Early Childhood Human Capital and Development*

Todd Schoellman†

November 2014

Abstract

A growing literature stresses the importance of early childhood for human capital formation. I ask whether variation in early childhood investments can help explain cross-country income differences. I first provide new empirical evidence: the adult outcomes of refugees are independent of age at arrival to the US up to age 6, despite dramatic improvements in income and environment upon arrival. I then ask how an off-the-shelf model of early childhood human capital can be consistent with this finding. The model can replicate the data if growing up in a poor country is not a significant disadvantage, or if early childhood disadvantages are easy to remediate. Given the current consensus that remediation is difficult, I conclude that early childhood plays a limited role for development accounting.

JEL Classification: O11, E24

*Thanks to Rody Manuelli, John Knowles, Lance Lochner, Matt Wiswall, Chris Robinson, Kevin Donovan, Gustavo Ventura, Berthold Herrendorf and audiences at Ohio State, UC – Davis, Western Ontario, Toronto, York University, the University of British Columbia, UC – Berkeley, Purdue, the 2012 Midwest Macroeconomics Meetings, the 2013 Society for Economic Dynamics, the 2013 Tsinghua Workshop in Macroeconomics, the 2013 NBER Growth meetings, the 2013 Growth and Development Conference, the 2013 Vienna Macroeconomics Workshop, and the ASU Conference on the Development of Human Capital for helpful comments. The usual disclaimer applies.

†Address: Department of Economics, Arizona State University, Tempe, AZ 85287. E-mail: todd.schoellman@gmail.com

1 Introduction

One of the central challenges for economists is to explain the large differences in gross domestic product (GDP) per worker across countries. Development accounting provides a useful first step towards this goal. This technique measures the relative contribution of physical capital, human capital, and total factor productivity (TFP) in accounting for cross-country income differences. The main finding in the development accounting literature is that the quantitative role for physical and human capital is limited, implying a large role for TFP (Klenow and Rodriguez-Clare, 1997; Hall and Jones, 1999; Caselli, 2005; Hsieh and Klenow, 2010).

The benchmark approach to constructing human capital stocks uses each country's average years of schooling valued at the market return to schooling. In this paper I consider whether allowing for a broader notion of human capital that includes early childhood human capital – human capital acquired even before school starts, or roughly by age 5 – would imply a larger role for human capital. Early childhood human capital seems a promising avenue to explore in light of the large recent literature that stresses that early childhood is a critical period for human capital formation. This literature has documented numerous investments and interventions that produce economically and statistically significant effects, some of which have been shown to last into adulthood.[1] Given the cross-country variation in the quality and quantity of nutrition, medical expenditures per child, or disease prevalence, it seems natural to expect that young children in rich countries have received much more investment and attained much higher human capital than those in poor countries. However, it is difficult to translate the results of a typical early childhood intervention into implications for cross-country development. Manuelli and Seshadri (2014) offer the first quantitative evaluation. They calibrate a model of human capital investments over the life cycle to US data and then study the model's cross-country predictions. Their model predicts that even before school starts, human capital already varies by nearly a factor of four between the richest and poorest countries. If true, this finding would imply that a substantial shift of priorities is warranted for development policy.

An important challenge to quantifying the cross-country importance of early childhood is that there are currently no comparable cross-country data on investments or outcomes during early childhood. The empirical contribution of this paper is to provide new evidence to overcome this challenge. To do so, I study the adult outcomes of refugees who immigrated

[1] See Almond and Currie (2011) or Cunha et al. (2006) for some recent overviews.

to the US during early childhood as a function of their age at arrival. Refugees who arrived at older ages spent more of their early childhood in their country of birth or in refugee camps and less of their early childhood in the US. I exploit this variation to estimate the marginal cost of spending a year of early childhood in a poor instead of a rich country. This statistic has a natural economic interpretation and is also less prone to bias from the selection of refugees because it involves comparing two groups of refugees who differ only in the age of early childhood at which they arrived to the US. The historical details of refugee resettlement imply that children who arrived at different ages were not selected differently, in which case the estimated coefficient can be interpreted as the causal impact of one year of poor country early childhood.

The main analysis uses Indochinese refugees, who fled Vietnam, Cambodia, and Laos in the late 1970s and 1980s. Average incomes in these three countries were extremely low, roughly 3 percent of US levels. Additionally, the children were exposed to environments that one might expect to be disastrous for human capital accumulation, having survived the Vietnam War, the Khmer Rouge, refugee camps, and other hardship. In the face of these many disadvantages, the main empirical result is striking: there is no relationship between wages or other socioeconomic outcomes and age at arrival for the Indochinese refugees. This result is robust to the details of the estimation or sample selection. It also extends to other immigrants from poor countries. On the other hand, it is special to early childhood: a strong relationship between age at arrival and outcomes does emerge, but only after early childhood.

Although this experiment is quite different from the typical early childhood intervention, it is useful to understand how the result can be consistent with large results from most interventions. I use a standard model of human capital formation in the spirit of Cunha et al. (2010) and Del Boca et al. (2014) to help interpret this finding. The heart of the model is a two-stage human capital production function that maps human capital at birth and early childhood investments into human capital at 6, then human capital at 6 and investments made during school into adult human capital. Investments are allowed to take a variety of forms, including parental time or quality, child time, country environment, and purchased goods. I show that one of three strong restrictions on the human capital production function is required to make the model's implications consistent with the data. The first restriction is that there is no role for goods or country environment in early childhood human capital formation. The second restriction is that early childhood human capital is irrelevant for subsequent outcomes. The third restriction is that there is a high elasticity of substitution between human capital at age 5 and investments made during schooling. The last restriction

allows refugees to remediate any early childhood human capital deficits through high levels of post-migration investment.

I use additional evidence from refugees, as well as the broader literature on early childhood human capital formation, to help discriminate among these possibilities. The wealth of evidence in that literature suggests that early childhood human capital is important, ruling out the second case. Tentative evidence in that literature also suggests remediation is difficult. In a companion paper I document that the experience of the Indochinese refugees is also inconsistent with remediation (Schoellman, 2014). These facts suggest that the third case is less likely. The conclusion is that the first case is the most promising: goods and country environment are not important inputs to the production of early childhood human capital. This is consistent not only with the experience of the Indochinese refugees, but with the broader literature on early childhood human capital. It also implies that cross-country differences in early childhood human capital are likely to be small. On the other hand, my finding that outcomes do decline in age at arrival after early childhood is consistent with the importance of goods for later outcomes, offering support for a central mechanism in several papers in the literature (Manuelli and Seshadri, 2014; Erosa et al., 2010; Cordoba and Ripoll, 2013; Cubas et al., 2013).

My paper is most closely related to a recent literature that re-considers the importance of human capital for development. Two sets of papers are notable. The first takes a broader view of human capital by incorporating health, education quality, and experience.[2] This paper takes a broader view of human capital by incorporating early childhood. The primary difference is that whereas most of these previous papers have found large effects from broadening the scope of human capital, I do not. The second set of papers explore the idea of using immigrants to measure cross-country human capital differences. Hendricks (2002) first formalized this idea and clarified the essential tradeoff of using immigrants: on the one hand they bring only their birth-country human capital stock with them, leaving behind other factors such as physical capital and TFP, which makes them useful for isolating human capital differences. On the other hand immigrants are selected, which complicates the process of drawing inferences about the representative non-migrant. An advantage of this paper is that my identification scheme avoids comparing immigrants to natives and hence reduces selection concerns.

The rest of the paper proceeds as follows. Section 2 gives background on the Indochinese

[2]For example, Weil (2007), Shastry and Weil (2003), and Acemoglu and Johnson (2007) quantify the role of health; Schoellman (2012) and Kaarsen (2014) quantify the role of education quality; and Lucas Jr and Moll (forthcoming), Lagakos et al. (2014a) and Lagakos et al. (2014b) quantify the role of experience.

refugees while section 3 presents the empirical results. Section 4 formulates a model and and compares its predictions to the data. Section 5 presents the relevant evidence from the literature on early childhood human capital formation. Section 6 considers the implications for cross-country human capital differences. Section 7 considers alternative hypotheses for the empirical finding and section 8 concludes.

2 Background on the Indochinese Refugees

The empirical analysis of the paper exploits the variation in socioeconomic outcomes by age at arrival for Indochinese refugees who migrated to the US during childhood. Children who arrived at older ages spent more years of life in their birth country and fewer years in the US. This variation is useful because the disparities between the US and the Indochinese countries would seem to imply that older-arriving refugees were at a significant disadvantage. In the next two subsections I document the changes associated with moving to the US.

2.1 Pre-Migration Conditions

Years spent in Indochinese countries presented two main disadvantages. First, the Indochinese countries were poor. Their PPP-adjusted GDP per worker was roughly 3 percent of US levels in 1980, and did not exceed 5 percent of US levels until 2005 (Heston et al., 2012). Along with poverty came the usual associated problems such as poor or inadequate nutrition and endemicity of diseases including malaria, tuberculosis, and various intestinal parasites (United States General Accounting Office, 1982).

The second source of disadvantage was the ongoing conflict in the region between Communist and non-Communist forces, including the Vietnam War as well as similar wars in Cambodia and Laos. War ended in 1975 with Communists taking control in all three countries. Communist takeovers in Vietnam and Laos led to politically and ethnically motivated discrimination or "re-education" in labor camps (Robinson, 1998). The Khmer Rouge regime in Cambodia forced the population to switch to an extreme, agriculturally oriented form of Marxism that brought widespread famine, while at the same time killing an estimated 20 percent of the Cambodian population directly in the killing fields (Robinson, 1998). Conflict in the region was reignited when Vietnam invaded Cambodia in 1978.

Faced with these circumstances, roughly three million citizens fled to nearby countries. Refugees from Cambodia and Laos primarily traveled to Thailand by land, while refugees

from Vietnam traveled primarily to Hong Kong, Indonesia, and Malaysia by boat.[3] Violence and piracy en route were common, as were delays or "push-backs" of refugees by the countries of asylum.

1.4 million of those who fled were placed in refugee camps where they waited for an opportunity to settle within the country of first asylum, repatriate to their former home, or resettle to a third country. Conditions in the camps ranged from poor to adequate. They were the worst in Thailand, which practiced an official policy of "humane deterrence". The goal of this policy was to worsen conditions in camps so as to deter potential future migrants. Thai camps were subject to violence, regular shortages of food, poor sanitation, and inadequate medical staff and treatment (United States General Accounting Office, 1980; Robinson, 1998). Seventy percent of children in one Thai camp were malnourished (United States General Accounting Office, 1980). Conditions elsewhere were better but not ideal. Even in Hong Kong, where the treatment of refugees was considered to be the best, a nutritional survey by the Save the Children Fund revealed that roughly half of all children suffered from some form of malnourishment (Thomas, 2000).

Camps did present refugees with the time and opportunity to make investments in their skills. The most common opportunities were formal or informal classes in English and rudimentary educational facilities for early grades. Space for school and materials were typically scarce and were sometimes withheld as part of the policies of humane deterrence. Teachers came from within the refugee population, leading to variable quality and high turnover. Absenteeism was also a significant problem (United States General Accounting Office, 1980). Refugees were often discouraged and frustrated by the resettlement process, which manifested itself as lethargy rather than investment. One set of visitors reported an "overriding impression of camp life [of] people standing around doing nothing" (Davis, 1991). A time-use survey performed in the Thai refugee camps revealed that "doing nothing" and "work" were more common responses than "studying", except for the Vietnamese; the Hmong reported spending 52 percent of their time doing nothing (Hein, 1995). Surveys of newly-settled refugees revealed that less than twelve percent had taken English-language lessons in the refugee camps, one of the most common forms of investment (Reder and Cohn, 1984).

[3] A first wave of refugees in 1975 transited directly from these countries to the US. This wave consisted primarily of those with close political or military connections to the US, most notably the 130,000 people flown out of Saigon in the final days before the US withdrew from South Vietnam (Hung and Haines, 1996). These immigrants were highly selected on many dimensions, typically spoke English well, and had easy paths to resettlement in the US. I exclude 1975 arrivals from all subsequent analysis in order to focus on immigrants who were less selected and had more typical refugee and resettlement experiences.

2.2 Post-Migration Conditions

1.3 million refugees were permanently resettled to new homes in countries around the world. The US was by far the most common destination, accepting over 800,000 Indochinese refugees (United Nations High Commissioner for Refugees, 2001). The US government instituted a number of policies that ensured that resettlement implied a large, immediate increase in living standards for refugees. First, the government worked with a network of voluntary agencies to provide each refugee with a local sponsor. That sponsor received $500 to provide initial accommodations for the refugee, which was typically used to secure housing, furniture, and provide the refugee with some initial cash. The sponsor also helped guide the refugees to medical care, which the government considered essential for public health reasons, and to training programs, language programs, or employment services, as necessary. Uptake of these programs was substantial; three-fourths of refugees participated in English classes during their first few years, and two-thirds of refugees reported having seen a doctor in the last year (Reder and Cohn, 1984; Rumbaut, 1989).

The federal government also made all refugees immediately eligible for both the standard welfare programs and a few refugee-specific programs. The most important for the families with young children studied here are AFDC and Medicaid. The payments from these programs varied by state and year, but to give some context the payments for a family of three were 55 percent of the poverty line in the median US state in July of 1980. The two most common states for refugees – California and Texas – display the wide variation, with payments 91 and 22 percent of the poverty line (Haines, 1989; U.S. House of Representatives, 1996; Social Security Administration, 2012). When combined with labor market income, this assistance allowed refugees to obtain a living standard that was low by US standards but high by the standards of their birth countries. For example, refugees achieved an income of 46 percent of the poverty line after four months in the US and rose rapidly to 96 percent after eight months, trending gradually upward from that point onward (Caplan et al., 1991; Rumbaut, 1989). As of 1980, the US poverty line was twenty times Vietnam's PPP GDP p.c. (Heston et al., 2012). Hence, the available evidence suggests that refugee families saw large increases in family income and access to social services, especially health services, upon arrival to the US.

A final sign of the improvement in environment for newly arrived refugees comes from an extensive medical and psychological literature that documents the lingering physical and mental health issues among refugees. Groups of refugee children tested around the country consistently averaged between the 5th and 25th percentiles in height-for-age and

weight-for-age, with a greatly elevated number scoring below the 5th percentile (Barry et al., 1983; Peck et al., 1981). Follow-up studies showed that the gap in height-for-age did not diminish with time in the US (Dewey et al., 1986). Early-arriving Indochinese refugees suffered from tuberculosis, hepatitis B, malaria, and intestinal parasites at rates more than an order of magnitude higher than the general US population, although policy changes around 1982 increased the efficacy of medical screening of refugees (United States General Accounting Office, 1982; Barry et al., 1983; Goldenring et al., 1982). Likewise, psychological studies indicated the presence of post-traumatic stress disorder, depression, and other mental illnesses, with persistence in follow-up studies conducted as much as twenty years after migration (Marshall et al., 2005).

The Indochinese refugees provide useful information because of the change in environment and income that they experienced upon arrival in the US. Of course, many other immigrants to the US have similar experiences. What makes the Indochinese refugees more unique and useful for the question at hand is that they are much less selected than typical immigrants. The historical circumstances of their immigration make it unlikely that they were selected on age at arrival, which is critical to this paper's identification strategy. I discuss this issue at length in the next subsection.

2.3 Selection and Identification

Selection can bias inferences in two ways. First, if immigrant families are highly selected, then immigration may not change the family's income or environment that much. This concern is greatly mitigated in the case of refugees. The difficult environments in their home countries created a strong push factor for emigration by refugees. The US government applied liberal standards for immigration under the policy of "presumptive eligibility", which meant that immigrants from these Communist countries were presumed to be refugees and hence eligible for resettlement as long as they were not actively ill or dangerous to others. Observers from other countries generally considered the US to be the least selective in admitting refugees (Hawthorne, ed, 1982).

I provide evidence on the selection of Indochinese refugees in Appendix A. That evidence is consistent with a modest degree of positive selection. For example, refugees were 16–24% less likely to have no schooling and 8–22% more likely to have a high school degree as compared to non-migrants. Few immigrants spoke English before arrival and many came from farming, fishing, or military backgrounds that offered few skills of value in the US. In

short, the post-1975 refugee families do not represent the very elite of their birth countries, implying that resettlement led to an improvement in their day-to-day circumstances.

The second way selection can bias inferences is if immigrants who arrive at different ages are selected differently. Even if one imagines a rationale why families might wish to base their immigration decision on the age of their young children, the historical circumstances of the Indochinese refugees make this type of selection unlikely. The reason is that Indochinese refugees chose when to flee their home country, not when they would arrive to the US. The time elapsed in transit was long and variable. To help illustrate this point I make use of data from the Indochinese Health and Adaptation Research Project (IHARP), which collected a wealth of detailed data from a sample of 599 Indochinese refugees in the San Diego area in the 1980s.[4] The dataset includes the month and year a refugee left home and the month and year they arrived to the US, which allows me to provide concrete evidence on the long and variable transit times of refugees. Figure 1 shows a histogram of the time spent in transit for four of the main ethnic groups among Indochinese refugees. The median transit time ranged from 9 to 30.5 months, with the Khmer and Hmong facing longer transits. Just as importantly, there was large within-ethnic group variation in transit times: the standard deviation of transit times ranged from 6.1 to 22 months for this sample. Similar surveys undertaken in other regions of the US also support long waits in refugee camps for non-ethnic Vietnamese (Caplan et al., 1991).

The long and variable waits experienced by refugees were for the most part the result of changes in the policies of the involved countries. Origin countries sometimes actively promoted and sometimes hindered attempts to escape. Many refugees needed multiple attempts to escape Vietnam in particular, because arranging transit by boat was difficult once the government worked to prevent refugee departures (Hung and Haines, 1996; Scott, 1989). It was common for refugees to be denied asylum in the first country they reached because countries of asylum worried about the moral hazard of offering easy asylum and worked to redistribute the burden of refugees to neighboring countries. For refugees who reached the camps, the wait was determined mostly by US policy.[5] The total refugee quota

[4]This research used the Indochinese Health and Adaptation Research Project, 1982–1984 dataset (made accessible in 2005 in paper and computer data form). These data were collected by Rubén Rumbaut, and are available through the archive of the Henry A. Murray Research Center at the Radcliffe Institute for Advanced Study, Harvard University, Cambridge, Massachusetts (Producer and Distributor).

[5]Some refugees from Vietnam were able to bypass waiting in refugee camps entirely by participating in the Orderly Departure Program, which was an agreement for direct resettlement between the US and Vietnamese governments. However, waiting for the entry permit and exit visa necessary for resettlement was an equally long and variable process, taking between two and five years for most refugees (Hung and Haines, 1996).

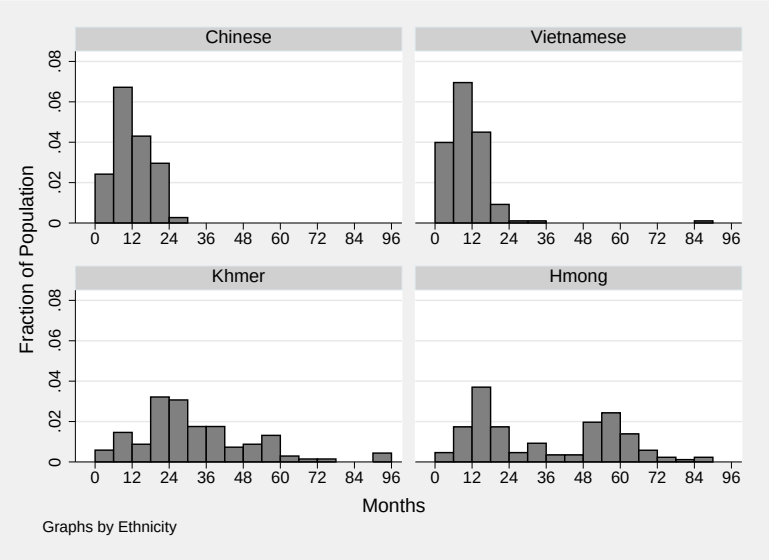


Figure 1: Time Elapsed in Transit by Ethnic Group

for the US was set annually by the President. Domestic politics played an important role here as concerns about the economic and health impact of refugees began to mount, particularly in light of an influx of Cuban refugees and the severe 1981–82 recession. International politics were even more important. A constant theme of discussions was to allocate enough resettlements to each country of first asylum to encourage their continued acceptance of future refugees. For example, a temporary break-down of the principle of asylum in 1979 led to a doubling of refugee resettlement to the US, and an especially large increase in resettlements from Thailand (Haines, 1989).

In light of the long, variable, and exogenous delays faced by refugees, it seems implausible that families were able to select when they entered the US. In Appendix A I test this hypothesis directly by examining the relationship between age at arrival of the child and a broad set of family characteristics in the 1990 US Census (when children still lived with their families and such comparisons are possible). I find correlations of mixed signs that are statistically indistinguishable from zero in most cases. The two exceptions are that that older-arriving children are more likely to be the eldest child for all ethnic groups, and that there is some evidence that ethnic Vietnamese families differed in their pre-migration characteristics or their post-migration investments.

A final reason to rule out this form of selection comes from the large size of Indochinese refugee families. Using the 1990 US Population Census, I find that conditional on having any foreign-born children, the average refugee family had three foreign-born children. In

this case, it is not even clear what it means for the family to be selected based on the age of "the" child. In light of this evidence I will interpret my findings as measures of acquired human capital and not selection on the innate characteristics of the children or their families. I now turn to the results.

3 Empirical Results

In this section I provide the data source, empirical methodology, and results for estimating the effect of age at arrival on adult wages and other socioeconomic outcomes for Indochinese refugees.

3.1 Data and Methodology

I study refugees' adult outcomes in the 2000 Population Census and the 2005–2012 American Community Surveys (ACSs), available online through Ruggles et al. (2010). When combined, these datasets provide a large sample with fairly consistent questions and responses. I identify Indochinese refugees as those born in Vietnam, Cambodia, or Laos who immigrated to the US during the years of heavy refugee flows.[6] Although the Census does not have a variable that separately identifies refugees, independent sources indicate that essentially all immigrants arriving from these countries during these years were refugees (U.S. Immigration and Naturalization Service, 1980–2000). I construct age at arrival using the year the data were collected, the respondent's age at that time, and the year of immigration. I include natives in the sample to identify age effects; the effect of age at arrival thus captures the differential effect of a year spent abroad instead of in the US, consistent with the intuition discussed in the introduction. See Appendix B for an extended discussion that follows closely in the footsteps of Friedberg (1992).

The main outcome of interest is the wages of refugees. For this outcome I restrict the sample to include only those workers who would typically be used in a wage regression: those aged 23–65, not enrolled in school, who work for wages, usually work at least 30 hours a week and worked at least 30 weeks the previous year, have between 0 and 40 years of potential experience, and report positive wage and salary income the previous year. I

[6]For Vietnam, the years are 1976–1990; for Cambodia, 1976–1994; and for Laos, 1976–1996. As a robustness check I have also experimented with adding children who were born in refugee camps and find similar results.

construct hourly wage W as annual wage and salary income divided by the product of hours worked per week and weeks worked in the previous year.[7] I then regress

$$\log(W) = \beta X + \sum_a \alpha_a d_a + \sum_y \omega_y d_y + \sum_{aa} \sum_e \phi_{aa,e} d_{aa,e} + \varepsilon, \tag{1}$$

where X is a vector of control variables that consists of state of residence and gender dummies. d_a and d_y are age and time dummies, while $d_{aa,e}$ is an interaction between age at arrival aa and ethnic group e. Indochinese refugees in the US come primarily from five distinct ethnic groups. The Vietnamese, Lao, and Khmer are the main ethnic groups of the respective countries, but there are also a sizable number of ethnic Chinese from Vietnam and ethnic Hmong from Laos in the US. Adult refugees from these ethnic groups differed greatly in their pre-migration characteristics such as education, occupational background, rural-urban status, and so on; see Appendix A. Given these large between-group differences, I estimate age at arrival effects separately for each ethnic group to avoid compositional biases. The coefficient of interest is $\phi_{aa,e}$, which captures the log-wage of a refugee of ethnic group e who arrived at age aa, relative to a native with the same age, gender, and state of residence.[8]

3.2 Results for Indochinese Refugees

Figure 2 shows the results in the format used for the rest of the paper. It plots the estimated coefficients $\phi_{aa,e}$ and the corresponding 95 percent confidence interval against age at arrival aa, which ranges from 0 (arrive before first birthday) to 5. Refugees with higher age at arrival spent more years of their early childhood in their birth country or in refugee camps and fewer years in the US. Given the poverty and difficult circumstances to which Indochinese refugees were exposed, it is perhaps natural to expect that refugees who were older upon arrival and who had longer exposure to their birth country would have worse outcomes. Graphically, this expectation would imply that $\phi_{aa,e}$ declines in aa, or a negative slope to the plotted lines. The striking finding is that there is instead no trend for any of the five ethnic groups.

I explore the refugees' years of completed schooling as a second outcome. The main reason

[7] From 2008 onward, weeks worked are reported in categories. I use the 2007 ACS data to estimate the mean value of weeks worked within each category, which I apply to the 2008–2012 ACSs.

[8] The immigration literature often considers more general forms of equation (1); Appendix B shows that the coefficients on age at arrival are identified, even in a more general model that allows for cohort or assimilation effects.

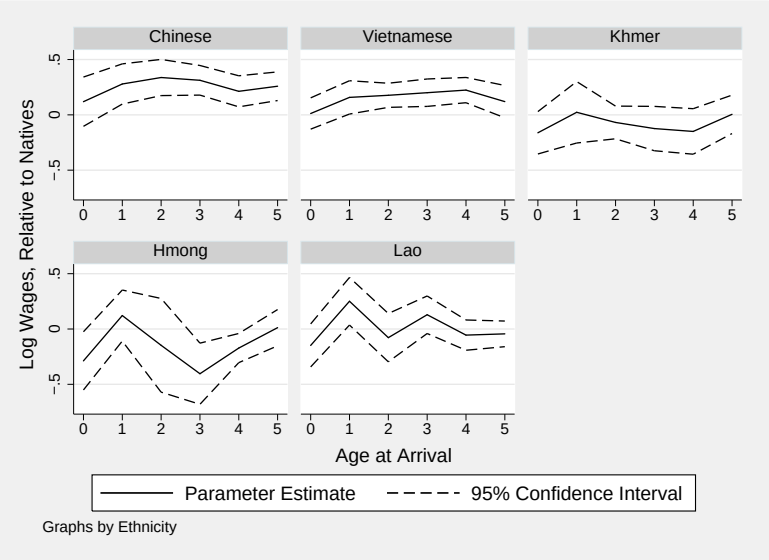


Figure 2: Log-Wages by Age at Arrival

to do so is that fewer sample selection criteria are required to study schooling. Here I limit the sample only to those ages 23–65 who are no longer enrolled in school. The provided educational attainment variable is transformed into years of schooling in the usual way. I then regress the years of schooling on the same set of control variables as in equation (1). Figure 3 again plots the coefficients $\phi_{aa,e}$ against the age at arrival. The same basic finding obtains: there is no trend relationship between age at arrival and schooling for Indochinese refugees.

Table 1 gives estimated results for the simplified case where the effect of age at arrival is restricted to be linear. In columns (2)–(6) I estimate a separate intercept and slope for each ethnic group, in line with figures 2 and 3. The results are as likely to be positive as negative, and none are statistically significant at conventional levels. I then estimate the same regression under the assumption that the slope (but not the intercepts) are common across ethnic groups, to create the largest sample size possible and smallest standard errors possible. Even in this case the estimated effect of age at arrival goes in the wrong direction. Further, the small standard errors make it possible to reject even modest negative effects of age at arrival with a reasonable degree of confidence.

These are the benchmark empirical results of the paper. Appendix C includes a number of robustness checks. There is also no trend by age at arrival for a number of other socioeconomic outcomes, including probability of being employed, total earnings, probability of graduating college, and so on. There is little difference in outcomes between first-generation

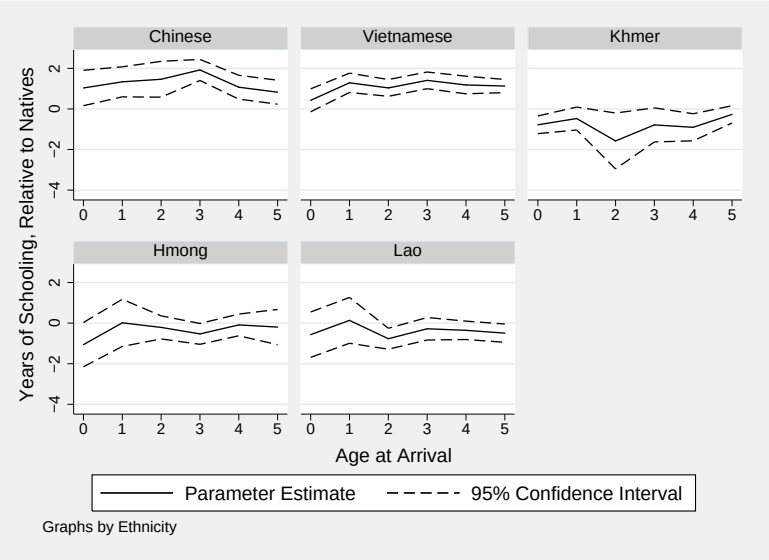


Figure 3: Years of Schooling by Age at Arrival

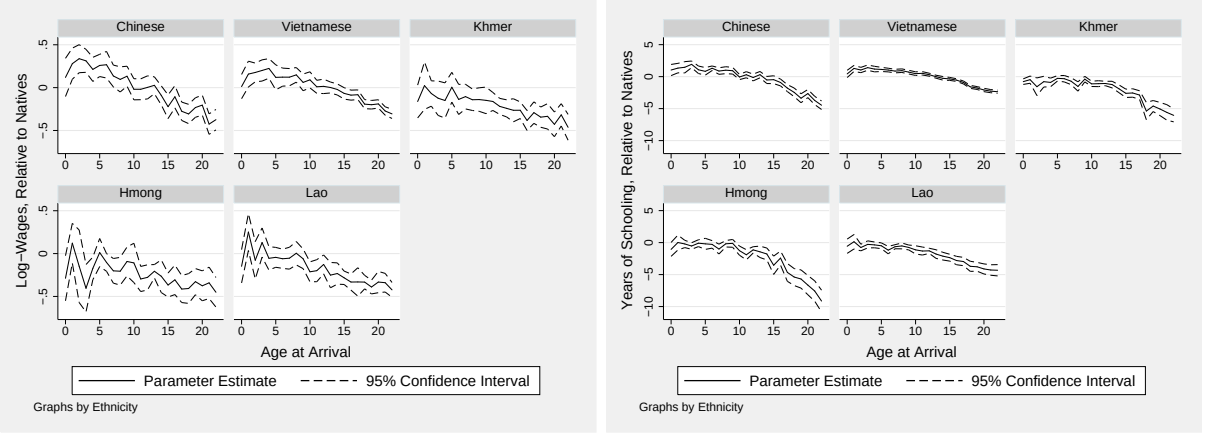
refugees born shortly before resettlement and second-generation refugees born after resettlement, suggesting that effects of in utero exposure to poor countries are also not critical. The result is not driven by breaking out the five ethnic groups; the same results obtain if I instead break results out by country of birth or language spoken at home. In general, there seems to be little difference between the adult outcomes of Indochinese refugees who arrived to the US at the beginning and at the end of early childhood. This impression agrees with existing work in the literature using other outcomes and other groups of immigrants with less dramatic moves (Myers et al., 2009; Lee and Edmonston, 2011; Gonzalez, 2003; Bleakley and Chin, 2004, 2010).

Finally, I emphasize that the results presented here are special to early childhood. To make this point I re-estimate wages and schooling as a function of age at arrival using a sample of refugees who arrived as late as age 22. Figure 4 shows the results. Outcomes begin to decline in age at arrival somewhere in late childhood.[9] The magnitude of the decline is quantitatively large: refugees who arrive at age 22 have 5–10 fewer years of schooling and 30–60 percent lower wages than refugees who arrive in early childhood, with the magnitude varying somewhat by ethnic group.

[9]I estimate a spline regression with a single endogenously chosen breakpoint using nonlinear least squares on all the ethnic groups pooled. The estimated breakpoint is at age 8.0 for wages and 11.7 for years of schooling. The estimated breakpoint for other immigrant groups (discussed below) is sometimes shortly after the end of early childhood.

Table 1: Effect of Age at Arrival on Wages and Schooling

	Vietnamese	Chinese	Lao	Hmong	Khmer	Pooled
Dependent Variable: Log-Wages						
Estimate	0.0064	-0.0043	-0.0090	0.0454	0.0116	0.0070
Standard Error	(0.0156)	(0.0204)	(0.0224)	(0.0268)	(0.0228)	(0.0097)
N	1202	500	384	267	220	2573
Dependent Variable: Years of Schooling						
Estimate	0.058	-0.113	0.000	0.137	0.125	0.043
Standard Error	(0.050)	(0.083)	(0.089)	(0.103)	(0.063)	(0.033)
N	1984	757	673	362	461	4237



(a) Log-Wages

(b) Years of Schooling

Figure 4: Outcomes by Age at Arrival for Older Ages

3.3 Results from Other Immigrants

Indochinese refugees are useful because they experienced a large change in income and environment upon migration and because it is implausible that they were selected on age at arrival. A disadvantage of studying Indochinese refugees is that the conditions they faced were sufficiently extreme to raise concerns about selective mortality risk. This section provides the same analysis of outcomes as a function of age at arrival for four other groups of immigrants from poor countries that were not exposed to such extreme events. The results are similar, suggesting that selective mortality (or selection more generally) does not explain the results.

The first group of immigrants consists of refugees from Afghanistan and Ethiopia; although these countries were quite poor, there were fewer immigrants and their parents are generally considered to have been more selected.[10] Since the strategy is to look at the adult outcomes of refugees who entered as children, these are the most recent groups of poor-country refugees that can be used in the analysis; more recent groups such as the Somali or Bosnian refugees are not old enough yet. The second group is Cuban immigrants. Cuba is much richer than Vietnam, Cambodia, or Laos, but still only one-fourth as productive as the US on a PPP GDP p.w. basis in 2005. It is not possible to disentangle whether Cuban immigrants were refugees, family reunification cases, or other categories. The third group is all immigrants from countries with 2005 PPP GDP p.w. less than 5 percent of the US.[11] These immigrants come from poor countries with worse environments for human capital formation, but are likely to be quite selected. Finally, the fourth sample is immigrants from Mexico.

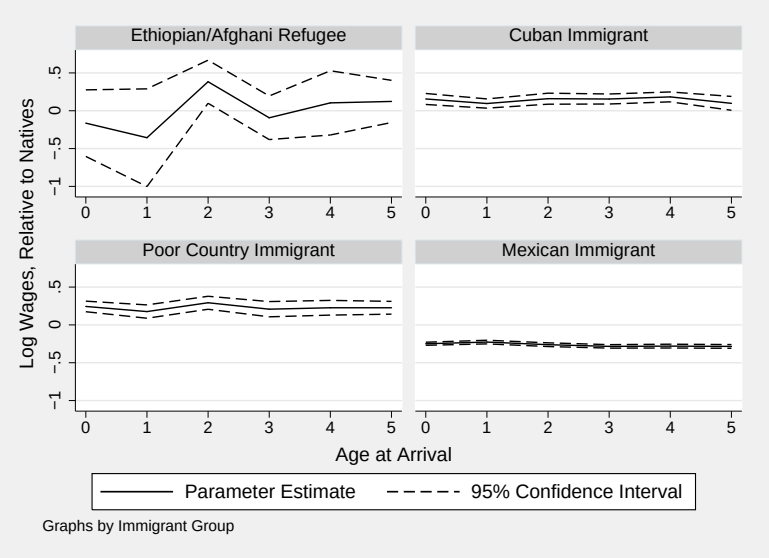


Figure 5: Log-Wages by Age at Arrival for Other Immigrant Groups

Figure 5 shows the results for the log-wage patterns by age at arrival for these four groups. The same basic pattern applies as for the baseline group: there is no trend in log-wages by age at arrival up to age 5. Similar results obtain for other socioeconomic outcomes. At this point it seems that the weak relationship between socioeconomic outcomes and age

[10]The sample includes immigrants who entered between 1980 and 1993, which were generally years of high refugee flows (U.S. Immigration and Naturalization Service, 1980–2000).

[11]The countries distinguished in the Census and so included are: Haiti, Bangladesh, Nepal, Ghana, Guinea, Liberia, Senegal, Sierra Leone, Kenya, Somalia, Tanzania, Uganda, Zimbabwe, and Eritrea.

at arrival is quite general.The next goal is to understand how this fact can be reconciled with the large literature that establishes the importance of early childhood human capital within countries.

4 A Model of Early Childhood Human Capital

In this section I introduce a model of early and late childhood human capital accumulation in the spirit of Cunha et al. (2010) and Del Boca et al. (2014). I study the model's predictions for refugees and how those predictions can be made consistent with the findings of the previous section.

4.1 Life-Cycle Human Capital Accumulation

The model describes the human capital accumulation and labor market decisions of one cohort of children and their families in countries $i \in \{1, 2, ..., I\}$. Within each country there is a continuum of heterogeneous families consisting of one working adult and one child who is newly born at time 0. There are two forms of ex-ante heterogeneity in the model. First, children are endowed with human capital at birth h_0. Second, parents are endowed with human capital h_p, which is determined by events that happen before the child is born. The two-dimensional human capital endowment is drawn from a joint distribution $F_i(h_0, h_p)$ defined on $(0, \infty)^2$, which may vary across countries.

Time is continuous and the parent and child are both infinitely lived. Each is endowed with a single unit of time at each instant. The child's life is split into three periods: early childhood, school, and work. Early childhood includes the first five years of the child's life, before schooling. School lasts from age 5 until an endogenously chosen graduation date. After graduation the child joins his or her parent in the labor force and works into the infinite future.

The heart of the model is the human capital production function, which explains how human capital is generated given the family's endowment (h_0, h_p), the country i that they live in, and the investments they make in the child during early childhood and school. Following the literature, human capital production occurs in two distinct stages. Early childhood human capital h_1 is determined by human capital at birth h_0 and a composite of the investments

made during early childhood x_1, combined using a CES production function:

$$h_1 = \left[\lambda_1 h_0^{\frac{\sigma_1-1}{\sigma_1}} + (1-\lambda_1) x_1^{\frac{\sigma_1-1}{\sigma_1}}\right]^{\frac{\sigma_1}{\sigma_1-1}}. \tag{2}$$

λ_1 is the weight on human capital at birth and $1-\lambda_1$ the weight on the composite investment, while σ_1 is the elasticity of substitution between the two.

The composite investment is a weighted combination of four different underlying inputs. The first input is an exogenous term that captures the quality of country i's environment for human capital formation z_i, including for example the prevalence of diseases or the quality of medical infrastructure. The second input is the goods purchased for the child g_1, which includes books or vaccines. The final two inputs are the time spent by parents with their children p_1 and the human capital of the parents h_p, which enter the composite investment separately. Given these inputs, x_1 is determined using a power function as in Del Boca et al. (2014):

$$x_1 = z_i^{\omega_1} g_1^{\omega_2} h_p^{\omega_3} p_1^{\omega_4}. \tag{3}$$

The ω_is are weight parameters that determine the relative importance of the different inputs.

At age 5 children start school. They remain in school for an endogenously chosen period of time S, graduating at age $5+S$. Their human capital at graduation is again a function of their human capital at the start of school h_1 and the composite investment made during the school years x_2:

$$h_2 = \left[\lambda_2 h_1^{\frac{\sigma_2-1}{\sigma_2}} + (1-\lambda_2) x_2^{\frac{\sigma_2-1}{\sigma_2}}\right]^{\frac{\sigma_2}{\sigma_2-1}}. \tag{4}$$

λ_2 is the weight on early childhood human capital and $1-\lambda_2$ is the weight on composite investment, while σ_2 is the elasticity of substitution between the two.

The composite investment during schooling is a weighted combination of five inputs. The first four are similar to early childhood: country environment z_i; goods g_2; and parental inputs p_2 and h_p. Finally, the human capital also depends on how long children stay in school, S. The terms are again combined using a power function,

$$x_2 = z_i^{\eta_1} g_2^{\eta_2} h_p^{\eta_3} p_2^{\eta_4} S^{\eta_5}, \tag{5}$$

where the η are the weight parameters that determine the relative importance of the different inputs. I use the notation η and ω to draw attention to the fact that the weights on the inputs may vary between early childhood and schooling.

The human capital production function in equations (2) – (5) borrows heavily from the microeconomic literature on early childhood human capital. Two key insights from that literature will be useful for understanding how the model works. First is the importance of the elasticity parameters σ_1 and σ_2 (Cunha and Heckman, 2007; Cunha et al., 2010). Their role in this model is to determine the extent to which a disadvantage in the form of low human capital at birth or low early childhood human capital can be remediated by subsequent investments. Cunha et al. (2010) estimate that $\sigma_1 > 1$ and $\sigma_2 < 1$, indicating that it is relatively easy to remediate low human capital at birth but much more difficult to remediate low early childhood human capital. Second is the importance of allowing for multiple types of inputs and allowing their importance to vary over the life cycle. Del Boca et al. (2014) estimate that the relative role for goods inputs and parental inputs changes over the life cycle. This flexibility is important in matching the empirical findings of the previous section.

4.2 Labor Market

I embed the human capital production function into a simple lifetime income maximization problem. Human capital is fixed after graduation. The Indochinese refugees studied in the empirical implementation were quite young, which makes this assumption less restrictive. After graduation, children enter the labor market and work full time. They are endowed with a linear production technology that turns h_2 units of human capital into $A_i(t)h_2$ units of the single output good, which can be used for consumption or as the goods inputs in human capital production. $A_i(t)$ is an exogenous TFP term. $A_i(t)$ grows at a rate γ that is common across countries, but the level of initial productivity $A_i \equiv A_i(0)$ differs across countries.

Finally, I assume that each family behaves altruistically and that children have no preferences over whether they work or study. In this case it is possible to focus on the family's income maximization problem and ignore the (trivial) utility maximization problem given optimal income. Families then choose the duration of schooling for children and the quantity of the two types of inputs at the two stages of the life cycle to maximize the present discounted value of lifetime earnings net of the cost of goods and the foregone earnings of

the parents. I assume that parents purchase all goods inputs and forego all labor earnings at the start of each of the respective stages of the life cycle, that is, at date 0 for early childhood and at date 5 for school. This assumption simplifies the discounting of investments without foregoing any insights. In this case the family's problem is to maximize lifetime earnings,

$$\max_{g_1,g_2,p_1,p_2,S} \qquad \int_{5+S}^{\infty} e^{-rt} A_i(t) h_2 dt - g_1 - e^{-5r} g_2 - A_i h_p p_1 - A_i e^{5(\gamma - r)} h_p p_2 \qquad (6)$$

where h_2 is derived from equations (2) – (5) and r is the exogenous interest rate. The next step is to consider the model's predictions.

4.3 A Simple Case: Isoelastic Production

To build insight, it is useful to start with the special case of an isoelastic production function, $\sigma_1 = \sigma_2 = 1$. This case is useful because the expression for optimal investment and human capital can be derived in closed form. I discuss the properties of the solutions here; derivations are available in Appendix D.

The model can generate differences in human capital from each of the four exogenous factors: z_i, h_0, h_p, and A_i. The appropriate elasticities of h_2 and h_1 with respect to these factors are given in Table 2. Two of the effects in Table 2 are straightforward. Children who grow up in countries that offer a better environment z_i, or children who have more human capital at birth h_0, will also have higher human capital at the end of early childhood and into adulthood. The quantitative importance of these effects is controlled by two sets of parameters, which are easiest to understand in the expressions for h_2. The numerator of these expressions is simply the technological coefficient on h_0 or z_i; the more weight the technology puts on these factors, the more important they are for accumulated human capital. This effect is amplified because families allocate more of the endogenously chosen inputs (goods and parental time) to children who have better environments or who are more able. This amplification effect is reflected in the markup $(1 - \Psi_2 - \Psi_4)^{-1}$ for adult human capital, where Ψ_2 and Ψ_4 are the technological weights on goods and parental time and hence $\Psi_2 + \Psi_4$ is the total return to scale of endogenous factors. $\Psi_2 + \Psi_4 < 1$ is a necessary condition for an interior solution and is assumed throughout.

The latter two effects are more subtle. Children whose parents have more human capital enjoy a direct advantage. However, higher human capital also enables their parents to

Table 2: Elasticity of Human Capital with Respect to Exogenous Characteristics

	Adult Human Capital, h_2	Early Childhood Human Capital, h_1
h_0	$\frac{\lambda_1\lambda_2}{1-\Psi_2-\Psi_4}$	$\lambda_1 + \frac{\lambda_1\lambda_2(\Phi_2+\Phi_4)}{1-\Psi_2-\Psi_4}$
z_i	$\frac{\Psi_1}{1-\Psi_2-\Psi_4}$	$\Phi_1 + \frac{\Psi_1(\Phi_2+\Phi_4)}{1-\Psi_2-\Psi_4}$
A_i	$\frac{\Psi_2}{1-\Psi_2-\Psi_4}$	$\Phi_2 + \frac{\Psi_2(\Phi_2+\Phi_4)}{1-\Psi_2-\Psi_4}$
h_p	$\frac{\Psi_3-\Psi_4}{1-\Psi_2-\Psi_4}$	$\Phi_2 + \Phi_3 + \frac{(\Psi_2+\Psi_3-1)(\Phi_2+\Phi_4)}{1-\Psi_2-\Psi_4}$

Note: $\Psi_i \equiv \omega_i\lambda_2(1-\lambda_1) + \eta_i(1-\lambda_2)$ and $\Phi_i \equiv \omega_i(1-\lambda_1)$ are the technological weight parameters on factor i in h_2 and h_1 in the isoelastic case.

earn more in the labor market. The effect of parental human capital on parental time allocation depends on parameters. In some cases the indirect effect on time allocation is sufficiently large that children whose parents have more human capital actually receive less total parental inputs. Finally, A_i provides the TFP multiplier effect that is well-known in the literature (Manuelli and Seshadri, 2014; Erosa et al., 2010). The logic of the multiplier is that higher A_i lowers the cost of goods relative to wages. The lower relative price of goods leads families in high-TFP countries to allocate endogenously more market goods to their children, which in turn raises average human capital.

The main focus here is on the model's predictions for refugees. A refugee is defined as someone who, at age aa, moves to another country with different A_i and z_i. Generically, I assume that these values rise with the move. I allow refugees to have distributions of h_0 and h_p that differ from natives, but I assume that refugees who arrive at different ages have the same distribution of h_0 and h_p. This is the model equivalent of the identifying assumption that refugees can be selected but that refugees who arrive at different ages are not selected differently. The main comparison of interest is between those who arrive at ages 0 and 5, that is, before or after early childhood investments are set and h_1 is determined.[12]

An important question arises as to the pre-migration beliefs of immigrants. I start with the simplifying assumption that this move is entirely unanticipated. In this case, families in poor countries make optimal investments under the assumption that they will live their entire lives in their native country. When the child is aa years old, they are unexpectedly

[12] An alternative would be to discretize the life cycle or at least early childhood more finely. Doing so would take the model further away from the existing literature, making it harder to compare the formulation to existing results (Cunha et al., 2010). The main advantage of this is to consider the elasticity of substitution across sub periods within early childhood. Given the empirical finding that there are no negative consequences of time spent abroad until after early childhood is complete, this extension does not seem promising.

moved to the US, at which point the family re-optimizes all future investments, taking past investments as given. These beliefs are the natural starting point in the analysis of refugees, but I consider alternatives in Section 7.

The model generally predicts that refugees' human capital and wages as adults decline in age at arrival. This prediction can be understood in two steps. First, the model predicts that the late-arriving refugees will have lower early childhood human capital because they had lower values of z_i and A_i in early childhood. The expressions in Table 2 show that h_1 is increasing in z_i and A_i under fairly general conditions on the human capital production function. For example, it suffices to assume that $\lambda_1 > 0$ (positive technological weight on investments in early childhood) and either $\omega_1 > 0$ or $\omega_2 > 0$ (positive technological weight on environment or goods in early childhood). Second, the deficit in early childhood human capital persists through to adulthood if $\lambda_2 > 0$ (positive technological weight on early childhood human capital in subsequent human capital formation). Hence, the isoelastic model can be consistent with the data only by setting some of the technological parameters to zero, which shuts down some mechanisms for human capital formation. Next I consider the results when the elasticity of substitution is allowed to deviate from unity.

4.4 Post-Migration Remediation

The isoelastic model is convenient because it admits closed forms, but the restriction $\sigma_2 = 1$ is not innocuous. $\lambda_2 > 0$ ensures that children with higher human capital at age 5 have higher human capital in adulthood if all other investments are held equal. Of course, all other investments need not be held equal; parents can choose to invest more (or less) in their children in response to low human capital at age 5. σ_2 controls the extent to which investments made during schooling can be used to remediate deficits in early childhood human capital.[13] It also indirectly controls the endogenous willingness of families to engage in remediation. Larger values of σ_2 imply more effective remediation; more willingness of families to remediate; and smaller human capital gaps in adulthood for later-arriving refugees.

For σ_2 sufficiently large, human capital gaps in adulthood are small enough to be difficult to detect statistically and economically unimportant. I need to restrict and calibrate the model to make this statement more precise. I focus on a special case that shares two

[13]On the other hand, σ_1 controls the ability of early childhood investments to remediate low human capital at birth. As long as refugees who arrive at different ages have the same underlying distribution of h_0, the value of σ_1 is innocuous.

common features with much of the existing macroeconomic literature. First, I focus on the case where goods and time of children are the only inputs to the production of human capital. Second, I calibrate my model to generate a gap in early childhood human capital between poor and rich countries in line with Manuelli and Seshadri (2014). Given this special case, I can ask how large σ_2 would need to be for these large differences in early childhood human capital to be remediated by adulthood.

The simplified model suggested by this discussion is:

$$h_1 = g_1^{\theta_1},$$
$$h_2 = g_2^{\theta_1}(Sh_1)^{\theta_2}.$$

I normalize A_{US} to 1. I then calibrate a TFP for "Indochina" A_{IND} as well as θ_1 and θ_2 to fit a factor of 33 income difference between the US and the Indochinese countries, consistent with PWT data; a factor of 3.6 difference in early childhood human capital, consistent with Manuelli and Seshadri (2014); and average years of schooling of 12 in the US, consistent with Barro and Lee (2012).[14]

The predictions of interest concern adult wages of refugees as a function of age at arrival. The isoelastic model predicts that wages decline strongly in age at arrival. This is shown as the solid black line in Figure 6. Refugees who arrive at the end instead of the beginning of early childhood have nearly 50 percent lower log-wages. This equates to a slope coefficient of -0.1, and clearly contradicts the empirical results of Figure 2. Note that the isoelastic model does predict substantial remediation, since the log-human capital difference at age 5 was -1.26.

I then vary σ_2 to study its effects on remediation and the relationship between adult wages and age at arrival. All the remaining parameters are kept fixed to isolate the role of the changing elasticity of substitution. A convenient feature of this model is that the predicted human capital gap at age 5 is independent of σ_2, so that figure is consistent with Manuelli and Seshadri throughout.

The results of this experiment are shown using the broken lines in Figure 6. Cunha et al. (2010) estimate $\sigma_2 \approx 0.5$. In this case it is even more difficult for immigrant families to

[14]The resulting human capital production function parameters are $\theta_1 = 0.35$ and $\theta_2 = 0.24$. In terms of the notation above, the parameters required are $\sigma_1 = \sigma_2 = 1$, $\lambda_2 = 0.24$, $\omega_2 = 0.35$, $\eta_2 = 0.46$, and $\eta_5 = 0.32$, with the remaining technological weight parameters set to 0. I also fix $r = 0.04$ and $g = 0.02$, but these parameters are less essential for the results. The mapping between my calibration and the one in Manuelli and Seshadri (2014) is not clear because they use a continuous time human capital accumulation model from age 5 onward.

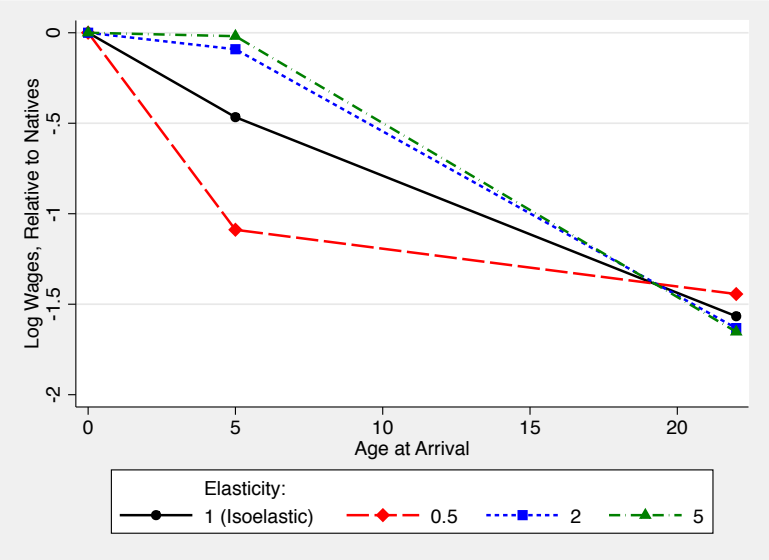


Figure 6: Model-Predicted Log-Wage Patterns for Immigrants

remediate the low early childhood human capital of their children with investments made after arrival, leading them to make fewer such investments. This logic leads the model to predict an even stronger relationship between outcomes and age at arrival. If their estimate of σ_2 is in the right ballpark then there is little hope for the remediation hypothesis. On the other hand, higher values of σ_2 imply greater remediation for children who come in early childhood. For example, $\sigma_2 = 2$ implies that age-5 arrivals have a log-wage gap as adults of a little more than 9 percent; the coefficient on age at arrival would be roughly -0.02. Larger coefficients, such as $\sigma_2 = 5$, imply so much remediation, and so little variation in human capital as adults, that it would be difficult to detect a relationship empirically. The critical question taken up in the next subsection is whether the evidence support such a high elasticity of substitution.

In sum, the generic prediction of the model is that wages of refugees decline in their age at arrival. This prediction rests on three logical steps. First, there is a disadvantage of growing up in poor rather than rich countries: either the poor country environment directly lowers human capital, or low incomes lead parents to spend less on their children ($\omega_1 > 0$ or $\omega_2 > 0$). Second, early childhood human capital differences persist to adulthood if all other factors are held equal ($\lambda_2 > 0$). Finally, remediation of early childhood human capital gaps is sufficiently difficult that families' endogenously chosen investments do not close such gaps entirely (small σ_2). Given this logic, it is clear that there are three particular parameterizations of the model that are capable of matching the data:

1. $\omega_1 = \omega_2 = 0$. This case implies that goods and country environment are not important inputs in early childhood, which also implies that there is no strong disadvantage to growing up in a poor country.[15]

2. $\lambda_2 = 0$. This case implies that any differences in early childhood human capital are irrelevant for subsequent outcomes. It corresponds to the idea that early childhood human capital does not matter.

3. $\sigma_2 > 2$. This case implies that there are disadvantages to growing up in poor countries, but that refugees and their families find it easy to remediate deficits in early childhood human capital after arrival to the US.

It is important to differentiate between these three cases because they have very different implications for the human capital production function as well as cross-country human capital differences. Cases 1 and 2 both suggest that early childhood human capital is unlikely to be a significant contributor to cross-country income differences. On the other hand, case 3 is largely silent about how large such differences may be, except to note that they can be remediated by refugees using investments made while in school. In the next section I consider the evidence on these three cases.

5 Evidence on Human Capital Production Function

5.1 Other Evidence from Refugees

The remediation hypothesis has two additional, testable implications. First, it implies that outcomes should decline in age at arrival if they are observed at younger ages, before remediation happens. Second, it implies that post-migration investments should increase in age at arrival, since this is the force that allows disadvantaged older-arriving children to catch up by adulthood. Each of these predictions is testable using data from refugees.

The census data are somewhat limited on this dimension. The main form of investment available is completed education. Figure 3 showed that completed years of schooling are independent of age at arrival; similar results apply for probability of completing high school or

[15]Inspection of Table 2 shows that it is also necessary to set $\eta_1 = \eta_2 = 0$ for the appropriate elasticities to be exactly zero. The reason is that the model admits a second-order effect whereby parents spend more time with children in early childhood in rich countries in anticipation of the advantages the children will enjoy in their school years stemming from $\eta_1 > 0$ or $\eta_2 > 0$. This effect is quantitatively small and is abstracted from.

college. In a companion paper, I use alternative specialized data sets that sampled refugees when they were adolescents (Schoellman, 2014). These data measure both investments such as time spent studying or time spent by parents helping students study as well as outcomes such as grades, reading and math test scores or mental well-being. I show in that paper that all of the investments and most outcomes show no variation in age at arrival. The main exceptions are reading test scores and self-esteem, both of which do decline in age at arrival, especially when measured in early adolescence. This is not surprising given the limited English language training received by immigrants while in refugee camps. The lack of pattern along other dimensions is evidence against remediation hypotheses.

5.2 Evidence from Early Childhood Literature

The literature on early childhood human capital provides substantial evidence on the question at hand. First, that literature establishes clearly that early childhood human capital matters ($\lambda_2 > 0$). The literature reviews cited in the introduction provide numerous examples of early childhood interventions and investments that have lasting effects on children. For example, the Perry Preschool program has been running long enough to verify that treated children had higher test scores, graduation rates, and earnings as adults (Cunha et al., 2006). This type of evidence is inconsistent with $\lambda_2 = 0$.

Of the remaining two cases, the evidence seems to favor case 1, goods and environment are not critical inputs in early childhood. A number of early papers in this literature found some positive role for goods, particularly nutritional supplements (Alderman et al., 2006; Maluccio et al., 2009). However, two recent randomized trials have tested precisely the question of whether goods or adult interactions are responsible for long-term improvements in outcomes. Each found that the treatment group exposed to adult interaction saw positive gains, whereas the treatment group exposed to nutritional supplements saw no gains or gains that faded over time (Walker et al., 2005; Attanasio et al., 2012). The only paper that structurally estimates these parameters finds that the relative importance of goods rises over the life cycle, although it is hard to interpret the level of importance for one versus the other (Del Boca et al., 2014).

On the other hand, studies that have investigated the possibility of remediation have suggested that the elasticity of substitution is likely to be low. As was mentioned before, Cunha et al. (2010) structurally estimate $\sigma_2 \approx 0.5$. A low elasticity of substitution is also consistent with the large literature that studies the effectiveness of intervention programs

based on when they are administered. The general finding is that programs administered in early childhood show large gains; those administered in early adolescence, smaller but still positive gains; and those administered in late adolescence often show no impact (Cunha et al., 2006).

A final useful result comes from the studies of Romanian orphans. A combination of pro-natalist policies and economic stagnation under Ceauşescu led parents to abandon large numbers of children to state-owned orphanages in Communist Romania. Conditions in the orphanages were dire: children were mostly confined to cots, given few toys, spoken to or allowed to play rarely, and fed primarily gruel. When Ceauşescu's regime fell in 1989, some of these orphans were adopted abroad. Thus, Romanian orphans provide a natural experiment of children who experience a dramatic increase in country environment and income similar to the Indochinese refugees. The critical difference is that Romanian orphans received upgrades in their parental input (h_p and p_1), trading essentially no parental input in the orphanages for the inputs of their adoptive parents in Britain. On the other hand, 95 percent of Indochinese refugees in my sample migrated with at least one biological parent and so had similar parental inputs before and after migration.[16] Thus, the important distinction between the refugees and Romanian orphans is that the refugees *only* changed their country environment and labor market productivity z_i and A_i, while the Romanian orphans also changed their parental input h_p and p_1.

A team of researchers has intensively studied the ongoing progress of a sample of Romanian orphans adopted into Britain. To date they have completed surveys of the orphans as early as age 4 and now through age 15 (Rutter and The English and Romanian Adoptees (ERA) study team, 1998; Beckett et al., 2010). Their most striking result is a consistent, negative relationship between age at arrival/adoption and outcomes, exactly the opposite of the finding for refugees. Romanian orphans who arrived before they were six months old do as well as British-born adoptees and better than Romanian orphans who are adopted at older ages. They also find smaller and less persistent differences between those who were adopted at 6–24 months and those who were adopted at 24–42 months, with the former sometimes doing better and sometimes roughly the same. Recent work has established that it was psychosocial deprivation and not lack of nutrition that is strongly linked to negative outcomes among orphans (Rutter et al., 2012). The difference in outcomes between Romanian orphans and Indochinese refugees suggests that parental inputs, rather than goods or country environment, are the critical input in early childhood. The inability of

[16] Computed using the 1990 US Census.

adoptive parents to remediate the deficits in Romanian orphans also establish evidence in favor of a low σ_2.

In summary, the evidence from a large number of studies supports long-lasting effects of early childhood that are inconsistent with the hypothesis that $\lambda_2 = 0$. A smaller number of papers have addressed the potential for remediation of early childhood human capital deficits and found that it is difficult, which implies a low value of σ_2. A low value of σ_2 is also consistent with the post-migration schooling investments of refugees documented in this paper. On the other hand, a small role for goods or country environment effects in early childhood ($\omega_1 = \omega_2 = 0$) provides the only explanation that is consistent with the experience of Indochinese refugees as well as the existing literature on early childhood human capital. In the next section I address what a model parameterized along these lines implies for cross-country differences in early childhood human capital.

6 Implications for Cross-Country Human Capital Differences

The findings of the previous section have strong implications for cross-country differences in early childhood human capital. Recent work by Manuelli and Seshadri (2014) and Erosa et al. (2010) have shown that if goods are important inputs to human capital formation, then this can act as an important mechanism for generating large cross-country differences in human capital. Manuelli and Seshadri (2014) suggested that this mechanism could be important already in early childhood, and indeed that early childhood human capital differences might be nearly a factor of four. The evidence presented here greatly limits the scope for such a mechanism. If low income has a direct effect on human capital formation, then older-arriving refugees should have lower human capital as adults, or at least lower human capital and higher investment as adolescents. Neither appears to be the case. Thus, it would seem that the most promising mechanism for generating large cross-country differences in early childhood human capital is ruled out by the evidence. The next question is whether alternative mechanisms offer more promise.

6.1 Parents and Early Childhood Human Capital

The paper's main empirical contribution is to show that there is no consistent relationship between adult outcomes and age at arrival for refugees or other immigrants. The model shows that this moment is informative about the role of country-specific factors that immigrants would have left behind, such as income levels or country environment. The lack of relationship suggests that these channels are likely not important. However, there is one remaining channel: parents. The relationship between outcomes and age at arrival is not informative about parents because most immigrants had the same parents before and after the move. Thus, it is worth asking whether it is possible for cross-country differences in parental investments (through h_p or p_1) to generate cross-country differences in early childhood human capital. This possibility has not been as well-studied in the literature.

To discipline this analysis requires alternative empirical evidence. One possible source of evidence comes from the Romanian orphans, who experience a dramatic change in parental inputs at an exogenously given age. Rutter et al. (2012) focus on children who experienced perhaps the lowest "parental" investment imaginable: orphans who spent their entire pre-adoption life in institutional care with almost no adult interaction. Children who spent more than six months in such care showed signs of significantly lower brain formation and head growth, as compared to a control group who exited before six months. Unfortunately the children are not yet old enough to have worked and so wage data are not available to use as a common metric. However, there is information on IQ as of age 15. That evidence suggests that children who were in institutional care for more than six months score 15–20 points lower on IQ tests, as compared to those adopted earlier. Bowles et al. (2001) collect a number of studies that estimate the log-wage return to test score or IQ and find a median return of 0.07 per standard deviation. This value would imply a wage or human capital impact of 7–10 percentage points, for severe deprivation of more than six (but less than 42) months.

An alternative way to judge the importance of parents is to look at the level difference between immigrants and natives. Of course, this statistic is biased to the extent that immigrants' parents are selected, which is precisely the reason it was avoided in the first place. However, a useful suggestive case is Mexican immigrants, because the existing evidence suggests that adult Mexican immigrants are extremely comparable on both observed and unobserved dimensions to non-migrants (Moraga, 2011). If this is the case, then the level of performance of Mexican immigrants relative to natives is an indicator of the difference in parental inputs. That difference is 25–30 percent, which can be seen as the level in Figure

5.

Although the evidence is admittedly incomplete, both statistics point in the same direction. Differences in parental inputs seem to be a more promising mechanism for generating cross-country differences in early childhood human capital, as compared to goods or country environment effects. The evidence also suggests that the plausible differences generated in this way are likely to be much smaller than what previous work had suggested using goods as the mechanism. Differences in the range of 7–30 percent lower human capital are much smaller than the factor of 4 implied in Manuelli and Seshadri (2014).

6.2 Human Capital Formation after Early Childhood

The findings for human capital formed during the schooling years are clear, but the interpretation is not. Figure 4 shows that both wages and school completion decline in age at arrival for children who arrive after early childhood. The exact turning point varies somewhat by country and by outcome, but is generally between ages five and ten. There are a number of possible explanations for this finding. In the context of the model it could be the case that the production function parameters do change after early childhood, allowing for η_1 or η_2 to be positive. This is consistent with the assumption of Manuelli and Seshadri (2014) and Erosa et al. (2010) for the school period and intuitively is consistent with the idea that the role of parents is diminished and the role of other factors enhanced once children start school. However, a number of alternative explanations seem equally promising. For example, age 10 is approximately when cognitive plasticity in children seems to decline, implying a lower ability to adapt (Lenneberg, 1967). This explanation is particularly promising in light of the finding of Bleakley and Chin (2010) that English language ability of immigrants declines in age at arrival past age 10 or so. Alternatively, these patterns could reflect low exposure to formal education, low education quality, or a limited ability to learn from peers in these countries (Schoellman, 2012; Lucas Jr and Moll, forthcoming). The evidence at hand is not well-suited to discriminate between these theories.

7 Alternative Explanations

In this section I consider two alternative explanations of the paper's empirical findings. First, I allow immigration to be anticipated, contrary to the simplifying assumption in the baseline model. Second, I consider whether measurement error provides a plausible

alternative explanation.

7.1 Anticipated Immigration

In the baseline model immigration is unanticipated. For the early waves of the Indochinese refugees this is probably the natural assumption, because the possibility of immigrating from these countries before the US started accepting refugees was trivial. For example, over the entire 1960s the US accepted 1200 immigrants from Cambodia, 100 from Laos, and 4600 from Vietnam. These figures were minuscule compared either to the 3.3 million immigrants the US accepted that decade or the nearly 50 million person population of the three countries in 1965 (Heston et al., 2012). However, the Indochinese refugee flows were protracted, and the consensus is that by the mid-1980s Indochinese citizens became aware of a significant probability of third-country resettlement for refugees. The relevant question is: to what extent does that change the model's predictions about the relationship between age at arrival and adult outcomes?

Qualitatively, the answer is that it allows the model to come closer to matching the data, even without imposing the strong technological restrictions discussed in the previous section. The reason is that if immigration is anticipated, forward-looking parents who anticipate that their children will work in the US understand that the benefits to investment are higher and so invest more. The strength of this effect depends on how well parents can predict resettlement and the parameterization of the model. For example, if refugees can perfectly predict their future resettlement in the parameterized isoelastic model from section 4.4 and figure 6, then future migrant families invest as much in their children as US natives. After resettlement, refugees would have the same adult outcomes as children born in the US, with no relationship between age at arrival and adult outcomes. In other words, allowing future refugees to anticipate their resettlement generates predictions consistent with the data.

There are at least three reasons to be skeptical of this extreme case. First, the probability of resettlement was much lower than this calculation suggests. As was mentioned earlier, less than half of those who fled their country of birth during this period were resettled abroad; many others were settled locally or repatriated to their home country. If refugees place a positive but less than one probability on resettlement, then the model's predicted negative relationship between adult outcomes and age at arrival is qualitatively restored. Generating a lack of relationship between age at arrival and outcomes relies on all refugees anticipating perfectly their immigration from birth.

A second reason to be skeptical of this extreme case is that it requires implausible expenditures for the necessary early childhood investments. For example, the US Department of Agriculture produces an annual report that estimates the expenditures of US families on their children (Lino, 2010). They also report several estimates from other sources. The lowest estimate from any source in the 2010 report is that families spend 21 percent of their budget on a child. To make this calculation conservative, I focus my attention on expenditures on food, health care, child care, and education, which comprise about 40 percent of the total spending on children, or a little more than 8 percent of the family's expenditures. Given the 33:1 income difference between the US and Indochinese countries, the implication is that future refugee families would have to spend about 250 percent of their annual income per child to match US spending levels. It is difficult to see how families could have borrowed or spent this much, particularly in the context of refugee camps.

Finally Appendix D.1 shows that the patterns for those who arrived before or after 1981 are roughly similar, despite the fact that earlier refugees must have known much less about how and where they could be resettled. In sum, allowing immigration to be anticipated changes the quantitative predictions of the model by flattening the predicted relationship between adult outcomes and age at arrival. However, for anticipated immigration to explain the findings entirely would require both an implausible amount of foresight on the part of refugee families and extraordinary borrowing and spending power while in their country of birth or refugee camps. Further, it would seem to contradict the experiences of early and late-arriving refugees.

7.2 Measurement Error in Age at Arrival

A second alternative explanation for the empirical findings is that refugees may not accurately report their arrival year, which is used to construct age at arrival. Measurement error in the right-hand side variable tends to attenuate the estimated coefficient, so if arrival year is measured with sufficient noise then this could explain the findings. This is potentially important given the existing evidence that year of arrival may not be well-measured (Lubotsky, 2007). However, this evidence is unlikely to play a significant role for the findings. The major problem documented in Lubotsky (2007) is that the standard question on year of arrival is not well-designed to capture the experience of immigrants who enter and exit the US multiple times for extended spells. This issue is unlikely to apply to refugees. Further, while refugees might have some recall bias for their year of immigration, the required magnitude of this bias is extraordinarily large for such an important event in their

lives.

To quantify this point, I return to the parameterized model of section 4.3, which makes predictions about the log-wages of refugees who arrive at age 0 and age 5. If I regressed log-wage on age at arrival in that parameterized model, I would find a coefficient of -0.091, or a wage decline of roughly 10 percent per year. I then simulate measurement error in that model and run the same regression. Simulated measurement error takes a simple form: a fraction π of each arrival group (age 0 or age 5) misreports that they are in the other group. In this case, $\pi = 0$ means no measurement error, and $\pi = 0.5$ implies that reported age at arrival is pure noise. It would take $\pi = 0.433$ to generate a coefficient greater than -0.01, which would be difficult to distinguish from zero in the data. I conclude that measurement error is unlikely to explain the findings.

8 Conclusion

In this paper I use the adult outcomes of refugees who arrived to the US in early childhood to quantify cross-country differences in early childhood human capital. The key finding is that outcomes for such immigrants are unrelated to the age at which they arrived to the US. Although this experiment is quite different from the usual early childhood intervention, it is useful to understand how this result can be consistent with those more commonly found in the literature. To this end, I embed a standard human capital production function into a cross-country model of skill acquisition. There are three parameterizations of that model that are consistent with the data. The most promising in light of the work here and existing findings in the literature is to assume that there is little role for goods or country environment as inputs to early childhood human capital. Given this, it is unlikely that early childhood human capital explains much of cross-country income differences. It is hard to reconcile large cross-country differences in early childhood human capital with the fact that refugees who immigrated from Vietnam, Cambodia, and Laos to the US had similar outcomes whether they arrived at age 0 or age 5.

The paper is less conclusive about human capital gaps at later ages. Outcomes do decline in age at arrival for years spent abroad in late childhood and adolescence. However, there are many theories consistent with this empirical finding. The evidence presented here does not discriminate between these alternatives. Further work on the mechanisms for human capital development is needed.

References

Acemoglu, Daron and Simon Johnson, "Disease and Development: The Effect of Life Expectancy on Economic Growth," *Journal of Political Economy*, 2007, *115* (6), 925–985.

Alderman, Harold, John Hoddinott, and Bill Kinsey, "Long term consequences of early childhood malnutrition," *Oxford Economic Papers*, 2006, *58* (3), 450–474.

Almond, Douglas and Janet Currie, "Human Capital Development before Age Five," in David Card and Orley Ashenfelter, eds., *Handbook of Labor Economics*, Vol. 4b, Elsevier, 2011, chapter 15, pp. 1315–1486.

Attanasio, Orazio, Emla Fitzsimons, Sally Grantham-McGregor, Costas Meghir, and Marta Rubio-Codina, "Early Childhood Stimulation, Nutrition and Development: A Randomised Control Trial," 2012. mimeo, Yale University.

Barro, Robert J and Jong-Wha Lee, "A new data set of educational attainment in the world, 1950–2010," *Journal of Development Economics*, 2012.

Barry, Michele, Joseph Craft, David Coleman, Holy O. Coulter, and Ralph Horwitz, "Clinical Findings in Southeast Asian Refugees," *Journal of the American Medical Association*, 1983, *249* (23), 3200–3203.

Beckett, Celia, Jennifer Castle, Michael Rutter, and Edmund J. Sonuga-Barke, "Institutional Deprivation, Specific Cognitive Functions, and Scholastic Achievement: English and Romanian Adoptee (ERA) Study Findings," *Monographs of the Society for Research in Child Development*, 2010, *75* (1), 125–142.

Bleakley, Hoyt and Aimee Chin, "Language Skills and Earnings: Evidence from Childhood Immigrants*," *Review of Economics and Statistics*, 2004, *86* (2), 481–496.

_ **and** _ , "Age at Arrival, English Proficiency, and Social Assimilation Among U.S. Immigrants," *American Economic Journal: Applied Economics*, 2010, *2* (1), 165–192.

Borjas, George J., "The Economic Analysis of Immigration," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3A, Elsevier Science, North-Holland Publishers, 1999, pp. 1697–1760.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne, "The Determinants of Earnings: A Behavioral Approach," *Journal of Economic Literature*, 2001, *39* (4), 1137–1176.

Caplan, Nathan S, Marcella H Choy, and John K Whitmore, *Children of the boat people: A study of educational success*, University of Michigan Press, 1991.

Caselli, Francesco, "Accounting for Cross-Country Income Differences," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 1A, Elsevier, 2005, chapter 9, pp. 679–741.

Cordoba, Juan Carlos and Marla Ripoll, "What Explains Schooling Differences Across Countries?," *Journal of Monetary Economics*, March 2013, *60* (2), 184–202.

Cubas, German, B Ravikumar, and Gustavo Ventura, "Talent, labor quality, and economic development," 2013. Federal Reserve Bank of St. Louis Working Paper Series.

Cunha, Flavio and James Heckman, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

_ , James J. Heckman, and Susanne M. Schennach, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78* (3), 883–931.

_ , _ , Lance Lochner, and Dimitriy V. Masterov, "Interpreting the Evidence on Life Cycle Skill Formation," in Eric Hanushek and Finis Welch, eds., *Handbook of the Economics of Education*, Vol. 2, Amsterdam: North Holland, 2006, chapter 12, pp. 697–812.

Davis, Leonard, *Hong Kong and the Asylum-Seekers from Vietnam*, London: Macmillan Academic and Professional Ltd, 1991.

Del Boca, Daniela, Christopher Flinn, and Matthew Wiswall, "Household Choices and Child Development," *Review of Economic Studies*, 2014, *81* (1), 137–185.

Dewey, Kathryn G., Johanne Daniels, Kiok S. Teo, Elissa Hassel, and June Otow, "Height and Weight of Southeast Asian Preschool Children in Northern California," *American Journal of Public Health*, 1986, *76* (7), 806–808.

Erosa, Andrès, Tatyana Koreshkova, and Diego Restuccia, "How Important is Human Capital? A Quantitative Theory Assessment of World Income Inequality," *Review of Economic Studies*, October 2010, *77* (4), 1421–1449.

Friedberg, Rachel M., "The Labor Market Assimilation of Immigrants in the United States: The Role of Age at Arrival," December 1992. mimeo, Brown University.

Goldenring, John M., Jeffrey Davis, and Margaret McChesney, "Pediatric Screening of Southeast Asian Immigrants," *Clinical Pediatrics*, 1982, *21* (10), 613–616.

Gonzalez, Arturo, "The Education and Wages of Immigrant Children: The Impact of Age at Arrival," *Economics of Education Review*, 2003, *22*, 203–212.

Haines, David W., "Introduction," in David W. Haines, ed., *Refugees as Immigrants: Cambodians, Laotians, and Vietnamese in America*, Rowman & Littlefield Publishers, Inc., 1989.

Hall, Robert E. and Charlies I. Jones, "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, February 1999, *114* (1), 83–116.

Hawthorne, Lesleyanne, ed., *Refugee: the Vietnamese Experience*, Melbourne: Oxford University Press, 1982.

Hein, Jeremy, *From Vietnam, Laos, and Cambodia: A Refugee Experience in the United States*, New York: Twayne Publishers, 1995.

Hendricks, Lutz, "How Important Is Human Capital for Development? Evidence from Immigrant Earnings," *American Economic Review*, March 2002, *92* (1), 198–219.

Heston, Alan, Robert Summers, and Bettina Aten, *Penn World Table Version 7.1*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, November 2012.

Hsieh, Chang-Tai and Peter J Klenow, "Development accounting," *American Economic Journal: Macroeconomics*, 2010, *2* (1), 207–223.

Hung, Nguyen Manh and David W. Haines, "Vietnamese," in David W. Haines, ed., *Refugees in America in the 1990s: A Reference Handbook*, Greenwood Press, 1996, pp. 305–327.

Kaarsen, Nicolai, "Cross-Country Differences in the Quality of Human Capital," *Journal of Development Economics*, 2014, *107* (2), 214–224.

Klenow, Peter J. and Andres Rodriguez-Clare, "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?," in Ben S. Bernanke and Julio J. Rotemberg, eds., *NBER Macroeconomics Annual 1997*, The MIT Press, 1997, pp. 83–103.

Lagakos, David, Ben Moll, Tommaso Porzio, and Nancy Qian, "Experience Matters: Human Capital and Development Accounting," 2014. mimeo, UCSD.

_ , Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman, "Lifecyle Human Capital Accumulation Across Countries: Lessons from U.S. Immigrants," 2014. mimeo, Arizona State University.

Lee, Sharon M. and Barry Edmonston, "Age-at-Arrival's Effects on Asian Immigrants' Socioeconomic Outcomes in Canada and the U.S.," *International Migration Review*, 2011, *45* (3), 527–561.

Lenneberg, Eric H., *Biological Foundations of Language*, New York: Wiley & Sons, 1967.

Lino, Mark, "Expenditures on Children By Families, 2010," 2010. U.S. Department of Agriculture, Center for Nutrition Policy and Promotion. Miscellaneous Publication No. 1528-2010.

Lubotsky, Darren, "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings," *Journal of Political Economy*, 2007, *115* (5), 820–867.

Lucas Jr, Robert E and Benjamin Moll, "Knowledge growth and the allocation of time," *Journal of Political Economy*, forthcoming.

Maluccio, John A, John Hoddinott, Jere R Behrman, Reynaldo Martorell, Agnes R Quisumbing, and Aryeh D Stein, "The Impact of Improving Nutrition During Early Childhood on Education among Guatemalan Adults," *The Economic Journal*, 2009, *119* (537), 734–763.

Manuelli, Rodolfo E. and Ananth Seshadri, "Human Capital and the Wealth of Nations," *American Economic Review*, 2014. forthcoming.

Marshall, Grant N., Terry L. Schell, Marc N. Elliott, S. Megan Berthold, and Chi-Ah Chun, "Mental Health of Cambodian Refugees 2 Decades After Resettlement in the United States," *Journal of the American Medical Association*, 2005, *294* (5), 571–579.

Minnesota Population Center, "Integrated Public Use Microdata Series — International: Version 6. [Machine-readable database].," Minneapolis: University of Minnesota. 2010. http://www.ipums.org.

Moraga, Jesus Fernandez-Huertas, "New evidence on emigrant selection," *The Review of Economics and Statistics*, 2011, *93* (1), 72–96.

Myers, Dowell, Xin Gao, and Amon Emeka, "The Gradient of Immigrant Age-at-Arrival Effects on Socioeconomic Outcomes in the U.S.," *International Migration Review*, 2009, *43* (1), 205–229.

Office of Refugee Resettlement, "Report to the Congress: Refugee Resettlement Program," 1980–1995. Washington DC.

Peck, Richard E., Margaret Chuang, Gordon E. Robbins, and Milton Z. Nichaman, "Nutritional Status of Southeast Asian Refugee Children," *American Journal of Public Health*, 1981, *71* (10), 1144–1148.

Reder, Stephen and Mary Cohn, "A Study of the Extent and Effect of English Language Training for Refugees. Phase II: Classroom Observation and Community Survey.," Technical Report, Northwest Regional Educational Laboratory, Portland, Oregon 1984.

Robinson, W. Courtland, *Terms of Refuge: The Indochinese Exodus and the International Response*, New York, NY: Zed Books Limited, 1998.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek, "Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]," Minneapolis: University of Minnesota 2010.

Rumbaut, Rubén G., "Portraits, Patterns, and Predictors of the Refugee Adaptation Process: Results and Reflections from the IHARP Panel Study," in David W. Haines, ed., *Refugees as Immigrants: Cambodians, Laotians, and Vietnamese in America*, Rowman & Littlefield Publishers, Inc., 1989.

_ **and John R. Weeks**, "Fertility and Adaptation: Indochinese Refugees in the United States," *International Migration Review*, 1986, *20* (2), 428–466.

Rutter, Michael and The English and Romanian Adoptees (ERA) study team, "Developmental Catch-up, and Deficit, Following Adoption after Severe Global Early Privation," *Journal of Child Psychology and Psychiatry*, 1998, *39* (4), 465–476.

_ **, Robert Kumsta, Wolff Schlotz, and Edmund Sonuga-Barke**, "Longitudinal studies using a natural experiment design: The case of adoptees from Romanian institutions," *Journal of the American Academy of Child & Adolescent Psychiatry*, 2012, *51* (8), 762–770.

Schoellman, Todd, "Education Quality and Development Accounting," *Review of Economic Studies*, 2012, *79* (1), 388–417.

_ , "The Adolescent Investments and Outcomes of Indochinese Refugees," 2014. Working Paper, Arizona State University.

Scott, Joanna C., *Indochina's Refugees: Oral Histories from Laos, Cambodia and Vietnam*, Jefferson, North Carolina: McFarland & Company, Inc., 1989.

Shastry, Gauri Kartini and David N Weil, "How Much of Cross-Country Income Variation is Explained by Health?," *Journal of the European Economic Association*, 2003, *1* (2-3), 387–396.

Social Security Administration, *Annual Statistical Supplement to the Social Security Bulletin* 2012.

Strand, Paul J. and Woodrow Jones Jr., *Indochinese Refugees in America: Problems of Adaptation and Assimilation*, Durham, North Carolina: Duke University Press, 1985.

Thomas, Joe, *Ethnocide: A Cultural Narrative of Refugee Detention in Hong Kong*, Aldershot: Ashgate, 2000.

United Nations High Commissioner for Refugees, "The State of the World's Refugees: Fifty Years of Humanitarian Action," Oxford University Press 2001. New York.

United States General Accounting Office, "Indochinese Refugees: Protection, Care, and Processing Can Be Improved," 1980. Washington DC.

_ , "Improved Overseas Medical Examinations and Treatment Can Reduce Serious Diseases in Indochinese Refugees Entering the United States," 1982. Washington DC.

U.S. House of Representatives, *Green Book* 1996.

U.S. Immigration and Naturalization Service, *Statistical Yearbook of the Immigration and Naturalization Service*, Washington, D.C.: U.S. Government Printing Office, 1980–2000.

Walker, Susan P, Susan M Chang, Christine A Powell, and Sally M Grantham-McGregor, "Effects of Early Childhood Psychosocial Stimulation and Nutritional Supplementation on Cognition and Education in Growth-Stunted Jamaican Children: A Prospective Cohort Study," *The Lancet*, 2005, *366* (9499), 1804–1807.

Weil, David N, "Accounting for the effect of health on economic growth," *The Quarterly Journal of Economics*, 2007, *122* (3), 1265–1306.

A Online Appendix: Selection of Refugees

Refugees are typically considered to be less selected than other forms of migrants, for two reasons. First, in refugee situations there is typically a large "push" factor that leads refugees to leave, mitigating the effects of self-selection. In the Indochinese case, the main elements were political persecution, ethnic discrimination, and in some cases ongoing conflict (such as between the Vietnamese government and the Khmer Rouge). Second, once refugees are formally labeled as such, they face a different set of immigration standards from potential host countries. In essence, they are accepted on humanitarian grounds even if they lack the usual marketable skills that countries desire in their immigrants.

Anecdotal evidence generally tends to point towards modest selection for Indochinese refugees after 1975. For example, Robinson (1998) gives a common characterization of the post-1975 refugees: "Beyond that, this second wave of refugees bore scant resemblance to the relatively homogeneous, well-educated Vietnamese of the first wave. These were peasant Khmer fresh from the 'killing fields' of Cambodia; they were pre-literate Hmong from the highlands of Laos; they were ethnic Chinese and Vietnamese traumatized by perilous boat journeys, push-backs, and pirate attacks." On the other hand, two margins of selection are well-known for Indochinese refugees. First, most refugees from Vietnam fled by boat. Doing so required paying a fare and, in some cases, a bribe to Vietnamese officials (Robinson, 1998). For this reason, boat refugees were probably selected based on family income. Second, the Indochinese refugee flows were protracted. While the initial refugees fled persecution, there was widespread agreement by the mid-1980s that many migrants were making a conscious decision to seek resettlement for the sake of improved economic opportunity. This process again likely implies a degree of self-selection.

In this appendix I review evidence on the extent of selection from two different sources. First, I use a number of studies that were conducted in the 1980s by or on behalf of the Office of Refugee Resettlement. Their primary goal was to quantify the challenges faced by refugees and their rate of assimilation, particularly with respect to finding work and leaving public assistance. The Office of Refugee Resettlement conducted an annual, nationally representative survey of newly-arrived refugees. They asked a few basic questions about the refugees' backgrounds before arriving, without distinguishing between subcategories of Indochinese refugees. As a whole, refugees averaged between 4 and 6 years of schooling over this time period, with roughly half of the new arrivals speaking no English and only a few percent reporting speaking English well or fluently. These studies also asked about the refugees' occupational backgrounds in their home country; roughly one-third of refugees

from the years of interest report having been farmers or fishermen, with about one-quarter reporting sales or clerical jobs and the rest distributed among managerial, technical, and blue collar jobs (Office of Refugee Resettlement, 1980–1995).

The Office of Refugee Resettlement also sponsored several studies that were more detailed in the questions they asked, but focused on a narrow geographic region at the expense of national representativeness. They are particularly likely to over-represent areas where refugees clustered since these were areas where it was cost-effective to sample refugees. Table 3 gives some of the basic descriptive statistics of refugees taken a few years after their arrival. I focus on refugees' education, their ability to write their own language and English, their occupational background, and the characteristics of how they arrived to the US. There are two key messages from this table. First, there are dramatic differences in the pre-arrival background of refugees of different ethnic groups. Vietnamese and Chinese refugees in particular were well-educated, wrote their own language well and sometimes wrote English, and had professional occupational backgrounds. The Hmong and Khmer, on the other hand, had very little education, were unlikely to be literate even in their own language, and were almost all farmers, fishermen, or soldiers in their home country. The Lao performed somewhere in the middle. Second, the Hmong in the US in particular come from disadvantaged backgrounds that show little evidence of families with a large degree of acquired human capital. Their experiences are consistent with their background as isolated, rural farmers with no written language until the 1950s.

A.1 Comparing Immigrants to Non-Migrants

As a second source of evidence I compare the characteristics of refugee migrants before and after they immigrated to the characteristics of non-migrants from the same country. I focus on education because it is the most useful variable that is easily compared across countries. The data come from three sources. The first is the IHARP study introduced in the text in Section 2.1. That study asked a small sample of refugees about their pre-migration characteristics, including their education level as of 1975. The second is the 1990 US Census. This dataset includes the post-migration educational outcomes of a large and representative sample of refugees. The third dataset is the population censuses from Cambodia and Vietnam. Unlike Laos, these countries have conducted censuses that collect information on age and education, with the earliest census taking place in 1989 in Vietnam and 1998 in Cambodia.[17]

[17] Available online at Minnesota Population Center (2010).

Table 3: Characteristics of Refugees

	Vietnamese	Chinese	Lao	Hmong	Khmer
Schooling					
Average Years[a]	9.8	6.7	4.9	1.6	5.0
Percent with None[b]	14.0		13.8	41.0	11.8
Percent with Primary – Some High School[b]	45.8		70.1	58.0	84.5
Percent with High School Degree or More[b]	39.8		16.3	1.0	3.6
Literacy					
Native Language (Pre-Migration)[c]	98.8	81.6		26.6	65.9
Native Language (Post-Migration)[b]	93.0		91.3	54.0	77.3
English (Pre-Migration)[c]	33.8	8.8		0.9	4.2
English (Post-Migration)[b]	52.6		42.8	31.0	24.5
Occupation					
Percent Military, Farming, or Fishing[c]	36.1	20.4		90.9	59.9
Migration					
Average Months in Camps[a]	7.8	10.3	23.0	34.3	25.5
Percent Paying Bribes[c]	32.7	71.7		21.3	19.3

[a] Source: Rumbaut and Weeks (1986).
[b] Source: Strand and Jones Jr. (1985). Education responses are for household heads. Ethnic Chinese are included with Vietnamese. Literacy figures are as of the study time and presumably include some learning of English since arrival.
[c] Source: Rumbaut (1989). Refugees were asked about their literacy and occupation as of 1975. Figures for ethnic Lao not reported.

Table 4: Schooling Comparison: Refugees and Non-Migrants

	0 Years	1–11 Years	12+ Years
Panel A: Vietnam			
Non-Migrants	22%	69%	9%
Refugees, 1975	6%	63%	31%
Refugees, 1990	12%	25%	63%
Panel B: Cambodia			
Non-Migrants	48%	50%	2%
Refugees, 1975	24%	66%	10%
Refugees, 1990	32%	28%	40%

Table notes: Results represent author's calculation using data from censuses of Vietnam and Cambodia (for non-migrants); from the 1982–83 wave of the Indochinese Health and Adaptation Research Project (for refugees in 1975); and from the 1990 US population census (refugees in 1990). See text for details.

In each case the sample is restricted to adults who were born before 1957, with the idea that they would ordinarily have completed any schooling by 1975. This cutoff is imposed because the retrospective pre-migration educational outcomes are measured as of 1975 in the IHARP. The samples of immigrants are restricted to those who arrived to the US after 1975. Educational attainment is coded into three categories: no schooling; 1–11 years (e.g., less than a high school degree) and 12+ years (e.g., at least a high school degree). Table 4 gives the figures. There are two main results. First, refugees were modestly selected already as of 1975. Refugees from both countries were somewhat less likely to be entirely uneducated and more likely to have at least a high school degree. Second, refugees increased their educational attainment substantially between the end of the Vietnam War in 1975 and the 1990 Census.

I also use the 1990 US Census to provide empirical evidence on the plausibility of the hypothesis that refugees who arrived at different ages were endowed with different family backgrounds in terms of family size, parental education, family income, and so on. To do so, I test whether there is a correlation between age at arrival of the child and any of the observable family attributes.[18] Differences in outcomes could arise if there was differential selection in age at arrival (despite the historical evidence that it was unlikely); or if parents of late-arriving children make systematically different choices in how to allocate their time between investing in their children, investing in themselves, and working in the

[18] I can connect 87 percent to their biological mother and 81 percent to their biological father.

labor market. It is not necessarily clear which direction of correlation would be more worrying. A positive correlation indicates, for example, that later-arriving children grow up in families with higher income, which has a direct benefit. On the other hand, a negative correlation could indicate that parents are foregoing the labor market and investments in their own human capital in favor of investing more in their children.

To implement the test I use a host of family attributes: number of siblings; probability of having the biological mother or father in the house; birth order; family income; hourly wages and earnings of the mother and father; education of the mother and father, measured several different ways; and English language ability of the mother and father. In general, there is not much evidence of differences in family attributes by age at arrival. The correlation is not statistically significant and is as likely to be negative as positive. There are two exceptions to this rule. First, older-arriving children of all ethnic groups are statistically more likely to be eldest children and less likely to be youngest children of their family. Averaging across groups, children who arrive at age 5 are 13 percentage points more likely to be the oldest child in their family than children who arrive at age 0. Second, there is a statistically significant (95% level) and positive correlation between parental characteristics and age at arrival for ethnic Vietnamese for seven out of the fifteen outcomes explored. For the remaining four groups in total there are only three such significant correlations, an outcome that could easily be generated by chance under multiple hypothesis testing. Hence, for four of the five ethnic groups there is not much room for a story of differential selection or differences in investment patterns by age at arrival of the children. For the fifth group the evidence suggests that there may be some differences in either pre-migration characteristics or post-migration investments.

B Online Appendix: Identification of Age at Arrival Effects

In the empirical section of the paper I estimate the effects of age at arrival by regressing outcomes such as years of schooling or wages on full sets of dummies for age, census year, and age at arrival using a pooled sample of natives and immigrants. Identification of the effect on age at arrival requires some assumptions, which are explicitly formulated and justified here.

To simplify the discussion, I specialize to the case where all time variables enter the regres-

sion equations in linear fashion; the same insights apply to the dummy variable specifications used in the paper. With linear time effects, the estimation model for the determination of some outcome of interest y is

$$y = \beta X + \alpha A + \omega Y + \phi AA + \varepsilon,$$

where the right hand side includes a vector of controls X, the age A, the year of the Census Y, and (for immigrants) the age at arrival, AA. Greek letters denote the corresponding coefficients.

Research in the literature often proposes a more general model. Adapting from Friedberg (1992) and Borjas (1999), native outcomes y^N and immigrant outcomes y^I are given by:

$$\begin{aligned} y^N &= \beta^N X^N + \alpha^N A^N + \omega^N Y^N + \varepsilon^N, \\ y^I &= \beta^I X^I + \alpha^I A^I + \omega^I Y^I + \phi^I AA^I + \gamma^I C^I + \delta^I YUS^I + \varepsilon^I. \end{aligned}$$

This specification is more general in two ways. First, it allows the effect of the controls, age, and year to be different for immigrants and natives. Second, immigrant outcomes are affected by year-of-immigration cohort effects C^I, which are intended to capture changes in the composition of immigrants by year of entry, and the assimilation term YUS^I which measures the number of years an immigrant has spent in the US

It is well-known that some of the coefficients in this general model are not identified without further assumptions. The problem arises from two linear dependencies in the immigrant equation, namely $YUS^I + C^I = Y^I$ and $AA^I + YUS^I = A^I$. The latter dependency is the problem for the analysis here, since it means that the coefficient on age at arrival is not identified without further assumptions. Friedberg (1992) proposes imposing the restriction $\alpha^I = \alpha^N$ to resolve this dependency. In words, the assumption is that immigrants and natives share the same age effects, which can be identified using the natives. The effect of age at arrival on immigrant outcomes is thus identified as the differential effect of a year spent abroad for immigrants as opposed to a year spent in the US for natives. This idea of age at arrival effects exactly captures the spirit of the estimation exercise. To implement this strategy I pool natives and immigrants and impose the further restriction $\beta^N = \beta^I$. In this case, a general model for the outcome y is

$$y = \beta X + \alpha A + \omega^N Y^N + \omega^I Y^I + \phi^I AA^I + \gamma^I C^I + \delta^I YUS^I + \varepsilon.$$

There is still a linear dependency in this model, but it is irrelevant for the coefficient of interest, ϕ^I; this can be seen by plugging in for the year effects for immigrants:

$$y = \beta X + \alpha A + \omega^N Y^N + \phi^I A A^I + (\gamma^I + \omega^I) C^I + (\delta^I + \omega^I) YUS^I + \varepsilon. \tag{7}$$

The effect of age at arrival is identified, although cohort effects and assimilation effects are not. This estimation model is more general than the one used in the text, because it also includes cohort effects as a regressor (even though the estimated coefficient does not measure "true" cohort effects). Nonetheless, implementing this equation produces essentially the same results, which are available upon request.

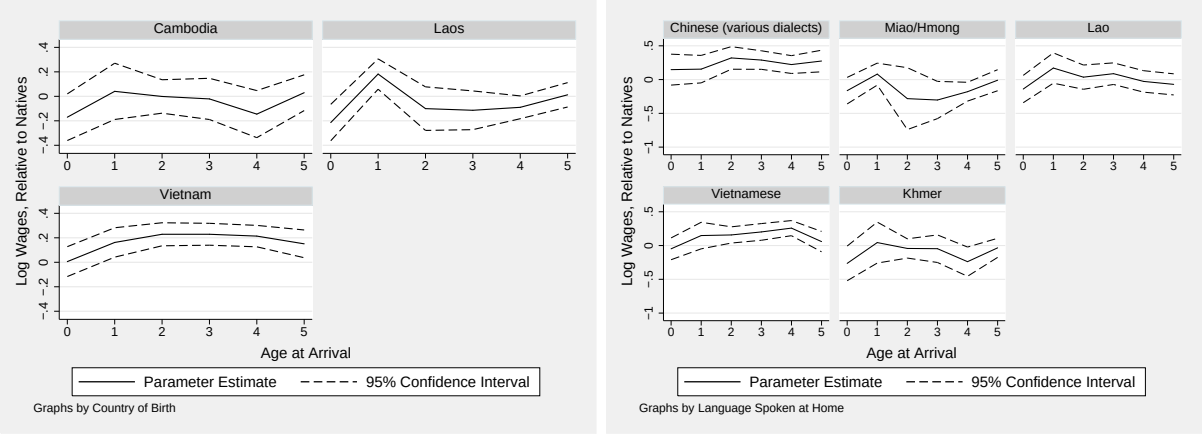
C Online Appendix: Robustness

C.1 Alternative Decompositions

In this subsection I explore alternative decompositions of the Indochinese refugees into subgroups. The main idea is that self-reported ethnicity may not appropriately capture the different groups of refugees. Also, some Indochinese refugees report ethnicities that do not fall neatly into the five major categories, and so are excluded from earlier figures. As a check on the baseline results, I also decompose refugees by their country of birth, which captures all Indochinese refugees; and by their reported language spoken at home, in case language rather than ethnicity is a better way of grouping immigrants. Figures 7a and 7b show that the patterns for wages are similar to those for the decomposition based on ethnicity, with no trend in outcomes by age at arrival.

As mentioned in the text, US immigration policy towards Indochinese refugees shifted in 1982. A report by the General Accounting Office documented that as of 1981 the required health screenings were cursory (lasting roughly 20 seconds per person); that children under 15 were not routinely screened; and that the results of examinations did not play a part in admissions decisions. The report led to much stricter screening after it was issued. Hence, one might suspect that pre-1982 refugees are less selected on health status, and post-1982 refugees more so. Figure 8 shows that the lack of a trend is consistent for the less selected, pre-1982 refugees, although the more selected, post-1982 refugees do display a more mixed pattern.

Finally, I consider two limits on the sample of interest. First, I exclude from the sample



(a) Birth Country (b) Language Spoken at Home

Figure 7: Log-Wages by Age at Arrival for Alternative Decompositions

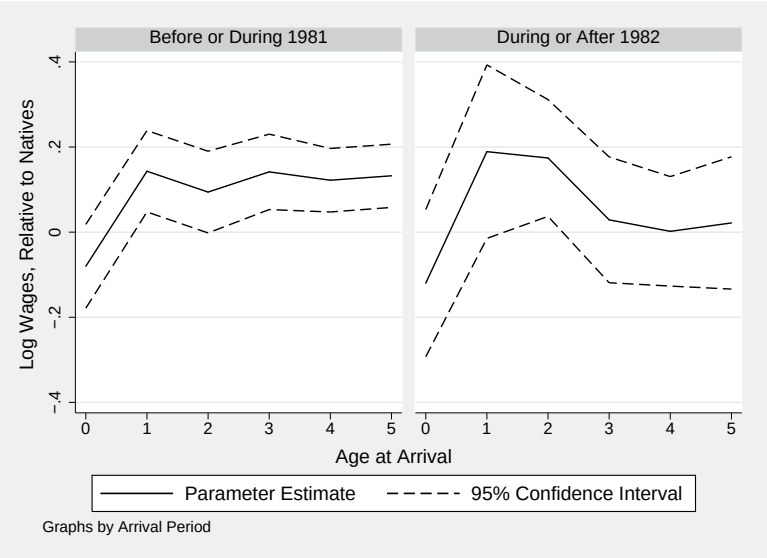


Figure 8: Log-Wage by Age at Arrival for Early and Late Arrivals

refugees who live in "ethnic enclaves", areas with high concentrations of other residents of the same ethnicity. I define a person as living in an ethnic enclave if they live in a Public Use Microdata Area (PUMA) where more than 5 percent of the population shares their ethnicity or if they live in a metropolitan statistical area (MSA) where more than 2.5 percent of the population shares their ethnicity. The PUMA is the smallest geographic region publicly available in the Census and includes between 100,000 and 200,000 people, corresponding typically to a portion of a city; MSAs are cities and the surrounding areas.

This definition of ethnic enclaves excludes roughly 30 percent of refugees from the sample. Figure 9 shows that the wage patterns are similar for those who live outside of enclaves. These findings suggest that the results are not driven by the ability of refugees to live and work in areas with others who share a similar cultural background or language.

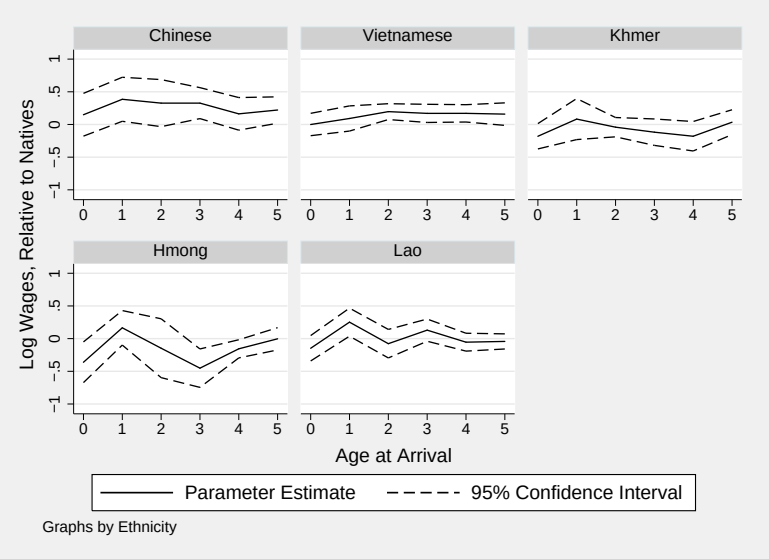


Figure 9: Log-Wage by Age at Arrival for Refugees Living Outside Enclaves

As a second sample restriction, I re-estimate my key regressions using only workers who are 23–26 years of age. The model abstracts from post-graduation human capital accumulation. Although most refugees in the sample are young some are older, and hence may have invested significantly in their human capital since graduation. Figure 10 shows that similar results obtain for very young workers who are unlikely to have made significant post-graduation investments.

C.2 Alternative Outcomes

Finally, I consider alternative socioeconomic outcomes. There is no relationship between age at arrival and probability of employment, log earnings (instead of log wages), or probability of having graduated college.[19]

The analysis can also be extended to look at the outcomes of children born in the US to

[19]For this and subsequent binary outcomes, estimation is performed via a probit model. The reported coefficient is the model-predicted change in average enrollment for each age-at-arrival group if they had instead been native-born children of non-refugee parents. Standard errors are simulated via Monte Carlo.

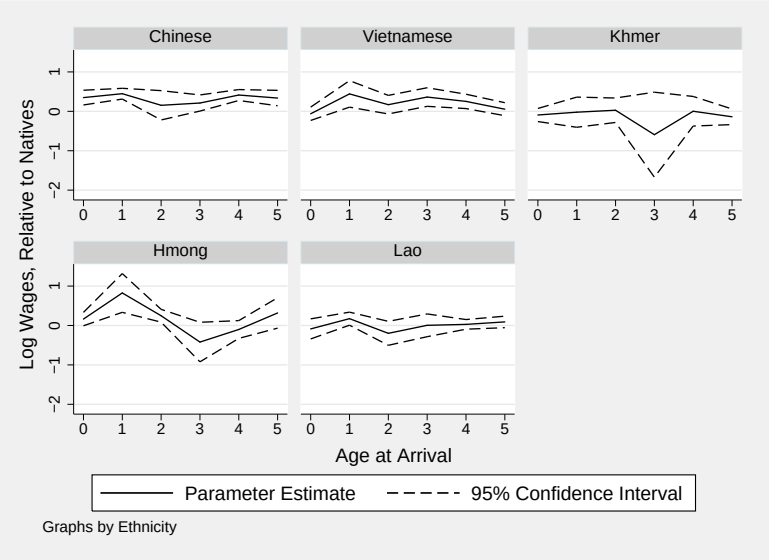
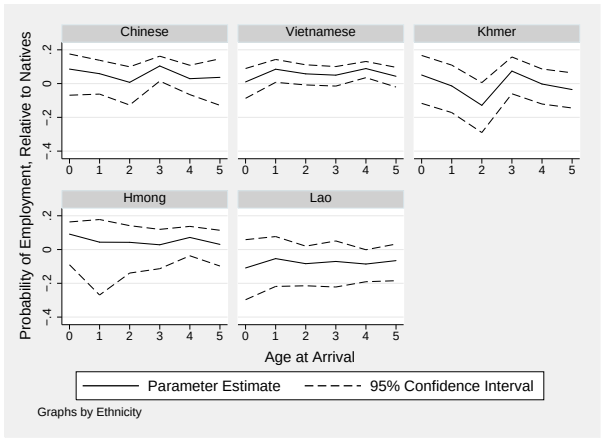


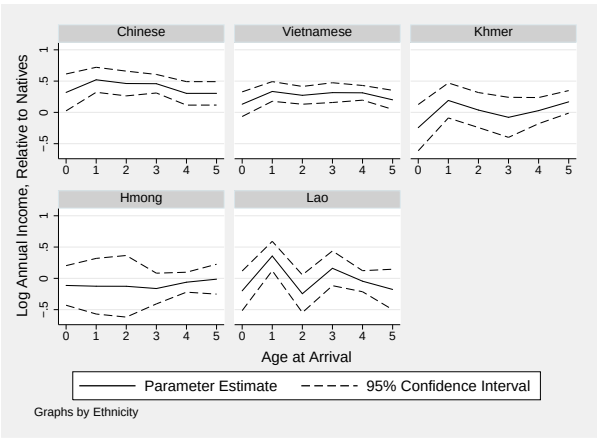
Figure 10: Log-Wage by Age at Arrival for Refugees 23–26 Years Old

refugee parents. This test is useful if there are important effects of exposure to adverse conditions while in utero. The test is somewhat more difficult to conduct because children can only be linked to their parents while they still live in the same household. Thus, it is necessary to find an outcome more relevant to the experience of those living at home than completed schooling or wages. One such outcome is the probability of being enrolled in school for 16–18 year olds. I include natives, as a control group; child refugees; and children born in the US to Indochinese refugee parents. I identify children as having refugee parents if both parents immigrated from Vietnam, Cambodia, or Laos during the refugee period as defined above and, additionally, the parents were born in the same country and immigrated in the same year. I then regress a school attendance dummy on the same set of controls as in equation (1).

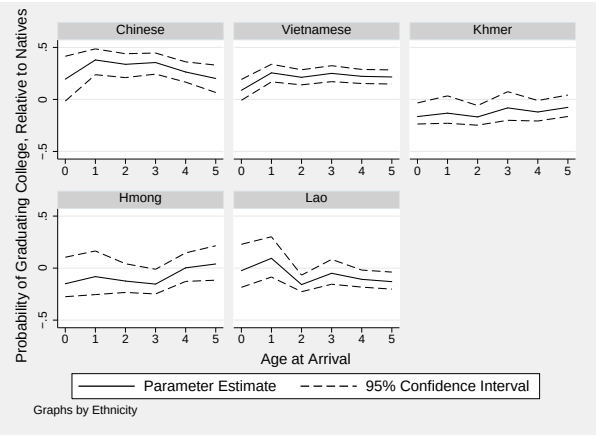
Figure 12 shows the results. Since the sample of 16–18 year olds is smaller, all Indochinese refugees are pooled together. Panel (b) shows the results for first-generation immigrants as a function of their age at arrival, similar to previous graphs. Panel (a) shows the results for second-generation immigrants as a function of their parents' arrival year minus their birth year. Children with values less than -1 in this panel have no exposure to the Indochinese countries or refugee camps, even while in utero. Their outcomes are similar to those of first-generation child refugees, suggesting that in utero exposure is not an important part of the story.



(a) Probability of Employment



(b) Log-Income



(c) Probability of Graduating College

Figure 11: Alternative Outcomes by Age at Arrival

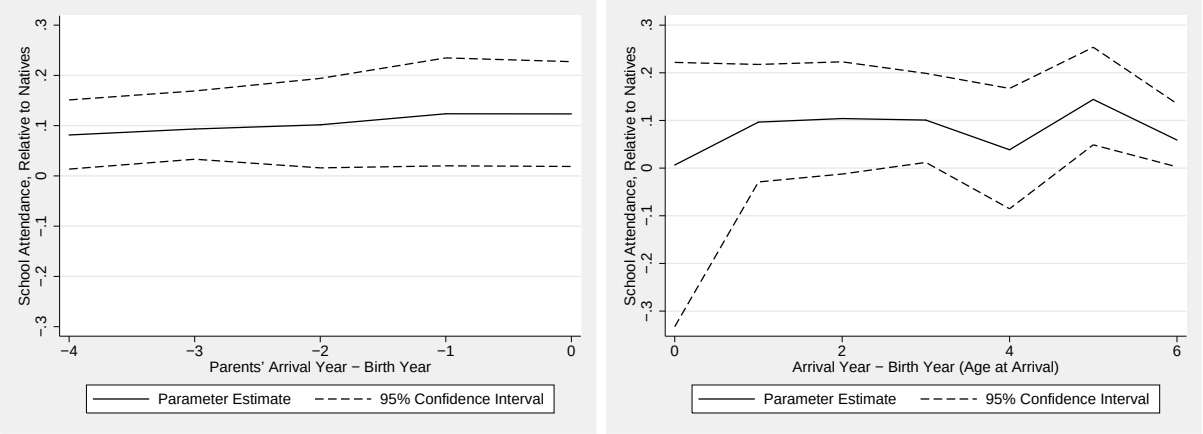
D Online Appendix: Derivations for the Isoelastic Model

From the text, the family's problem is:

$$\max_{g_1,g_2,p_1,p_2,S} \quad \int_{5+S}^{\infty} e^{-rt} A_i(t) h_2 dt - g_1 - e^{-5r} g_2 - A_i h_p p_1 - A_i e^{5(\gamma - r)} h_p p_2$$

$$\text{s.t.} \quad h_2 = \left[\lambda_2 h_1^{\frac{\sigma_2 - 1}{\sigma_2}} + (1 - \lambda_2) \left(z_i^{\omega_1} g_2^{\omega_2} h_p^{\omega_3} p_2^{\omega_4} S^{\omega_5} \right)^{\frac{\sigma_2 - 1}{\sigma_2}} \right]^{\frac{\sigma_2}{\sigma_2 - 1}}$$

$$h_1 = \left[\lambda_1 h_0^{\frac{\sigma_1 - 1}{\sigma_1}} + (1 - \lambda_1) \left(z_i^{\eta_1} g_1^{\eta_2} h_p^{\eta_3} p_1^{\eta_4} \right)^{\frac{\sigma_1 - 1}{\sigma_1}} \right]^{\frac{\sigma_1}{\sigma_1 - 1}}.$$



(a) Second Generation Immigrants (b) First Generation Immigrants

Figure 12: Probability of Attending School at Ages 16–18

Integrating out and substituting in the isoelastic case yields:

$$\max_{g_1,g_2,p_1,p_2,S} \quad \frac{A_i e^{(\gamma-r)(S+5)}}{r-\gamma} h_2 - g_1 - e^{-5r} g_2 - A_i h_p p_1 - A_i e^{5(\gamma-r)} h_p p_2 \tag{D1}$$

$$\text{s.t.} \quad h_2 = h_0^{\lambda_1\lambda_2} \left(z_i^{\omega_1} g_1^{\omega_2} h_p^{\omega_3} p_1^{\omega_4}\right)^{\lambda_2(1-\lambda_1)} \left(z_i^{\eta_1} g_2^{\eta_2} h_p^{\eta_3} p_2^{\eta_4} S^{\eta_5}\right)^{1-\lambda_2} \tag{D2}$$

with $h_1 = h_0^{\lambda_1} \left(z_i^{\omega_1} g_1^{\omega_2} h_p^{\omega_3} p_1^{\omega_4}\right)^{1-\lambda_1}$ already substituted out.

The first-order conditions for the problem are:

$$S: \quad A_i e^{(\gamma-r)(S+5)} h_2 = \frac{A_i e^{(\gamma-r)(S+5)}}{r-\gamma} \eta_5 (1-\lambda_2) \frac{h_2}{S} \tag{D3}$$

$$g_1: \quad \frac{A_i e^{(\gamma-r)(S+5)}}{r-\gamma} \omega_2 \lambda_2 (1-\lambda_1) \frac{h_2}{g_1} = 1 \tag{D4}$$

$$g_2: \quad \frac{A_i e^{(\gamma-r)(S+5)}}{r-\gamma} \eta_2 (1-\lambda_2) \frac{h_2}{g_2} = e^{-5r} \tag{D5}$$

$$p_1: \quad \frac{A_i e^{(\gamma-r)(S+5)}}{r-\gamma} \omega_4 \lambda_2 (1-\lambda_1) \frac{h_2}{p_1} = A_i h_p \tag{D6}$$

$$p_2: \quad \frac{A_i e^{(\gamma-r)(S+5)}}{r-\gamma} \eta_4 (1-\lambda_2) \frac{h_2}{p_2} = A_i e^{5(\gamma-r)} h_p. \tag{D7}$$

Inspection of (D3) reveals that it pins down $S = \frac{\eta_5(1-\lambda_2)}{r-\gamma}$, which implies that S does not vary within or across countries. Equations (D4)–(D7) link together the optimal market goods and parental investments in the two periods. Inspection shows that the model predicts

$g_1 \propto g_2 \propto A_i h_p p_1 \propto A_i h_p p_2$, where the proportionality factors are functions of the share parameters (ω, η, and λ) as well as discount and growth rates (e^{-5r} and $e^{5(\gamma-r)}$), and so do not vary within or across countries.

Using proportionality, it is possible to rewrite (D6) as:

$$\kappa_1 h_0^{\lambda_1\lambda_2} z_i^{\omega_1\lambda_2(1-\lambda_1)+\eta_1(1-\lambda_2)} \left(A_i h_p p_1\right)^{\omega_2\lambda_2(1-\lambda_1)+\eta_2(1-\lambda_2)} h_p^{\omega_3\lambda_2(1-\lambda_1)+\eta_3(1-\lambda_2)} p_1^{\omega_4\lambda_2(1-\lambda_1)+\eta_4(1-\lambda_2)} = h_p p_1$$

where κ_1 captures functions of parameters and discount and growth rates that do not vary within or across countries. Solve for the time parents spend with their children p_1 in terms of exogenous parameters to find:

$$p_1 = \kappa_2 h_0^{\frac{\lambda_1\lambda_2}{1-\Psi_2-\Psi_4}} z_i^{\frac{\Psi_1}{1-\Psi_2-\Psi_4}} A_i^{\frac{\Psi_2}{1-\Psi_2-\Psi_4}} h_p^{\frac{\Psi_2+\Psi_3-1}{1-\Psi_2-\Psi_4}} \tag{D8}$$

where κ_2 is again a constant and $\Psi_i \equiv \omega_i\lambda_2(1-\lambda_1) + \eta_i(1-\lambda_2)$. Finally, use the proportionality relationship again as well as (D8) to substitute in for h_2 to find:

$$h_2 = \kappa_3 h_0^{\frac{\lambda_1\lambda_2}{1-\Psi_2-\Psi_4}} z_i^{\frac{\Psi_1}{1-\Psi_2-\Psi_4}} A_i^{\frac{\Psi_2}{1-\Psi_2-\Psi_4}} h_p^{\frac{\Psi_3-\Psi_4}{1-\Psi_2-\Psi_4}} \tag{D9}$$

Likewise, taking equation (D8) and the proportionality relationship and plugging in for h_1 yields:

$$h_1 = \kappa_4 h_0^{\lambda_1+\frac{\lambda_1\lambda_2(\Phi_2+\Phi_4)}{1-\Psi_2-\Psi_4}} z_i^{\Phi_1+\frac{\Psi_1(\Phi_2+\Phi_4)}{1-\Psi_2-\Psi_4}} A_i^{\Phi_2+\frac{\Psi_2(\Phi_2+\Phi_4)}{1-\Psi_2-\Psi_4}} h_p^{\Phi_2+\Phi_3+\frac{(\Psi_2+\Psi_3-1)(\Phi_2+\Phi_4)}{1-\Psi_2-\Psi_4}} \tag{D10}$$

with $\Phi_i \equiv \omega_i(1-\lambda_1)$. The elasticity properties in Table 2 follow directly.

Last, I characterize the problem of the refugee who moves after early childhood. They take their level of h_1 as given and choose subsequent investments g_2, p_2, and S. Their problem then is:

$$\max_{g_2,p_2,S} \quad \frac{A_i e^{(\gamma-r)S+5\gamma}}{r-\gamma} h_2 - g_2 - A_i e^{5\gamma} h_p p_2 \tag{D11}$$

$$\text{s.t.} \quad h_2 = h_1^{\lambda_2}\left(z_i^{\eta_1} g_2^{\eta_2} h_p^{\eta_3} p_2^{\eta_4} S^{\eta_5}\right)^{1-\lambda_2} \tag{D12}$$

The first-order conditions for the problem are:

$$S: \qquad A_i e^{(\gamma-r)S+5\gamma} h_2 = \frac{A_i e^{(\gamma-r)S+5g}}{r-\gamma} \eta_5 (1-\lambda_2) \frac{h_2}{S} \tag{D13}$$

$$g_2: \qquad \frac{A_i e^{(\gamma-r)S+5\gamma}}{r-\gamma} \eta_2 (1-\lambda_2) \frac{h_2}{g_2} = 1 \tag{D14}$$

$$p_2: \qquad \frac{A_i e^{(\gamma-r)S+5\gamma}}{r-\gamma} \eta_4 (1-\lambda_2) \frac{h_2}{p_2} = A_i e^{5\gamma} h_p. \tag{D15}$$

It is still the case that $S = \frac{\eta_5(1-\lambda_2)}{r-\gamma}$. Likewise, it is still the case that there is a proportionality relationship between the remaining two inputs, with $g_2 \propto A_i h_p p_2$. Plugging this information into (D15) yields:

$$p_2 = \kappa_5 h_1^{\frac{\lambda_2}{1-(\eta_2+\eta_4)(1-\lambda_2)}} z_i^{\frac{\eta_1(1-\lambda_2)}{1-(\eta_2+\eta_4)(1-\lambda_2)}} A_i^{\frac{\eta_2(1-\lambda_2)}{1-(\eta_2+\eta_4)(1-\lambda_2)}} h_p^{\frac{(\eta_2+\eta_3)(1-\lambda_2)-1}{1-(\eta_2+\eta_4)(1-\lambda_2)}}$$

where κ_5 is a function of share parameters and other constants. Substitution yields an expression for h_2:

$$h_2 = \kappa_6 h_1^{\frac{\lambda_2}{1-(\eta_2+\eta_4)(1-\lambda_2)}} z_i^{\frac{\eta_1(1-\lambda_2)}{1-(\eta_2+\eta_4)(1-\lambda_2)}} A_i^{\frac{\eta_2(1-\lambda_2)}{1-(\eta_2+\eta_4)(1-\lambda_2)}} h_p^{\frac{(\eta_3-\eta_4)(1-\lambda_2)}{1-(\eta_2+\eta_4)(1-\lambda_2)}}$$

where κ_6 is a final constant. It follows that human capital in the labor force is increasing in early childhood human capital for $\lambda_2 > 0$.