

NBER WORKING PAPER SERIES

LIFE-CYCLE HUMAN CAPITAL ACCUMULATION ACROSS COUNTRIES:
LESSONS FROM U.S. IMMIGRANTS

David Lagakos
Benjamin Moll
Tommaso Porzio
Nancy Qian
Todd Schoellman

Working Paper 21914
<http://www.nber.org/papers/w21914>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2016

Email: lagakos@ucsd.edu, moll@princeton.edu, tommaso.porzio@yale.edu, nancy.qian@yale.edu, todd.schoellman@gmail.com. For helpful comments we thank Rachel Friedberg, Gordon Hanson, Pete Klenow, Gary Becker, Bob Lucas, Christos Makridis and seminar participants at Houston, MIT, Ottawa, Konstanz, Mannheim, Oxford, and especially the conference honoring the life and work of Gary Becker at the Becker Friedman Institute in Chicago. We thank Caleb Johnson for excellent research assistance. All potential errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by David Lagakos, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Life-Cycle Human Capital Accumulation Across Countries: Lessons From U.S. Immigrants
David Lagakos, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman
NBER Working Paper No. 21914
January 2016
JEL No. E24,J61,O11,O15

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.

David Lagakos
Department of Economics, 0508
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093
and NBER
lagakos@ucsd.edu

Nancy Qian
Department of Economics
Yale University
27 Hillhouse Avenue
New Haven, CT 06520-8269
and NBER
nancy.qian@yale.edu

Benjamin Moll
Department of Economics
Princeton University
106 Fisher Hall
Princeton, NJ 08544
and NBER
moll@princeton.edu

Todd Schoellman
Department of Economics
W.P. Carey School of Business
Arizona State University
P.O. Box 873806
Tempe, AZ 85287
todd.schoellman@gmail.com

Tommaso Porzio
Yale University
28 Hillhouse Ave.
New Haven, CT 06520
tommaso.porzio@yale.edu

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 \(2005\)](#) finds 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

¹This 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 key 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, which says 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 from poor countries tend to lose a larger fraction of their skills after migrating. The third theory is differential human capital accumulation, which says that workers in poor countries accumulate less human capital. 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. Note that in this exercise we are using immigration as an opportunity to measure and account for cross-country differences in experience human capital. This exercise is fundamentally different from quantifying the development or growth implications of migration; see [Ehrlich and Kim \(2015\)](#) for work on this latter point.

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. Combining the data is straightforward because most questions and responses were maintained in the transition.

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 public 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.² 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 $\text{age} - \text{schooling} - 6$. A small subset of our sample reports very low levels of schooling. Following [Lagakos et al. \(2015\)](#), we define experience as $\text{age} - 18$ 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 have a largely but not entirely U.S education, defined as those arrive to the U.S. after age 12 but at least two years before their expected age of graduation. Table 1 shows the ten countries with the most immigrants in these two samples and the decomposition by age at arrival group. Even within 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

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 13.

year, usual hours worked per week, and weeks worked in the prior year.³ In 1980 income was top 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. natives, 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:

$$\log(w_{it}) = \alpha + \theta s_{it} + \sum_{x \in X} \phi_x D_{it}^x + \mu_t + \varepsilon_{it} \quad (1)$$

where D_{it}^x is a dummy variable that takes the value of one if a worker is in experience group $x \in X = \{5 - 9, 10 - 14, \dots\}$; 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 ϕ_x captures the average wage of workers in experience group x relative to workers with less than five years of experience. The coefficient θ captures the return to schooling and μ_t controls for time effects, since we have pooled multiple cross-sections. The regression coefficients (α, θ, ϕ_x) naturally differ across countries, but we

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 worked per category in 2007 and impose this on the 2008–2013 data.

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 have large samples of such new immigrants: the United Kingdom, Canada, 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 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 higher wages. 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

$$\log(w_{it}) = \alpha + \beta z_{it} + \theta s_{it} + \sum_{x \in X} \phi_x D_{it}^x + \mu_t + \sum_c \omega_{ic} D_{ic} + \varepsilon_{it} \quad (2)$$

where α is a country fixed-effect, z_{it} is a vector of controls for state, gender and English ability, θ is country-specific return to schooling, the ϕ_x are the country-specific returns to experience group x , and D_{ic} 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 coefficients ϕ_x capture the wages of an individual in experience group x relative to an immigrant from the same country with 0–4 years of experience.

In Figure 2 we plot our estimated returns to experience from equation (2) using one simple summary statistic: the returns to 20–24 years of foreign experience. We plot this statistic for each country

⁴We have also estimated equation (1) with immigrants that arrived within two years of a census. We find similar results, available upon request.

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 experience are higher for immigrants from richer countries than immigrants from poorer countries.⁵

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 country 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.⁶

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

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

⁵One 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 from 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.

⁶By 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.

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 $x \in X = \{5 - 9, 10 - 14, \dots\}$, 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 of foreign experience, and second, the returns to 20–24 years of 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. 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 related. 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. Here, we explore whether similar results apply for immigrants.

First, we repeat the standard analysis 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 higher 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 the 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 immigrants with U.S. education have higher returns to U.S. experience, which is shown as the level difference in the figure. This finding suggests a complementarity between location of schooling and returns to experience. The relationship between height of the profile and GDP per capita is again similar across the two groups; we cannot reject that the slopes are the same at even the 10 percent level. The fact that poor and rich country workers receive similar gains from U.S. education 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 results 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 life-cycle wage growth 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 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, defined on the basis of industry codes in IPUMS. Third, we show that the results are if anything stronger for immigrants from English-speaking countries or who self-report that they speak excellent English. Fourth, we show that the results are very similar if we exclude immigrants who live in ethnic enclaves, defined as a public use microdata area where more than five percent of the population is from the same birth country or a metropolitan statistical area where more than 2.5 percent of the population is from the same birth country; these restrictions exclude roughly one-third of the immigrant population. 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. education, 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 says 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 human capital η_{ic} and their learning ability z_{ic} . We assume that, on average, individuals are equally able in all countries $\mathbb{E}[\eta_{ic}] = \mathbb{E}[z_{ic}] = 1$ 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

$$\dot{h}_{ic}(t) = z_{ic}\phi_c(x_{ic}(t))h_{ic}(t), \quad (4)$$

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

$$\dot{h}_{ic}^*(t) = z_{ic}\phi_c^*(x_{ic}^*(t))h_{ic}^*(t). \quad (5)$$

In particular, we allow for the possibility that, upon arrival in the U.S. immigrants gain access to a “human capital accumulation function” ϕ_c^* that is different from the one in their birth country ϕ_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 $h_{ic}^* = m_c(h_{ic})h_{ic}$. 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

$$\mathbb{E}[\eta_{ic}|i \in M_c(x)] \neq \mathbb{E}[\eta_{ic}] = 1, \quad \mathbb{E}[z_{ic}|i \in M_c(x)] \neq \mathbb{E}[z_{ic}] = 1$$

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

$$w_{ic}^*(x, x^*) = \omega_c^* h_{ic}^*(x, x^*) e^{\varepsilon_{ic}} \quad (6)$$

where ω_c^* is the skill price earned by immigrants from country c in the U.S. and ε_{ic} is an error term. Given our assumptions, the immigrant’s human capital can be solved for in closed form and satisfies⁷

$$\log h_{ic}^*(x, x^*) = \log \gamma_c + \log \eta_{ic} + \theta_c z_{ic} \int_0^x \phi_c(y) dy + z_{ic} \int_x^{x+x^*} \phi_c^*(y) dy \quad (7)$$

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

$$\log w_{ic}^*(x, 0) = \log \omega_c^* + \log \gamma_c + \log \eta_{ic} + \theta_c z_{ic} f_c(x) + \varepsilon_{ic} \quad (8)$$

where we denote by $f_c(x) = \int_0^x \phi_c(y) dy$ the cumulative returns to foreign experience. The regression

⁷To 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).

we run using data on new immigrants only is therefore

$$\log w_{ic}^* = \alpha_c + R_c(x_{ic}) + \varepsilon_{ic}, \quad R_c(x) = \mathbb{E}[\log \eta_{ic} | i \in M_c(x)] + \theta_c \mathbb{E}[z_{ic} | i \in M_c(x)] f_c(x) \quad (9)$$

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. $\mathbb{E}[\log \eta_{ic} | i \in M_c(x)]$ decreases with x . Third, there may be selection on learning ability (both experience-dependent and standard selection are a problem, i.e. $\mathbb{E}[z_{ic} | i \in M_c(x)]$ 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 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 α_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 to compare the 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 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 strong positive relationship between the two sets of estimates; the correlation coefficient between the two estimates is 0.797 with a P-value of 0.004. Countries like Germany, the UK, and Australia have high returns 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 the returns to experience among non-migrants to be 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 the returns to experience among non-migrants to 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. Comparing Other Characteristics of Immigrants and Stayers

Previous work in the immigration literature has considered two additional factors that may affect returns to experience for immigrants: selection and skill loss. Our main concern is that selection or skill loss works differently for immigrants from poor and rich countries, and that this differential selection or skill loss explains why returns to experience vary with GDP per capita. To address each of these possibilities, we combine evidence from immigrants with data on non-migrants (stayers) from a large set of countries for which appropriate data are available. In particular, we use data on education and occupation from as many countries as possible from nationally representative surveys from IPUMS ([Minnesota Population Center, 2011](#)). This data source is ideal because the creators have devoted substantial effort to harmonizing variables across countries in a way that is also compatible with our data on immigrants. To further maximize this benefit, we use a much broader sample in this section, including any adults with valid responses to the pertinent variables.

We begin by addressing the hypothesis of differential selection. In short, this theory states that immigrants from rich countries are more positively selected (or less negatively selected) on ability to learn than immigrants from poor countries, where ability to learn is an individual trait that affects the

human capital generated per year of potential experience. To test this hypothesis, we consider the average years of schooling for immigrants and non-migrants. Our underlying assumption is that ability to learn will be positively correlated with duration of schooling, which allows us to make inferences about ability selection from data on school selection.

Figure 8 shows the results. The left-hand panel shows that average years of schooling among stayers is strongly correlated with log GDP per capita, with less than five years of schooling on average in the poorest countries and more than twelve years on average in the richest. In contrast, the right-hand panel shows that immigrants from countries of all income levels are highly educated on average, with the majority having roughly twelve years of schooling. These data do not support the differential selection hypothesis, because immigrants from poorer countries are actually much more positively selected on schooling attainment than are immigrants from richer countries.

These findings are consistent with 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 Hanson \(2011\)](#) 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. [Ehrlich and Kim \(2015\)](#) shows a similar result for migrants to a wide range of countries.

We now turn to the hypothesis of differential skill loss. Briefly, this theory says that immigrants from rich countries can transfer more of their experience human capital to the U.S. than can immigrants from poor countries. To test this hypothesis, we compare rates of skill loss for immigrants and non-migrants. We restrict the sample to workers and define them as experiencing skill loss if they have a college education (our notion of "skilled") but work in a low-skilled occupation.⁸ For each country, we calculate the fraction of all college-educated immigrants and stayers that work in high-skilled occupations.

Figure 9 plots the results. The left-hand panel shows that among stayers, a high fraction of college-educated workers are in high-skill jobs in countries of all income levels. The fraction is increasing in GDP per capita, meaning that college-educated workers in rich countries tend to work at high-skilled jobs with higher frequency. The right-hand panel shows that a large fraction of college-educated immigrants are employed at high-skilled jobs as well, and that the relationship is also increasing in the GDP per capita of the birth country. It is clear that immigrants are less likely to work at high skilled jobs than stayers in countries of any income level. This is consistent with the presence of

⁸IPUMS has standardized occupation codes across all our data sources. We define high skilled to be "professionals," "technicians and associate professionals," and "legislators, senior officials and managers," and 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.

skill loss. However, the slopes for immigrants and stayers appear similar, suggesting that skill loss is present to a similar degree in countries of all income levels. This is evidence against the possibility that our findings are explained by differential skill loss.

We also investigate a more subtle form of differential selection or skill loss that operates through an association with experience. The idea here is that immigrants with more experience may be selected differently than those with less experience, and that the difference in how selected they are may be correlated with GDP per capita; a similar story works for skill loss. To test this, we compared selection or skill loss (as defined above) between two discrete groups, those with low and high experience (defined as less than ten and ten or more years of experience). The results are shown in Figures 9 and 10. In each figure a comparison of the upper left and bottom left figures shows the unsurprising result that low and high experience stayers have similar patterns of educational attainment and skill loss. More importantly, a comparison of the upper right and bottom right figures shows that the extent of selection on education and skill loss is remarkably similar for less and more experienced immigrants. In particular, there is little evidence that the relationship between educational selection or skill loss and GDP per capita varies between less and more experienced workers. These checks, when combined with the fact that returns to experience patterns are strongly correlated with the same patterns for non-migrants, suggest that cross-country differences in human capital accumulation are the most plausible interpretation of the data.

4.3. 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 richer countries on a variety of outcomes, including education, occupation, wages, and earnings. This finding is consistent with our evidence above that the differences in the life-cycle wage 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 \(2005\)](#).

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

$$success_1 = \frac{var(\log Y_{KH,c})}{var(\log Y_c)}$$

where $Y_{KH,c} = K_c^\alpha H_c^{1-\alpha}$ is the component of output due to factors of production. Values of $success_1$ 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 $success_1$ is that measurement error in $Y_{KH,c}$ could increase $success_1$, 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. Instead of assuming that this profile is common across countries, we use estimated profiles similar to those from Section 2, but utilizing the broadest possible sample, including women, part-time workers, and public employees. Our logic is that development accounting results should reflect the full labor force, but as we showed in [Table 2](#), the estimated results are similar with or without these groups. Thus, we find similar development accounting results if we use instead the baseline results of Section 2.

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 and a decrease in time spent investing at work 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.⁹ 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, $H_c = h_c(\bar{s}_c, \bar{x}_c)$. As discussed above, in the case of active accumulation of human capital the return to foreign experience measured in the U.S. $R_c^*(x, 0)$ captures both the increase in human capital over the life cycle $f_c(x)$ and a term due to changes in the amount of time allocated towards human capital accumulation:

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

See Appendix B for details of the derivation.

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, which assumes that the investment time allocation, $\ell_c(x)$, is constant across experience levels for each country. This assumption allows us to measure human capital accumulation directly from the experience-wage profiles as $f_c(x) = R_c(x, 0)$. Given that most countries have roughly 17 years of experience on average, we use the estimated returns to 15–19 years of experience (Caselli, 2005). From the perspective of a passive investment model this is the correct measure of human capital because the investment time allocation is constant at 0. From the perspective of a Ben-Porath model this overstates the importance of experience human capital because time devoted to human capital investment is decreasing over the life cycle in all countries but more so in richer countries, implying that $\log \left(\frac{1 - \ell_c^*(x)}{1 - \ell_c(0)} \right)$ is positive and increasing in GDP per capita. By abstracting from this we have overstated cross-country human capital differences from the perspective of a Ben-Porath model.

We plot our estimated human capital stocks against GDP per capita in Figure 12. This figure plots the human capital stocks implied by our upper bound, 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.

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 0.48. When they are both considered, $success_1$ rises to 0.74, meaning that now almost three-fourths of income differences are accounted

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

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 or just experience, the slope coefficient from a regression of $\log(Y_{KH})$ on $\log(GDP)$ is 0.63 and 0.65. With both schooling and experience used to compute human capital stocks, the slope coefficient rises to 0.81. Thus, under this upper bound at least, the importance of human capital increases substantially when we include experience human capital estimated 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. The challenge to providing a lower bound is bounding the endogenous changes in life-cycle human capital investment. In Appendix B, we show that in a simple Ben-Porath model the time devoted to human capital investment for immigrants depends only on the remaining working life and the exogenous efficiency of their human capital accumulation in the U.S. ($T - t$ and B_c^* in the notation of Appendix B; see Lemma 1). We show there that a useful intermediate step is to study the difference between the returns to $\bar{x}_c = 15 - 19$ years of foreign experience and $\bar{x}_c^* = 15 - 19$ years of U.S. for immigrants from c . This is useful because both groups of immigrants face the same remaining working life and the same efficiency of human capital investment going forward; hence, they are predicted to invest the same fraction of their time in human capital accumulation. By taking the difference between the two we can focus on the difference in human capital stocks.

While the level of the difference in human capital stocks is not useful, its variance and correlation with GDP per capita are. To see why, note that returns to U.S. experience are weakly increasing in GDP per capita, which implies that human capital stocks are weakly increasing in GDP per capita. In the simple case where the U.S.-acquired human capital stock is constant, we are subtracting a constant from all countries. Hence, the variance and correlation would capture exactly the variance and correlation of foreign human capital stocks and the lower bound would be exact. If the U.S.-acquired human capital stock is strictly increasing in GDP per capita then we are biasing downward the variance and correlation, implying that we have found a lower bound on the importance of foreign-acquired human capital stocks.

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 0.66, up from 0.44 when only schooling is considered. The slope coefficient from a regression of $\log(Y_{KH})$ on $\log(GDP)$ is 0.76, up from 0.63 when only schooling is considered as human capital. It is also important to note that our bounding exercise produces a relatively tight range on the importance of human capital for development accounting: between 0.66 and 0.74 by the first criteria and between 0.76 and 0.81 by the second criteria. We conclude that the importance of human capital increases greatly when experience is included, regardless of whether life-cycle wage growth is driven by passive or active investment in human capital accumulation.

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 be 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 the quantity and type of schooling results in less “learning how to learn” among individuals who attend school in poor countries. We have found support for this hypothesis by documenting that the returns to U.S. experience among foreign-educated workers are lower than the returns to U.S. experience for natives.

At the same time, we have also documented that the returns to U.S. experience among U.S.-educated workers are very similar to those of natives. Combined, these two facts suggest a complementarity between both quantity and type of education and subsequent human capital accumulation that may be worth exploring further in the future.

References

- Bils, M. and P. J. Klenow (1998). Does schooling cause growth or the other way around? NBER Working Papers 6393, National Bureau of Economic Research.
- Bils, M. and P. J. Klenow (2000, December). Does schooling cause growth? *American Economic Review* 90(5), 1160–83.
- Borjas, G. J. (1985). Assimilation, changes in cohort quality and the earnings of immigrants. *Journal of Labor Economics* 3, 463–489.
- Caselli, F. (2005). Accounting for cross-country income differences. In P. Aghion and S. Durlauf. (Eds.), *Handbook of Economic Growth*, 679-741. Elsevier.
- Chiquiar, D. and G. H. Hanson (2005). International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of Political Economy* 113(2), 239–281.
- Chiswick, B. R. (1978). The effect of Americanization on the earnings of foreign-born men. *Journal of Political Economy* 86(5), 897–921.
- Coulombe, S., G. Grenier, and S. Nadeau (2014). Quality of work experience and economic development - estimates using Canadian data. *Journal of Human Capital* 8(3), 199–234.
- Ehrlich, I. and J. Kim (2015, November). Immigration, human capital formation and endogenous economic immigration, human capital formation and endogenous economic growth. NBER Working Paper 21699.
- Erosa, A., T. Koreshkova, and D. Restuccia (2010). How important is human capital? a quantitative theory assessment of world income inequality. *Review of Economic Studies* 77(4), 1421–49.
- Friedberg, R. M. (1992). The labor market assimilation of immigrants in the United States: The role of age at arrival. Unpublished Manuscript, Brown University.
- Grogger, J. and G. H. Hanson (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics* 95(1), 42–57.
- Hall, R. E. and C. I. Jones (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics* 114(1), 83–116.
- Hendricks, L. (2002). How important is human capital for development? evidence from immigrant earnings. *American Economic Review* 92(1), 198–219.
- Hendricks, L. and T. Schoellman (2015). Human capital and development accounting: New evidence from immigrant earnings. mimeo, University of North Carolina - Chapel Hill.

- Hsieh, C.-T. and P. J. Klenow (2010). Development accounting. *American Economic Journal: Macroeconomics* 2(1), 207–23.
- Klenow, P. J. and A. Rodríguez-Clare (1997). The neoclassical revival in growth economics: Has it gone too far? In B. S. Bernanke and J. Rotemberg (Eds.), *NBER Macroeconomics Annual 1997*. Cambridge: MIT Press.
- Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman (2015). Life-cycle wage growth across countries. Unpublished Manuscript, Princeton University.
- Lemieux, T. (2006). The “mincer equation” thirty years after *Schooling, Experience, and Earnings*. In S. Grossbard (Ed.), *Jacob Mincer: A Pioneer of Modern Labor Economics*, Chapter 11, pp. 127–145. Spring.
- Manuelli, R. E. and A. Seshadri (2015). Human capital and the wealth of nations. *American Economic Review* 104(9), 2736–2762.
- Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: Columbia University Press.
- Minnesota Population Center (2011). Integrated public use microdata series-international: Version 6.1. Minneapolis: University of Minnesota.
- Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development* 22(9), 1325–43.
- Reinhold, S. and K. Thom (2013). Migration experience and earnings in the mexican labor market. *Journal of Human Resources* 48(3), 768–820.
- Ruggles, S., K. Genadek, J. Grover, and M. Sobek (2015). *Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]*. Minneapolis: University of Minnesota.
- Schoellman, T. (2012). Education quality and development accounting. *Review of Economic Studies* 79(1), 388–417.

Table 1: Ten Largest Countries by Number of Immigrants

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

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 2: Returns to Experience and GDP per capita: Robustness of Results

	Schooling: Experience:	Foreign Foreign	Foreign U.S.	U.S. U.S.
Sample:				
Baseline		20.0***	5.61*	10.5
35–39 Years Experience		23.3**	7.2	11.0
Average Height of Profile		16.4***	5.5**	7.9
Discounted Average Height of Profile		8.4***	2.7**	3.9
Include Women		20.1***	3.6	2.3
Include Some Part-Time		20.5***	5.4*	10.8
Include All Part-Time		23.8***	7.8***	5.3
Include Public Sector		21.5***	6.3**	15.6**
Experience at 16		20.9***	3.3	1.5
≥College Graduate		38.6***	5.5	15.1
≤H.S. Graduate		13.8***	4.9	10.2
No Advanced Degrees		12.5***	4.8**	11.9**
Manufacturing		15.7***	5.4	N/A
Service Industry		23.2***	11.6***	16.0**
Excellent English		26.4***	6.7*	7.8
English-Speaking Country		14.4***	2.6	N/A
No Ethnic Enclaves		22.6***	6.3*	17.0**
Year 2000+		21.9***	6.4**	22.0***

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

Human Capital Measure:	Success ₁ (1)	Slope ($\log(Y_{KH}), \log(GDP)$) (2)
(a) Upper Bound		
Schooling	0.44	0.63
Experience	0.48	0.65
Schooling + Experience	0.74	0.81
(b) Lower Bound		
Schooling	0.44	0.63
Experience	0.42	0.60
Schooling + Experience	0.66	0.76

Note: Success₁ defined as the ratio of the variance of $\log(Y_{KH})$ to the variance of $\log(Y)$. Slope ($\log(Y_{KH}), \log(GDP)$) is defined as the slope coefficient from a regression of $\log(Y_{KH})$ on $\log(GDP)$ and a constant.

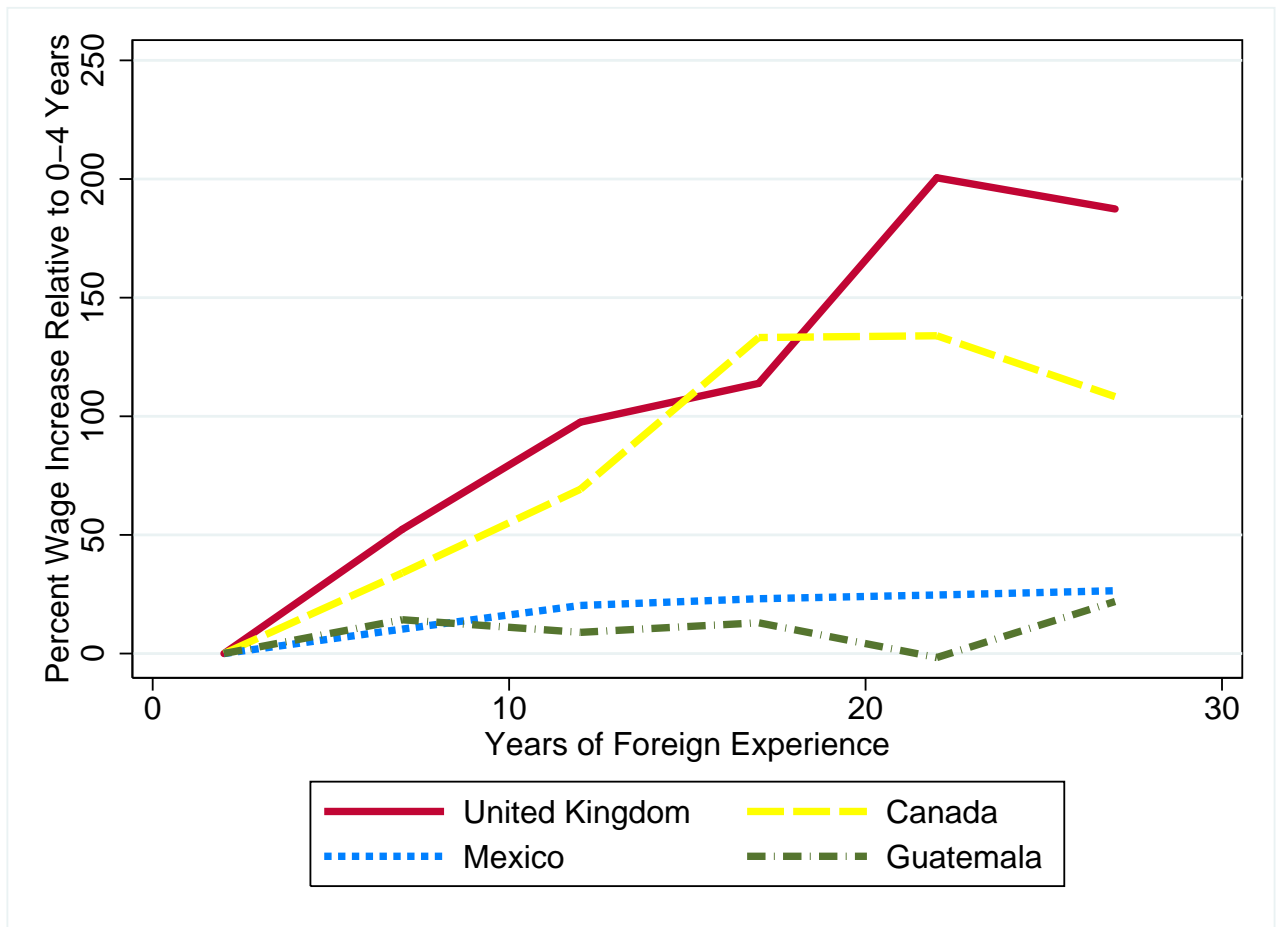


Figure 1: Returns to Foreign Experience Among New Immigrants, Select Countries

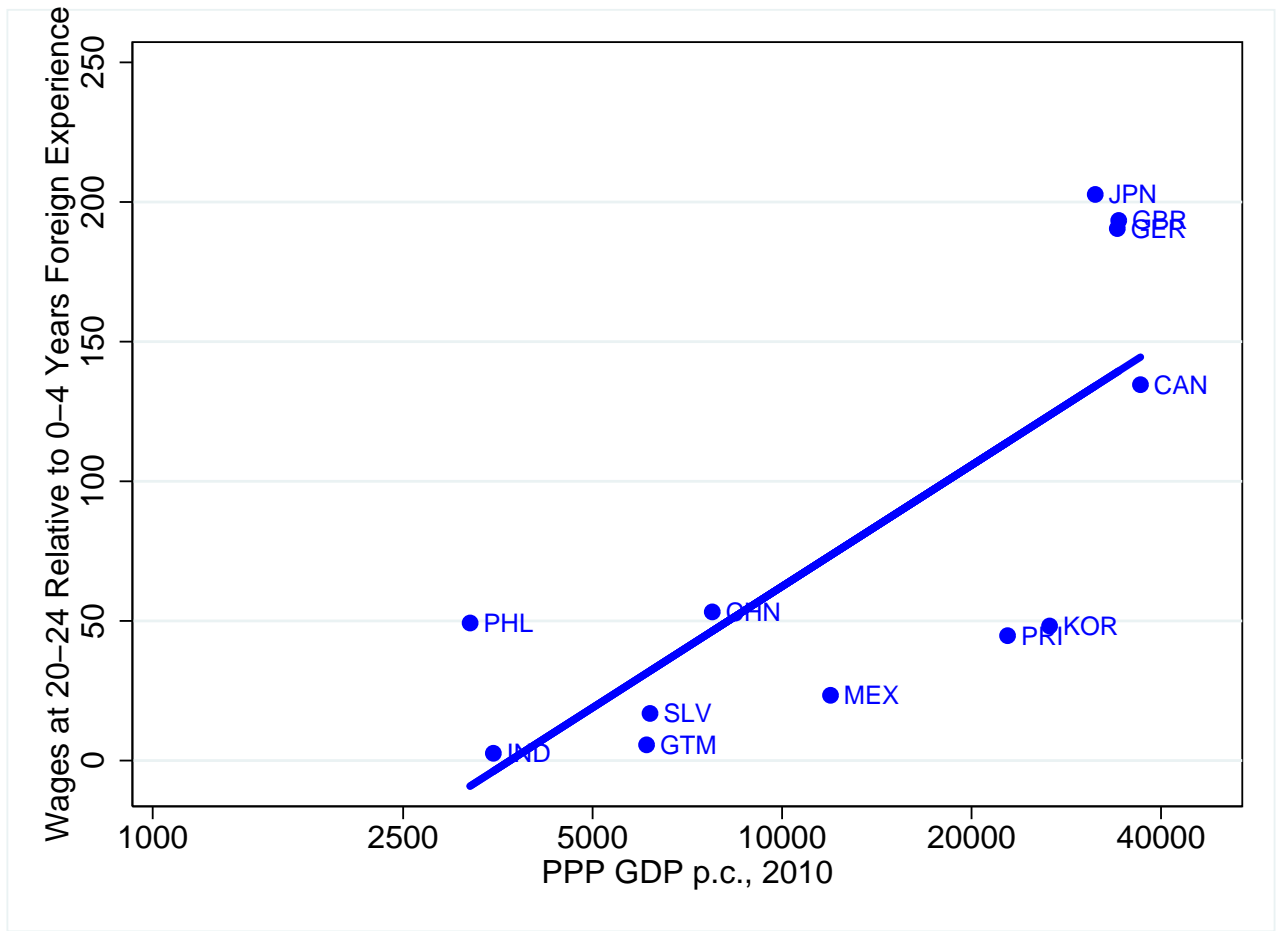


Figure 2: Returns to Foreign Experience Among New Immigrants by GDP p.c. of Birth Country

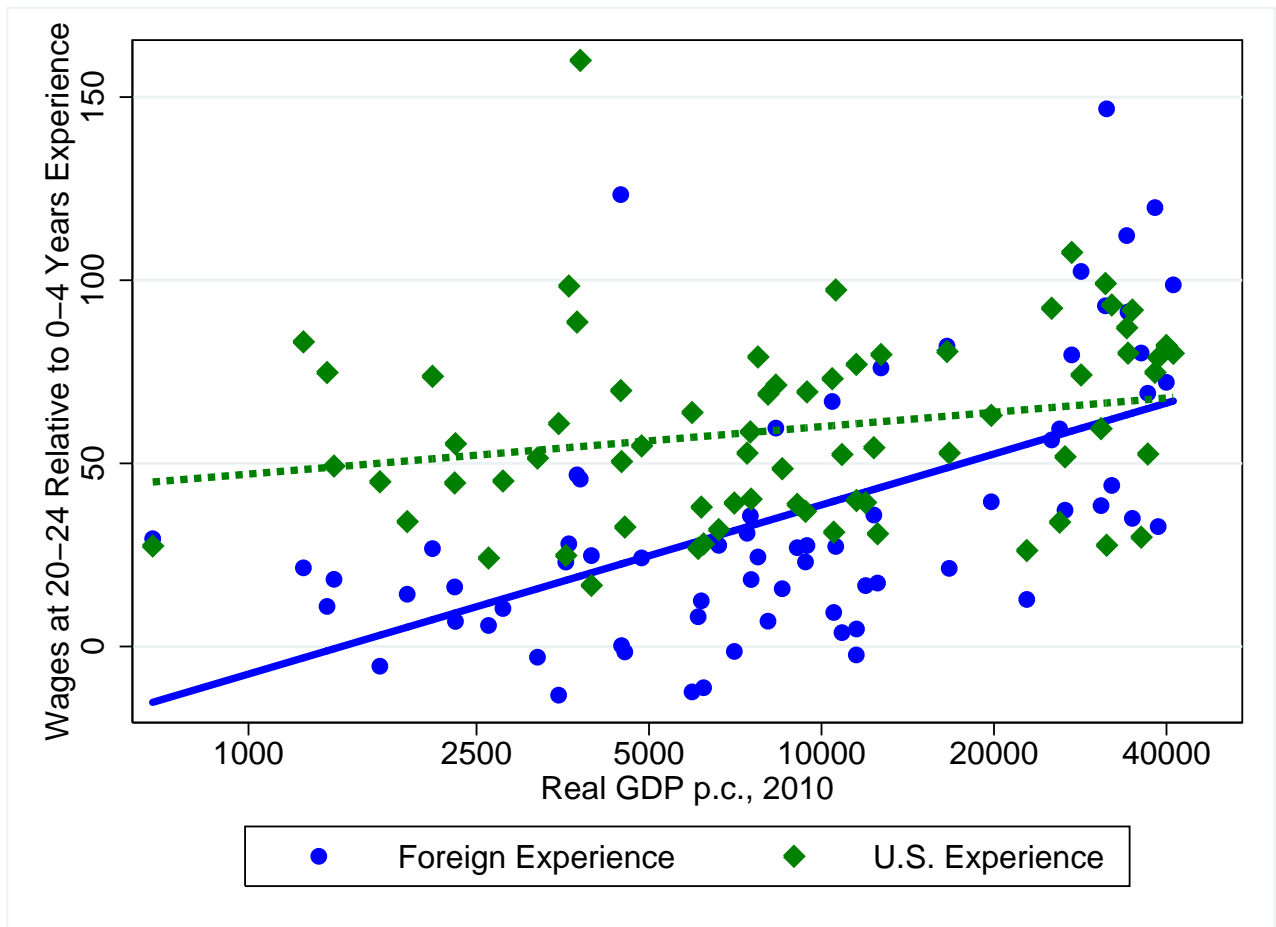


Figure 3: Returns to Foreign and U.S. Experience by GDP p.c. of Birth Country

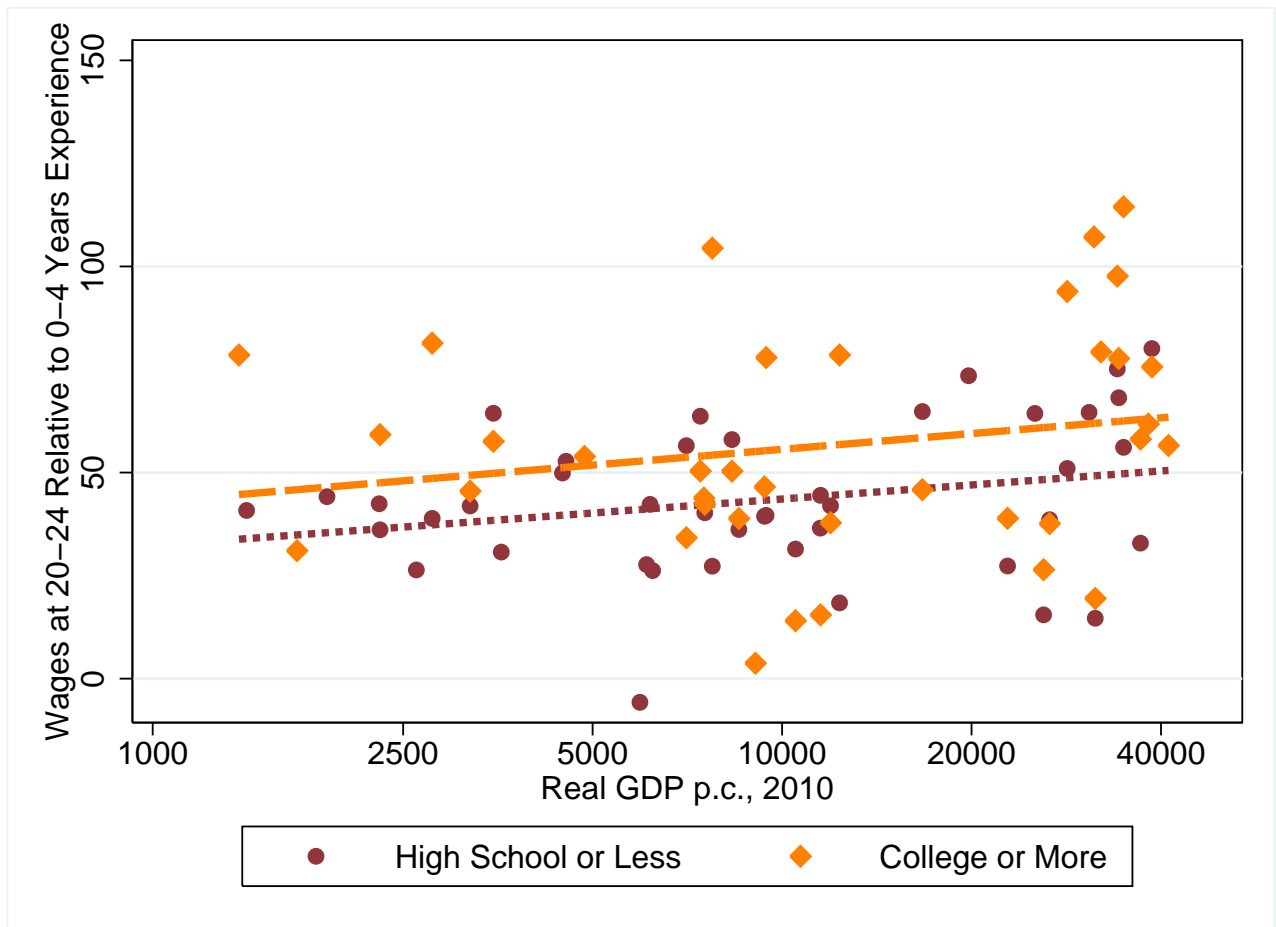


Figure 4: Returns to Experience by Education Level and GDP p.c. of Birth Country

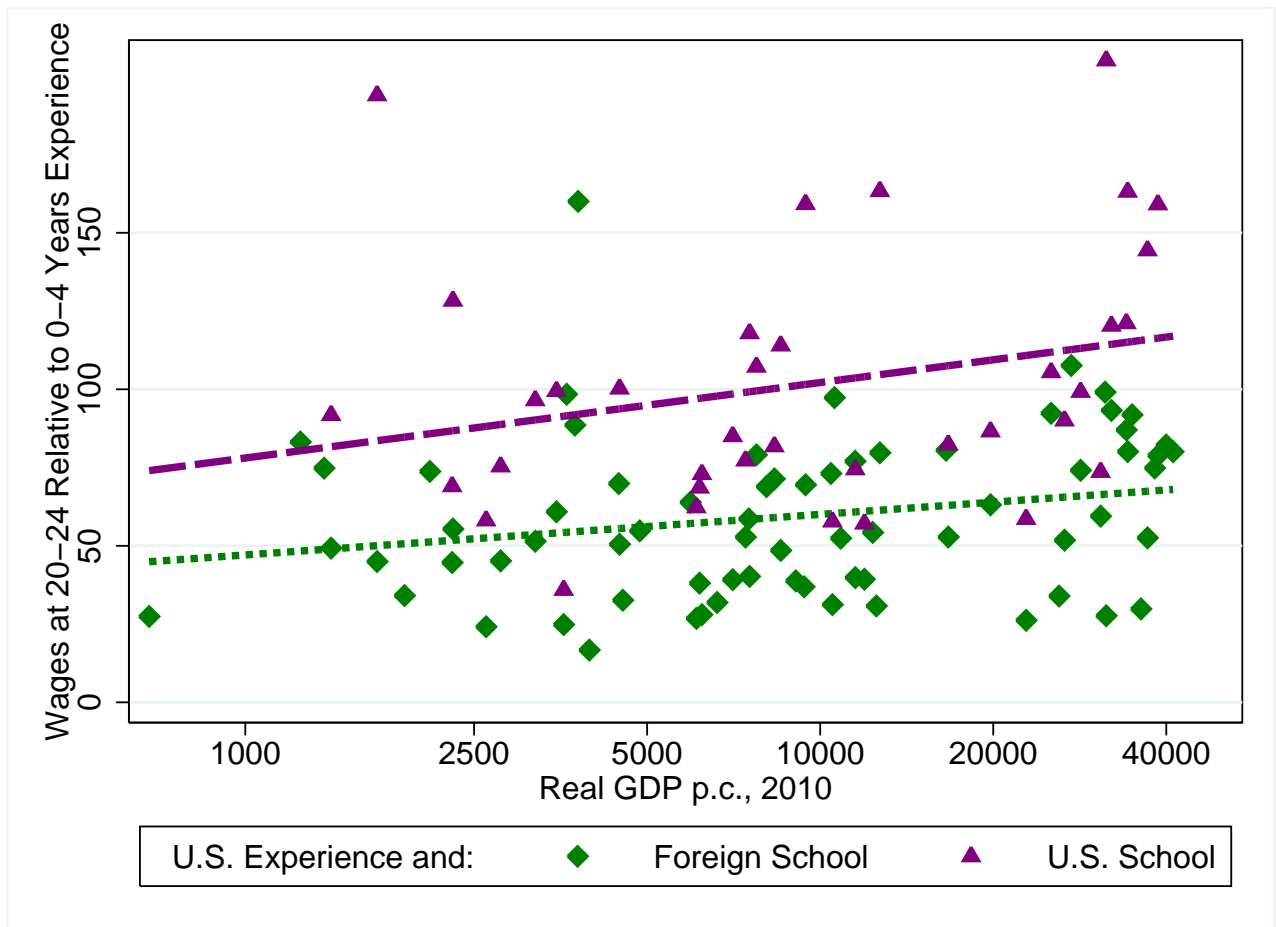


Figure 5: Returns to U.S. Experience by Education Location and GDP p.c. of Birth Country

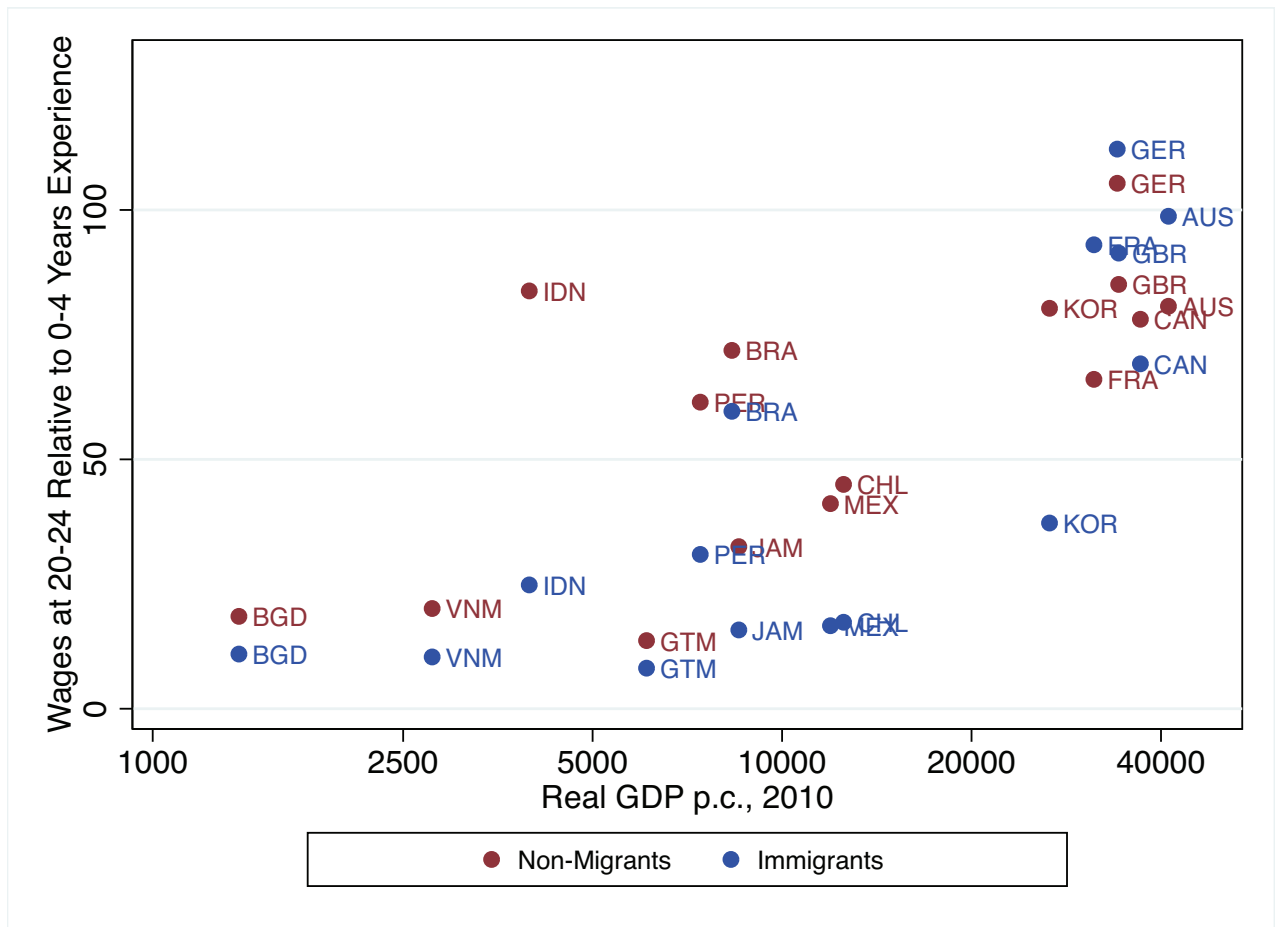


Figure 6: Returns to Foreign Experience of Immigrants and Non-Migrants by GDP p.c. of Birth Country

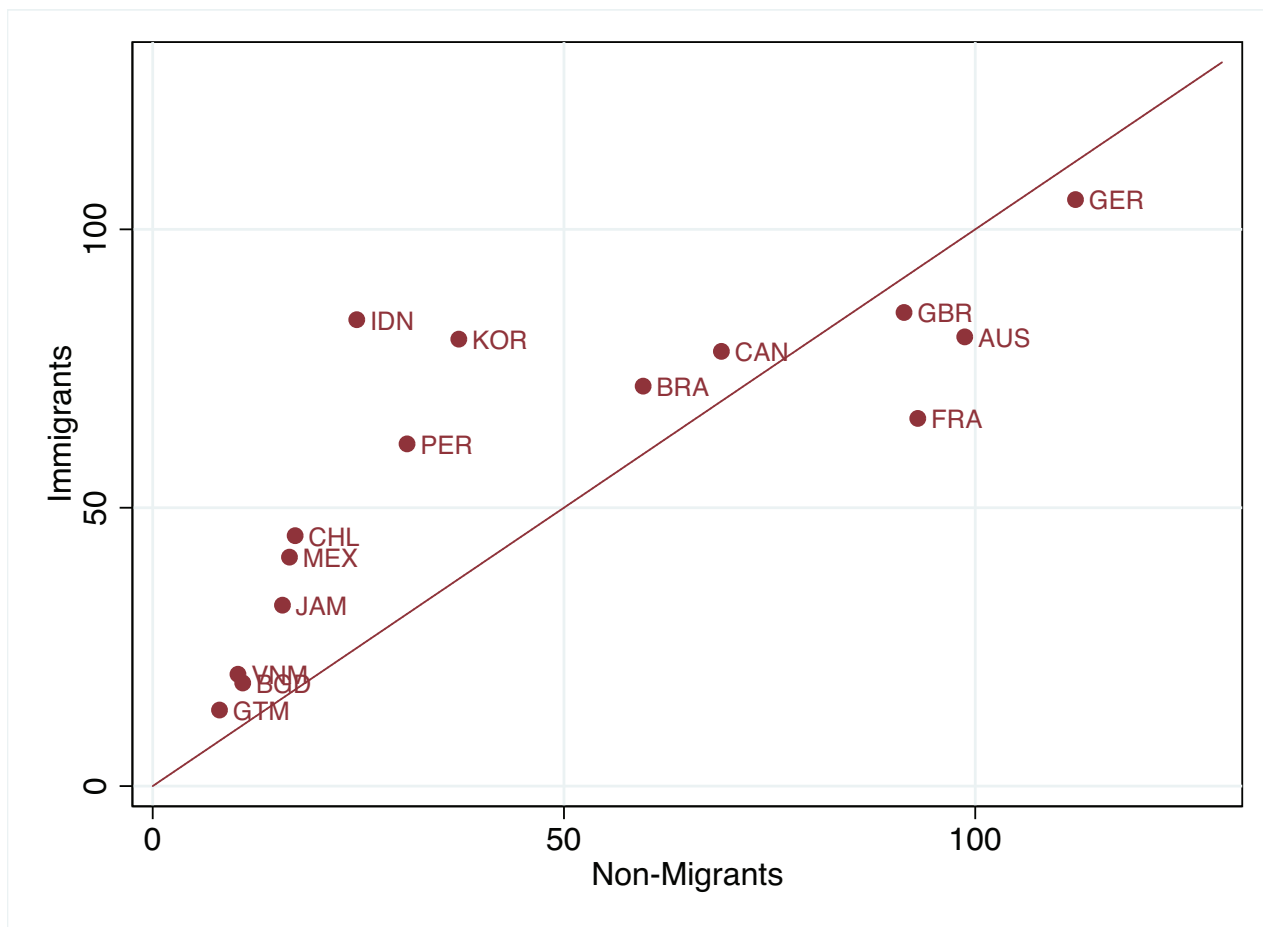


Figure 7: Returns to Foreign Experience, Immigrants vs Non-migrants

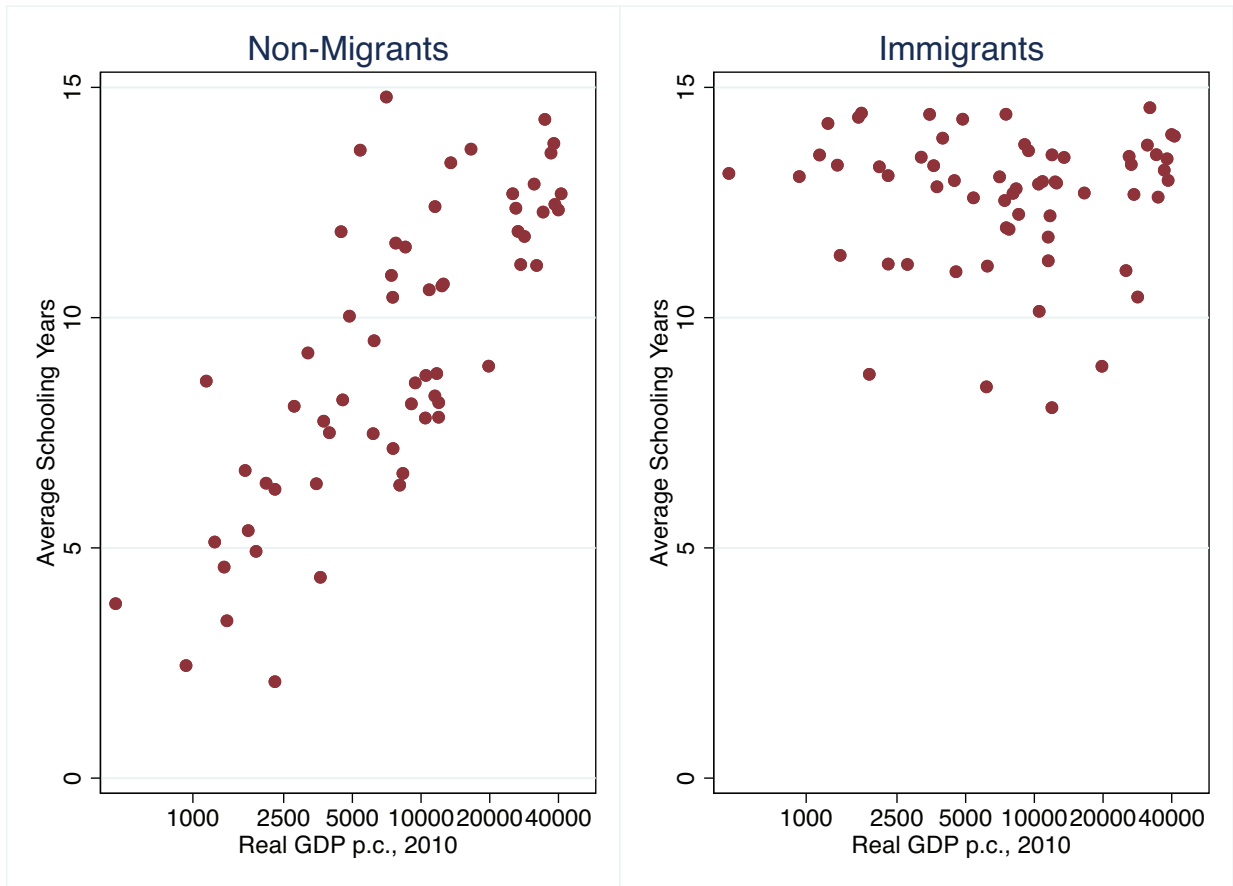


Figure 8: Years of Schooling Completed Among Immigrants and Non-migrants

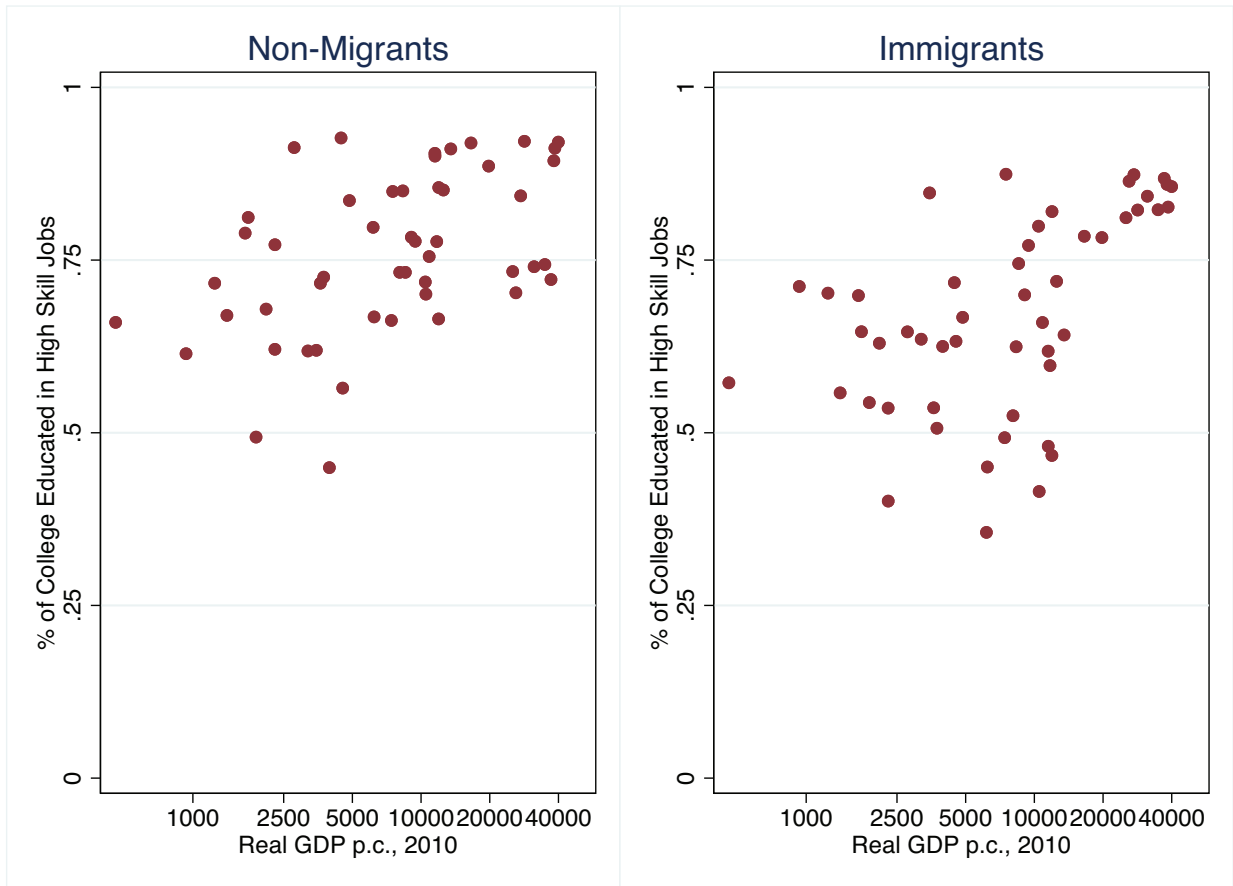


Figure 9: Percent of Educated Workers in High-Skilled Occupations, Immigrants and Non-migrants

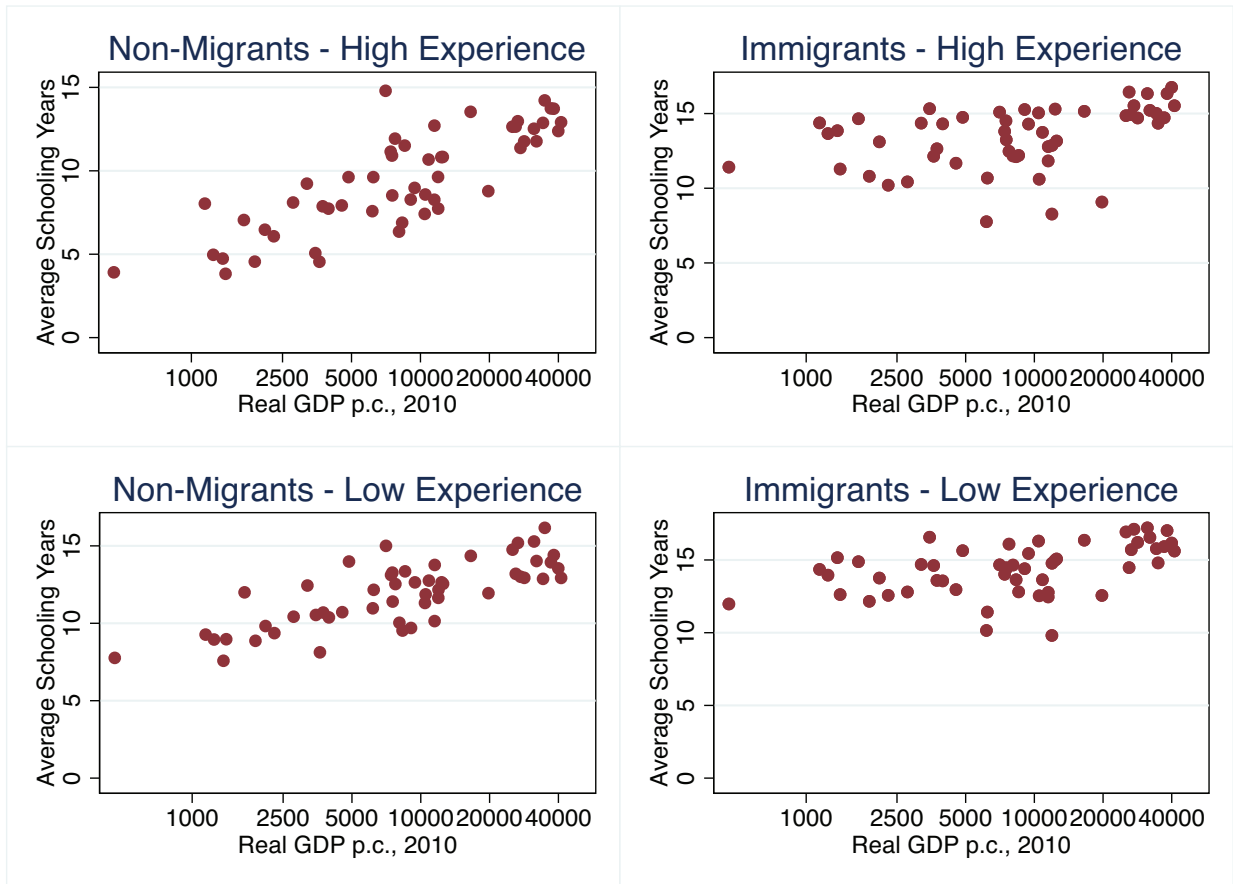


Figure 10: Years of Schooling Completed Among Immigrants and Non-migrants by Foreign Experience Level

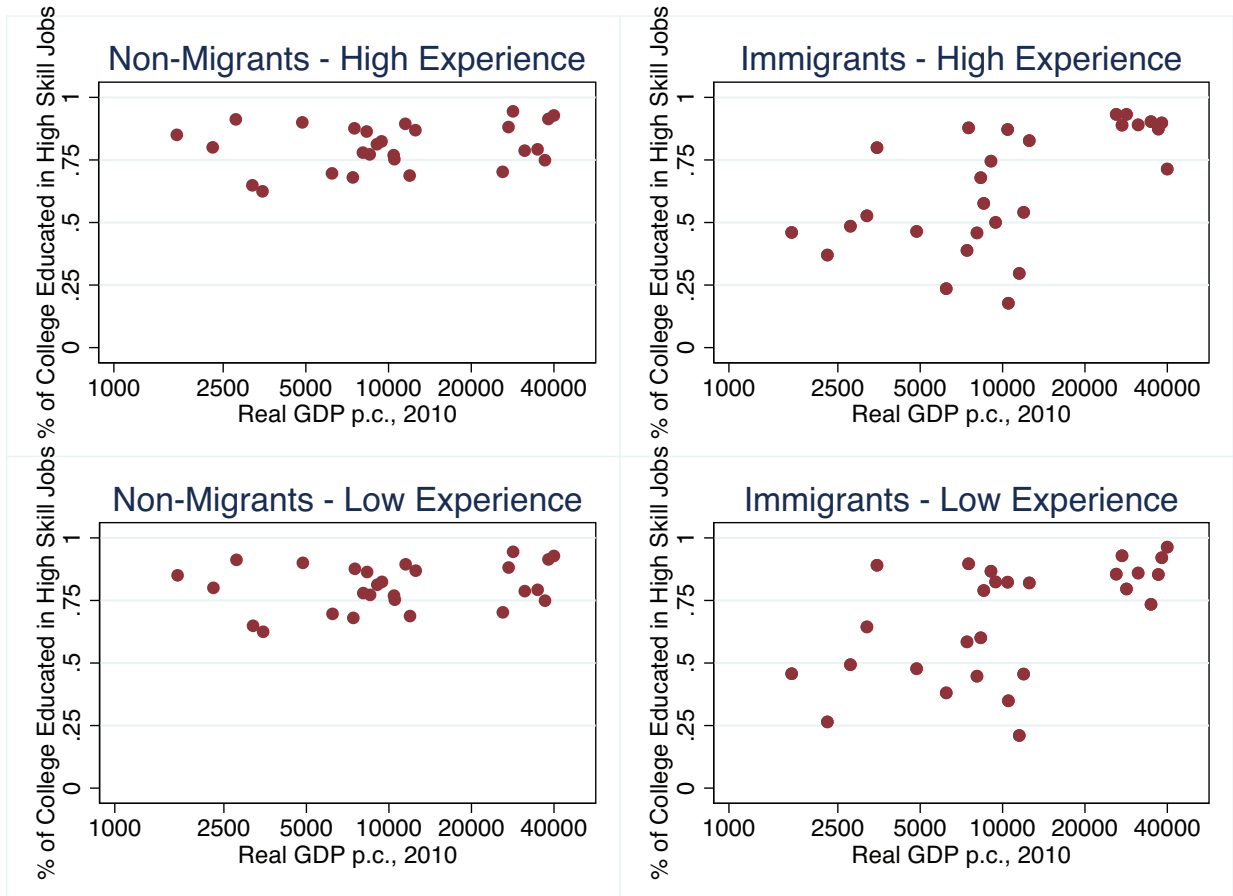


Figure 11: Percent of Educated Workers in High-Skilled Occupations, Immigrants and Non-migrants by Foreign Experience Level

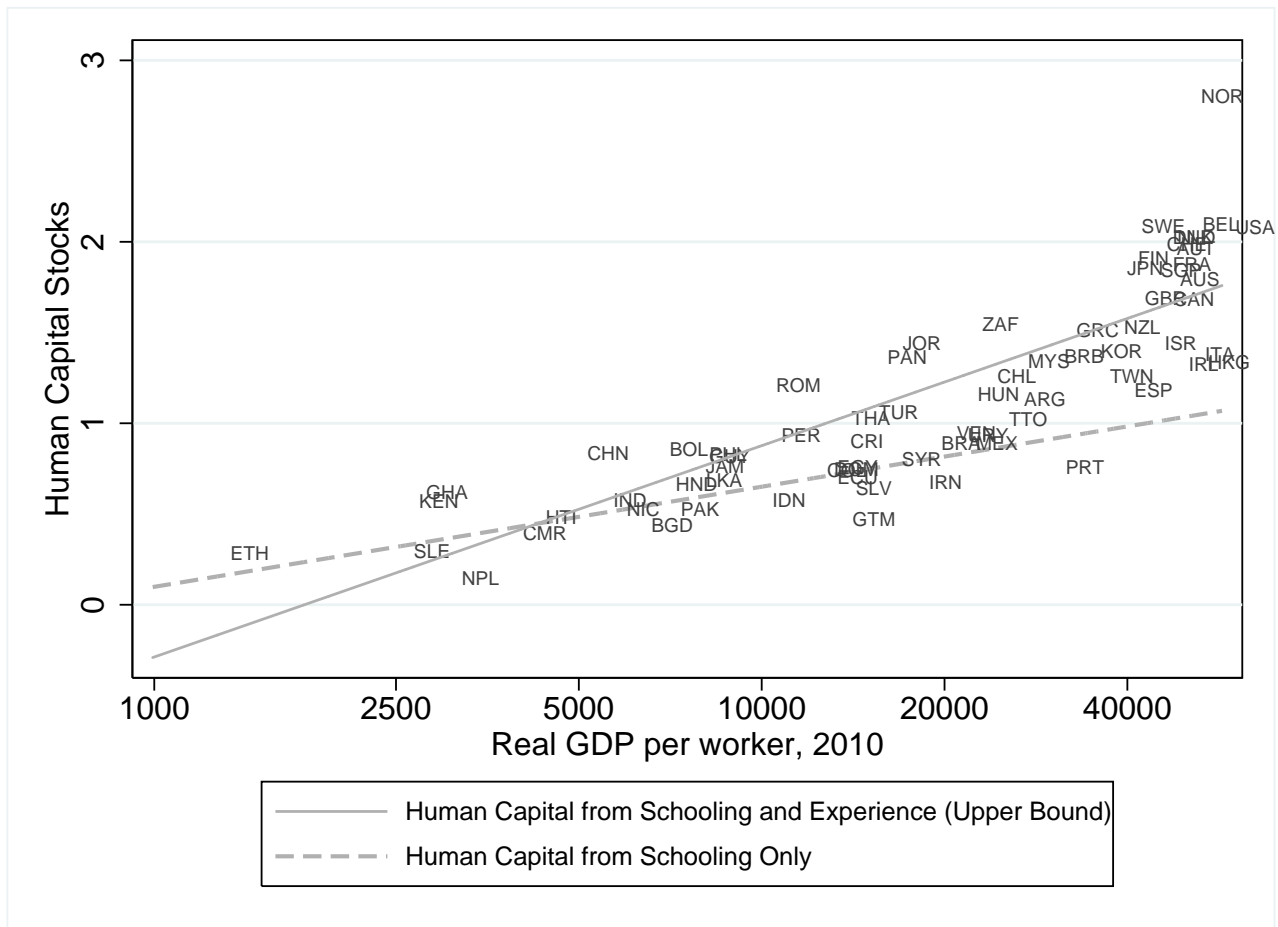


Figure 12: Upper Bound Human Capital Stocks

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. Given these assumptions the Mincerian regression equation for natives is then:

$$\log(w^N) = \beta^N X^N + \phi^N E^N + \omega^N Y^N + \mu^N S^N + \varepsilon^N \quad (10)$$

where Greek variables denote the coefficients and ε 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 Mincerian regression equation for immigrants:

$$\log(w^I) = \beta^I X^I + \phi_1^I FE^I + \phi_2^I DE^I + \omega^I Y^I + \mu^I S^I + \alpha^I AA^I + \gamma^I C^I + \delta^I YUS^I + \varepsilon^I \quad (11)$$

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:

$$1. FE^I = A^I - S^I - 6 - (Y^I - YI^I)$$

$$2. DE^I = Y^I - YI^I$$

$$3. AA^I = A^I - (Y^I - YI^I)$$

$$4. C^I = YI^I$$

$$5. YUS^I = Y^I - YI^I$$

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, we believe that our results support assigning a dominant role to experience human capital.

B. Ben-Porath Model of Human Capital Accumulation

In this Appendix we extend the simple learning-by-doing model of Section 3 to feature endogenous human capital accumulation. That is, we show that essentially the same insights about differences in experience-wage profiles for immigrants continue to go through in a Ben-Porath model of human capital accumulation. We also show how we derive the upper and lower bounds on human capital accumulation used in Section 5.

B.1. Setup

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 in the U.S. with asterisk superscripts, and those observed abroad 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

$$\dot{h}_c = \phi_c(\ell_c)h_c - \delta h_c,$$

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

$$\dot{h}_c = \phi_c^*(\ell_c)h_c - \delta h_c,$$

where $\delta \geq 0$ is the depreciation rate of human capital. We assume that $\phi_c(\ell) = B_c \ell^\sigma$ and $\phi_c^*(\ell) = B_c^* \ell^\sigma$, with $\sigma < 1$. 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 migration to the U.S.

The wage in an individual’s home country is $w_c(t) = \omega_c(1 - \ell_c(t))h_c(t)$ and analogously for the U.S. wage. At some level of home-country 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.¹⁰ Finally, we assume that when workers migrate, they take with them their entire human capital stock so that $h_c(x^*) = h_c^*(x^*)$.

As in the simple learning-by-doing model in Section 3, one can 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. Similarly, can also extend the model to feature “skill loss” upon migration. All the insights discussed there would go through.

¹⁰This assumption can be relaxed at the expense of some extra notation.

An individual who lives in his home country solves

$$\begin{aligned} \max_{\{\ell_c(t)\}} \int_0^T e^{-rt} w_c(t) dt \quad \text{s.t.} \\ w_c(t) = \omega(1 - \ell_c(t))h_c(t) \\ \dot{h}_c(t) = \phi_c(\ell_c(t))h_c(t) - \delta h_c(t) \\ 0 \leq \ell_c(t) \leq 1 \end{aligned} \quad (\text{P})$$

where the human capital at the beginning of the work life, $h_c(0)$, is given. 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 $[0, T]$. 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.

If an individual with x years of foreign experience migrates to the U.S., he thereafter solves

$$\begin{aligned} \max_{\{\ell_c^*(t)\}} \int_x^T e^{-r(t-x)} w_c^*(t, x) dt \quad \text{s.t.} \\ w_c^*(t, x) = \omega(1 - \ell_c^*(t))h_c^*(t) \\ \dot{h}_c^*(t) = \phi_c^*(\ell_c^*(t))h_c^*(t) - \delta h_c^*(t), \quad h_c^*(x) = h_c(x) \\ 0 \leq \ell_c^*(t) \leq 1 \end{aligned} \quad (\text{P}^*)$$

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 worker right before migration satisfies:

$$\log w_c(x) = \log \omega_c + \log(1 - \ell_c(x)) + \log h_c(x), \quad \log h_c(x) = \int_0^x (\phi_c(\ell_c(t)) - \delta) dt.$$

Hence running a regression using individuals observed in their home country would identify:

$$\log w_c(x) = \alpha_c + R_c(x), \quad R_c(x) = L_c(x) + \log h_c(x).$$

where $L_c(x) = \log[(1 - \ell_c(x))/(1 - \ell_c(0))]$ reflects the non-migrant's time allocation decision. Hence, *measured non-migrant returns to experience* $R_c(x)$ reflect both the non-migrant's current time allocation decision and his accumulated human capital stock.

Similarly, denote by $w_c^*(x, x^*)$ the wage of a migrant to the U.S. who immigrates with x years of home-country experience and accumulates x^* years of U.S. experience (and so has $x + x^*$ years of

total experience). Given our assumptions, this wage can be written as

$$\begin{aligned}\log w_c^*(x, x^*) &= \log \omega^* + \log(1 - \ell_c^*(x + x^*)) + \log h_c^*(x, x^*), \\ \log h_c^*(x, x^*) &= \log h_c(x) + \int_x^{x+x^*} (\phi_c^*(\ell_c^*(t)) - \delta) dt\end{aligned}$$

Hence, running a regression among the set of individuals with $x^* = 0$ years of U.S. experience identifies

$$\log w_c^*(x, 0) = \alpha_c^* + R_c^*(x, 0), \quad R_c^*(x, 0) = L_c^*(x) + \log h_c^*(x, 0) \quad (12)$$

where $L_c^*(x) = \log[(1 - \ell_c^*(x))/(1 - \ell_c^*(0))]$. Similarly to above, *measured immigrant returns to experience* $R_c^*(x)$ reflect the immigrant's human capital stock as well as his time allocation decision.

The following Lemma (proved at the end of this section) will be useful below.

Lemma 1 *An individual's optimal time allocation in the U.S. $\ell_c^*(t)$ satisfies*

$$\ell_c^*(t) = \min \left\{ 1, (\sigma \lambda_c^*(t) B_c^*)^{\frac{1}{1-\sigma}} \right\} \quad (13)$$

where $\lambda_c^*(t)$ satisfies the differential equation

$$\dot{\lambda}_c^*(t) = (r + \delta - B_c^* \ell_c^*(t)^\sigma) \lambda_c^*(t) - (1 - \ell_c^*(t)), \quad \lambda_c^*(T) = 0. \quad (14)$$

Therefore, an individual's optimal time allocation is independent of her human capital stock $h_c^*(t)$. Furthermore, the optimal time allocation $\ell_c^*(t)$, the optimal human capital stock and the instantaneous returns to experience $\dot{w}_c^*(t)/w_c^*(t)$ are all monotonically increasing in B_c^* at each time t .

A key implication of the Lemma is that, once an immigrant arrives in the U.S., her time allocation decision and wage path are defined uniquely from her productivity of human capital production while in the U.S., B_c^* , and from the number of years of birth country experience. In contrast, it does not matter how much human capital an immigrant has accumulated in his birth country before migration.

B.2. Upper and Lower Bounds for Development Accounting

The development accounting exercise in Section 5 requires estimates of a country's human capital stocks and correlates these with GDP per capita. As discussed there, we approximate a country's human capital stock as $H_c = h_c(\bar{x}_c, \bar{s}_c)$ where \bar{x}_c and \bar{s}_c are a country's average experience and schooling attainments.

A difficulty is that, through the lens of a Ben-Porath model, one cannot identify immigrants' human capital stocks simply from regression estimates of their returns to experience. This is because the

presence of the time allocation decision complicates the identification of human capital stocks. Ignoring schooling for the moment, an estimate of the immigrant returns to experience from the regression (12) does not allow one to identify an immigrant's human capital stock $h_c^*(x)$. This can be seen from the fact that $R_c^*(x, 0) = L_c^*(x) + \log h_c^*(x)$ where $L_c^*(x)$ is unknown.

We now show how one can make use of the data we observe together with the structure of our model to derive upper and lower *bounds* on a key object of interest, namely the slope coefficient from a regression of countries' human capital stocks on their GDP per capita (see Figure 12):

$$\beta = \frac{\text{Cov}(\log H_c, \log Y_c)}{\text{Var}(\log Y_c)}$$

Upper Bound. The derivation of the upper bound is straightforward. The idea is simply to count the entire immigrant returns to experience $R_c^*(\bar{x}_c, 0)$ as human capital $h_c^*(\bar{x}_c)$, i.e. to assume that the experience-wage profile is not affected by the time allocation changing over the life cycle. From (12), since time allocation is declining over the life cycle $\ell_c^*(x) \leq \ell_c(0), x > 0$ and hence $L_c^*(x) = (1 - \ell_c^*(x))/(1 - \ell_c(0)) \geq 0$, we have

$$\log H_c = \log h_c^*(\bar{x}_c) \leq R_c^*(\bar{x}_c, 0)$$

One can further show that $\text{Cov}(\log H_c, \log Y_c) \leq \text{Cov}(R_c^*(\bar{x}_c, 0), \log Y_c)$ and hence

$$\beta = \frac{\text{Cov}(\log H_c, \log Y_c)}{\text{Var}(\log Y_c)} \leq \frac{\text{Cov}(R_c^*(\bar{x}_c, 0), \log Y_c)}{\text{Var}(\log Y_c)}$$

Lower Bound. The derivation of the lower bound is somewhat more subtle. The key idea is to take two sets of individuals, both with the same amount of total experience $x + x^*$ equal to some number e . The first group are newly arrived immigrants from country c with $x = e$ years of home-country experience and $x^* = 0$ years of U.S. experience; second, immigrants from country c who arrived straight after high school and therefore have $x = 0$ years of home-country experience and $x^* = e$. The key insight is that it follows from Lemma 1 that the two individuals will choose *the same* time allocation.

Hence consider Mincerian regressions for these two groups

$$\log w_c^*(e, 0) = \alpha_c^* + R_c^*(e, 0), \quad R_c^*(e, 0) = L_c^*(e) + \log h_c^*(e, 0). \quad (15)$$

$$\log w_c^*(0, e) = \alpha_c^* + R_c^*(0, e), \quad R_c^*(0, e) = L_c^*(e) + \log h_c^*(0, e). \quad (16)$$

Here $R_c^*(e, 0)$ are the estimated returns to e years of experience of newly arrived immigrants, and $R_c^*(0, e)$ are the returns to e years of experience for immigrants who arrived straight after high school.

Importantly note that, due to Lemma 1, the term reflecting the time allocation $L_c^*(e)$ is identical for the two types of individuals.

Lemma 2 *We have that*

$$\beta = \frac{\text{Cov}(\log H_c, \log Y_c)}{\text{Var}(\log Y_c)} \geq \frac{\text{Cov}(R_c^*(\bar{x}_c, 0) - R_c^*(0, \bar{x}_c), \log Y_c)}{\text{Var}(\log Y_c)} \quad (17)$$

where $R_c^*(\bar{x}_c, 0)$ are the estimated returns to \bar{x}_c years of experience of newly arrived immigrants, and $R_c^*(0, \bar{x}_c)$ are the returns to \bar{x}_c years of experience for immigrants who arrived straight after high school.

Proof of Lemma 2 From the two regressions (15) and (16), we have $R_c^*(e, 0) - R_c^*(0, e) = \log h_c^*(e, 0) - \log h_c^*(0, e)$. Further by the assumption that there is no skill loss when migrating $h_c^*(e, 0) = h_c(e)$ and hence

$$R_c^*(e, 0) - R_c^*(0, e) = \log h_c(e) - \log h_c^*(0, e) \equiv \Delta \log h_c(e)$$

That is, the difference between returns to experience of newly arrived immigrants, $R_c^*(e, 0)$ and those of immigrants who arrived straight after high school, $R_c^*(0, e)$ identifies the difference between the human capital stock of individuals who stayed in their home country and those that migrated straight after high school.

Next, we explain how this can be used to obtain the lower bound (17). To fix ideas, consider first the case $B_c^* = B^*$ for all c , i.e. immigrants from all country learn at the same pace once they arrive in the U.S. In that case, the human capital stock of individuals who spent their entire working life in the U.S. $h_c^*(0, e)$ does not depend on their country of origin c . Therefore $\text{Cov}(\log h_c^*(0, e), \log Y_c) = 0$ for all e and hence $\text{Cov}(R_c^*(e, 0) - R_c^*(0, e), \log Y_c) = \text{Cov}(\Delta \log h_c(e), \log Y_c) = \text{Cov}(\log h_c(e), \log Y_c)$. Given that Caselli (2005) documents that \bar{x}_c is approximately the same for all countries, then if we assume $B_c^* = B^*$ for all c , (17) would not only be a lower bound on the regression coefficient but an exact expression.

More generally, we allow for the case $B_c^* \neq B^*$ and in particular that immigrants from richer countries face a higher B_c^* . In that case, Lemma 1 implies that immigrants from richer countries accumulate more human capital and hence $\text{Cov}(\log h_c^*(0, e), \log Y_c) \geq 0$. In that case, one can only obtain a lower bound:

$$\text{Cov}(R_c^*(e, 0) - R_c^*(0, e), \log Y_c) = \text{Cov}(\log h_c(e) - \log h_c^*(0, e), \log Y_c) \leq \text{Cov}(\log h_c(e), \log Y_c)$$

Using that $H_c = h_c(\bar{x}_c)$, we obtain (17). \square

B.3. Proof of Lemma 1

We drop subscripts to ease notation, since we want to establish results for arbitrary B_c and B_c^* . Without loss of generality, we normalize $\omega = \omega^* = 1$.

The Hamiltonian of the model is given by

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

and the condition for optimality are

$$\begin{aligned} \mathcal{H}_\ell &\leq 0, \quad \mathcal{H}_\ell(\ell - 1) = 0 \\ \dot{\lambda} &= r\lambda - \mathcal{H}_h \end{aligned}$$

plus the terminal 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

$$\begin{aligned} h &\leq \lambda \sigma B \ell^{\sigma-1} h, \quad [h - \lambda B \sigma \ell^{\sigma-1} h] (1 - \ell) \leq 0 \\ \dot{\lambda} &= (r + \delta - B \ell^\sigma) \lambda - (1 - \ell) \\ \lambda(T) &= 0 \end{aligned}$$

and we notice that¹¹ h cancels out in the optimality equation and so the solution is fully characterized by (13) and (14).

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

$$\dot{\ell} = \left(\frac{1}{1 - \sigma} \right) [(r + \delta) \ell - ((1 - \sigma) B) \ell^{\sigma+1} - \sigma B \ell^\sigma] \quad (18)$$

together with the terminal condition $\ell(T) = 0$. From (18), 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 ℓ is going to decrease 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 ℓ 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 18 that defines the path for $\ell(t)$ is independent of $h(t)$. \square

¹¹This is due to the assumption of constant returns in the human capital accumulation technology.

