i$ and greater than 0 even for the cases where $q_2^i < q_1^i$

As before, we have a dynamic system characterized by the joint evolution of q_1^i, q_2^i and x_1. The dynamics equations can be derived in exactly the same way as in the case of linear socialization technology, described in the previous appendix section. The only difference is in the migration probabilities and in socialization probability functions. $(q_2^i, q_2^j = 0, 1)$ and $(q_1^i, q_1^j = 1, 0)$ are equilibrium points of this system based on the fact $\frac{\partial q_1^i(q_1^i, q_2^i, x_1)}{\partial t}, \frac{\partial q_2^i(q_1^i, q_2^i, x_1)}{\partial t}, \frac{\partial x_1(q_1^i, q_2^i, x_1)}{\partial t}$ are all equal to 0 at these points and the jacobian yields negative eigenvalues. To understand the intuition, suppose there is a slight deviation from the equilibrium points $(q_1^i, q_2^i = 1, 0)$ and $q_1^i = 1 - \epsilon_1$ where $0 < \epsilon_1 < .5$ and and $q_2^i = 0 + \epsilon_2$ where $.5 < \epsilon_2 < 1$. In so far as the deviations are small, share of i in n_1 would still exceed share of i in n_2 and $\lambda_{21}^i > \lambda_{12}^i$. In the next period, share of trait i would increase in n_1 and decline in n_2. Socialization effort would decline but due to high horizontal socialization, fraction of trait i would increase[30].

In n_2, probability of socialization to i would decrease due to lower direct socialization effort and low horizontal socialization. The process would continue till all $i's$ are again in n_1 and $j's$ are in n_2. If suppose the deviation is large and q_1^i drops to less than .5 and q_2^i increases to a share $> .5$.Then, λ_{12}^i would exceed λ_{21}^i in each period and the dynamics would shift the equilibrium to $(q_1^i, q_2^i) = (0, 1)$. Figure below illustrates the dynamics and equilibrium points of this system.

[30]Note that the optimum effort of parent is increasing(decreasing) in share of own trait when that share is less (greater) than 1/2.

- *Parents with cultural transmission preferences are less likely to move out, the higher the fraction of own type in origin.*

Probability of out migration of parent i in n_1 is given by:

$$\lambda_{12}^i = \frac{q_1^i}{q_2^i + q_1^i}$$

and that of parent of i in n_2 is

$$\lambda_{21}^i = \frac{q_2^i}{q_2^i + q_1^i}$$

$\frac{\partial \lambda_{12}^i}{\partial q_1^i} \leq 0$ and $\frac{\partial \lambda_{21}^i}{\partial q_1^i} \leq 0$.

- *Parents are more likely to move into neighborhoods the higher fraction of own type in that neighborhood.*

Probability of migration of parent i into n_1 is given by:

$$\lambda_{21}^i = \frac{q_1^i}{q_2^i + q_1^i}$$

and that of parent of i into n_2 is

$$\lambda_{12}^i = \frac{q_2^i}{q_2^i + q_1^i}$$

$\frac{\partial \lambda_{21}^i}{\partial q_1^i} \geq 0$ and $\frac{\partial \lambda_{12}^i}{\partial q_2^i} \geq 0$.

The sorting pattern of parents with cultural transmission preferences under non linear socialization technologies and migration probabilities is intuitively consistent with the linear case.

Chapter 2

Intergenerational Language Transmission and Assimilation: Evidence from the U.K

2.1 Introduction

In nations where immigrants constitute a significant share of the population, immigrant assimilation is an issue of policy relevance. U.K is one such country that has witnessed a substantial increase in the share of it's foreign born population. Between 1991-2017, immigration increased by almost 91% from 329,000 in 1991 to 630,000 in 2017[1]. An important determinant of immigrant integration is language. There is a vast literature that has demonstrated the impact of language proficiency on social assimilation and other educational, economic and health outcomes[2]. For U.K in particular, studies have found English proficiency

[1]Taken from the Briefing Paper SN06077 (2018); Source: ONS, Long-Term International Migration Estimates, 2 series (LTIM calendar year); Long-Term International Migration, Quality of Long-Term International Migration Estimates from 2001 to 2011.

[2] For language and education outcomes, see for instance Portes and MacLeod (1999); Card, DiNardo, and Estes (2000); Hirschman (2001); and Grogger and Trejo (2002); Chin and Bleakely (2004, 2010); Glick and White (2003).

For language and labor market outcomes, see for instance Bleakley and Chin (2004); Dustmann and Fabbri (2003); Chiswick and Miller (1995); Dustmann and van Soest (2002); Casey and Dustmann (2008). For a review, also see Aoki and Santiago (2017).

to have a negative impact on residential deprivation (Aoki and Santiago (2017)) and a positive impact on labor market outcomes such as increase in immigrant earnings (Dustman and Fabri(2003)), reduction in the native-immigrant wage gap (Miranda and Zhu (2013)) and increase in likelihood of self-employment (Aoki and Santiago (2017)). Improvements in language skills have also been shown to improve health outcomes including reduction in fertility, teenage pregnancy and age at which a woman has her first child (Aoki and Santiago (2015)).

The significant impact of English proficiency on individual's own outcomes suggests that if transmitted over generations, linguistic constraints can have important implications for the cultural assimilation and upward mobility of future generations as well[3].

In this paper, I evaluate the impact of immigrant parent's English proficiency on outcomes for second generation immigrants in the U.K. An important challenge in the estimation of parental language effect on children's outcomes stems from the endogeneity of language. Parent's language skills are correlated with various non-language factors such as income and education of the parents that also directly affect children's outcomes. Furthermore, reverse causality is a possibility given that English proficient children could also positively impact the language skills of their parents. Given these endogeneity issues, I identify the causal effect of parent's English proficiency on second generation outcomes by employing a difference in difference strategy using the first wave (2009) of the U.K Household Longitudinal Survey. I use as the explanatory variable, the interaction between age at arrival and dummy to indicate whether the parent is from a non-English speaking country of birth. This is based on the instrumen-

For marriage outcomes, see Chin and Bleakely (2004); Guven and Islam (2015); Gillian Stevens and Gray Swicegoo (1987); Alberto Dávila and Marie T. Mora (2001); Xin Meng and Robert G. Gregory (2005); Brian Duncan and Stephen Trejo (2007); Lazear (2007). For residential location outcomes, see Chin and Bleakely (2004); Edward Funkhouser and Fernando A. Ramos (1993); Maude Toussaint-Comeau and Sherry L.W. Rhine (2004) and Aoki and Santiago (2017) for a review. For a range of other social outcomes and review of language effect on health outcomes, see for instance Guven and Islam (2007) and Aoki and Santiago (2015).

[3]For intergenerational progress of immigrants, look at for example Card, DiNardo and Estes (2000), Grogger and Trejo (2002) and Smith (2003).

tal variable strategy suggested by Bleakley and Chin (2004; 2008; 2010) who employ the aforementioned interaction variable as the instrument. However, in this paper, due to measurement issues with the English proficiency variable in the U.K Household Longitudinal Survey, I only present reduced form estimates. The empirical strategy is based on the phenomenon that young children learn languages more easily than older children (Newport(1990), Lenneberg(1967)).

However, using only arrival age as the dependent variable and comparing early arrivers to later arrivers could be problematic as the outcomes of these two groups could differ due to non-language factors as well. For instance, early immigration could lead to better knowledge of the country's institutions and culture and in turn better economic outcomes among immigrants and their U.K born children. This problem is alleviated with the additional interaction of age at arrival with the country of birth dummy under the assumption that immigrants from both English and non-English speaking countries encounter similar (non-language) assimilation issues of cultural and institutional familiarity upon arrival. Then, any difference in outcomes between early and late arrivers that are due non-language factors will be the same for those from the two country groups and get differenced out with this additional interaction.

This paper contributes by looking at intergenerational language transmission in the context of a new country[4]. Additionally, the focus of this paper is on adult educational and attitudinal outcomes that have not previously investigated but have implications for assimilation and policy. Some of these implications are underlined by the vast literature that has explored these outcomes in other settings[5]. Gruber (2005) for example finds religious market density to have a positive impact on religious participation, reduction in welfare receipts, reduction in divorce rates and improvements in educational outcomes.

[4]Wang et al. (2017) examine returns to english proficiency in China. Azam et al. (2013) examine returns to English language skills in India.

[5]See for instance Rachel Mc Cleary and Rober J. Barro (2006) for a review of studies that study the interactions between religious beliefs and outcomes. See Sapienza (2013) for a review of studies on social attitude. Also see Handbook of Social Economics (2011) for a review of literature on identity formation.

Since Arrow (1972), there have also been innumerable papers demonstrating the impact of attitudes related to risk and trust on a country's economic development as well as on individual economic outcomes (Delhey and Newton (2003)). The tradeoffs between retaining own identity and cultural integration has been demonstrated in Abramitzsky, Boustan and Eriksson (2016).

Studies trying to understand the origin of social preferences have found the environment and home country culture to be important factors[6]. Others have found evidence of intergenerational transmission (Pattachini and Zenou (2016); Dohmen et al. (2012); Ljunge (2012); Algan and Cahuc (2010); Tabellini (2008)). This paper adds to the literature by providing evidence on parental language proficiency as another important channel through which such attitudes can be transmitted across generations.

The rest of the paper is organized as follows. Section 2.2 includes the literature review. Section 2.3 discusses the data and empirical strategy. Section 2.4 provides estimates of the first stage equation and impact of parental english proficiency on language skills of the second generation immigrants. Section 2.5 includes results for other outcomes. Section 2.6 discusses the channels and section 2.7 concludes.

2.2 Relevant Literature

There are several papers that study the effect of parental attributes on children's outcomes[7]. One strand of literature focuses on intergenerational transmission of income and wealth and finds evidence of intergenerational mobility (Becker and Tomes (1986); Solon (1992); Zimmerman (1992); Dearden, Machin, and Reed (1997); Bjorklund and Jantti (1997); Wiegand (1997)). In addition to income, studies also find evidence of parental education as an important explanatory factor in children's academic outcomes (Portes and MacLeod (1999); Card, DiNardo, and Estes (2000); Hirschman (2001); and Grogger and Trejo

[6] See Moshion and Tabasso (2014) for a review.

[7] Also see Guven and Islam (2015) for a review

(2002)). Currie and Moretti (2003), Black, Devereux, and Salvanes (2005) and Oreopoulos, Page, and Stevens (2006) focus on the intergenerational effects of parental education.

There is extensive literature on the impact of individual's English proficiency on their own outcomes. Studies that use OLS include McManus, Gould and Welch (1983), Kassoudji (1988), Tanier (1988) and Chiswick (1991). Angrist and Lavy (1997) use a policy change in the schooling system in Morocco to estimate the return to speaking French in the country. Other papers that attempt to address the endogeneity of language and earnings using other instrumental variable techniques include Chiswick and Miller (1995) and Dustmann and van Soest (2002). Bleakley and Chin (2003) are the first to exploit age at arrival as an instrumental variable and document a strong correlation between age at arrival and English proficiency and find a positive impact of English proficiency on own earnings. Bleakley and Chin (2008) show English proficiency to also have an impact on social assimilation, in particular an increase in the likelihood of divorce and intermarriage, decrease in fertility and, for some, decrease in ethnic enclave residence.

Among studies in the U.K, Miranda and Zhu (2013) estimate the effect of English deficiency on the native-immigrant wage gap for employees in the U.K using the first wave of the U.K Household Longitudinal Survey (Understanding Society). They employ Bleakley and Chin's instrumental variable strategy and find significant effects of English proficiency on the native-immigrant wage gap. As previously discussed, both Aoki and Santiago (2015, 2017) also use the Bleakely and Chin instrument to look at educational, health, social and labor market outcomes in the U.K. Other papers that look at the impact of English proficiency on earnings include Dustmann and Fabri (2003) who use the Family and Working Lives Survey (FWLS) as well as the U.K Fourth National Survey of Ethnic Minorities (U.K FNSEM). They find significant positive result on earnings using OLS. However, after employing an instrumental variable strategy, their results become insignificant[8].

Some papers have also looked at the determinants of language acquisition among children of immigrants. Portes and Hao(1998) for instance focus on the extent to which children of immigrants maintain their home country/heritage language. Others such as Hernandez and Charney (1998) and Suro and Passel (2003)) focus on the rate at which members of certain ethnic groups acquire English. Grogger and Trejo(2002) and Glick and White(2003) focus on children's outcomes but using their home language background as one of the explanatory variables. Chiswick, Lee and Miller (2005) using bivariate and multivariate probit analysis study the correlation in language acquisition and learning among household members using the Australian Census (1996).

Few papers have focused on intergenerational transmission of language. Casey and Dustmann (2008) using the German Socio-Economic Panel, study the transmission of language capital among immigrants and the effect of language deficiencies of parents on labor market outcomes of second generation immigrants in particular[9]. They rely on 'selection on observables' as their identification strategy and condition on characteristics related to family background that could potentially confound the impact of parental language and children's English proficiency and outcomes. They find that language deficiencies among second generation individuals leads to poorer labour market outcomes, but for females only. Bleakley and Chin (2010) using U.S Census data (2000) and two-stage least squares estimate the effect of parent's English speaking proficiency on children's schooling outcomes. They find positive and statistically significant effect on children's English-speaking proficiency and pre-school attendance as well as significant and negative results for dropping out of high

[8] To some extent this may be attributable to their small sample size (around 250 in FWLS and 920 in FNSEM).

[9] Also see for example Borland (2006), Houle (2011) and R Caminal et al. (2018) for a review and for intergenerational language transmission in Canada, among Australian migrant communities and Catalonia respectively.

school and being below age-appropriate grade. Although the main focus of the paper is on individual's own outcomes, Guven and Islam (2015) using a similar empirical strategy as Bleakely and Chin (2004; 2008; 2010) also report positive estimates of the effect of parent's English proficiency in Australia on the second generation's English proficiency, self reported math scores, probability of 'club' membership, probability of being a supervisor at a job and how social the individual is.

2.3 Data and Empirical Strategy

For a second generation immigrant i with parent j who arrived in the U.K at age c, I estimate the following equation:

$$y_{ijc} = \alpha + \beta_1(arrival\,age_{0-c} X non\,english\,speaking\,country_j)$$
$$+ \beta_2 arrival\,age_{0-c} + \beta_3 non\,english\,speaking\,country_j + X'_{ijc}\beta_4 + \omega_{ijc} \quad [1]$$

where y_{ijc} is the outcome of interest for the U.K born individual i. The arrival age dummy is given by $arrival\,age_{0-c}$ that equals 1 if a parent arrived prior to age c and 0 otherwise. $non\,english\,speaking\,country_j$ is a dummy that takes the value 1 if immigrant parent j is from a non English speaking country of birth. X'_{ijc} is a vector of individual controls such as parent's and child' age and sex. η_c is a full set of age at arrival dummies and γ_j is a full set of country of birth dummies.

The empirical strategy takes advantage of the 'sensitive' period phenomenon' for language acquisition. The hypothesis is that with the onset of puberty, the ease of learning an additional language declines significantly. Thus, children who are exposed to a second language when they are young and in their critical period will have higher proficiency in that language compared to those with later exposure[10]. Looking only at the difference between young and later arrivers from non English speaking countries could be endogenous. Additional years since arrival that impact language proficiency are also correlated with higher assimilation for instance through a greater familiarity with a country's institutions. However, individuals from English speaking countries presumably encounter similar such non language familiarity issues. Then, any difference in the outcomes between late and early arrivers in U.K that is unrelated to English proficiency should be the same for those from non English speaking as well English speaking countries. Under this assumption, interaction of arrival age with a dummy for non English speaking country will difference out any non language effects and provide a causal effect of parent's English proficiency on second generation immigrant outcomes.

I implement the strategy using the first wave of the U.K Household Longitudinal Survey (Understanding Society). Understanding Society is a survey of 32,000 households conducted between 2009-2015. Additional 8000 households from the British Household Panel are added in wave 2. The survey over samples ethnic minorities with a total of 4000 households from among ethnic minorities. The main survey interviews individuals over age 16 and these individuals can be matched with their parents as long as they are living in the same household. I match individuals with their parents and end up with 5968 unique individual-parent matches. 10% of this sample are individuals born abroad. Finally after dropping missing values, keeping immigrant parents only, keeping only one parent when both parents are present, I end up with 970 unique matches of individuals born in the U.K and parents born abroad. Given measurement errors and missing values in outcome variables, the sample size might vary across regressions for different outcomes.

In terms of arrival age c, since there is not one exact critical period, I will discuss estimates based on a range of age cut offs from 5-11. Following Bleakely

[10]This phenomenon is discussed in Newport(1990) and Lenneberg(1967).

and Chin (2004; 2008; 2010), I keep immigrant parents who arrive in the U.K when they were less than age 18[11]. This alleviates the issue of selection into the timing of migration since for childhood immigrants this decision was presumably taken by adults in the household and was not a choice for them. I define English speaking countries as those where English was a dominant or official language. Immigrants from these countries are likely to have had exposure to English prior to arriving in the U.K. List of English speaking countries include U.S.A, Australia, Republic of Ireland, Canada, New Zealand, Kenya, South Africa, Nigeria, Uganda, Ghana, Pakistan, India and Jamaica. List on non-English speaking countries include Germany, Italy, Bangladesh, Sri Lanka, Cyprus, China, Hong Kong, Spain, Poland, Turkey and France.

2.4 English Proficiency

2.4.1 Graphical Evidence

Individuals are asked if they find difficulty in English while speaking day to day, reading, speaking on the phone or while filling forms. However, based on these questions, there are almost negligible number of second generation individuals and after keeping childhood immigrants, almost negligible number of parents, who report any difficulty in English.

Individuals are also asked if they consider English to be their first language or a second/additional language. Given measurement issues in the previous questions, I follow Alfonso and Zu (2013) and create a dummy = 1 if individuals report English to be their second language and 0 otherwise.

In figure 2.1, I provide graphical evidence of age at arrival effect on English proficiency. The figure is based on the parent sample only. The x axis includes parent's age at arrival in the U.K for immigrant parents and the $y$axis includes the average number of individuals who report no difficulty in speaking

[11]In some cases, I report estimates based on even younger childhood immigrants. Unless otherwise specified, the sample is based on parents who moved when they were less than 18 years of age.

day to day English. The figure demonstrates that for immigrant parents from non- English speaking countries, English proficiency as measured by the ease of speaking the language on a daily basis declines significantly as age of arrival increases. For those from English speaking countries, who would have had exposure to English prior to immigrating, reported ease of speaking English does not change with arrival age. The plot for the English speaking group is nearly flat.

Figure 2.2 shows how the mean probability of reporting English as a second language for U.K born individuals varies with their parent's arrival age. This probability increases for those whose parents immigrated as children from non-English speaking countries and as before remains almost unchanged for those with parents from English speaking countries[12].

2.4.2 Estimates

I keep mother's information when that of both parents is available. Estimates of the first stage equation 2 are shown in table 2.3. The dependent variable is a dummy that takes the value of 1 if parents report English to be a second language. I regress this on the interaction between a dummy to indicate a non-English speaking country of birth and arrival age dummy. Arrival age dummy takes the the value of 1 if the individual immigrated before a certain age cut off. I allow the age cut off to vary from 5 to 11 in columns 1 to 6. I keep parents who moved to the U.K before turning 18. Under the identifying assumption that non- language age at arrival effects are similar for individuals from English and non-English speaking country of birth, the interaction term measures parent's English proficiency. Note that the number of countries in the sample are limited and standard errors are large when clustering. Panel A reports results with clustering at parent country of birth level and panel B reports them without clustering.

First stage estimates are negative and statistically significant for all age cut

[12]Note that these plots are not regression adjusted.

offs. Individuals from non-English speaking countries are also more likely to report English as a second language as indicated by the positive and significant coefficient for the non-English speaking country of birth dummy. Although when standard errors are clustered, the estimates are significant only for arrival age cut offs, 10 and 11. Those who arrive at younger ages in the U.K are also less likely to report English as a second language.

In table 2.4, I report estimates of the effect of parent English proficiency on the probability of the child reporting English as a second language. I find the effects to be negative and statistically significant for age cut offs 10 and 11. Increase in parent's English proficiency reduces the reported probability among second generation individuals by approximately 9 percentage points. According to the OLS estimates, the probability of reporting English as a second language for those whose parents also report English as a second language is 14.5 percentage points higher relative to those who report English as a first language.

2.5 Other Outcomes

'English as a second language' is an imperfect measure of English proficiency and for that reason, I will provide reduced form estimates in this paper. Wherever possible, I provide additional graphical evidence to attribute the results to English proficiency. Specifically, I regress second generation outcomes on parent's arrival age (with controls for age and sex of first and second generation immigrants) for both English and non-English speaking countries. I plot these regression adjusted estimates by parent arrival age and country of birth. The intuition is that if the effect of parent arrival age on outcomes is driven by language only, the plot of the estimates for English speaking countries by arrival age should be flat. I should also observe an increase in differences between estimates of English and non-English speaking countries with increase in arrival age.

However, there could be other non-language effects as well due to which early arrivers could differ from late arrivers from both type of countries. In that case, the regression adjusted plot of estimates by parent arrival age for those from English speaking countries need not be flat. But I should still observe a few patterns to determine the validity of the identifying assumption.

Note that the identifying assumption is that non-language arrival age effects should be the same for individuals from the two country groups. We also know that the difference in outcomes attributable to language increase with arrival age. Then for early arrival ages, the difference in outcomes (due to either language, non-language factors or both) between those from English and non-English speaking countries should be smaller than the differences for later arrival ages. Secondly, the plot for anglophone countries should at least trend in the same direction as that of non-English speaking countries. An upward sloping plot of estimates for one country group and downward sloping plot for another seems to suggest a positive non-language age at arrival effect for the former and a negative non-language age at arrival effect for the latter. Such a plot would imply a violation of the identifying assumption.

Finally, I also explore what those non-language age at arrival effects might be. In some cases, controlling for one or two variables immediately leads to an almost 0 slope in the plot for English speaking countries. I know then that after controlling for those variables, the estimates can be attributed to language. In fact, in so far as those variables have similar age at arrival effects for English and non-English speaking countries, they will get differenced out in the actual regression. But as an additional robustness check, I provide estimates for such cases, with and without those controls.

2.5.1 Risk Preparedness

Individuals are asked the following question - *"On a scale of 0-10, are you generally a person who is fully prepared to take risks or do you try to avoid taking risks; 0 if avoid taking risks and 10 if fully prepared to take risks ?"*

Estimates of the impact of parent's English proficiency on the risk preparedness of their U.K born children is reported in table 2.5. I find an increase in parent's English proficiency to have a negative impact on the degree to which individuals report being prepared to take risk. In columns 1a and 1b, dependent variable is defined as a dummy=1 if individuals report the degree to which they are prepared to take risks to be > 7. For arrival age dummy $= 1$ if individuals moved before age 10, increase in parent' english proficiency reduces the probability of reporting risk preparedness of greater than 7 by about 26 percentage points. The estimates are similar (around 27 percentage points) when age 11 is used to define parent's arrival age dummy. In columns 2a and 2b, I allow the dependent variable to be continuous. Based on arrival age cut off 10, an increase in parental English proficiency reduces risk preferences among children by 1.128%. Similar estimates for cut off 11 is about 1.14%. I find these effects to be significant only when I keep the father's information for cases when both parent's information is available.

A possible story to explain the results is that improved outcomes (better income and educational outcomes for example) associated with English proficiency also increase aversion to making risky choices that might alter those outcomes. This is in line with evidence in the literature that finds an increase in education to have a positive impact on risk aversion (Seeun Jung (2015)).

Figure 2.3 plots the regression adjusted estimates (based on the dummy definition) for English and non-English speaking country of birth. There is almost no age at arrival effect for those whose parents are from English speaking countries. Among those from non-English speaking countries, later arrivers are more likely to report a higher degree of risk preparedness. The plot suggests that age at arrival effect on risk preparedness is largely driven by English proficiency. If other dominant non-language effects were driving the estimates, the plot for those from English speaking countries may have been less flat.

2.5.2 Importance of being British

Tables 2.6 and 2.7 provide estimates of the impact of parental English proficiency on the extent to which individuals identity themselves as British using the following question- *"On a scale of 0 to 10 where 0 means 'not at all important' and 10 means 'extremely important,' how important is being British to you ?"*

I find the effects to be significant only when I keep father's information for cases when both parent's information is available. For arrival age dummy defined as 1 if parents arrived at age 7 or earlier, I find that an improvement in English proficiency by 1 unit increases the importance of being British by 2%. Similar estimates for cut offs 8 and 9 are 1.63% and 1.48% respectively.

However, the plot of estimates from the regression of the dependent variable by parent arrival age for English and non-English speaking country is noisy and uninformative.

Next, I replace the previous dependent variable with a dummy variable that takes the value of 1 if individuals report the importance of being British to be at least 5. Interestingly, I find graphical evidence of the English language effect when I: (1) focus on a sample of parents who moved to the U.K when they were less than 17 years of age and (2) include controls for parental income and qualification (in addition to parent and child's age and sex) in the regression of the Dep. Var. on parent arrival age. Figure 3 plots these estimates for English and non-English speaking country of birth. For English speaking countries, probability of reporting importance of being British among second generation immigrants does not change with arrival age. For those from non -English speaking countries, the stated probability declines.

Although not included here, I also plot the estimates based on a regression without controls for parental income and qualification. Doing so suggests that these variables could have different age at arrival effects for English and non-English speaking countries. To account for that, I provide estimates with and without these additional controls in table 2.7, panel A. The results are qualitatively similar to estimates in table 2.6 but only for arrival age cutoff 7 and younger. Increase in English proficiency increases the probability of reporting importance of being British of 5 and above by 31 percentage points without controls for parental income and qualification and 35 percentage points after the controls are included. As shown by panel B of table 2.7, estimates are not significant when I define the dependent variable as continuous (but keep the same sample as in panel A or figure 2.4).

2.5.3 Trustworthiness of Others

Language proficiency of the parents by shaping their experiences and that of their children might also impact the degree of trust individuals have in others. In this section, I estimate the effect of parent's English proficiency on the probability that their U.K born children answer 'yes' to the question of trust in people. Specifically, the outcome variable is a dummy = 1 if individuals say that most people can be trusted and 0 if they say that it depends or that one cannot be too careful.

I keep mother's information when that of both parents is available. I estimate the effect of parental English proficiency on trust based on arrival age dummy that takes the value of 1 if the parents who moved to the U.K were 10 and younger or 11 and younger. Columns 1 and 2 of table 2.8 include estimates for the two arrival age cutoffs. I find that a unit increase in English proficiency reduces the probability of reporting high trust in people by about 30 percentage points. To the extent that English proficiency might improve overall understanding of society better, the negative effect may be driven by individuals who report a more balanced answer and say that the degree of trust depends. Although this could also be the result of a drop in sample size, estimates fall in magnitude and significance when individuals who report 'it depends' are dropped from the sample.

As further graphical evidence of the language effect, figure 2.5 shows regression adjusted estimates by parent's arrival ages for immigrant parents from English and non-English speaking countries. After controlling for parent's qualification, I find that the increase in arrival age seems to have a negligible impact on the mean probability of reporting trust in others for those with parents from English speaking countries and a positive impact for those from the latter.

If the arrival age effect of parent's qualification is not similar for individuals with parents from English and non-English speaking countries, the interaction of arrival age dummy and non English speaking country of birth dummy would be endogenous and estimates would be biased. In that case, estimates with and without parental qualification would differ. To check, I first report estimates without controls for parent's qualification in columns 1 and 2 of table 2.8. In columns 3 and 4, I report estimates with parent's qualification as an additional control and find the estimates to be qualitatively robust and quantitatively similar to those in the first two columns. For parent arrival age dummy = 1 if between 0-10, improvement in English proficiency reduces the probability of reporting 'high degree of trust in others' by 30 percentage points. For arrival age cut off 11, the probability reduces by 33 percentage points.

2.5.4 Educational Outcomes

To evaluate the effect on educational outcomes, I look at school leaving age for those who no longer report to be in school. This includes children less than equal to 18 years who dropped out before the high school completion age of 18 and older individuals who are well beyond their schooling years. I also look at the probability of pursuing higher education by defining a dummy = 1 if individual is currently in university or reports his or her highest qualification to be a university degree such as a Masters or above.

Results are shown in table 2.9. English proficiency has a statistically significant and positive impact on both years of schooling as well the probability of pursuing higher education. For years of schooling, the estimates are significant only if I use a cut off of 9 to define arrival age dummy. Parental qualification appears to be an important channel to explain the effects on school leaving age. As indicated by column 2 of panel B, estimates lose their significance when parental qualification is controlled for.

An increase in English proficiency also increases the likelihood of pursuing higher education by about 44 percentage points when arrival age dummy is defined as 1 if parents arrived before age 7. Estimates are similar for arrival age dummy = 1 if between 0-8. The effect drops to about 37 percentage points for cut off 9 and to 23 percentage points for cut off 10.

The estimates are only significant when I keep mother's information for households with information on both parents. This suggests that for educational outcomes, mother's English proficiency matters more.

2.5.5 Importance of Religion

Individuals are asked to report on a sale of 1-4, if they consider religion to make a big difference to life. The variable takes the value of 1 if individuals report that religion makes no difference to life and 4 if they report that it makes a great difference to life. Estimates based on this outcome variable are reported in table 2.10.

In Panel A, I define the dependent variable as a dummy that takes the value of 1 if individuals report that religion makes a great difference to life. Columns 1 and 2 are based on father's information for cases when information on both parents is available. In columns 3 and 4, I keep mother's information. Estimates are significant only for younger arrival age cut offs, of 5 and 6 years. Keeping father's information and arrival age dummy = 1 if parents moved before age 5, I find that a 1 unit increase in English proficiency reduces the probability of reporting that religion makes a great difference to life by about 20 percentage points. Similar estimates for the case when I use 6 years as the arrival age cut off to define the arrival age dummy is 30 percentage points. Keeping mother's information and similar arrival age dummy definitions, I find that an improvement in English proficiency reduces the probability of reporting that religion makes a great difference to life by about 46 percentage points.

Figure 2.6 plots estimates from the regression of this outcome variable on parent arrival age for English and non-English speaking countries. I plot estimates for the sample with mother's information as well as the sample with father's information among cases with both parent's information. Increase in arrival age increases the probability of reporting religion to have a great influence to life for individuals with parents form both English as well as non-English speaking countries. The upward sloping line for the non-English speaking country case suggests that there is an important non-language arrival age effect as well. It is unclear what that effect might be. However in so far as this effect is the same for English and non-English speaking countries, the estimates will be valid. The fact that the lines are at least trending in similar directions and that the difference in outcome for the two types of countries increase with arrival age lends some support to the validity of the identifying assumption.

In Panel B of table 2.8, I allow the dependent variable to be continuous and between 1-4. Estimates are qualitatively similar when I keep mother's information as in columns 3 and 4. An increase in English proficiency reduces the reported importance of religion by about 1.12% when arrival age of 5 is used as a cut off to define the arrival age dummy and by 1.09 percent when arrival age of 6 is used. When father's information is kept, the estimates are significant and negative only for the case when arrival age of 5 is used as cut off.

2.5.6 Level of Political Interest

Level of interest in politics is defined on a scale of 1-4 with 1 indicating no interest and 4 indicating a high level of interest. I find a statistically significant and negative effect of English proficiency on political interest. Estimates are shown in table 2.11. In panel A, I keep father's information when both parents are present and in panel B, I keep mother's information. In column 1, I define arrival age dummy as 1 if the parent moved before age 5 and in column 2 as

1 if the parent moved before age 6. Keeping father's information, I find that for a 1 unit increase in English proficiency, level of interest in politics declines by .9% at arrival age cut off 5 and by 1.155% for cut off 6. Keeping mother's information, the decline in political interest is 1.1.5% and 1.3% for cut offs 5 and 6 respectively.

A plot of the regression estimates suggests that age at arrival effect of parental income and qualification on political interest is also important. After controlling for these variables, I find that age at arrival has almost no impact on political interest for individuals with parents from English speaking countries and a positive impact on those from non-English speaking countries.

As long as the arrival age effects of parental income and qualification are similar for English and non-English speaking countries, estimates will be valid. But to be sure, I also provide estimates by including controls for the same. These estimates are included in columns 2 and 3 of table 2.11. I find that when I keep father's information (panel A), the estimates are no longer statistically significant. However, when I look at the sample with mothers as in panel B, I find the estimates to be qualitatively robust to the inclusion of these additional controls. This seems to suggests that father's income and qualification is an important channels to explains the effect of English proficiency on political interest due to which the effect shows up in a sample with mostly fathers.

One might also argue that the arrival age effects of income and qualification are different for English and non-English speaking countries and the estimates are biased. However, there is no particular reason for the estimates to be robust to the inclusion of these controls for the sample with mother's information but not father's.

Regardless, I do find statistically significant effects on political interest with and without additional controls at least for the sample with mother's information. Figure 2.7 lends support to the role of English proficiency in explaining the results.

2.5.7 Job Satisfaction

Job satisfaction is defined on a scale of $1 - 7$ where the outcome variable takes the value of 1 if individual is completely dissatisfied and 7 if completely satisfied. Figure 2.8 plots the regression adjusted mean probability of reporting job satisfaction of 6 or above against parent arrival age for those from English and non-English speaking countries. The negligible arrival age effect for English speaking countries and a negative effect for non English speaking countries, suggest that the estimates can be attributable to English proficiency and that the age at arrival effect of non-language factors on job satisfaction is not dominant.

I report estimates for job satisfaction in table 2.12. In column 1, I define arrival age dummy as 1 if parents moved before age 5 and in column 2 as 1 if they moved before age 6. In both cases, I find that a unit increase in English proficiency increases job satisfaction by about 1.138%. Estimates are lower for arrival age cut offs 8 and 9. The effect of English proficiency on job satisfaction is positive but only by about .80% and .87% for the higher arrival age cut offs. For job satisfaction, father's English proficiency seems to be more important. The estimates are not significant when mothers information is kept in cases when that of both parents is available.

2.6 CHANNELS

Here I discuss channels that could possibly explain the language effects on various outcomes. I find that most of the effects are mediated through three important channels-parent region of residence, educational qualification and job status. Parental job status refers to the economic activity that parents are currently involved in. This includes self employed, paid employed, unemployed, retired, maternity leave, full time student, long term sick/disabled, in government training, unpaid worker in family business or looking after family/home.

As table 2.13 shows, the effect of English proficiency on the probability of reporting being British as important is driven by parent's region of residence. Table 2.13 includes all the same estimates as in table 2.6 and 2.7 but with a full set of parent region of residence dummies. None of the effects are now significant.

Region of residence also seems to be an important channel to explain the effect of English proficiency on political interest. Recall that controlling for parental income and qualification diminished the effect of English proficiency on political interest only for the case when father's information was kept among two parent households. Table 2.17 provides estimates with controls for region of residence for the sample with mother's information. The estimates are no longer significant. This suggests that father's income and qualification and mother's region of residence are important channels in explaining the effect of English proficiency on the level of political interest among second generation immigrants.

The effect of English proficiency on the probability of pursuing higher education among second generation is mediated through both parent's residence as well as parental qualification. Table 2.15 provides estimates with these additional controls. For cases when arrival age dummy is defined based on cut off 7 and 8, the estimates are not statistically significant. For cut offs 9 and 10, estimates of the effect of English proficiency on the probability of pursuing higher education are still statistically significant but reduce in magnitude from about 37 and 32 percentage points respectively to about 25 percentage points.

Parental qualification and job status seem to explain the effect of English proficiency on job satisfaction. Estimates for job satisfaction after controlling for these two variables are shown in table 2.18. English proficiency no longer has a statistically significant effect on job satisfaction. This is consistent with a simple story of intergenerational transmission of inequality. Individuals born to parents with lower qualification might also experience lower occupational and educational outcomes and in turn lower overall job satisfaction.

Table 2.14 and 2.16 demonstrate that parental job status is an important channel in explaining language effects on reported degree of 'high' trust in others and also on how important second generation immigrants consider religion to be. There is evidence in the literature that finds a decline in religiousness with economic development and education[13]. To the extent that parent's English proficiency improves their economic status and lowers their religiousness, children in such households may not be socialized to strong religious beliefs. This could explain the negative parent proficiency effect on the probability of reporting religion to be of great importance to life and the fact that the effect disappears when parent's economic status is controlled for.

Note that the decline in the probability of reporting high trust in people with improved English proficiency and job status may not necessarily reflect a decline in preference for trust in others but the (less-naive) belief that trust in people should depend on situations / circumstances (Paola Sapienza Anna Toldra Luigi Zingales (2013))[14]. This belief may come from a larger set of experiences that English proficient parents with better economic opportunities might have, that they transmit to the next generation (Dohmen at al. (2011)).

For risk preparedness, the channels are not clear. Perhaps a better measure of risk preferences and a larger data set is required to be able to investigate the channels for risk related outcomes better.

2.7 CONCLUSION

I estimate the causal effect of first generation's English proficiency on outcomes of the second generation in the U.K. This is the first study to look at adult outcomes such as the probability of enrolling in higher education and job satisfaction. I find parent's English proficiency to have a positive and significant impact on years of schooling, higher education enrollment and job satisfaction.

[13] See Thomas Buser(2015) for a review.

[14] recall the trust variable takes the value of 1 if most people can be trusted and 0 of they cannot or if it depends.

This paper also contributes to the literature by demonstrating the impact of language on social preferences and attitudes. I find English proficiency to be associated with lower trustworthiness and lower risk preparedness among second generation immigrants. Interestingly, I also find a decline in political interest with parent's English proficiency. Religion is perceived to be less important and probability of reporting being British as important is higher among those with English proficient parents. With the exception of risk preparedness, language effects on all other outcomes can be explained by region of residence, educational qualification and economic status of the first generation.

Although the data is not without its limitations, the results point to important interdependencies between language and intergenerational outcomes that can persist even among adults. By impacting attitudes and identity, these effects can have potentially large and long run implications for immigrant assimilation. Therefore, from both a short run and long run policy perspective, investments in institutions and programs that increase English proficiency are worthwhile goals.

2.8 Figures

Figure 2.1: Parent English Proficiency and Parent Arrival Age

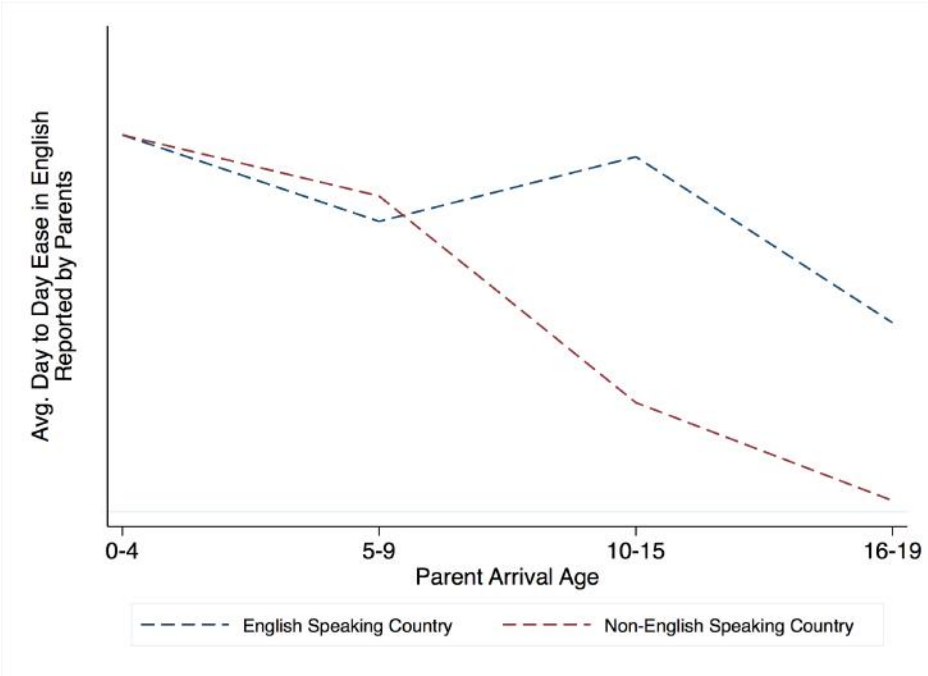


Figure 2.2: Mean Probability of reporting English as First Language among Second Generation Immigrants and Parent Arrival Age

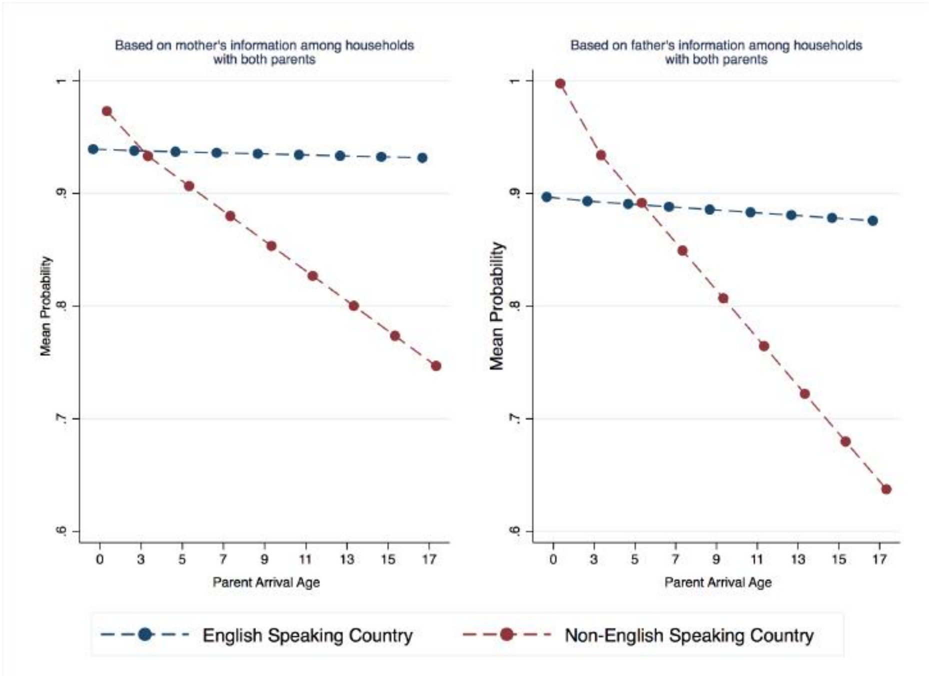
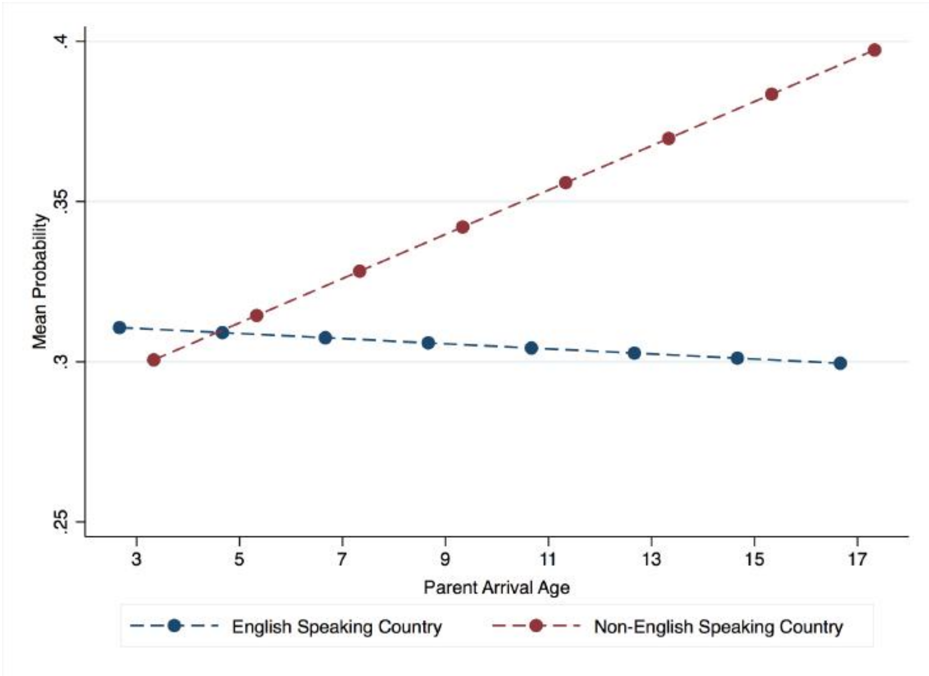
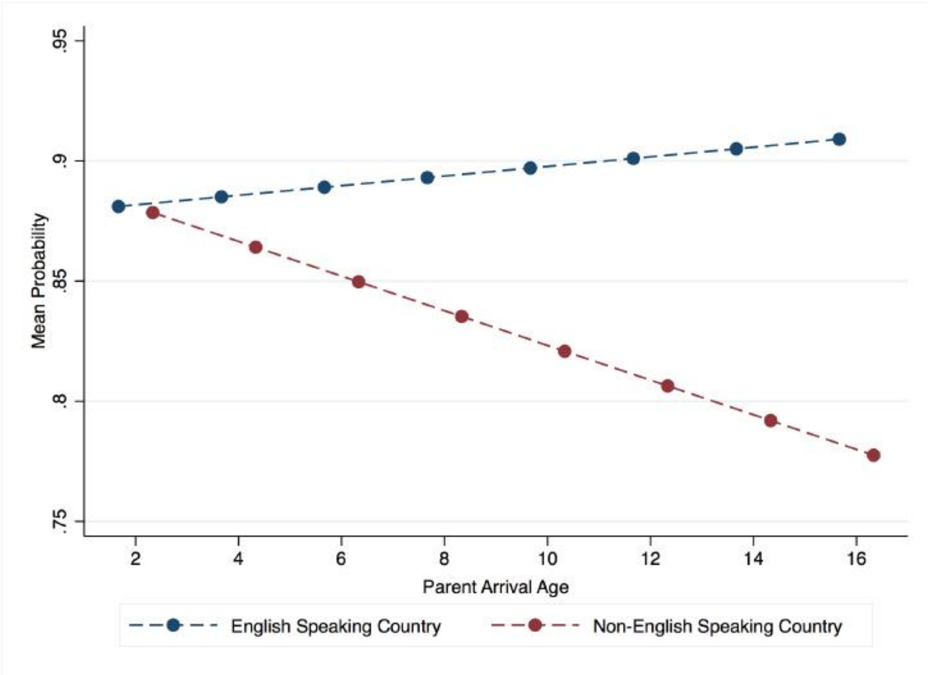


Figure 2.3: Regression Adjusted Mean Probability of reporting Risk Preparedness among Second Generation Immigrants and Parent Arrival Age



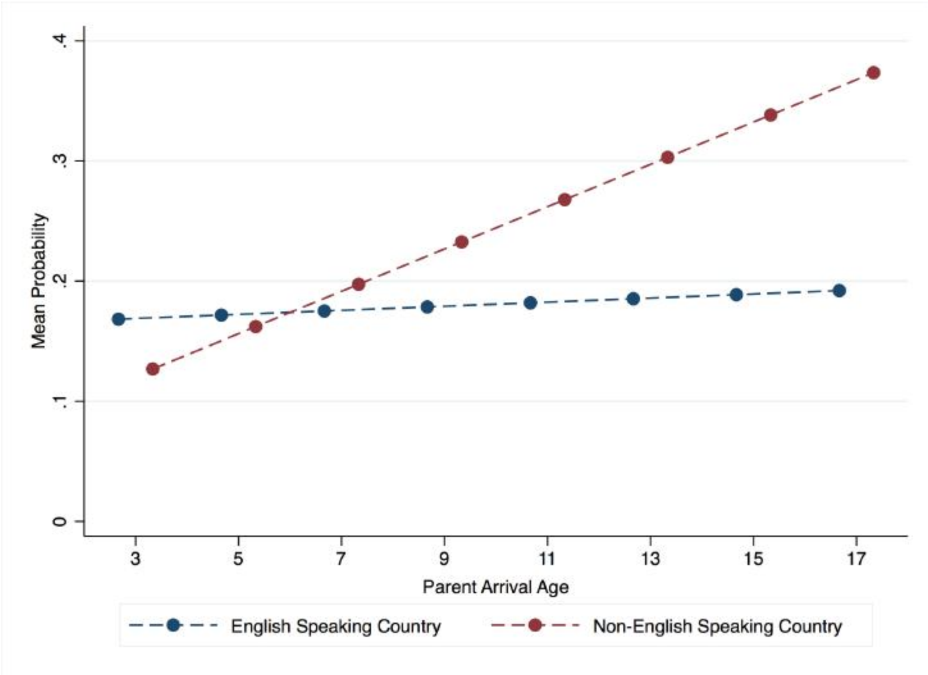
Note: 'Risk Preparedness' is defined as a dummy = 1 if the degree to which individuals report being prepared to take risk is greater than 7. The measure is on a scale of 0-10, with 0 being the least prepared and 10 being the most prepared. Sample is based on keeping father's information when both parent's information is provided in the household. Regression based on which the probabilities are plotted includes controls for second generation age and sex and parent's age and sex.

Figure 2.4: Regression Adjusted Mean Probability of reporting British Identity to be Important among Second Generation Immigrants and Parent Arrival Age



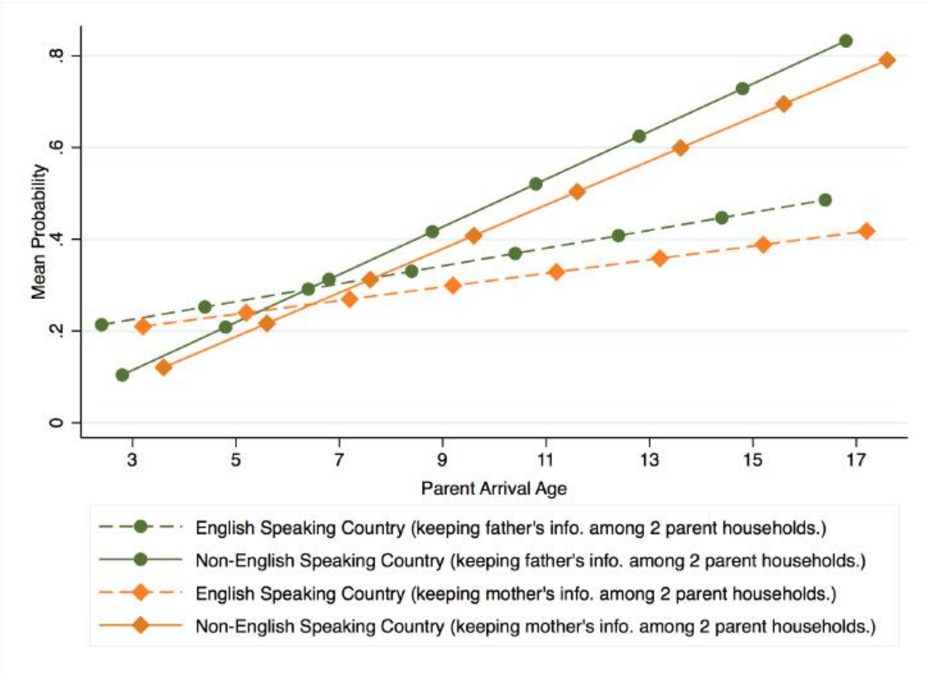
Note: 'Importance of being British' is defined as a dummy = 1 if individuals on a scale of 0-10 report the importance to be >= 5. 0 implies that being British is not at all important and 10 is very important. Sample is based on keeping father's information when both parent's information is provided in the household, age 18 and over and parents who moved to U.K at age 16 and earlier. Regression based on which the probabilities are plotted includes controls for second generation age and sex and parent's age, sex, income and qualification.

Figure 2.5: Regression Adjusted Mean Probability of reporting High Trustworthiness among Second Generation Immigrants and Parent Arrival Age



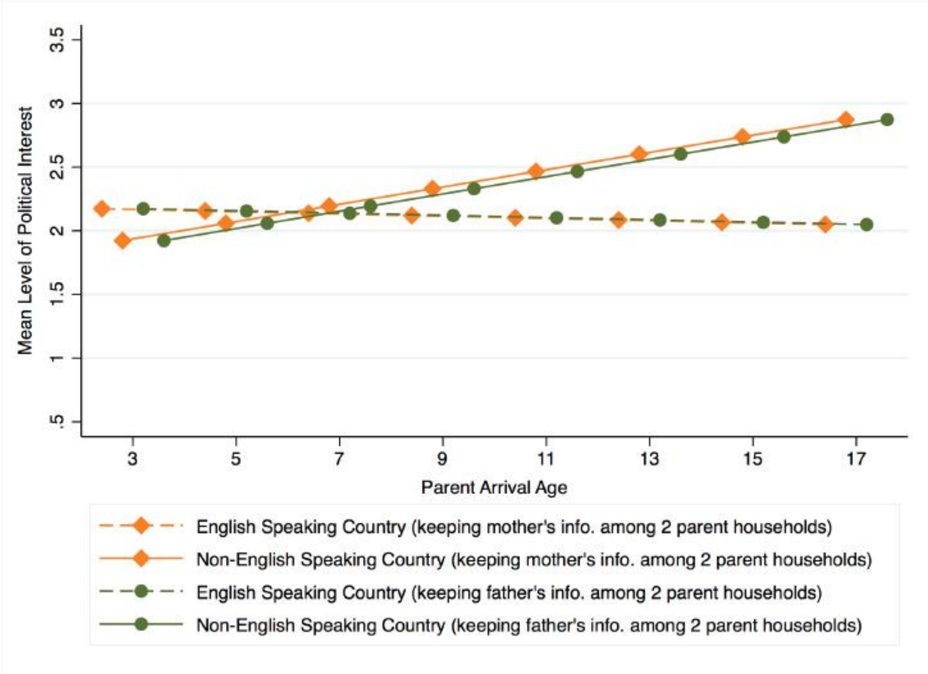
Note: 'High Trustworthiness' is defined as a dummy = 1 if individuals report that most people can be trusted. It takes the value 0 if individuals report that it depends or if they report that most people cannot be trusted. Sample is based on keeping mother's information when both parent's information is provided in the household. Regression based on which the probabilities are plotted includes controls for second generation age and sex and parent's age, sex and qualification.

Figure 2.6: Regression Adjusted Mean Probability of reporting Religion to be of Great Importance to Life among Second Generation Immigrants and Parent Arrival Age



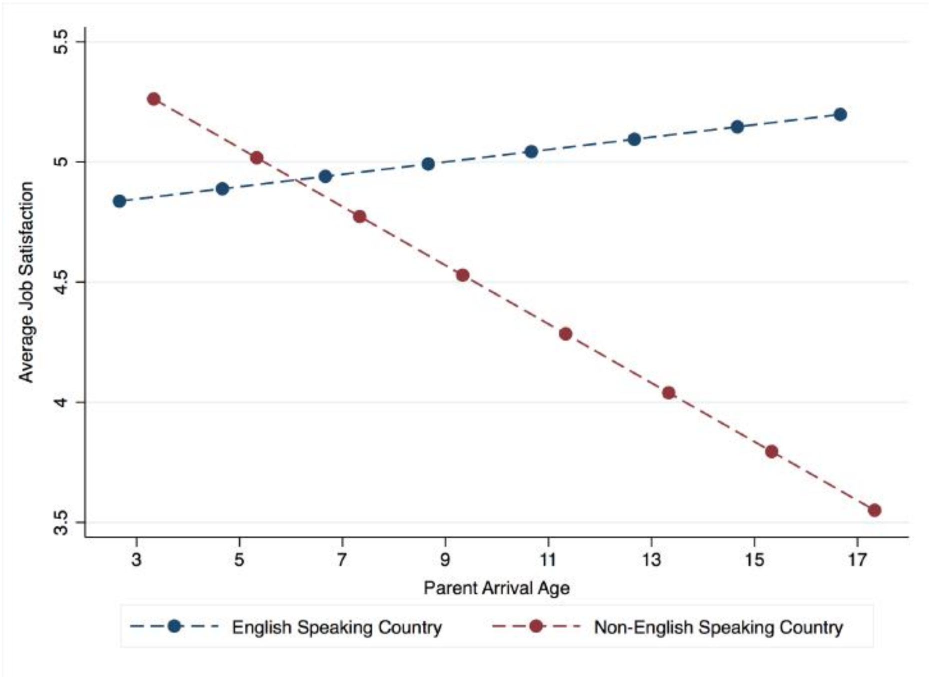
Note: The outcome variable is defined as a dummy = 1 if individuals report that religion makes a great difference to life and 0 if they report that it makes some to no difference at all. Sample is based on individuals 18 years and over. Regression based on which the probabilities are plotted includes controls for second generation age and sex and parent's age and sex.

Figure 2.7: Mean Reported Interest in Politics



Note: 'Political interest' is defined on a scale of 1-4. 1 implies no interest in politics at all and 4 implies a high level of interest.and 4 being the most interested. Regression based on which the probabilities are plotted includes controls for second generation age and sex and parent's age, sex, income and qualification.

Figure 2.8: Average Degree of Reported Job Satisfaction



Note: 'Job Satisfaction' is defined on a scale of 1-7 with 1 being the the lowest level of satisfaction and 7 indicating complete satisfaction. Sample is based on keeping father's information when both parent's information is provided in the household. Regression based on which the probabilities are plotted includes controls for second generation age and sex and parent's age and sex.

2.9 Tables

Table 2.1: Summary Statistics, First Generation Immigrants

	Mean	S.D	Mean	S.D
Age	51.0	9.43	53.0	9.41
Female	0.83	0.38	0.41	0.49
Years Since Arrival	32.1	9.96	33.8	10.3
Arrival Age in U.K	18.9	8.85	19.2	8.90
Arrival Age (0-5)	0.10	0.30	0.079	0.27
Arrival Age (0-6)	0.11	0.31	0.091	0.29
Arrival Age (0-7)	0.12	0.33	0.11	0.31
Arrival Age (0-8)	0.14	0.35	0.12	0.33
Arrival Age (0-9)	0.17	0.37	0.14	0.35
Arrival Age (0-10)	0.20	0.40	0.16	0.37
Arrival Age (0-11)	0.21	0.41	0.19	0.39
Non English Speaking Country	0.30	0.46	0.30	0.46
Parent has a Masters/Higher Degree	0.12	0.33	0.11	0.32
Paid Employee/Self Employed	0.39	0.49	0.50	0.50
Parent Monthly Personal Gross Income	1073.9	1320.0	1433.2	1728.2
English 2nd language	0.72	0.45	0.70	0.46

Note: Statistics are based on immigrant parents. Columns 1 and 2 are based on mother's information when both parents are present. Columns 3 and 4 are based on father's information when both parents are present.

Table 2.2: Summary Statistics, Second Generation Immigrants

	Mean	S.D	Mean	S.D
Age	22.4	6.66	22.4	6.66
Female	0.47	0.50	0.47	0.50
English 2nd language	0.21	0.41	0.21	0.41
Risk Preference; scale (0-10) (0: Avoid Risks; 10: Fully prepared to take risks)	0.30	0.46	0.30	0.46
Imp. of being British; scale (0-10) (0: Not imp. at all; 10: Extremely Imp.)	7.80	2.08	7.80	2.08
Can trust most People	0.16	0.37	0.16	0.37
Attended University	0.20	0.40	0.20	0.40
School Leaving Age	0.85	0.35	0.85	0.35
Religion Makes a difference to life scale (1-4); (1: great diff.; 4: no diff)	0.85	0.36	0.85	0.36
Interest in Politics; scale (1-4) (1: very interested; 0: no interest)	0.85	0.36	0.85	0.36
Job Satisfaction; scale (1-7) (1: fully dissatisfied; 7: fully satisfied)	0.37	0.48	0.37	0.48

Note: Statistics are based on second generation immigrants in U.K. Columns 1 and 2 are based on mother's information when both parents are present. Columns 3 and 4 are based on father's information when both parents are present.

Table 2.3: Parent English Proficiency and Parent Arrival Age

Dep. Var.: Dummy=1 if parents report English as 2nd language

Panel A	0-5	0-6	0-7	0-8	0-9	0-10	0-11
Arrival Age(0-c)X Non English Speaking Country	-0.421**	-0.382**	-0.319**	-0.279**	-0.313***	-0.373***	-0.407***
	(0.148)	(0.138)	(0.105)	(0.0771)	(0.0758)	(0.0958)	(0.0960)
Non English Speaking Country	0.248	0.235	0.197	0.162	0.144	0.177**	0.187**
	(0.157)	(0.158)	(0.139)	(0.134)	(0.123)	(0.0656)	(0.0675)
Arrival Age(0-c)	-0.320**	-0.332**	-0.345**	-0.402**	-0.394***	-0.411***	-0.367***
	(0.143)	(0.141)	(0.134)	(0.116)	(0.0866)	(0.0744)	(0.0762)
Panel B	0-5	0-6	0-7	0-8	0-9	0-10	0-11
Arrival Age(0-c)X Non English Speaking Country	-0.421***	-0.382**	-0.319**	-0.279**	-0.313***	-0.373***	-0.407***
	(0.120)	(0.120)	(0.106)	(0.0958)	(0.0916)	(0.0859)	(0.0868)
Non English Speaking Country	0.248**	0.235**	0.197**	0.162*	0.144*	0.177**	0.187**
	(0.0912)	(0.0930)	(0.0884)	(0.0828)	(0.0834)	(0.0823)	(0.0841)
Arrival Age(0-c)	-0.320***	-0.332***	-0.345***	-0.402***	-0.394***	-0.411***	-0.367***
	(0.0638)	(0.0631)	(0.0615)	(0.0570)	(0.0593)	(0.0594)	(0.0622)
Parent Age	✓	✓	✓	✓	✓	✓	✓
Parent Sex	✓	✓	✓	✓	✓	✓	✓
n	315	315	315	315	315	315	315

Note: Standard errors are in parentheses. In panel A, standared errors are clustered on parent country of birth. In panel B, standard errors are not clustered. I Keep mothers among two parent households. I keep parents who moved to the U.K before age 18. * (p<.10), ** (p<.05), *** (p<.01).

Table 2.4: English Proficiency Among Second Generation and Parent Arrival Age

Dep. Var.: Dummy=1 if individual reports English as 2nd language	0-10	0-11
Arrival Age(0-c)X Non English Speaking Country	-0.0931*	-0.0951*
	(0.0533)	(0.0510)
Non English Speaking Country	0.0682	0.0663
	(0.0747)	(0.0717)
Arrival Age(0-c)	-0.0809*	-0.0819**
	(0.0389)	(0.0307)
Child Age and Sex	✓	✓
Parent Age and Sex	✓	✓
n	290	290

Note: Standard errors are in parentheses and clustered on parent country of birth. I keep mother's information among 2 parent households. I keep parents who moved to the U.K before age 18. * (p<.10), ** (p<.05), *** (p<.01).

Table 2.5: Risk Preparedness and Parent Arrival Age

	1a	1b	2a	2b
	0-10	0-11	0-10	0-11
Arrival Age(0-c)X Non English Speaking Country	-0.262**	-0.270**	-1.128**	-1.140**
	(0.0906)	(0.0905)	(0.417)	(0.430)
Non English Speaking Country	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Parent Age and Sex	✓	✓	✓	✓
n	282	282	282	282

Note: Standard errors are in parentheses and clustered on parent country of birth. I keep father's information when both parents are present. In columns 1a and 1b, dependent variable is defined as a dummy = 1 if individuals report the degree to which they are prepared to take risks to be > 7 with 0 being least prepared and 10 being the most prepared. In columns 2a and 2b, the dependent variable is continuous, (defined on a scale of 0-10). * (p<.10), ** (p<.05), *** (p<.01).

Table 2.6: Importance of being British and Parent Arrival Age

Panel A: Mother's Information	0-7	0-8	0-9
Arrival Age(0-c)X Non English Speaking Country	1.477	0.831	0.420
	(1.132)	(1.162)	(1.090)
Arrival Age(0-c)	-1.063	-1.040	-1.045
	(0.929)	(0.946)	(0.909)
Non English Speaking Country	-0.375	-0.176	-0.112
	(0.903)	(0.909)	(0.944)
n	217	217	217
Panel B: Father's Information	**0-7**	**0-8**	**0-9**
Arrival Age(0-c)X Non English Speaking Country	2.008*	1.632*	1.489*
	(1.034)	(0.908)	(0.804)
Non English Speaking Country	-0.188	-0.0878	-0.154
	(0.734)	(0.674)	(0.652)
Arrival Age(0-c)	-1.455	-1.606*	-1.785**
	(0.917)	(0.804)	(0.692)
n	241	241	241
Child Age and Sex	✓	✓	✓
Parent Age and Sex	✓	✓	✓

Note: Standard errors are in parentheses and clustered on parent counry of birth. Panel A is based on mother's information when both parents are present. Panel B is based on father's information when both parents are present. I keep parents who arrived in U.K before age 18. * (p<.10), ** (p<.05), *** (p<.01).

Table 2.7: Importance of being British and Parent Arrival Age

Panel A	(1)			(1)+ Additional		
	0-7	0-8	0-9	0-7	0-8	0-9
Arrival Age(0-c)X Non English Speaking Country	0.316*	0.207	0.238	0.350*	0.194	0.266
	(0.171)	(0.177)	(0.166)	(0.178)	(0.161)	(0.168)
Non English Speaking Country	-0.173	-0.140	-0.159	-0.196	-0.147	-0.182
	(0.141)	(0.141)	(0.140)	(0.154)	(0.144)	(0.154)
Arrival Age(0-c)	-0.199	-0.162	-0.184	-0.217	-0.149	-0.180
	(0.164)	(0.147)	(0.124)	(0.157)	(0.130)	(0.114)
n	164	164	164	164	164	164

Panel B				(1)+ Additional		
	0-7	0-8	0-9	0-7	0-8	0-9
Arrival Age(0-c)X Non English Speaking Country	1.961	1.414	0.944	2.685	1.874	1.553
	(1.286)	(1.111)	(1.078)	(1.643)	(1.234)	(1.180)
Non English Speaking Country	-0.530	-0.379	-0.303	-1.007	-0.772	-0.743
	(0.859)	(0.841)	(0.836)	(0.829)	(0.731)	(0.789)
Arrival Age(0-c)	-1.617	-1.727*	-1.589*	-1.995	-1.910*	-1.801*
	(1.172)	(0.947)	(0.902)	(1.387)	(1.028)	(0.987)
n	164	164	164	164	164	164
Child Age and Sex	✓	✓	✓	✓	✓	✓
Par Age and Sex	✓	✓	✓	✓	✓	✓
Parent Income	No	No	No	✓	✓	✓
Parent Qual.	No	No	No	✓	✓	✓

Note: Standard errors are in parentheses and clustered on parent country of birth. I keep father's information when both parents are present in household. In Panel A, depndent variable is defined as a dummy = 1 if reported 'importance of being British' is $\geq$ 5 on a scale of 0-10 with 0 being the lowest and 10 being the highest. In Panel B, dependent variable is continuous, defined as the degree to which individuals report being British to be important. I keep parents , $\geq$ 18 who immigrated before age 17. ** (p<.10), ** (p<.05), *** (p<.01).

Table 2.8: Trustworthiness of Others and Parent Arrival Age

	(1)		1+Additional Controls	
	0-10	0-11	0-10	0-11
Arrival Age(0-c)X Non English Speaking Country	-0.308**	-0.316**	-0.298**	-0.331*
	(0.118)	(0.124)	(0.142)	(0.164)
Country of Birth Dummies	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Parent Age and Sex	✓	✓	✓	✓
Parent Qual.	No	No	✓	✓
n	252	252	252	252

Note: Note: Standard errors are in parentheses and clustered on parent country of birth. I keep mother's information when both parents are present. Dependent variable is a dummy = 1 if individuals report that most people can be trusted. * (p<.10), ** (p<.05), *** (p<.01).

Table 2.9: Educational Outcomes and Parent Arrival Age

Panel A: Higher Education	0-7	0-8	0-9	0-10
Arrival Age(0-c)X Non English Speaking Country	0.449**	0.437**	0.371**	0.320*
	(0.163)	(0.147)	(0.136)	(0.180)
Country of Birth Dummies	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Parent Age and Sex	✓	✓	✓	✓
n	262	262	262	262

Panel B: School Leaving Age	(1)	(2)
Arrival Age(0-9)X Non English Speaking Country	0.308*	0.254
	(0.174)	(0.179)
Non English Speaking Country	0.0154	0.0398
	(0.281)	(0.270)
Arrival Age(0-9)	0.136	0.0612
	(0.190)	(0.204)
Child Age and Sex	✓	✓
Parent Age and Sex	✓	✓
Parent Qual.	No	✓
n	262	262

Note: Standard errors in parentheses and clustered at parent country of birth level. Based on mother's information when both parents are present in household. In Panel A, Dep. Var. is a dummy = 1 if individuals report to be currently enrolled in or having completed university. * p<.10, ** (p<.05), *** (p<.01).

Table 2.10: Importance of Religion and Parent Arrival Age

	Father's Info.		Monther's Info.	
Panel A	0-5	0-6	0-5	0-6
Arrival Age(0-c)X Non English Speaking Country	-0.205	-0.303**	-0.460**	-0.463**
	(0.173)	(0.142)	(0.212)	(0.210)
n	257	257	225	225
Panel B	0-5	0-6	0-5	0-6
Arrival Age(0-c)X Non English Speaking Country	-0.869*	-0.764	-1.123*	-1.096*
	(0.463)	(0.509)	(0.547)	(0.558)
n	257	257	225	225
Dummy for Non English Speaking Country	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Parent Age and Sex	✓	✓	✓	✓

Note: Standard errors are in parentheses and clustered on parent country of birth. In Panel A, dependent variable is a dummy = 1 if individuals report that religion makes a great difference to life. In Panel B, dependent variable is continuous and takes a value of 1 if religion makes no difference to life and 4 if it makes a great difference. Columns 1 and 2 are based on father's information; 3 and 4 on mother's information when both parents are present. * (p<.10), ** (p<.05), *** (p<.01).

Table 2.11: Level of Political Interest and Parent Arrival Age

	(1)		(1)+ Additional Controls	
Panel A: Father's Information	(1) 0-5	(2) 0-6	(3) 0-5	(4) 0-6
Arrival Age(0-c)X Non English Speaking Country	-0.944*	-1.155**	-0.956	-1.130
	(0.512)	(0.527)	(0.570)	(0.671)
n	344	344	344	344
Panel B: Mother's Information	0-5	0-6	0-5	0-6
Arrival Age(0-c)X Non English Speaking Country	-1.157**	-1.308**	-1.362*	-1.482**
	(0.529)	(0.504)	(0.687)	(0.672)
n	310	310	310	310
Dummy for Non English Speaking Country	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Parent Age and Sex	✓	✓	✓	✓
Parent Income	No	No	✓	✓
Parent Qual.	No	No	✓	✓

Note: Standard errors are in parentheses and clustered on parent country of birth. Dependent variable is defined on a scale of 1-4 and takes the value 1 if individuals report no interest and 4 if they report a very strong interest in politics. In panel A, I keep father's information and in panel B, I keep mother's information when both parents are present. * p<.10, ** (p<.05), *** (p<.01).

Table 2.12: Job Satisfaction and Parent Arrival Age

	0-5	0-6	0-7	0-8
Arrival Age(0-c)X Non English Speaking Country	1.138**	1.138**	0.807*	0.871*
	(0.320)	(0.320)	(0.392)	(0.453)
Dummy for Non English Speaking Country	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Parent Age and Sex	✓	✓	✓	✓
n	156	156	156	156

Note: Standard errors are in parentheses and clustered on parent country of birth. Dependent variable is defined on a scale of 1-7 and takes the value 1 if individuals report to be completely dissatisfied and 7 if completely satisfied with job. I keep father's information when both parents are present. * (p<.10), ** (p<.05), *** (p<.01).

2.10 Appendix Table

Table 2.13: Importance of being British and Parent Arrival Age

				(1)+ Additional Controls		
Panel A	0-7	0-8	0-9	0-7	0-8	0-9
Arrival Age(0-c)X Non English Speaking Country	0.108	0.0121	0.0223	0.166	0.0501	0.0214
	(0.0975)	(0.0793)	(0.0706)	(0.110)	(0.0695)	(0.0459)
Non English Speaking Country	-0.0233	0.00934	0.000268	-0.103	-0.0609	-0.0556
	(0.0486)	(0.0597)	(0.0555)	(0.0692)	(0.0605)	(0.0584)
Arrival Age(0-c)	-0.117	-0.0895	-0.0711	-0.140	-0.108	-0.0485
	(0.109)	(0.0859)	(0.0744)	(0.0922)	(0.0693)	(0.0503)
n	164	164	164	164	164	164
Panel B	0-7	0-8	0-9	0-7	0-8	0-9
Arrival Age(0-c)X Non English Speaking Country	0.597	0.341	-0.270	1.553	1.199	0.288
	(0.781)	(0.644)	(0.539)	(1.366)	(1.140)	(0.991)
Non English Speaking Country	-0.0630	0.0946	0.280	-0.830	-0.686	-0.351
	(0.524)	(0.530)	(0.547)	(0.827)	(0.705)	(0.668)
Arrival Age(0-c)	-0.275	-0.824	-0.429	-0.991	-1.414	-0.799
	(0.497)	(0.542)	(0.623)	(0.965)	(0.883)	(0.900)
n	164	164	164	164	164	164
Child Age and Sex	✓	✓	✓	✓	✓	✓
Par Age and Sex	✓	✓	✓	✓	✓	✓
Parent Income	No	No	No	✓	✓	✓
Parent Qual.	No	No	No	✓	✓	✓
Parent Region of Residence Dummies	✓	✓	✓	✓	✓	✓

Note: See table 2.6 and text for notes. * (p<.10), ** (p<.05), *** (p<.01).

Table 2.14: Trustworthiness of Others and Parent Arrival Age

	(1)		(1)+Additional Controls	
	0-10	0-11	0-10	0-11
belowcXnoneng1	-0.187	-0.200	-0.193*	-0.231*
	(0.116)	(0.124)	(0.107)	(0.133)
Country of Birth Dummies	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Parent Age and Sex	✓	✓	✓	✓
Parent Qual.	No	No	✓	✓
Parent Job Status	✓	✓	✓	✓
n	252	252	252	252

Note: See table 2.8 and text for notes. * (p<.10), ** (p<.05), *** (p<.01)

Table 2.15: Probability of Pursuing Higher Education and Parent Arrival Age

Panel A: Higher Education	0-7	0-8	0-9	0-10
Arrival Age(0-c)X Non English Speaking Country	0.227	0.212	0.261**	0.258**
	(0.135)	(0.130)	(0.0982)	(0.0957)
Country of Birth Dummies	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Par Age and Sex	✓	✓	✓	✓
Parent Region of Residence Dummies	✓	✓	✓	✓
Parent Qual.	✓	✓	✓	✓
n	262	262	262	262

Note: See table 2.9 and text for notes. * (p<.10), ** (p<.05), *** (p<.01)

Table 2.16: Importance of Religion and Parent Arrival Age

	Father's	Info.	Monther's	Info.
Panel A	0-5	0-6	0-5	0-6
Arrival Age(0-c)X Non English Speaking Country	-0.195	-0.328*	-0.335	-0.368
	(0.208)	(0.179)	(0.234)	(0.215)
Dummy for Non English Speaking Country	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Parent Age and Sex	✓	✓	✓	✓
Parent Job Status	✓	✓	✓	✓
n	257	257	225	225
Panel B	0-5	0-6	0-5	0-6
Arrival Age(0-c)X Non English Speaking Country	-0.957	-0.880	-0.895	-0.901
	(0.576)	(0.645)	(0.604)	(0.598)
Dummy for Non English Speaking Country	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Parent Age and Sex	✓	✓	✓	✓
Parent Job Status	✓	✓	✓	✓
n	257	257	225	225

Note: See table 2.10 and text for notes. * (p<.10), ** (p<.05), *** (p<.01).

Table 2.17: Level of Political Interest and Parent Arrival Age

	0-5	0-6
Arrival Age(0-c)X Non English Speaking Country	1.515	1.584
	(0.909)	(0.920)
Dummy for Non English Speaking Country	✓	✓
Age of Entry Dummies	✓	✓
Child Age and Sex	✓	✓
Par Age and Sex	✓	✓
Par Income	✓	✓
Par Qual	✓	✓
Parent Region of Residence Dummies	✓	✓
n	310	310

Note: See table 2.11 and text for notes. * (p<.10), ** (p<.05), *** (p<.01).

Table 2.18: Job Satisfaction and Parent Arrival Age

	0-5	0-6	0-7	0-8
Arrival Age(0-c)X Non English Speaking Country	0.551	0.551	0.0614	0.662
	(0.428)	(0.428)	(0.547)	(0.456)
Dummy for Non English Speaking Country	✓	✓	✓	✓
Age of Entry Dummies	✓	✓	✓	✓
Child Age and Sex	✓	✓	✓	✓
Parent Age and Sex	✓	✓	✓	✓
Parent Qual.	✓	✓	✓	✓
Parent Job Status	✓	✓	✓	✓
n	156	156	156	156

Note: See table 2.12 and text for notes. * (p<.10), ** (p<.05), *** (p<.01).

Bibliography

Abramitzky, R., Boustan, L. P., and Eriksson, K. (2016). Cultural assimilation during the age of mass migration. Technical report, National Bureau of Economic Research.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American economic review*, 91(5):1369–1401.

Agostinelli, F. (2017). Investing in children's skills: An equilibrium analysis of social interactions and parental investments.

Agostinelli, F. and Sorrenti, G. (2018). Money vs. time: family income, maternal labor supply, and child development.

Algan, Y., Mayer, T., and Thoenig, M. (2013). The economic incentives of cultural transmission: Spatial evidence from naming patterns across france.

Altonji, J. G. and Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market*, pages 201–234. University of Chicago Press.

Angrist, J. D. and Lavy, V. (1997). The effect of a change in language of instruction on the returns to schooling in morocco. *Journal of Labor Economics*, 15(1, Part 2):S48–S76.

Aoki, Y. and Santiago, L. (2015). Education, health and fertility of uk immigrants: The role of english language skills.

Aoki, Y. and Santiago, L. (2017). Speak well, do well? english proficiency and social segregration of uk immigrants.

Arrow, K. J. (1972). Gifts and exchanges. *Philosophy & Public Affairs*, pages 343–362.

Åslund, O., Böhlmark, A., and Skans, O. N. (2015). Childhood and family experiences and the social integration of young migrants. *Labour Economics*, 35:135–144.

Azam, M., Chin, A., and Prakash, N. (2013). The returns to english-language skills in india. *Economic Development and Cultural Change*, 61(2):335–367.

Baum-Snow, N. and Lutz, B. F. (2011). School desegregation, school choice, and changes in residential location patterns by race. *American Economic Review*, 101(7):3019–46.

Becker, G. S. and Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of labor economics*, 4(3, Part 2):S1–S39.

Benhabib, J., Bisin, A., and Jackson, M. O. (2010). *Handbook of social economics*. Elsevier.

Bester, H. and Güth, W. (1998). Is altruism evolutionarily stable? *Journal of Economic Behavior & Organization*, 34(2):193–209.

Bezin, E. and Moizeau, F. (2017). Cultural dynamics, social mobility and urban segregation. *Journal of Urban Economics*, 99:173–187.

Bisin, A., Patacchini, E., Verdier, T., and Zenou, Y. (2008). Are muslim immigrants different in terms of cultural integration? *Journal of the European Economic Association*, 6(2-3):445–456.

Bisin, A., Patacchini, E., Verdier, T., Zenou, Y., et al. (2006). *'Bend it Like Beckham': Identity, Socialization and Assimilation.* Centre for Economic Policy Research.

Bisin, A., Topa, G., and Verdier, T. (2004). Religious intermarriage and socialization in the united states. *Journal of political Economy*, 112(3):615–664.

Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory*, 97(2):298–319.

Bisin, A. and Verdier, T. (2011). The economics of cultural transmission and socialization. 1:339–416.

Björklund, A. and Jäntti, M. (1997). Intergenerational income mobility in sweden compared to the united states. *The American Economic Review*, 87(5):1009–1018.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American economic review*, 95(1):437–449.

Bleakley, H. and Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and statistics*, 86(2):481–496.

Bleakley, H. and Chin, A. (2008). What holds back the second generation? the intergenerational transmission of language human capital among immigrants. *Journal of human resources*, 43(2):267–298.

Bleakley, H. and Chin, A. (2010). Age at arrival, english proficiency, and social assimilation among us immigrants. *American Economic Journal: Applied Economics*, 2(1):165–92.

Borjas, G. J. (1992). Ethnic capital and intergenerational mobility. *The Quarterly journal of economics*, 107(1):123–150.

Borjas, G. J. (1998). To ghetto or not to ghetto: Ethnicity and residential segregation. *Journal of Urban Economics*, 44(2):228–253.

Borland, H. (2006). Intergenerational language transmission in an established australian migrant community: what makes the difference? *International journal of the sociology of language*, 2006(180):23–41.

Bosch, M., Carnero, M. A., and Farré, L. (2015). Rental housing discrimination and the persistence of ethnic enclaves. *SERIEs*, 6(2):129–152.

Buser, T. (2015). The effect of income on religiousness. *American Economic Journal: Applied Economics*, 7(3):178–95.

Calvó-Armengol, A., Patacchini, E., and Zenou, Y. (2009). Peer effects and social networks in education. *The Review of Economic Studies*, 76(4):1239–1267.

Caminal, R., Cappellari, L., and Di Paolo, A. (2018). Linguistic skills and the intergenerational transmission of language.

Card, D. (2009). Immigration and inequality. *American Economic Review*, 99(2):1–21.

Card, D., Dinardo, J., and Estes, E. (2000a). The more things change: Immigrant.

Card, D., DiNardo, J., Estes, E., and Borjas, G. J. (2000b). Issues in the economics of immigration.

Card, D., Mas, A., and Rothstein, J. (2008). Tipping and the dynamics of segregation. *The Quarterly Journal of Economics*, 123(1):177–218.

Cascio, E. U. and Lewis, E. G. (2012). Cracks in the melting pot: immigration, school choice, and segregation. *American Economic Journal: Economic Policy*, 4(3):91–117.

Casey, T. and Dustmann, C. (2008). Intergenerational transmission of language capital and economic outcomes. *Journal of Human Resources*, 43(3):660–687.

Cavalli-Sforza, L. L. and Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach*. Number 16. Princeton University Press.

Chetty, R. and Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *The Quarterly Journal of Economics*, 133(3):1107–1162.

Chetty, R., Hendren, N., and Katz, L. F. (2016a). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Chetty, R., Hendren, N., Lin, F., Majerovitz, J., and Scuderi, B. (2016b). Childhood environment and gender gaps in adulthood. *American Economic Review*, 106(5):282–88.

Chiswick, B. R. (1991). Speaking, reading, and earnings among low-skilled immigrants. *Journal of labor economics*, 9(2):149–170.

Chiswick, B. R., Lee, Y. L., and Miller, P. W. (2005a). Family matters: the role of the family in immigrants' destination language acquisition. *Journal of Population Economics*, 18(4):631–647.

Chiswick, B. R., Lee, Y. L., and Miller, P. W. (2005b). Parents and children talk: English language proficiency within immigrant families. *Review of Economics of the Household*, 3(3):243–268.

Chiswick, B. R. and Miller, P. W. (1995). The endogeneity between language and earnings: International analyses. *Journal of labor economics*, 13(2):246–288.

Chiswick, B. R. and Miller, P. W. (1996). Ethnic networks and language proficiency among immigrants. *Journal of Population Economics*, 9(1):19–35.

Clark, C. A., Worthington Jr, E. L., and Danser, D. B. (1988). The transmission of religious beliefs and practices from parents to firstborn early adolescent sons. *Journal of Marriage and the Family*, pages 463–472.

Cornwall, M. (1988). The influence of three agents of religious socialization: Family, church, and peers. *The religion and family connection: Social science perspectives*, 16(2):207–231.

Currie, J. and Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4):1495–1532.

Cutler, D. M., Glaeser, E. L., and Jacob, L. (2005a). Vigdor. 2002. ghettos and the transmission of economic capital. department of economics, harvard university. *Unpublished manuscript. Retrieved January*, 21.

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (2005b). Ghettos and the transmission of ethnic capital. *Ethnicity, Social Mobility and Public Policy: Comparing the US and UK*, pages 204–222.

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (2008a). Is the melting pot still hot? explaining the resurgence of immigrant segregation. *The Review of Economics and Statistics*, 90(3):478–497.

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (2008b). When are ghettos bad? lessons from immigrant segregation in the united states. *Journal of Urban Economics*, 63(3):759–774.

Davila, A. and Mora, M. T. (2001). Hispanic ethnicity, english-skill investments, and earnings. *Industrial Relations: A Journal of Economy and Society*, 40(1):83–88.

Dearden, L., Machin, S., and Reed, H. (1997). Intergenerational mobility in britain. *The Economic Journal*, pages 47–66.

Delhey, J. and Newton, K. (2003). Who trusts?: The origins of social trust in seven societies. *European Societies*, 5(2):93–137.

Dixit, A. (2009). Governance institutions and economic activity. *American economic review*, 99(1):5–24.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2011). The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies*, 79(2):645–677.

Duncan, B. and Trejo, S. J. (2007). Ethnic identification, intermarriage, and unmeasured progress by mexican americans. In *Mexican immigration to the United States*, pages 229–268. University of Chicago Press.

Dustmann, C. and Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the uk. *The Economic Journal*, 113(489):695–717.

Dustmann, C. and Van Soest, A. (2002). Language and the earnings of immigrants. *ILR Review*, 55(3):473–492.

Echenique, F., Fryer Jr, R. G., and Kaufman, A. (2006). Is school segregation good or bad? *American Economic Review*, 96(2):265–269.

Edin, P. A., Fredriksson, P., and Aslund, O. (2003). Ethnic enclaves and the economic success of immigrants evidence from a natural experiment. *The quarterly journal of economics*, 118(1):329–357.

Erickson, F. (1992). Ethnographic microanalysis of interaction. *The handbook of qualitative research in education*, pages 201–225.

Eshel, I., Samuelson, L., and Shaked, A. (1998). Altruists, egoists, and hooligans in a local interaction model. *American Economic Review*, pages 157–179.

Fershtman, C. and Weiss, Y. (1998). Social rewards, externalities and stable preferences. *Journal of Public Economics*, 70(1):53–73.

Funkhouser, E. and Ramos, F. A. (1993). The choice of migration destination: Dominican and cuban immigrants to the mainland united states and puerto rico. *International migration review*, 27(3):537–556.

George, L. and Waldfogel, J. (2003). Who affects whom in daily newspaper markets? *Journal of Political Economy*, 111(4):765–784.

Glick, J. E. and White, M. J. (2003). Academic trajectories of immigrant youths: Analysis within and across cohorts. *Demography*, 40(4):759–783.

Gradstein, M. and Justman, M. (2002). Education, social cohesion, and economic growth. *American Economic Review*, 92(4):1192–1204.

Gradstein, M. and Justman, M. (2005). The melting pot and school choice. *Journal of Public Economics*, 89(5-6):871–896.

Gradstein, M., Justman, M., and Meier, V. (2005). The political economy of education. *The CESifo Book Series*.

Grogger, J. and Trejo, S. J. (2002). Falling behind or moving up. *The Intergenera*.

Gruber, J. H. (2005). Religious market structure, religious participation, and outcomes: Is religion good for you? *The BE Journal of Economic Analysis & Policy*, 5(1).

Guiso, L., Sapienza, P., and Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic perspectives*, 20(2):23–48.

Guiso, L., Sapienza, P., and Zingales, L. (2016). Long-term persistence. *Journal of the European Economic Association*, 14(6):1401–1436.

Guven, C. and Islam, A. (2015). Age at migration, language proficiency, and socioeconomic outcomes: evidence from australia. *Demography*, 52(2):513–542.

Hayes, B. C. and Pittelkow, Y. (1993). Religious belief, transmission, and the family: An australian study. *Journal of Marriage and the Family*, pages 755–766.

Hernandez, D. J., Charney, E., et al. (1998). From generation to generation. *The Health and Well Being of Children in Immigrant Families*, page 1998.

Hirschman, C. (2001). The educational enrollment of immigrant youth: A test of the segmented-assimilation hypothesis. *Demography*, 38(3):317–336.

Houle, R. (2011). Recent evolution of immigrant-language transmission in canada. *Canadian Social Trends*, 92.

Ioannides, Y. M., Zanella, G., et al. (2008). *Searching for the Best Neighborhood: Mobility and Social Interactions*. Università di Siena.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018). Shift-share instruments and the impact of immigration. Technical report, National Bureau of Economic Research.

Jia, R. and Persson, T. (2017). Individual vs. social motives in identity choice: Theory and evidence from china.

Jung, S. (2015). Does education affect risk aversion? evidence from the british education reform. *Applied Economics*, 47(28):2924–2938.

Kassoudji, S. (1988). English language abilities and the labor market opportunities of hispanic and east asian men. *Journal of Labor Economics*, 6(2):205–28.

Koçkesen, L., Ok, E. A., and Sethi, R. (2000). The strategic advantage of negatively interdependent preferences. *Journal of Economic Theory*, 92(2):274–299.

Laan Bouma-Doff, W. v. d. (2007). Confined contact: Residential segregation and ethnic bridges in the netherlands. *Urban Studies*, 44(5-6):997–1017.

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, 2(12):59–67.

Ljunge, M. (2012). Cultural transmission of civicness. *Economics Letters*, 117(1):291–294.

McCleary, R. M. and Barro, R. J. (2006). Religion and economy. *Journal of Economic perspectives*, 20(2):49–72.

McConnell, E. D. (2012). House poor in los angeles: Examining patterns of housing-induced poverty by race, nativity, and legal status. *Housing Policy Debate*, 22(4):605–631.

McManus, W., Gould, W., and Welch, F. (1983). Earnings of hispanic men: The role of english language proficiency. *Journal of Labor Economics*, 1(2):101–130.

Meng, X. and Gregory, R. G. (2005). Intermarriage and the economic assimilation of immigrants. *Journal of Labor economics*, 23(1):135–174.

Miranda, A. and Zhu, Y. (2013a). The causal effect of deficiency at english on female immigrants' labor market outcomes in the uk.

Miranda, A. and Zhu, Y. (2013b). English deficiency and the native–immigrant wage gap. *Economics Letters*, 118(1):38–41.

Montgomery, J. D. (2010). Intergenerational cultural transmission as an evolutionary game. *American Economic Journal: Microeconomics*, 2(4):115–36.

Moschion, J. and Tabasso, D. (2014). Trust of second-generation immigrants: intergenerational transmission or cultural assimilation? *IZA Journal of Migration*, 3(1):10.

Mulligan, C. B. (1997). *Parental priorities and economic inequality.* University of Chicago Press.

Oreopoulos, P., Page, M. E., and Stevens, A. H. (2006). The intergenerational effects of compulsory schooling. *Journal of Labor Economics*, 24(4):729–760.

Patacchini, E. and Zenou, Y. (2007). Intergenerational education transmission: neighborhood quality and/or parents' involvement?

Patacchini, E. and Zenou, Y. (2012). Ethnic networks and employment outcomes. *Regional Science and Urban Economics*, 42(6):938–949.

Patacchini, E. and Zenou, Y. (2016). Social networks and parental behavior in the intergenerational transmission of religion. *Quantitative Economics*, 7(3):969–995.

Portes, A. and Hao, L. (1998). E pluribus unum: Bilingualism and loss of language in the second generation. *Sociology of Education*, pages 269–294.

Portes, A. and MacLeod, D. (1999). Educating the second generation: Determinants of academic achievement among children of immigrants in the united states. *Journal of ethnic and migration studies*, 25(3):373–396.

Portes, A. and Manning, R. D. (1986). *The immigrant enclave: Theory and empirical examples.* na.

Reber, S. J. (2005). Court-ordered desegregation successes and failures integrating american schools since brown versus board of education. *Journal of Human resources*, 40(3):559–590.

Rindos, D., Carneiro, R. L., Cooper, E., Drechsel, P., Dunnell, R. C., Ellen, R., Gullick, C., Hackenberg, R. A., Hartung, J., Kunkel, J. H., et al. (1985). Darwinian selection, symbolic variation, and the evolution of culture [and comments and reply]. *Current anthropology*, 26(1):65–88.

Robson, A. J. (1996). A biological basis for expected and non-expected utility. *Journal of economic theory*, 68(2):397–424.

Ruggles, S., Goeken, R., Grover, J., and Sobek, M. (2017). Integrated public use microdata series: Version 7.0 [dataset].

Sáez-Martí, M. and Sjögren, A. (2008). Peers and culture. *Scandinavian Journal of Economics*, 110(1):73–92.

Sapienza, P., Toldra-Simats, A., and Zingales, L. (2013). Understanding trust. *The Economic Journal*, 123(573):1313–1332.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology*, 1(2):143–186.

Smith, J. P. (2003). Assimilation across the latino generations. *American Economic Review*, 93(2):315–319.

Solon, G. (1992). Intergenerational income mobility in the united states. *The American Economic Review*, pages 393–408.

Stevens, G. and Swicegood, G. (1987). The linguistic context of ethnic endogamy. *American sociological review*, pages 73–82.

Suro, R. and Passel, J. S. (2003). The rise of the second generation: Changing patterns in hispanic population growth.

Tabellini, G. (2008a). Institutions and culture. *Journal of the European Economic Association*, 6(2-3):255–294.

Tabellini, G. (2008b). The scope of cooperation: Values and incentives. *The Quarterly Journal of Economics*, 123(3):905–950.

Tainer, E. (1988). English language proficiency and the determination of earnings among foreign-born men. *Journal of Human Resources*, pages 108–122.

Van der Laan Bouma-Doff, W. (2007). Involuntary isolation: Ethnic preferences and residential segregation. *Journal of Urban Affairs*, 29(3):289–309.

Wiegand, J. (1997). Intergenerational earnings mobility in germany. *University College London, Mimeo*.

Wong, M. (2008). Estimating the impact of the ethnic housing quotas in singapore. *V MIT Thesis*.

Zimmerman, D. J. (1992). Regression toward mediocrity in economic stature. *The American Economic Review*, pages 409–429.