Life-Cycle Human Capital Accumulation Across Countries: Lessons From U.S. Immigrants

David Lagakos, UCSD and NBER
Benjamin Moll, Princeton and NBER
Tommaso Porzio, Yale
Nancy Qian, Yale and NBER
Todd Schoellman, ASU

October, 2015

Abstract

How much does life-cycle human capital accumulation vary across countries? This paper seeks to answer this question by studying U.S. immigrants, who come from a wide variety of countries but work in a common labor market. We document that returns to potential experience among U.S. immigrants are higher on average for workers coming from rich countries than for those coming from poor countries. To understand this fact we build a model of life-cycle human capital accumulation that features three potential theories, working respectively through cross-country differences in: selection, skill loss, and human capital accumulation. To distinguish between theories, we use new data on the characteristics of immigrants and non-migrants from a large set of countries. We conclude that the most likely theory is that immigrants from poor countries accumulate relatively less human capital in their birth countries before migrating. Our findings imply that life-cycle human capital stocks are on average much larger in rich countries than poor countries.

Email: lagakos@ucsd.edu, moll@princeton.edu, nancy.qian@yale.edu, tommaso.porzio@yale.edu, todd.schoellman@gmail.com. For helpful comments we thank Rachel Friedberg, Gordon Hanson, Pete Klenow, Bob Lucas, Christos Makridis and seminar participants at Houston, MIT, Ottawa, Konstanz, Mannheim, and Oxford. We thank Caleb Johnson for excellent research assistance. All potential errors are our own.
1. Introduction

How important is human capital in accounting for aggregate income differences across countries? A large literature on development accounting has concluded that the answer is “only somewhat.” Specifically, the seminal work of Klenow and Rodríguez-Clare (1997), Hall and Jones (1999) and Caselli (2005) find that human capital stocks vary by roughly a factor of two between the richest and poorest countries, whereas actual output per worker varies by a factor of more than twenty.

One reason the existing literature has found such a modest role for human capital is that it has focused largely on human capital arising through schooling. Several previous studies have included human capital arising over the life cycle, i.e. after finishing schooling, but have found that it did not improve the explanatory power of human capital (Klenow and Rodríguez-Clare, 1997; Bils and Klenow, 2000, 1998). The data underlying this conclusion came from the Mincer estimates of Psacharopoulos (1994), which show no systematic variation across countries in either the returns to potential experience or the average level of potential experience. As a result, researchers using these data concluded that human capital differences arising through potential experience must be negligible.\(^1\)

In contrast, a recent literature has argued that workers in rich countries accumulate much more human capital over the life cycle than their counterparts in poor countries. Manuelli and Seshadri (2015) show that this conclusion arises out of a standard Ben-Porath model of human capital accumulation, as workers in rich countries are able to devote more goods inputs (e.g. books and computers) to their time spent accumulating human capital. Empirically, Lagakos, Moll, Porzio, Qian, and Schoellman (2015) use micro-level wage data from a large set of countries to document that returns to potential experience are generally higher in rich countries than in poor countries. They note that this evidence is consistent with the hypothesis that workers in poorer countries accumulate less human capital while working. However, they also discuss alternative explanations such as search frictions, credit constraints, or other country-specific wage-setting institutions that break the link between wages and the marginal product of labor. Finally, they note some concern that data quality and measurement concepts could vary across countries in ways that would explain their empirical findings.

In this paper we turn to U.S. immigrants to help measure and understand differences in life-cycle human capital accumulation across countries. Studying U.S. immigrants offers several advantages. First, the workers are all observed in a common labor market, as opposed to a diverse set of economies with varying labor market conditions and institutions. Second, data for all workers come from a common data source, the U.S. census, thus minimizing worries about international data comparability. Finally, the data span more than three decades in time and cover U.S. natives, allowing us to isolate cohort-of-migration and time effects consistently for workers from a large set of countries. The insight

\(^1\)This conclusion has been arrived at by others as well, including Caselli (2005) and Erosa, Koreshkova, and Restuccia (2010). See the summary of Hsieh and Klenow (2010) for a clear overview of the developing accounting literature.
of using immigrants to study human capital accumulation across countries is based on the work of Hendricks (2002) and Schoellman (2012), though the current paper is the first to measure and explain stocks of human capital from experience using U.S. immigrants.

We begin by documenting a new fact about immigrant returns to experience: returns to experience are lower among immigrants from poor countries than immigrants from rich countries. We find that this is true both for returns to foreign experience, acquired before migrating, and returns to U.S. experience, acquired in the United States after migrating. We reach this conclusion in several versions of a standard Mincerian wage regression. The first version looks only at new immigrants, who have been in the United States less than one year, and considers returns only to foreign experience (which is essentially all they have.) The second version considers all U.S. immigrants and estimates the return to foreign and U.S. experience, accounting for possible interactions between the two. Both versions show that returns to foreign experience are strongly increasing in GDP per capita of the birth country. The second version shows that returns to U.S. experience are increasing in GDP per capita of the birth country, but not as sharply as for foreign experience.

To understand these facts we consider a simple model of life-cycle human capital accumulation. The model captures three basic theories of why returns to experience would be lower for immigrants from poorer countries. The first theory is differential selection, and states that immigrants from poor countries are less strongly selected on learning ability than their counterparts in rich countries. The second theory is differential skill loss, which says that immigrants tend to lose a larger fraction of their skills after migrating. The third theory is differential human capital accumulation, which says that the efficiency of human capital accumulation is lower in poor countries than richer countries. All three theories are consistent with lower measured returns to foreign experience among immigrants from poor countries, and all three make different predictions along other dimensions.

To distinguish between theories we turn to new data we construct that compares immigrants to non-migrants in a large set of countries. The data contains the average years of school completed by immigrants and non-migrants, and the fraction of both groups working at “high-skilled” occupations, both of which are taken from national census data from around the world (Ruggles, Genadek, Grover, and Sobek, 2015). The data also contain the returns to experience for immigrants and non-migrants, taken from the current study and Lagakos et al. (2015), respectively.

The data on immigrants and non-migrants are most consistent with the theory that low life-cycle human capital accumulation before migrating is the proximate cause of low returns to experience among U.S. immigrants. The reasons are as follows. First, returns to experience among non-migrants look quite similar to returns to foreign experience among immigrants for most countries. This is inconsistent with theories centered around differential skill loss or differential selection, which imply that returns to experience should differ between the groups. Second, evidence on years of schooling com-
pleted and pre-migration wages suggest that immigrants from poorer countries are more selected than immigrants from richer countries. This provides evidence that weaker selection of immigrants from poor countries is unlikely to explain our results. Finally, the fraction of educated immigrants who are working at low-skill jobs varies little between rich and poor countries. This provides evidence against the theory that immigrants from poor countries lose disproportionately more skills after migrating.

We conclude by illustrating how our results help better account for income differences across countries. We follow the development accounting literature, which measures human and physical capital across countries, and computes the implied income variance in a world where countries only differ in these capital stocks. We depart from the literature in that we use our estimated returns to experience among U.S. immigrants to construct stocks of human capital from experience in each source country (where our data allow). We conclude that experience human capital stocks are substantially larger in rich countries than poor countries, and that incorporating these stocks into development accounting substantially increases the importance of human capital.

The rest of this paper is structured as follows. In Section 2 we describe the facts that we document about returns to experience among U.S. immigrants. In Section 3 we present a model capturing the three different theories of the facts described above, and in Section 4 we draw on evidence comparing immigrants and non-migrants to help distinguish between the theories. In Section 5 we illustrate what our empirical findings imply for development accounting. In Section 6 we conclude.

2. Immigrant Returns to Experience: The Facts

2.1. Sample and Data

Our data on immigrants draw on the 1980–2000 U.S. Population Censuses as well as the 2005–2013 American Community Surveys (ACSs), downloaded via IPUMS. Each of these data sets includes a large, representative cross-section of the U.S. population in a particular year. We choose not to use data from earlier Censuses because their sample size were smaller (1 percent instead of 5 percent) and immigrants were a much smaller share of the population before 1980. The 2000 Census was the last to include a long form with detailed questionnaires sent to a subset of the population; the ACS, an annual 1 percent sample of the American population, is the successor to the Census long form. Most questions and responses were maintained in the transition, so that combining the data is straightforward.

Our basic sample selection is very similar to Lagakos et al. (2015). We focus on men age 16 or older who work full time, for wages, in the private sector. The restriction to male full-time workers is made because we measure potential rather than actual experience; for women and part-time workers the relationship between the two is less clear. We exclude the self-employed and private sector workers
because it is more difficult or requires more assumptions to measure their marginal product given their reported income. See Lagakos et al. (2015) for further discussion and robustness analysis for these choices; we also show our results when we relax them below. We also exclude workers who have missing or zero responses to the key variables, primarily work intensity, labor income, and education; such people are relatively rare in the Census.

We identify immigrants using country of birth. The Census and ACSs provide detailed responses that code the country of birth for most of the major source countries of U.S. immigrants. Our datasets also include information on the year of immigration. In the 1980 and 1990 Censuses this information was provided in ranges (e.g. 1975–1979). This category coding is unfortunate for our analysis because we want to compute years of foreign and domestic potential experience. We experiment with coding these ranges to the midpoint and using them in our analysis. We also provide results for the case where we use only data from 2000 onward, where the exact year of immigration is recorded.

We construct potential experience (henceforth: experience) using information on age and educational attainment. In the 1980 Census the raw data was years of schooling, while from 1990 onward it was recorded as educational attainment (e.g., high school graduate). We recode educational attainment into years in the standard fashion. We then define experience as age – schooling – 6. A small subset of our sample reports very low levels of schooling. Following Lagakos et al. (2015), we define experience as age – 18 for anyone with less than twelve years of schooling, under the assumption that no one acquires significant useful experience before age 18. Given this variable, we focus our attention on the subsample with between 0 and 40 years of experience, inclusive. For immigrants we split their experience into foreign (birth country) and domestic (U.S.) experience.

For immigrants, we also distinguish between two different age at arrival groups. Our baseline results are for immigrants who enter after their expected age of graduation. However, we also present results for immigrants who arrive to the U.S. after age 12 but by two years before their expected age of graduation. The latter group allows us to study the value of experience for a group received a large fraction (but not all) of their education in the U.S. Table 1 shows the ten countries with the most immigrants in our sample in total and breaks down the totals by age at arrival category. The table shows that even for the top ten source countries there is a reasonable mixture of rich and poor birth countries, with the income per capita range from roughly 3,200 to 43,000 dollars in 2010 (PPP GDP p.c., PWT 7.1, Vietnam to Canada).

We construct the hourly wage using information on annual wage and salary income for the prior year, usual hours worked per week, and weeks worked in the prior year. In 1980 income was top

2We find that most immigrants report being in their country of birth right before migrating: 87% report being in their birth country five years before migrating and 83% report being their one year before migrating. There also appears to be no systematic relationship between this secondary migration and GDP per capita: see Appendix Figure 12.

3Weeks worked is coded into categories in 1980 and from 2008 onward. We use 1990 data to compute the average weeks worked per category in 1990 and impose this on the 1980 data; we use the 2007 data to compute the average weeks
coded; we multiply all top-coded values by 1.4, in line with the literature. From 1990 onward the Census replaces all top-coded values with the mean of state income within the top-coded group, so no adjustment is needed.

Finally, we use two Census-provided controls in our analysis. The first is state of residence, which is designed to help capture the large cross-state differences in cost of living that would otherwise bias our results. The second is English-language ability. The Census has included a self-reported measure of English language ability throughout this time, with five options ranging from “Does not speak English” to “Yes, speaks only English.” Given that we study immigrants this is a useful control. We further parse the data by creating a sixth category for U.S. born persons, so that the remaining categories all capture variation within the immigrant population.

2.2. New Immigrants

This section illustrates the main spirit of our exercise in the simplest possible way by focusing on new immigrants, which we define as immigrants that arrived in the United States in the year prior to a census. The advantage of looking at new immigrants is that they have a negligible amount of U.S. work experience. Thus we can estimate the returns to foreign experience, for each country, without having to consider interaction effects between foreign and U.S. experience.

2.2.1. Simplest Specification

We begin by estimating returns to foreign experience among immigrants in the simplest possible specification, motivated by the classic approach of Mincer (1974). Also for simplicity, we estimate the returns one country at a time. Letting \( w_{it} \) be the wage of worker \( i \) in time period \( t \) and \( s_{it} \) be their years of schooling, we estimate for each country:

\[
\log(w_{it}) = \alpha + \theta s_{it} + \sum_{x \in X} \phi_x D^x_{it} + \mu_t + \epsilon_{it}
\]

where \( D^x_{it} \) is a dummy variable that takes the value of one if a worker is in experience group \( x \in X = \{5 - 9, 10 - 14, ..\} \); the omitted category is less than five years. This specification allows us to capture non-linearities in the return to experience in a flexible way. The coefficient \( \phi_x \) captures the average wage of workers in experience group \( x \) relative to workers with less than five years of experience. The coefficient \( \theta \) captures the return to schooling and \( \mu_t \) controls for time effects, since we have pooled multiple cross-sections. The regression coefficients \( (\alpha, \theta, \phi_x) \) naturally differ across countries, but we suppress country indices for simplicity.

For each country we focus only on new immigrants, who arrived in the United States in the year prior to a census. For illustrative purposes we begin by presenting the results for four select countries that worked per category in 2007 and impose this on the 2008–2013 data.
have large samples of such new immigrants: Canada, the United Kingdom, Mexico, and Guatemala each have more than 500 new immigrants in our sample. In the subsequent section we present our findings for all countries for which we have sufficient numbers of new immigrants.

Figure 1 presents the estimated returns to foreign experience for these four countries. Note that although we estimate the regression for log-wages, we report the resulting coefficients in percentage change in the level of wages from the omitted category, 0–4 years of experience. Notably, returns to foreign experience are high for immigrants from Canada and the United Kingdom and are much more modest for immigrants from Mexico and Guatemala. Relative to a new immigrant with 0–4 years of foreign potential experience (i.e. one that worked little in his birth country), an immigrant from the United Kingdom or Canada with 20–24 years of foreign experience earns 125–200 percent higher wages. For Mexico and Guatemala, immigrants with 20–24 years of potential experience earn roughly 10–30 percent wage premiums. These findings suggest that returns to experience can vary dramatically across immigrants from different countries.\(^4\)

2.2.2. Richer Specification

We now consider a richer specification that allows for cohort-of-immigration effects, following the work of Borjas (1985), to capture the idea that immigrants who enter in different years may be drawn from different parts of the income or talent distribution in their birth country. We also pool all countries for which we have at least 500 new immigrants, include native-born workers, and add controls for state of residence, gender and English-language ability. We now estimate

\[
\log(w_{it}) = \alpha + \beta z_{it} + \theta s_{it} + \sum_{x \in X} \phi_x D_{it}^x + \mu_t + \sum_c \omega_{ic} D_{ic} + \epsilon_{it} \tag{2}
\]

where \(\alpha\) is a country fixed-effect, \(z_{it}\) is a vector of controls for state, gender and english ability, \(\theta\) is country-specific return to schooling, the \(\phi_x\) are the country-specific returns to experience group \(x\), and \(D_{it}^x\) is a dummy for decadal cohort of immigration. As before, each of the estimated coefficients is country specific, but we suppress country indices for simplicity. Note that since we include country fixed effects, the “base person” in each country is an immigrant from that country in question with 0–4 years of potential experience.

In Figure 2 we plot our estimated returns to experience, using (2), using one simple summary statistic: the returns to 20–24 years of foreign experience. We plot this statistic for each country against the country’s GDP per capita in 2010. One can see that the returns to foreign experience vary positively with GDP per capita. The simple linear regression line (drawn in solid blue) has a slope of 62.5 and is significant at the one percent level. We conclude that among new immigrants, returns to foreign

\(^4\)We have also estimated equation (1) with immigrants that arrived within two years of a census. We find similar results, available upon request.
experience are higher for immigrants from richer countries than immigrants from poorer countries.\footnote{One potential source of bias in our calculations comes from selection on which types of immigrants obtain jobs within a year of migrating. This would drive our results if the selection is such that those with low ability from poor countries are more likely to land jobs when they first arrive, while those with high ability from rich countries are more likely to land jobs when they first arrive. In fact we find that virtually all immigrants are employed within a year of migrating, casting doubt on this possible bias. Another possibility is that the types of jobs that are taken by new immigrants are better reflective of their skills for immigrants rich countries than immigrants from poor countries. In Section 4 we compare the occupations of immigrants and non-migrants, country by country, and find little support for this possibility.}

While our paper is the first to estimate the returns to U.S. immigrant experience by income level of the birth country, our findings build on several prior studies. Chiswick (1978) uses earlier U.S. data and finds that returns to experience tend to be lower for immigrants from poorer regions of the world. Coulombe, Grenier, and Nadeau (2014) find that in Canada there are also lower returns to experience for Canadian immigrants from poorer countries. The primary innovation of our paper relative to these two is to provide new insight on why returns are lower for immigrants from poorer countries, using data from both immigrants and non-migrants. We return to this issue in Section 4 to follow.

2.3. Full Set of Immigrants

We now consider returns to experience using the entire sample of immigrants in our data. The main advantage to doing so is that it allows us to draw on more immigrants from more countries. However, their wages are somewhat more complicated because they have experience that accrued in their birth county and experience that accrued in the United States. This fact presents a challenge for estimation because the returns to experience are generally concave. Because of this, it is likely that the value of an immigrant’s U.S. experience will be affected by the amount of prior foreign experience he acquired before immigrating. Our preferred specification captures this by allowing for country-specific quadratic interactions between U.S. and foreign experience.\footnote{By this we mean controls for the product of U.S. and foreign experience; the product of U.S. and the square of foreign experience; and the product of foreign experience and the square of U.S. experience. This approach has been employed elsewhere in the literature. We also considered allowing for more polynomials and explored less parametric functional forms such as interactions between dummy terms. We found that these alternatives gave less precise estimates for many countries and offered little better fit. Details are available upon request.}

We restrict our attention to countries that have at least 1,000 immigrants who meet our sample criteria. We then estimate a parsimonious specification:

$$\log(w_{it}) = \alpha + \beta z_{it} + \theta s_{it} + \sum_{x \in X} \phi_{f,x} D_{it}^{f,x} + \sum_{x \in X} \phi_{u,x} D_{it}^{u,x} + g(x_f, x_u) + \mu_t + \sum_c \omega_{ic} D_{ic} + \epsilon_{it} \quad (3)$$

This semi-parametric specification allows us to estimate the returns to foreign and U.S. experience as before. Now $D_{it}^{x,f}$ is a dummy variable that takes the value of one if a worker is in foreign experience group $x \in X = \{5 - 9, 10 - 14, \ldots\}$, and $D_{it}^{x,u}$ is a similar dummy variable for U.S. experience. $g(x_f, x_u)$ is the polynomial that controls for interactions between foreign and U.S. experience, while
the remaining controls are similar to equation (2).

Figure 3 presents the results. For each country of origin we present two estimates: first, the returns to 20–24 years foreign experience, and second, the returns to 20–24 years U.S. experience. The first thing to note is that our sample size is much larger than in previous figures; we now have estimated returns to experience for 70 countries. The blue dots in the figure represent the returns to foreign experience. As the figure shows, these tend to be lower in the countries with lower GDP per capita than in the countries with higher GDP per capita. The slope coefficient from a regression of the return to 20–24 years foreign experience on log GDP per capita is 20.0, and is statistically significant at the one percent level. The green dots show the returns to U.S. experience. As can be seen, these are also higher in countries with higher GDP per capita, yet the relationship is weaker than for foreign experience. The slope coefficient from a regression of 20–24 years of U.S. experience on log GDP per capita is 5.61, which is significant only at the ten percent level. Finally, we find that the two slopes are significantly different from one another, also at the one percent level.

2.4. Education and Experience

In the previous subsection we documented that the returns to foreign experience were strongly related to birth country GDP per capita, while the returns to U.S. experience were weakly but statistically significantly related. These returns were estimated for all immigrants. Recent research has suggested an important complementary relationship between education and the returns to experience (Lemieux, 2006; Lagakos et al., 2015). For most papers, the evidence for this point comes from estimating the interaction between quantity of schooling and the returns to experience; the main finding is that more educated workers also have steeper life-cycle wage growth.

We can repeat this finding here for immigrants. To do so, we focus on two subsamples: immigrants with no more than a high school degree; and immigrants with at least a college degree (this excludes immigrants with some college or associate’s degrees). We restrict our attention to country-education pairs with at least 1,000 immigrants, and then re-estimate the returns to experience for country and education level. We focus on the returns to U.S. experience, since this holds fixed the country of experience and isolates the effect of quantity of schooling. The result is shown in Figure 4, which plots the estimated heights of the profiles at 20–24 years of experience against GDP per capita. The main finding is that immigrants with more education have steeper returns to U.S. experience, which is shown as the level difference in the figure. This difference corresponds to about 7.5 percent higher wages at 20–24 years of experience. The relationship between height of the profile and GDP per capita is similar across the two groups; we cannot reject that the slopes are the same at even the 10 percent level. This fact supports the idea of an education-experience complementarity that is common across a wide variety of countries.

Immigrants also present a novel opportunity for a second type of test: they allow us to study the
relationship between the country of schooling and the returns to experience. Here we exploit the subsample of immigrants who moved to the U.S. before their education was complete. We restrict our attention to countries with at least 1,000 immigrants in this subsample and estimate the returns to experience for each such country. Note that this is the returns to U.S. experience, since immigrants who move to the U.S. prior to graduation have only U.S. experience. The main finding is shown in Figure 5, which plots the estimated heights of the profiles at 20–24 years of U.S. experience against GDP per capita for immigrants who migrated before and after their expected age of graduation. The main finding is that U.S. education raises the returns to U.S. experience for all countries. This suggests an important complementarity between country of education and country of experience. The fact that this return is similar for poor and rich countries suggests that the effect is more about the type of education or cultural acclimation than a pure education quality effect.

2.5. Robustness

We now explore the robustness of our stylized results given above. We explore robustness along three dimensions. First, we explore whether the results are robust to using alternative metrics for the steepness of profiles. Second, we explore whether the results are robust to the sample selection criteria. Third, we explore whether the results are robust to controlling for possible confounding influences relevant for immigrants. Throughout, we focus on the relationship between the steepness of wage profiles and birth country PPP GDP per capita, in line with Figure 3. The results of our robustness checks are summarized in Table 2.

The first row of that table shows the baseline results for three types of experience: foreign experience; the U.S. experience of foreign-educated workers; and the U.S. experience of U.S.-educated workers. As discussed above the returns to experience are much more strongly related to birth country PPP GDP p.c. for foreign than for U.S. experience. For U.S. experience, it seems to matter little whether the immigrant was entirely educated abroad or was partially educated in the U.S.

The next three rows explore alternative metrics for the steepness of profiles. We see that the same results prevail if we focus on the height of profiles at 35–39 years of experience rather than 20–24. Likewise, the same results prevail if we focus on the average height of the profile or the discounted average of the profile, where future wage growth is discounted at 4 percent per year. The latter is interesting because it corresponds to a present discounted value of life-time earnings calculation in the spirit of what is often done in the education literature. For all possible metrics we find a strong relationship between the value of foreign experience and birth country PPP GDP p.c. The relationship for U.S. experience and PPP GDP p.c. is much weaker.

The next five rows explore alternative sample selection criteria and measures of experience. We see that we find similar results if we include women, part-time workers, or public sector workers. We also find very similar results if we allow experience to start from as early as age 16 rather than age 18 as in
the baseline. These results are very much in line with those in Lagakos et al. (2015) for cross-country comparisons: while we have imposed standard sample selection criteria for estimating life-cycle wage profiles, the estimated profiles are not particularly sensitive to those criteria.

The remaining checks explore robustness to possible concerns for an immigrant sample. First, we show that similar results apply for immigrants with different educational backgrounds, ranging from college graduates to those with at most a high school degree. Second, we show that the results are similar for workers who only work in manufacturing or service industries. Third, we show that the results are if anything stronger for immigrants who speak excellent English. Fourth, we show that the results are very similar if we exclude immigrants who live in ethnic enclaves. Finally, the last row shows that the results are similar if we focus on data from the year 2000+; for these years we can measure year of immigration exactly, rather than in ranges of years. We can see that the relationship is somewhat stronger for college graduates, immigrants who work in the service sector, and immigrants who speak English well. However, the relationship is strong and statistically significant across all of these checks for foreign experience.

Across all of these robustness checks, three common themes emerge. First, it is consistently true that there is a strong and statistically significant relationship between foreign experience and PPP GDP per capita. Second, this relationship is weaker and less statistically significant for U.S. experience. Third, if we further condition on immigrants who received some U.S. experience, we find essentially no significant relationship.

### 3. Model of Immigrant Returns to Experience

In the preceding section, we documented that returns to birth-country experience are lower for immigrants from poor countries. This raises the question why this may be the case, and in particular whether this tells us anything about cross-country differences in human capital accumulation. In this section, we present a simple model that encompasses three different theoretical explanations for this fact. The first of these is differential human capital accumulation in the birth country, and says that immigrants from poor countries accumulate less human capital over the life cycle than immigrants from rich countries. The second is differential selection, and states that immigrants from poor countries are less selected on learning ability on average than immigrants from rich countries or that the extent to which selection varies with experience differs across countries. The third is differential skill loss, and says that immigrants from poor countries lose a lot of skills after migrating, while immigrants from rich countries lose fewer skills. Note that through the lens of the second and third theories different returns to birth-country experience have nothing to do with human capital accumulation. In Section 4 we then bring additional evidence to the table to distinguish between the three different theories laid out in the present section.
In our model, there is a large number of individuals indexed by $i$, each of whom is born in a country indexed by $c$. An individual may work either in his country of origin, acquiring foreign experience or in the United States, acquiring U.S. experience. We denote variables observed in immigrants’ birth countries without superscripts and those observed in the U.S. with asterisk superscripts. For instance, the wage of an individual from country $c$ who works in his birth country is $w_c$ and if he works in the U.S. it is $w^*_c$. Within each country of origin, individuals are heterogeneous along two dimensions: their initial ability or human capital $\eta_{ic}$ and their learning ability $z_{ic}$. We assume that, on average, individuals are equally able in all countries $E[\eta_{ic}] = E[z_{ic}] = 1$ for all $c$. But as discussed below, migrants may come from a selected part of the population. We further denote by $x_{ic}(t)$ and $x^*_{ic}(t)$ the amount of birth-country and U.S. experience an individual has accumulated up to time $t$. If an individual works in his birth country his human capital accumulates passively according to

$$\dot{h}_{ic}(t) = z_{ic} \phi_c(x_{ic}(t)) h_{ic}(t),$$

with $h_{ic}(0) = \eta_{ic}$, and when he works in the U.S. it accumulates according to

$$\dot{h}^*_{ic}(t) = z_{ic} \phi^*_c(x^*_{ic}(t)) h^*_{ic}(t).$$

In particular, we allow for the possibility that, upon arrival in the U.S. immigrants gain access to a “human capital accumulation function” $\phi^*_c$ that is different from the one in their birth country $\phi_c$. We also allow these functions to differ across countries. For simplicity, we assume that individuals do not face any human capital investment decision in our benchmark model, although we show in Appendix B that similar results arise in a Ben-Porath model with endogenous human capital accumulation.

At some level of birth-country experience $x$, workers from country $c$ migrate to the U.S. For simplicity, we assume that individuals do not anticipate migration. When individuals migrate they take their human capital with them. However, some of their human capital may be country-specific and may hence be lost upon migration. In particular we assume that individuals keep only a fraction $m_c(h_{ic})$ of their human capital upon migration $h^*_c = m_c(h_{ic}) h_{ic}$. To make our argument in the most transparent way, it further turns out to be convenient to assume the functional form $m_c(h_{ic}) = \gamma_c h_{ic}^{\theta_c - 1}$ so that human capital upon arrival in the U.S. is

$$h^*_{ic} = \gamma_c h_{ic}^{\theta_c}$$

The parameter $\gamma_c > 0$ captures the average “skill loss” incurred by a migrant from country $c$. The parameter $\theta_c > 0$, in contrast captures whether skill loss is more of a problem for high human capital types. For instance, if $\theta_c < 1$ an immigrant with high human capital loses a larger fraction of his human capital than one with low human capital.
Migrants may also be selected to be different from “stayers.” More precisely, we denote the set of individuals with experience level \( x \) who migrate from country \( c \) to the U.S. by \( M_c(x) \) and allow for the possibility that

\[
\mathbb{E} \left[ \eta_i | i \in M_c(x) \right] \neq \mathbb{E} \left[ \eta_i \right] = 1, \quad \mathbb{E} \left[ z_i | i \in M_c(x) \right] \neq \mathbb{E} \left[ z_i \right] = 1
\]

For example, immigrants from country \( c \) with experience level \( x \) are positively selected on learning ability \( z_{ic} \) if \( \mathbb{E} \left[ z_{ic} | i \in M_c(x) \right] > \mathbb{E} \left[ z_{ic} \right] = 1 \). Note that we allow for fairly general types of selection: there may be selection on both initial ability \( \eta_{ic} \) and learning ability \( z_{ic} \), and furthermore both types of selection may differ with the level of experience.

The wage of immigrant \( i \) from country \( c \) with \( x \) years of birth-country experience and \( x^* \) years of U.S. experience is

\[
w^*_{ic} (x, x^*) = \omega^*_c h^*_ic (x, x^*) e^{\varepsilon_{ic}}
\]

where \( \omega^*_c \) is the skill price earned by immigrants from country \( c \) in the U.S. and \( \varepsilon_{ic} \) is an error term.

Given our assumptions, the immigrant’s human capital can be solved for in closed form and satisfies

\[
\log h^*_ic (x, x^*) = \log \gamma_c + \log \eta_{ic} + \theta_c z_{ic} \int_0^x \phi_c(y) dy + z_{ic} \int_x^{x+x^*} \phi_*(y) dy
\]

Combining with (6), the wage of a new immigrant, i.e. one with zero years of U.S. experience \( x^* = 0 \), is therefore

\[
\log w^*_{ic} (x, x) = \log \omega^*_c + \log \gamma_c + \log \eta_{ic} + \theta_c z_{ic} f_c (x) + \varepsilon_{ic}
\]

where we denote by \( f_c (x) = \int_0^x \phi_c (y) dy \) the cumulative returns to foreign experience. The regression we run using data on new immigrants only is therefore

\[
\log w^*_{ic} (x, x) = \alpha_c + R_c (x_{ic}) + \varepsilon_{ic}, \quad R_c (x) = \mathbb{E} \left[ \log \eta_i | i \in M_c(x) \right] + \theta_c \mathbb{E} \left[ z_i | i \in M_c(x) \right] f_c (x)
\]

The measured return to foreign experience \( R_c (x) \) may be low for one of four reasons. First, the true returns to experience \( f_c (x) \) may be low. Second, there may be experience-dependent selection on initial ability, i.e. \( \mathbb{E} \left[ \log \eta_i | i \in M_c(x) \right] \) decreases with \( x \). Third, there may be selection on learning ability (both experience-dependent and standard selection are a problem, i.e. \( \mathbb{E} \left[ z_i | i \in M_c(x) \right] \) is either less than one or decreasing). Finally, there may be experience-dependent skill loss, \( \theta_c < 1 \). Estimates from the regression (9) by themselves do not allow us to distinguish between these four determinants.

\[
\text{To see this note that for example (4) can be integrated to yield}
\]

\[
\log h_{ic} (x, x) = \log \eta_{ic} + z_{ic} \int_0^x \phi_c (y) dy.
\]

Following similar steps yields (7).
of low measured returns to foreign experience.

In contrast, note that two other potential issues do not show up as low measured returns to foreign experience: selection on initial ability and skill-loss that are not experience-dependent (i.e. $M_c(x) = M_c$ with $\mathbb{E}[\log \eta_{ic} | i \in M_c] < 1$ and $\gamma_c < 1$). These will simply be picked up the country fixed effects $\alpha_c$. In the next section, we bring additional evidence to the table to distinguish between the three different theories: differential human capital accumulation, differential skill loss, and differential selection.

4. Distinguishing Between Theories

In this section we draw on new data that compares characteristics of immigrants and non-migrants from a large set of countries. We draw on three basic facts that help us distinguish between the theories above. First, returns to foreign experience among immigrants are similar to returns to potential experience among non-migrants. Second, immigrants from poor countries tend to be more selected on pre-migration characteristics such as years of schooling. Third, educated immigrants tend to work in high-skilled occupations at a lower frequency than non-migrants, though at a similar rate in rich and poor countries alike.

4.1. Returns to Experience Among Immigrants and Non-Migrants

We begin by comparing our returns to foreign experience among immigrants to the returns among non-migrants estimated by Lagakos et al. (2015). We can make these comparisons in the 15 countries for which we have an estimate of immigrant returns, and for which Lagakos et al. (2015) calculate returns using a representative sample of non-migrants. Since we have followed the sample selection and variable construction of Lagakos et al. (2015) closely, the comparability of the results is informative about the extent to which life-cycle wage growth differs between immigrants and non-migrants. We begin by plotting the estimated returns to 20–24 years of experience against GDP per capita in Figure 6. As one can see from the figure, both estimates show a strong positive relationship with GDP per capita, with higher returns to experience, on average, in the economies with higher GDP per capita. Among immigrants, the slope coefficient in a regression of GDP per capita is 28.0 for the immigrants, with a P-value less than 0.001. Among non-migrants the slope coefficient is 19.2 and the P-value is 0.003.

Figure 7 plots the estimated returns to 20–24 years of experience for immigrants against the same estimated return for non-migrants. The 45-degree line is also plotted for reference. As one can see, there is a positive relationship between the two sets of estimates, though the relationship is far from completely linear. The correlation coefficient between the two estimates is 0.797 with a P-value of 0.004. Countries like Germany, the UK, and Australia are high among both immigrants and non-migrants, and most of the developing countries have low returns in both groups. Prominent outliers
include Indonesia and Korea.

The fact that estimated returns to experience from poor countries are low both for immigrants and non-migrants provides one piece of evidence against differential selection as a theory of the immigrant evidence. If low returns to experience among immigrants were driven solely by negative selection by immigrants from poor countries, one would expect that returns to experience among non-migrants were similar in countries of all income levels. As Figures 6 and 7 show, this is not the case. The broad similarity between returns to experience among immigrants and non-migrants is also evidence against differential skill loss as a theory of the immigrant returns. If low returns among immigrants from poor countries were solely due to skill loss, one would again expect that returns to experience among non-migrants would be similar in countries of all income levels. This prediction is not borne out in the figures. Instead, the figures suggest a world where workers in poor countries do not acquire much human capital while in their birth countries.

4.2. Years of Schooling Among Immigrants and Non-Migrants

We next compare years of schooling completed among immigrants to years of schooling for non-migrants. As the theory above shows, years of schooling are informative about learning ability, with higher ability individuals attending more school on average. We compare the average years of schooling among immigrants for each country in our data to the average years of schooling completed by country from Barro and Lee (2012). We can make this comparison for every country for which data is available from both sources.

Figure 8 shows the two data sets plotted against one another, with the 45-degree line for reference. As can be seen in the figure, the average schooling level of immigrants is higher than that of non-migrants in every country in the world (except for the U.S., which lies on the line by construction.) The schooling gaps are particularly large for the poorest countries, where immigrants average more than 12 years of schooling, and non-migrants average far less, in many cases less than 6 years.

Interestingly, the schooling of immigrants in some countries are not very high compared to non-migrants, and compared with other countries of similar income levels. Mexico, for example, has virtually identical schooling levels for immigrants and non-migrants. The immigrant schooling levels are low also in Guatemala, Laos, Cambodia, El Salvador, Honduras, Portugal and Yemen. This suggests that immigrants from these countries may not be very positively selected. This could account for the low estimated returns to experience among immigrants from Mexico and Guatemala.

We note that our conclusions in this section are consistent with findings of previous studies. Chiquiar and Hanson (2005) use census data from Mexico and the United States to argue that there is "intermediate" selection of immigrants from Mexico. Their key piece of evidence is that years of schooling attained are a bit higher among Mexican immigrants than Mexican non-migrants. Grogger and Han-
son (2011) show that, across a wide set of countries, the share of college educated workers among immigrants is substantially higher than the same share among all individuals. They argue that this implies positive selection among immigrants in general.

4.3. Occupations of Educated Immigrants and Non-Migrants

In this section we compare the occupations of college-educated immigrants and non-migrants. The goal is to understand “skill loss” after migrating, by which we mean having skills that are valued in the country of origin but that are not applied after migrating. This could be, for example, because the production technologies in which the immigrant was experienced pre-migration differed the ones used post-migration.

We proxy skill loss by comparing the fraction of college-educated immigrants that work at high-skilled occupations to the same fraction calculated for non-migrants. We focus on college educated individuals since presumably they are the ones who have the most skills to lose. As above, we ask specifically whether our proxies for skill loss are correlated with GDP per capita of the origin country.

We define occupations as either “high-skilled” or “low-skilled” using the international standard code of occupations constructed by IPUMS (Ruggles et al., 2015). We defined high skilled to be professionals, technicians and associate professionals, and legislators, senior officials and managers. We define low skilled to be clerks, service workers and shop and market sales, skilled agricultural and fishery workers, crafts and related trades workers, plant and machine operators and assemblers and elementary occupations. We omit individuals in the armed forces or other unspecified or unreported occupations.

Figure 9 shows that fraction of college-educated immigrants and non-migrants that work at high-skilled occupations. Not surprisingly, most countries lie below the 45-degree line, meaning that a larger fraction of educated non-migrants work at high-skilled occupations than educated immigrants. This confirms that skill loss is likely to be an important reality for immigrants. What matters for the current study, however, is whether skill loss is correlated with GDP per capita.

Figure 10 plots the ratio of the high-skilled employment rate for immigrants to the the high-skilled employment rate for non-migrants against GDP per capita. In other words, we look at the ratio of the y-value to x-value of Figure 9 against GDP per capita. Most countries have ratios between 0.5 and 1.0, meaning that on average a substantially smaller fraction of immigrants work at high-skilled occupations than non-migrants. However, this ratio seems to be largely uncorrelated with GDP per capita. This lack of correlation suggests that skill loss is not disproportionately a phenomenon pertaining to immigrants from poor countries. Instead, skill loss after migrating seems to be present in immigrants from countries of all income level, and on average the magnitudes appear similar across

---

8When we focus on high-school graduates or higher we research similar conclusions as the ones presented below.
the income distribution. Thus, it seems unlikely that the facts on immigrant returns to experience we document are explained simply by skill loss that mostly affects immigrants from poor countries. That skill loss affects immigrants from all countries, but without a systematic relationship with GDP per capita, is consistent with the fact documented in Figures 6 and 7 that profiles are on average flatter for immigrants than for non-migrants from the same country.

4.4. Additional Evidence from the Existing Literature

We now briefly discuss additional evidence from existing research that allows one to distinguish between the three theories that can potentially account for our facts. One additional type of evidence favoring the theory that experience human capital accumulation is higher in the U.S. than in developing countries comes from return migrants. Reinhold and Thom (2013) find that Mexican immigrants to the United States earn a large premium on their U.S. experience when returning to Mexico. A second is recent work by Hendricks and Schoellman (2015) which uses pre- and post-migration labor market outcomes for a sample of immigrants to document basic facts of selection. They find systematic evidence that immigrants from poorer countries are substantially more selected than immigrants from rich countries on a variety of outcomes, including education, occupation, wages, and earnings. This finding is consistent with our evidence above that the low age-earnings profiles of immigrants are unlikely to be explained by immigrants from poorer countries being less selected.

5. Development Accounting

In this section we use development accounting to quantify the economic importance of the empirical results shown in Section 2. To keep our findings as comparable as possible to the previous literature, we follow the accounting approach of Klenow and Rodríguez-Clare (1997), Hall and Jones (1999) and in particular Caselli (2005). We follow Lagakos et al. (2015) in focusing on the importance of human capital from experience, and Schoellman (2012) in using immigrants to construct human capital stocks.

The accounting procedure uses a Cobb-Douglas aggregate production function \( Y_c = K_c^\alpha (A_c H_c)^{1-\alpha} \), where \( Y_c \) is GDP per worker of country \( c \), \( K_c \) is physical capital per worker and \( H_c \) is human capital per worker. The capital share is assumed to equal one-third. As in Caselli (2005), we calculate the measure

\[
\text{success}_1 = \frac{\text{var} (\log Y_{KH,c})}{\text{var} (\log Y_c)}
\]

where \( Y_{KH,c} = K_c^\alpha H_c^{1-\alpha} \) is the component of output due to factors of production. Values of \( \text{success}_1 \) close to one suggest that cross-country differences in capital stocks account for nearly all of measured income differences. Values close to zero imply that capital stocks account for none of income differences. One limitation of the measure \( \text{success}_1 \) is that measurement error in \( Y_{KH,c} \) could increase
success, while clearly this does not imply a greater importance of capital stocks. Thus, to complement the successes metric, we also report the slope of a regression of $\log Y_{KH,c}$ on $\log Y_c$.

To highlight the difference between our findings and those of the previous literature, we use the same physical capital estimates as Caselli (2005), and assume that all individuals in a given country have the same levels of schooling and experience $\bar{s}_c$ and $\bar{x}_c$ (also taking these averages from Caselli (2005)). Our measure of the stock of human capital differs only in the assumed life-cycle profile of labor market productivity. We consider two assumptions on this profile, corresponding to whether we view human capital as the result of passive investment (simple learning-by-doing, as in Section 3) or active accumulation (Ben-Porath, as in Appendix B). These two models differ slightly in their interpretation of life-cycle increases in wages. The former attributes all of this increase to rising human capital over the life cycle. By contrast, the latter attributes some of this increase to an increase in time spent producing at work (resulting from a decrease in time spent investing in human capital accumulation over the life cycle).

Both of these formulations allow us to express the human capital of a worker with years of schooling $s$ and experience $x$ in country $c$ as

$$h_c(s,x) = \exp(g_c(s) + f_c(x)).$$

The functions $g_c$ and $f_c$ measure the human capital returns to schooling and experience. The aggregate human capital stock of country $c$ is then simply defined as the human capital of an individual with the average years of schooling and experience, $H_c = h_c(\bar{s}_c, \bar{x}_c)$.

We use our estimates from Section 2 to discipline $f_c$. In particular, note that the return to foreign experience measured in the U.S. $R_c(x,0)$, which we have measured there, identifies the sum of the true human capital return to experience $f_c(x)$ and a term due to changes in the amount of time allocated towards human capital accumulation:

$$R_c(x,0) = f_c(x) + \log \left( \frac{1 - \ell_c(x)}{1 - \ell_c(0)} \right).$$

We conduct two alternative accounting exercises, which provide an upper bound and lower bound on the importance of human capital in development accounting implied by our empirical results. We begin with the upper-bound exercise, which assumes that the investment time allocation, $\ell_c(x)$, is con-

---

9The connection to the model in the Appendix is a bit more subtle in this respect. Experience human capital accumulates according to $\dot{h}_c = B_c \phi(\ell_c)h_c - \delta h_c$ and hence its logarithm at experience level $x$ can be written as $f_c(x) = \int_0^x (B_c \phi(\ell(\bar{s})) - \delta) d\bar{s}$.

10We have also done an accounting exercise where we use the subset of countries for which we have micro data on the distributions of schooling and experience to relax the assumption that all individuals in each country have the average schooling and experience level. Specifically, we use the empirical distribution of schooling and experience in country $c$, $F_c(s,x)$ and define aggregate human capital as $H_c = \int h_c(s,x) dF_c(s,x)$. We find similar results in this exercise as the ones described in the text, and hence omit them for brevity.
stant across experience levels for each country (as in passive investment, where it is constant at zero). This assumption allows us to measure human capital accumulation directly from the experience-wage profiles we estimated in Section 2, as $f_c(x) = R_c(x,0)$. From the perspective of a Ben-Porath model this is an upper bound on the importance of experience human capital because workers from rich countries spend more time investing early in their life cycle than do workers from poor countries. Thus, by assuming constant investment time we overestimate the life-cycle human capital accumulation of workers from rich countries relative to that of workers from poor countries.

The accounting under this upper bound is presented in the top panel of Table 3. The first column presents our measures of $success_1$. When only schooling is taken into consideration, $success_1$ is 0.44, meaning that human and physical capital account for just under one half of income differences. When only experience is considered, $success_1$ is similar, at 0.53. When they are both considered, $success_1$ rises to 0.81, meaning that now more than three quarters of income differences are accounted for by measured capital stocks. The second column shows that the correlation of measured capital stocks and GDP per capita rises substantially as well. With just schooling, the slope coefficient from a regression of $\log(Y_{KH})$ and $\log(GDP)$ is 0.64. With both schooling and experience used to compute human capital stocks, the slope coefficient rises to 0.84. Thus, under this upper bound at least, the importance of human capital increases substantially when we include experience human capital estimating using immigrant returns to experience.

We turn now to our second accounting exercise, which provides a lower bound on the importance of human capital implied by our empirical findings. In this lower-bound exercise we estimate $f_c(\bar{x}_c)$ as the difference between the returns to $\bar{x}_c$ years of U.S. experience and $\bar{x}_c$ years of foreign experience. To understand why this provides a lower bound, note that time allocation decisions in the model are not driven by the location in which experience has been accumulated, but only by the total number of years left to be spent at work (see Lemma 1). Furthermore, for a given level of (total) potential experience, immigrants from poor countries will spend a greater fraction of their time working than immigrants from rich countries. Thus, we will tend to overestimate the life-cycle human capital accumulation of workers from poor countries relative to workers from rich countries.

The accounting under this lower bound is presented in the bottom panel of Table 3. This time, when human capital from both schooling and experience are taken into consideration, $success_1$ is 0.65, up from 0.44 when only schooling is considered. The slope coefficient from a regression of $\log(Y_{KH})$ on $\log(GDP)$ is 0.77, up from 0.64 when only schooling is considered as human capital. We conclude that under this lower bound, the importance of human capital increases greatly when experience is included.

We plot our estimated human capital stocks against GDP per capita in Figure 11. This figure plots the human capital stocks implied by our two accounting exercises, and the slope from a regression
of human capital stocks measured only using schooling on log GDP per capita. As the figure shows, our estimated human capital stocks are substantially larger in rich countries than poor countries once experience is included.

6. Conclusion

This paper seeks to understand whether workers in richer countries acquire more human capital over the life cycle than workers in poor countries. The answer has first-order implications for the literature that attempts to account for cross-country income differences using measured stocks of human and physical capital. Previous studies have concluded that cross-country differences in life-cycle human capital accumulation are negligible, and that the overall importance of human capital in accounting for income differences is modest (Klenow and Rodríguez-Clare, 1997; Bils and Klenow, 2000, 1998; Caselli, 2005). Yet more recent work claims that human capital plays a much more central role (Manuelli and Seshadri, 2015; Lagakos et al., 2015).

To address this question, this paper draws on evidence from U.S. immigrants, who come from countries of all income levels but work in a common labor market. We document that immigrants from richer countries tend to have higher returns to potential experience than immigrants coming from poor countries. We argue that the most likely explanation of this fact is that workers in rich countries simply acquire more human capital before migrating. Another logical possibility is that immigrants from rich countries are just better selected on learning ability than immigrants from the developing world. Yet this contrasts with the observation that immigrants from poor countries tend to much better educated than their counterparts that stayed behind, whereas immigrants from richer countries are only modestly more educated than non-migrants from the same countries. Yet another possibility is that immigrants from poor countries disproportionately lose skills after migrating. But this contrasts with evidence on the occupations of immigrants compared to non-migrants, which suggest similar skill loss across countries. Finally, the fact that returns to experience are similar between immigrants and non-migrants, in most countries, is most consistent with a model in which workers in poor countries simply accumulate less human capital during their working years.

Why are our findings relevant for macroeconomics? A large literature on development accounting has concluded that human capital accounts for at best a modest fraction of living standard differences across countries. This literature has concluded that including differences in life-cycle human capital accumulation (i.e. human capital from experience) doesn’t change the accounting. In contrast, our findings point to a very different conclusion, which is that life-cycle human capital differences are large. Our development accounting, based on our evidence from U.S. immigrants, suggests a much larger role for human capital in accounting for cross-country income differences.

A natural but challenging next step is to explain why life-cycle human capital accumulation tends
to be lower in poor countries than rich countries. One possible explanation is that a lower-quality schooling environment in poor countries results in less “learning how to learn” among individuals who attend school there. One piece of evidence we find supporting this theory is that immigrants that arrived in the U.S. during schooling have returns to subsequent experience that look very similar to those of natives. On the other hand, those that finished schooling in a poor country and then migrated to the United States have lower returns to U.S. experience. We plan to explore this finding more in future research.
References


<table>
<thead>
<tr>
<th>Country</th>
<th>Total</th>
<th>By Age at Arrival</th>
<th>Post-Graduation</th>
<th>Late School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico</td>
<td>329,284</td>
<td>280,955</td>
<td>49,630</td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>52,622</td>
<td>38,208</td>
<td>14,422</td>
<td></td>
</tr>
<tr>
<td>Philippines</td>
<td>38,093</td>
<td>29,905</td>
<td>8,199</td>
<td></td>
</tr>
<tr>
<td>El Salvador</td>
<td>31,443</td>
<td>26,476</td>
<td>5,039</td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>31,115</td>
<td>24,075</td>
<td>7,054</td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td>27,750</td>
<td>20,058</td>
<td>7,721</td>
<td></td>
</tr>
<tr>
<td>United Kingdom, ns</td>
<td>21,748</td>
<td>18,231</td>
<td>3,522</td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td>21,326</td>
<td>18,465</td>
<td>2,929</td>
<td></td>
</tr>
<tr>
<td>Cuba</td>
<td>20,698</td>
<td>16,266</td>
<td>4,454</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>20,010</td>
<td>15,567</td>
<td>4,449</td>
<td></td>
</tr>
</tbody>
</table>

Note: Total represents the total number of immigrants in our sample. The remaining columns represent the total number of immigrants by their age at arrival.
<table>
<thead>
<tr>
<th>Schooling:</th>
<th>Foreign</th>
<th>Foreign</th>
<th>U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>20.0***</td>
<td>5.61*</td>
<td>10.5</td>
</tr>
<tr>
<td>35–39 Years Experience</td>
<td>23.3**</td>
<td>7.2</td>
<td>11.0</td>
</tr>
<tr>
<td>Average Height of Profile</td>
<td>16.4***</td>
<td>5.5**</td>
<td>7.9</td>
</tr>
<tr>
<td>Discounted Average Height of Profile</td>
<td>8.4***</td>
<td>2.7**</td>
<td>3.9</td>
</tr>
<tr>
<td>Include Women</td>
<td>20.1***</td>
<td>3.6</td>
<td>2.3</td>
</tr>
<tr>
<td>Include Some Part-Time</td>
<td>20.5***</td>
<td>5.4*</td>
<td>10.8</td>
</tr>
<tr>
<td>Include All Part-Time</td>
<td>23.8***</td>
<td>7.8***</td>
<td>5.3</td>
</tr>
<tr>
<td>Include Public Sector</td>
<td>21.5***</td>
<td>6.3**</td>
<td>15.6**</td>
</tr>
<tr>
<td>Experience at 16</td>
<td>20.9***</td>
<td>3.3</td>
<td>1.5</td>
</tr>
<tr>
<td>≥College Graduate</td>
<td>38.6***</td>
<td>5.5</td>
<td>15.1</td>
</tr>
<tr>
<td>≤H.S. Graduate</td>
<td>13.8***</td>
<td>4.9</td>
<td>10.2</td>
</tr>
<tr>
<td>No Advanced Degrees</td>
<td>12.5***</td>
<td>4.8**</td>
<td>11.9**</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>15.7***</td>
<td>5.4</td>
<td>N/A</td>
</tr>
<tr>
<td>Service Industry</td>
<td>23.2***</td>
<td>11.6***</td>
<td>16.0**</td>
</tr>
<tr>
<td>Excellent English</td>
<td>26.4***</td>
<td>6.7*</td>
<td>7.8</td>
</tr>
<tr>
<td>No Ethnic Enclaves</td>
<td>22.6***</td>
<td>6.3*</td>
<td>17.0**</td>
</tr>
<tr>
<td>Year 2000+</td>
<td>21.9***</td>
<td>6.4**</td>
<td>22.0***</td>
</tr>
</tbody>
</table>

Note: Each estimate in the table represents the slope coefficient from a regression of the estimated returns to 20–24 years of potential experience on GDP per capita. Estimates are for the sample, school location, and experience location given. ***,**, and * denote results that are significant at the 99 percent, 95 percent, and 90 percent. N/A denotes fewer than fifteen countries with estimates of the relevant return to experience.
Table 3: Development Accounting

<table>
<thead>
<tr>
<th>Human Capital Measure</th>
<th>Success$_1$ (1)</th>
<th>Slope($\log(Y_{KH})$, $\log(GDP)$) (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(a) Upper Bound</td>
<td></td>
</tr>
<tr>
<td>Schooling</td>
<td>0.44</td>
<td>0.64</td>
</tr>
<tr>
<td>Experience</td>
<td>0.53</td>
<td>0.68</td>
</tr>
<tr>
<td>Schooling + Experience</td>
<td>0.81</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>(b) Lower Bound</td>
<td></td>
</tr>
<tr>
<td>Schooling</td>
<td>0.44</td>
<td>0.64</td>
</tr>
<tr>
<td>Experience</td>
<td>0.41</td>
<td>0.61</td>
</tr>
<tr>
<td>Schooling + Experience</td>
<td>0.65</td>
<td>0.77</td>
</tr>
</tbody>
</table>
Figure 1: Returns to Foreign Experience Among New Immigrants, Select Countries
Figure 2: Returns to Foreign Experience Among New Immigrants by GDP p.c. of Birth Country
Wages at 20–24 Relative to 0–4 Years Experience

Real GDP p.c., 2010

Foreign Experience

U.S. Experience

Figure 3: Returns to Foreign and U.S. Experience by GDP p.c. of Birth Country
Figure 4: Returns to Experience by Education Level and GDP p.c. of Birth Country
Figure 5: Returns to U.S. Experience by Education Location and GDP p.c. of Birth Country
Figure 6: Returns to Foreign Experience of Immigrants and Non-Migrants by GDP p.c. of Birth Country
Figure 7: Returns to Foreign Experience, Immigrants vs Non-migrants
Figure 8: Years of Schooling Completed Among Immigrants and Non-migrants
Figure 9: Percent of Educated Workers in High-Skilled Occupations, Immigrants and Non-migrants
Figure 10: Percent of Educated Workers in High-Skilled Occupations: Ratio of Immigrants to Non-migrants
Figure 11: Upper and Lower Bound Human Capital Stocks
Appendix

A. Estimating Returns to Experience Among Immigrants

The identification issues are mostly clearly explained when we assume that experience, schooling, and year all enter the regression equation linearly; the identification issues are most clearly explained in this case. Given these assumptions the regression equation is then:

$$\log(w^N) = \beta^N X^N + \phi^N E^N + \omega^N Y^N + \mu^N S^N + \epsilon^N$$  \hspace{1cm} (10)

where Greek variables denote the coefficients and $\epsilon$ is the error term. The superscript $N$ is used to denote natives.

Our primary goal is to study the determinants of immigrants’ earnings. Similar to Chiswick (1978) and Schoellman (2012), we want to allow the return to foreign-acquired schooling to differ from U.S.-acquired schooling. We also want to distinguish between the return to foreign (birth country) and domestic (U.S.) experience for immigrants $FE$ and $DE$. We will also allow the return to domestic experience to be different for immigrants and natives.

However, a by-now large literature has proposed alternative possible factors that may matter for the determinants of immigrants’ earnings, and raised some identification issues that need to be addressed. Borjas (1985) suggested allowing for year of immigration cohort effects, $C$, to capture the idea that immigrants who enter in different years may be drawn from different parts of the income or talent distribution in their birth country. Friedberg (1992) suggested allowing for an effect of age at arrival, $AA$. She hypothesizes that older immigrants will be more invested in their birth country and less able to adapt to the U.S. Finally, some authors have suggested allowing a role for years in the U.S. $YUS$ to capture the assimilation of immigrants. Combining all of these potential factors would suggest a regression equation of:

$$\log(w^I) = \beta^I X^I + \phi_1^I FE^I + \phi_2^I DE^I + \omega^I Y^I + \mu^I S^I + \alpha^I AA^I + \gamma^I C^I + \delta^I YUS^I + \epsilon^I$$  \hspace{1cm} (11)

where Greek variables denote again coefficients and the superscript $I$ denotes immigrants. Note that we have allowed the returns to common characteristics (such as $S$ and $X$) to vary between natives and immigrants.

A well-known problem in the literature is that a number of the terms on the right-hand side of equation (11) are linearly related to one another, in which case it is not possible to identify the corresponding coefficients. A useful way to express these dependencies is to show that seven of the right-hand side variables are actually constructed using linear combinations of four survey questions: age, years of schooling, dataset year, and year of immigration $YI^I$. Years of schooling and dataset year enter the
regression equation directly; five other variables in that equation are linear combinations of these four survey questions:

1. \( FE^I = A^I - S^I - 6 - (Y^I - Y^I) \)
2. \( DE^I = Y^I - Y^I \)
3. \( AA^I = A^I - (Y^I - Y^I) \)
4. \( C^I = Y^I \)
5. \( YUS^I = Y^I - Y^I \)

Equation (11) thus includes seven variables that are linear combinations of four survey questions. Three assumptions or restrictions are necessary to make estimation feasible.

Our first restriction comes from pooling immigrants and natives into a single regression and restricting \( \omega^N = \omega^I \). The assumption here is that time effects capture aggregate economic conditions such as recessions or inflation that affect immigrants and natives equally. In this case the time effects can be estimated for the natives and imposed on the immigrants, reducing the number of equations by one.

The remaining two restrictions are almost definitional in nature. First, note that U.S. experience and years in U.S. are in fact defined in the same manner. In this case it is impossible to identify separately the effect of U.S. potential experience from any other, more general effects of spending time in the U.S., including social assimilation. Hence, we can include only one of these two regressors. In general, it is not clear whether the resulting estimated coefficient captures the effect of U.S. experience or of other factors related to years since migration. The second restriction arises from the fact that foreign potential experience and age at arrival are almost identical: they differ only by the expected age at graduation, \( S^I + 6 \). Once again, the implication is that it is difficult to distinguish between the effects of foreign experience and a more general effect for age at arrival due to, say, adaptability. However, given that our estimated experience effects for immigrants look strikingly similar to those estimated in Lagakos et al. (2015) for non-migrants, it seems that we will be able to make a concrete contribution to the interpretation of these coefficients.

An alternative way to explain our identification is to compare the wages of hypothetical immigrants who are constructed to clarify the sources of identifying variation. Table 4 does this. Immigrant 1 is an arbitrarily constructed “baseline” immigrant with year, year of immigration, age, and years of schooling listed in Panel A. These are the primitive statistics available in the Census. Panel B then shows the statistics we would construct and use in the regression: foreign and domestic experience,

\[ ^{11} \text{A less restrictive assumption is to require that immigrants share the time effects of a particular subgroup of natives. I have never seen this idea actually implemented.} \]
schooling, and year of immigration cohort. Immigrants 2–5 are constructed to allow for the identification of a single parameter of interest. For example, immigrant 2 is identical to immigrant 1 in almost all characteristics and is observed in the same dataset year. The only difference is that this immigrant is a year older. Inspection reveals that immigrant 2 differs from immigrant 1 only in foreign experience, so comparing the wages of these two immigrants will allow us to identify the role of foreign experience. Likewise, immigrant 3 is a year older than immigrant 1, but also has a year more schooling. When we construct right-hand side variables this immigrant will be identical to immigrant 1, except with an additional year of schooling. Comparing these two immigrants allows us to identify the effects of foreign schooling.

The remaining two parameters are a bit trickier to identify. To capture domestic experience effects we want to study an otherwise identical immigrant who has one additional year in the U.S. To do so, we need to observe this immigrant a year later, in the 2011 data. If we assume that dataset year effects are the same for immigrants as for natives, then we can estimate this effect using natives and net it out. Then the comparison between immigrants 4 and 1 identifies the effect of domestic experience. Finally, a similar problem applies to cohort effects. The cleanest way to identify cohort effects is to consider immigrant 5, who is born and immigrates to the U.S. a year later. Again, we have to change two variables at once, but if we can estimate dataset year effects using natives then this is not a problem. Under the same assumption, the comparison between immigrants 5 and 1 identifies the effect of entering in a particular cohort. These latter two examples clarify the role of the assumption that year effects are shared between natives and immigrants.

An alternative approach is to study an immigrant who is a year older and immigrates a year earlier; but this again involves changing two variables simultaneously (in this case domestic experience and cohort).
A simple way to reduce these equations is as follows: we identify the effects of foreign experience using older immigrants; the effects of foreign schooling using older and better educated students; the effect of domestic experience using older immigrants observed in later years; and cohort effects using immigrants who enter later and are observed later. Now that issues of identification are clear, we turn to the data and estimation procedures.

**B. Ben-Porath Model of Human Capital Accumulation**

In this Appendix we show that differential selection, differential skill loss, and differential human capital can also explain differences in experience-wage profiles for immigrants in a model of endogenous human capital accumulation, namely the Ben-Porath model. We model the human capital accumulation decision of an individual from country \(c\) who may work either in his country of origin, acquiring foreign experience or in the United States, acquiring U.S. experience. We denote variables observed abroad with asterisk superscripts, and those observed in the U.S. without superscripts. For instance, the wage of an individual from country \(c\) who works in his country of origin is \(w^*_c(t)\) and if he works in the U.S. it is \(w_c(t)\). Individuals devote a fraction \(\ell_c(t)\) of their time to human capital accumulation. If they work in their birth country their human capital accumulates according to

\[
\dot{h}_c = B^*_c \phi(\ell_c) h_c - \delta h_c,
\]

and when they are in the U.S. it accumulates according to

\[
\dot{h}_c = B_c \phi(\ell_c) h_c - \delta h_c.
\]

We assume that \(\phi(\ell) = \ell^\gamma, \gamma < 1\) and that the depreciation rate \(\delta \geq 0\). The parameters \(B_c\) and \(B^*_c\) determine how quickly human capital accumulates for a given amount of time devoted to human capital accumulation. \(B^*_c\) may vary across countries and may be different from \(B_c\), capturing the idea that countries differ in the quality of their “learning environment.” We also allow the “learning environment” in the U.S. \(B_c\) to vary across countries so that it matters where an individual is born, even after he migrated to the U.S. In Sections B.2 and B.3 we extend the model to feature individual-specific heterogeneity in the parameters \(B_c\) and \(B^*_c\) so as to explore the issue of selection of migrants with different learning abilities.

The U.S. wage is \(w_c(t) = \omega_c (1 - \ell_c(t)) h_c(t)\) and analogously for the foreign wage. An individual
born in the U.S. solves

$$\max_{\{\ell_c(t)\}} \int_0^T e^{-rt} w_c(t) dt \quad \text{s.t.}$$

$$w_c(t) = \omega (1 - \ell_c(t)) h_c(t)$$

$$h_c(t) = B_c \phi (\ell_c(t)) h_c(t) - \delta h_c(t)$$

$$0 \leq \ell_c(t) \leq 1$$

where the human capital at the beginning of the work life, $h_c(0)$, is given. Workers abroad solve an analogous problem. At some level of foreign experience $x^*$, workers from country $c$ migrate to the U.S.. For simplicity, we assume that individuals do not anticipate migration so that their human capital accumulation decision before migration depends only on the environment in their birth country. In our benchmark model we assume that when workers migrate, they take with them their entire human capital stock so that $h_c(x^*) = h^*_c(x^*)$. In Section B.2 we extend the model to allow for “skill loss” upon migration, that is $h_c(x^*) < h^*_c(x^*)$. Finally, in our benchmark exercise we focus on parameter constellations such that there is an interior solution for the time allocation decision, $0 < \ell(t) < 1$ for all $t < T$ so that in particular individuals earn a strictly positive wage.

One useful feature of our Ben-Porath model is that it can be mapped very directly to the empirical model in section 2, in particular Mincer type regressions such as (1). To see this note that the wage of a non-migrant satisfies

$$\log w^*_c(t) = \log \omega^* + \tilde{f}_c^*(t)$$

where the non-migrant experience-wage profile $\tilde{f}_c^*(t)$ reflects both the non-migrant’s current time allocation decision and his accumulated human capital stock. Similarly, denote by $w_c(t;x^*)$ the wage of a migrant to the U.S. who immigrates with $x^*$ years of foreign experience. Given our assumptions, this wage can be written as

$$\log w_c(t;x^*) = \log \omega + \tilde{f}_c(t;x^*)$$

where the immigrant experience-wage profile $\tilde{f}_c(t;x^*)$ now reflects the immigrant’s current time allocation decision, his human capital stock upon arrival to the U.S. and the human capital he has accumulated since. It is also useful to define the following objects of interest. First, we define the return to foreign experience for non-migrants measured in their birth country as

$$R^*_c(x^*, y^*) = \log w^*_c(y^*) - \log w^*_c(x^*)$$

for any levels of foreign experience $y^* > x^*$. Second, we define the return to foreign experience

---

13 This assumption can be relaxed at the expense of some extra notation. Results are available upon request.

14 In particular, $\tilde{f}_c^*(t) = \log(1 - \ell^*_c(t)) + f^*_c(t)$ and $f_c^*(t) = \log h^*_c(t) = \int_0^t (B^*_c \phi (\ell^*_c(s)) - \delta) ds$.

15 $\tilde{f}_c(t;x^*) = \log(1 - \ell_c(t)) + \log h_c(x^*) + \int_0^t (B_c \phi (\ell_c(s)) - \delta) ds$. 

---
measured in the U.S. as
\[ R_c(x^*, y^*) = \log w_c(y^*; y^*) - \log w_c(x^*; x^*) \]
for any levels of foreign experience \( y^* > x^* \). Finally, define the return to U.S. experience of a migrant who immigrated with \( x^* \) years of foreign experience as
\[ \rho_c(x, y; x^*) = \log w_c(y; x^*) - \log w_c(x; x^*) \]
for any levels of U.S. experience \( y > x \) and foreign experience \( x^* \).

Using this notation, the facts we documented in Section 2 can be stated as follows.

**Empirical Fact 1** Returns to foreign experience between any foreign experience levels \( x^* \) and \( y^* \) measured in the U.S. immediately after arrival \( R_c(x^*, y^*) \) are lower the poorer is the birth country \( c \).

**Empirical Fact 2** Returns to U.S. experience \( \rho_c(x, y; x^*) \) between any U.S. experience levels \( x \) and \( y \) and for any level of foreign experience \( x^* \) are lower the poorer is the birth country \( c \).

Fact 1 summarizes what we documented in Figures 1 and 2: new immigrants from poor countries have lower returns to foreign experience. Fact 2 summarizes the empirical finding in Figure 3: returns to U.S. experience are lower for immigrants from poorer countries. In Sections B.1 to B.3, we now lay out three different versions of our Ben-Porath model and argue that all three can generate Fact 1 and two of them can generate Fact 2. In section 4 we then bring additional evidence to the table to distinguish between these three theories.

**B.1. Differential Human Capital Accumulation**

In this section we propose one particularly parsimonious theory that can explain Facts 1 and 2: different learning environments in different birth countries. In particular, in this section we make the following assumptions.

**Assumption 1** Both \( B_c \) and \( B^*_c \) are higher the richer is country \( c \): individuals born in richer countries have a higher ability to learn throughout their whole life.

**Assumption 2** \( B_c > B^*_c \) for all \( c \): the U.S. offers a better learning environment than all other countries (which are poorer).

Under these assumptions, we can show the following result (all proofs are in the Appendix C).

**Result 1** The Ben-Porath model with differential human capital accumulation across countries (Assumptions 1 and 2) generates Facts 1 and 2.
Different learning environments in different birth countries in a Ben-Porath model therefore provide a parsimonious theory for our empirical findings in Section 2.

**B.2. Differential Skill Loss**

In this section we present an extension of the model in section B that captures an alternative theory of Fact 1 (but not Fact 2): differential skill loss. We show that this theory is consistent with Fact 1 even under the assumption that there are no human capital difference across countries. Specifically we assume the following.

**Assumption 3** Ability to learn is identical in each country and for immigrants from each country: $B_c = B^*_c = B$, for all $c$.

To introduce the possibility of skill loss, we assume that an immigrant’s human capital he has accumulated through work experience in his birth country is less valuable in the United States. Moreover we assume that human capital accumulated in a poor country is less valuable than the one accumulated in a rich country. Specifically, we assume the following.

**Assumption 4** Immigrants lose a non-negative fraction of the logarithm of their human capital upon arrival in the United States: $h_c(x^*) = [h^*_c(x^*)]^\theta_c$, where $0 < \theta_c \leq 1$ for all $c$. Moreover if country $c$ is poorer than country $c'$, then $\theta_c < \theta_{c'}$.

**Result 2** The Ben-Porath model with differential skill loss across countries (Assumption 4) generates Fact 1 (but not Fact 2) even when there are no cross-country differences in human capital accumulation (Assumption 3).

Skill loss that differs across birth countries therefore provides a second theory for Fact 1 that the returns to foreign experience of new immigrants are lower for immigrants from poor birth countries. The intuition is simple: if individuals lose a large fraction of their human capital upon migration, an examination of the wages of new immigrants in the cross-section yields a flat experience-wage profile. It is, however, also easy to see that a theory of differential skill loss alone cannot generate Fact 2 that returns to U.S. experience are lower for immigrants from poor countries. This is because skill loss hurts immigrants only once exactly at the time they migrate, but not afterwards.

**B.3. Differential Selection**

We now show that there is another theory that can explain both Fact 1 and Fact 2 documented in section 2: selection that differs across birth countries. We continue to make Assumption 3 that learning ability is identical for immigrants from all countries. To allow for the possibility that immigrants are
selected we introduce heterogeneity in individual ability to learn. We assume that an individual $i$ in country $c$ accumulates human capital according to

$$ h_{ci} = z_{ci} B \phi (l_{ci}) h_{ci} - \delta h_{ci} $$

where $z_{ci}$ is the individual idiosyncratic ability to learn. Individual ability of individuals in country $c$ is distributed according to the CDF $F_c(z)$. Because the aim of this section is to argue that differential selection may explain our empirical findings even in the absence of any cross-country differences in learning environment or learning ability, we make the following assumption.

**Assumption 5** The learning ability distribution is the same in all countries, $F_c(z) = F(z)$ for all $c$.

What we mean by selection is that immigrants to the U.S. are not a random draw from $F$. Instead, we allow the distribution of immigrants which we denote by $G_c$ to differ from the distribution of non-migrants. Importantly, the distribution of migrant ability may also vary across countries. We say that immigrants from country $c$ are negatively selected if the distribution of non-migrants first order stochastically dominates the one of migrants, $F(z) \leq G_c(z)$ for all $z$, with strict inequality for some $z$. We make the following assumption.

**Assumption 6** Immigrants from poor countries are more negatively selected: if country $c$ is poorer than country $c'$, then $G_c$ first order stochastically dominates $G_c$.

The introduction of heterogeneity means that we need to adjust the definitions of the various returns measures in section B. Denoting the wage of an immigrant to the U.S. from country $c$ and with ability $z$ by $w_c(t; x^*, z)$ we adjust the definitions of the return to foreign experience measured in the U.S. $R_c(x^*, y^*)$ and the return to U.S. experience $\rho_c(x, y; x^*)$ as follows:

$$ R_c(x^*, y^*) \equiv \int \bar{R}_c(x^*, y^*, z) dG_c(z), \quad \bar{R}_c(x^*, y^*, z) \equiv \log w_c(y^*; y^*, z) - \log w_c(x^*; x^*, z) $$

$$ \rho_c(x, y; x^*) \equiv \int \bar{\rho}_c(z, x, y; x^*) dG_c(z), \quad \bar{\rho}_c(z, x, y; x^*) = \log w_c(y; x^*, z) - \log w_c(x; x^*, z) $$

Not surprisingly the introduction of heterogeneity means that these measures of returns to experience are now weighted averages of individual-level returns. With this adjusted definition in hand, we can show the following result.

**Result 3** The Ben-Porath model with differential selection on learning ability across countries (Assumption 6) generates Facts 1 and 2 even when there are no cross-country differences in human capital accumulation (Assumptions 3 and 5).
The intuition for this result is straightforward. Returns to experience for new immigrants are smaller because they are more negatively selected on the human capital they have accumulated up to their date of migration. Similarly, because immigrants from poorer countries have lower average $z$, they accumulate less human capital once they arrived in the U.S. and hence have lower returns to U.S. experience. Selection that differs across birth countries therefore provides a third theory for our empirical findings in Section 2. Each of these three possible forces operates in a Ben-Porath model of endogenous human capital investment.

C. Proofs

While in their birth country, individuals solve

$$\max \{\ell_c(t)\} \int_0^T e^{-rt} w^*_c(t) dt \quad \text{s.t.}$$

$$w^*_c(t) = \omega^*(1 - \ell^*_c(t)) h^*_c(t)$$

$$\dot{h}_c^*(t) = B^* \phi(\ell_c(t)) h_c^*(t) - \delta h_c^*(t)$$

$$0 \leq \ell^*_c(t) \leq 1$$

Note our assumption that individuals do not anticipate migrating to the United States so that they optimize assuming they will live in their country of origin over their entire time horizon $[0, T]$. If an individual with $x^*$ years of foreign experience migrates to the U.S., he thereafter solves

$$\max \{\ell_c(t)\} \int_{x^*}^T e^{-r(t-x^*)} w_c(t, x^*) dt \quad \text{s.t.}$$

$$w_c(t, x^*) = \omega (1 - \ell_c(t)) h_c(t)$$

$$\dot{h}_c(t) = B \phi(\ell_c(t)) h_c(t) - \delta h_c(t)$$

$$0 \leq \ell_c(t) \leq 1$$

where $h(x^*) = h^*(x^*)$, i.e. the human capital stock that he brought from his birth country.

We drop subscripts to ease notation, since we want to establish results for arbitrary $B$ and $B^*$. Without loss of generality, we normalize $\omega = \omega^* = 1$. We first prove a preliminary lemma. It is also useful to define the “instantaneous return to experience”

$$\frac{\dot{w}}{w} = \frac{\dot{h}}{h} - \frac{\dot{\ell}}{1 - \ell} = B \ell^r - \delta - \frac{\dot{\ell}}{1 - \ell} \quad (12)$$

**Lemma 1** The optimal time allocation $\ell(t)$, the optimal human capital stock and the instantaneous returns to experience $\dot{w}(t)/w(t)$ are all monotonically increasing in $B$ at each time $t$. Furthermore,
the optimal time allocation $\ell(t)$ is independent of an individual’s human capital stock.

A key implication of the Lemma is that, once an immigrant arrives in the U.S. his wage path is going to be defined uniquely from the number of years of birth country experience and from his productivity of human capital production function while in the U.S., $B$, and not from how much human capital he has accumulated in his birth country.

**C.1. Proof of Lemma 1**

The Hamiltonian of the model is given by

$$\mathcal{H} = h (1 - \ell) + \lambda (B\ell^\gamma - \delta) h$$

and the condition for optimality are

$$\mathcal{H}_\ell \leq 0, \quad \mathcal{H}_\ell (\ell - 1) = 0$$

$$\dot{\lambda} = r\lambda - \mathcal{H}_h$$

plus the transversality condition that the marginal value of human capital in the last period is equal to zero, $\lambda (T) = 0$. The solution is thus given by

$$h \leq \lambda \gamma B\ell^{\gamma-1} h, \quad [h - \lambda B\gamma\ell^{\gamma-1} h] (1 - \ell) \leq 0$$

$$\dot{\lambda} = (r + \delta - B\ell^\gamma) \lambda - (1 - \ell) \lambda$$

$$\lambda (T) = 0$$

and we notice that$^{16}$ $h$ cancels out in the optimality equation and so the solution is fully characterized by

$$\ell = \min \left\{ 1, \left( \gamma \lambda B \right)^{\frac{1}{\gamma-1}} \right\}$$

$$\dot{\lambda} = (r + \delta - B\ell^\gamma) \lambda - (1 - \ell) \lambda$$

$$\lambda (T) = 0$$

The optimal training time $\ell(t)$, as long as the time constraint $\ell(t) \leq 1$ does not bind, is the solution of the differential equation

$$\dot{\ell} = \left( \frac{1}{1 - \gamma} \right) \left[ (r + \delta) \ell - ((1 - \gamma) B) \ell^{\gamma+1} - \gamma B\ell^\gamma \right]$$

$^{16}$This is due to the assumption of constant returns in the human capital accumulation technology.
together with the terminal condition $\ell(T) = 0$. From (13), we have $\partial \dot{\ell}(t)/\partial B < 0$ for all $t$. Given the terminal condition $\ell(T) = 0$, therefore $\partial \ell(t)/\partial B$ for all $t$. Intuitively, the larger is $B$ the faster $\ell$ is going to decreases over time, and since we know that at time $T$, individual do not devote any time to training, then going backward it must be that the higher is $B$ the larger is $\ell$ at any point in time. That $h(t)$ and $\dot{w}(t)/w(t)$ are increasing in $B$ follows immediately from their definitions. Finally, the last part of the Lemma follows from the fact that the differential equation 13 that defines the path for $\ell(t)$ is independent of $h(t)$. □

C.2. Proof of Result 1

The goal is to compare $\log w_c(x^*_2;x^*_2) - \log w_c(x^*_1;x^*_1)$ for different levels of $x^*_2$ and $x^*_1$ where wage immediately after immigration satisfies

$$w_c(x^*,x^*) = h^*(x^*)(1 - \ell(x^*))$$

We have that

$$\frac{d \log w_c(x^*,x^*)}{dt} = \frac{\dot{h}^*_c(x^*)}{h^*_c(x^*)} - \frac{\dot{\ell}^*_c(x^*)}{1 - \ell^*_c(x^*)} = B^*_c \phi(\ell^*_c(x^*)) - \delta - \frac{\dot{\ell}^*_c(x^*)}{1 - \ell^*_c(x^*)}$$

This says that if an individual from country $c$ delays migration a little bit, his wage immediately after migration changes for two reasons: first he accumulates some more human capital before migration according to the technology in his country of origin $B^*_c$ and he may adjust his work hours after migration. The first term is increasing in $B^*_c$ from Lemma 1 and similarly the second term is increasing in $B_c$, also from Lemma 1. Since

$$\log w_c(x_2;x_2) - \log w_c(x_1;x_1) = \int_{x_1}^{x_2} \frac{d \log w_c(t,t)}{dt} dt$$

and from Assumption 1 richer countries have both higher $B_c$ and $B^*_c$, we obtain the desired result. The second part of the result follows directly from Lemma 1, Assumption 1 and that

$$\rho_c(x,y;x^*) = \log w_c(y;x^*) - \log w_c(x;x^*) = \int_x^y \frac{\dot{w}_c(t;x^*)}{w_c(t;x^*)} dt. □$$

C.3. Proof of Result 2

We have

$$R_c(x^*,y^*) = \theta_c \left[ \log h^*_c(y^*) - \log h^*_c(x^*) \right] + \left[ \log (1 - \ell^*_c(y^*)) - \log (1 - \ell^*_c(x^*)) \right]$$
Under assumption 3 both
\[ \log (1 - \ell_c (y^*)) - \log (1 - \ell_c (x^*)) \]
and
\[ \log h_c^* (y^*) - \log h_c^* (x^*) \]
are the same across countries. The result then follows directly from the fact that \( \theta_c \) is smaller in poor countries. □

C.4. Proof of Result 3

Let’s first consider \( R_c (z, x^*, y^*) \), which is given by
\[
R_c (z, x^*, y^*) = \log w_c (z, y^*; y^*) - \log w_c (z, y^*; y^*) \\
\quad \quad = [\log h^* (z, y^*) - \log h^* (z, x^*)] + [\log (1 - \ell (z, y^*)) - \log (1 - \ell (z, y^*))]
\]
where, from Lemma 1, we know that both the first and the second square brackets are increasing in \( z \).

Next, returns to foreign experience in country \( c \) are given by
\[
R_c (x^*, y^*) = \int [\log h^* (z, y^*) - \log h^* (z, x^*)] dG_c (z) + \int [\log (1 - \ell (z, y^*)) - \log (1 - \ell (z, y^*))] dG_c (z)
\]
and if country \( c \) is poorer than country \( c' \), then
\[
\int [\log h^* (z, y^*) - \log h^* (z, x^*)] dG_c (z) < \int [\log h^* (z, y^*) - \log h^* (z, x^*)] dG_{c'} (z)
\]
and similarly
\[
\int [\log (1 - \ell (z, y^*)) - \log (1 - \ell (z, y^*))] dG_c (z) < \int [\log (1 - \ell (z, y^*)) - \log (1 - \ell (z, y^*))] dG_{c'} (z)
\]
due to the fact that \( G_{c'} \) first order stochastic dominates \( G_c \).\(^{17}\) This concludes the proof of the first part.

Next, consider \( \rho_c (z, x, y; x^*) \), which is given by
\[
\rho_c (z, x, y; x^*) = \log w_c (z, y; x^*) - \log w_c (z, x; y^*) = \int_x^y \frac{w_c (z, t; x^*)}{w_c (z, t; x^*)} dt
\]
Lemma 1 implies that \( \rho_c (z, x, y; x^*) \) is increasing in \( z \). Then since
\[
\rho_c (x, y; x^*) = \int \rho_c (z, x, y; x^*) dG_c (z)
\]
\(^{17}\)This is due to the following result. Let’s consider two CDF, \( F (x) \) and \( G(x) \), where \( F (x) \) first order stochastic dominates \( G(x) \), and let’s consider a non-decreasing function \( h (x) \), then the following holds: \( \int h (x) dF > \int h (x) dG.\)
the fact that if country $c$ is poorer than country $c'$, then $G_{c'}$ first order stochastic dominates $G_c$, concludes the proof. $\square$
Figure 12: Secondary Migration Rate by GDP per capita