

Lifecycle Human Capital Accumulation Across Countries: Lessons From U.S. Immigrants

David Lagakos, UCSD and NBER

Benjamin Moll, Princeton and NBER

Tommaso Porzio, Yale

Nancy Qian, Yale and NBER

Todd Schoellman, ASU

February 16, 2014

Abstract

Does lifecycle human capital accumulation vary across countries? If so, why? This paper seeks to answer these questions 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 Ben-Porath model of lifecycle 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 home countries before migrating. Our data suggest that lower quality schooling in poor countries may be the proximate cause of their workers' lower lifecycle human capital accumulation.

Email: lagakos@ucsd.edu, moll@princeton.edu, nancy.qian@yale.edu, tommaso.porzio@yale.edu, tschoell@asu.edu. We thank Rachel Friedberg, Gordon Hanson, Pete Klenow and seminar participants at Houston and Ottawa for helpful comments. We thank Caleb Johnson for excellent research assistance. All potential errors are our own.

1. Introduction

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

One reason the existing literature has found such a modest role for human capital is that it has focused largely on human capital arising through schooling. Several previous studies have included human capital arising over the lifecycle, 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

A recent literature has argued, in contrast, that workers in rich countries accumulate much more human capital over the lifecycle than their counterparts in poor countries. [Manuelli and Seshadri \(2010\)](#) 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, and Qian \(2013\)](#) 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 show that through the lens of a standard development accounting exercise, the implied human capital stocks from experience are substantially larger in rich countries.

A key challenge to inferring human capital stocks from returns to potential experience across countries, as do both [Lagakos, Moll, Porzio, and Qian \(2013\)](#) and [Bils and Klenow \(2000, 1998\)](#), is that wages may reflect other factors than human capital, such as search frictions, credit constraints, or other country-specific wage-setting institutions. Thus, lower measured returns to potential experience in poor countries than rich countries may reflect other factors besides lower human capital accumulation. Furthermore, wages are often measured differently in different surveys, and the fraction of people working for wages differs systematically with income per capita. Thus, measured returns to experience may be influenced by data quality or sample selection differences.

In this paper we turn to U.S. immigrants to help measure and understand differences in lifecycle human capital accumulation across countries. Studying immigrants offers several advantages. First,

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.

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, studying immigrants allows us to compare those that left to those that stayed behind, which allows us to draw insight as to *why* wages may evolve differently across countries. The insight of using immigrants to study human capital accumulation across countries is based on the work of [Hendricks \(2002\)](#) and [Schoellman \(2012\)](#), though the current paper is the first to measure and explain stocks of human capital from experience using U.S. immigrants.

We begin by documenting a new fact about immigrant returns to experience: returns to experience are lower among immigrants from poor countries than immigrants from rich countries. We find that this is true both for returns to *foreign experience*, acquired before migrating, and returns to *U.S. experience*, accruing 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 a semi-parametric mixture specification between foreign and U.S. experience. Both versions predict that returns to foreign experience are strongly increasing in GDP per capita of the source country. The second version predicts that returns to U.S. experience are increasing in GDP per capita of the source country, but not as sharply as for foreign experience.

To understand these facts we build a Ben-Porath style model of lifecycle human capital accumulation, similar to the ones studied by [Manuelli and Seshadri \(2010\)](#) and [Erosa, Koreshkova, and Restuccia \(2010\)](#). The model captures three basic theories of why returns to experience would be lower for immigrants from poorer countries. The first theory is differential selection, and states that immigrants from poor countries are less strongly selected on learning ability than their counterparts in rich countries. The second theory is differential skill loss, which says that immigrants tend to lose a larger fraction of their skills after migrating. The third theory is differential human-capital accumulation, which says that the efficiency of human capital accumulation is lower in poor countries than richer countries. All three theories are consistent with lower measured returns to foreign experience among immigrants from poor countries, and all three make different predictions along other dimensions.

To distinguish between theories we turn to new data we construct that compares immigrants to non-migrants in a large set of countries. The data contains the average years of school completed by immigrants and non-migrant, and the fraction of both groups working at “high-skilled” occupations, both of which are taken from national census data from around the world ([Minnesota Population Center, 2011](#)). The data also contains the returns to experience for immigrants and non-migrants, taken from the current study and [Lagakos, Moll, Porzio, and Qian \(2013\)](#), respectively.

The data on immigrants and non-migrants are most consistent with the theory that low lifecycle 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, in which non-migrant returns to experience should look different for migrants and non-migrants. Second, for the poor countries, years of schooling completed among immigrants tend to be much higher than years of schooling for non-migrants, whereas for the rich countries, years of schooling are similar in the two groups. This provides evidence against the theory that immigrants from poor countries tend to be systematically negatively selected on learning ability, since those negatively selected on learning ability would tend to complete less schooling than average, not more.² Finally, immigrants tend to work at high-skilled occupations at a lower frequency than non-migrants of similar educational attainment, though the frequency does not appear to be correlated with GDP per capita. This provides evidence against the theory that immigrants from poor countries lose disproportionately more skills average migrating.

We conclude by asking *why* lifecycle human capital accumulation tends to be lower in poor countries than rich countries. We argue that one proximate cause may be lower quality schooling in poor countries, which leads to less “learning how to learn” among individuals who attended schooling there. The main piece of evidence supporting this conclusion is that immigrants that arrived in the U.S. during schooling have returns to subsequent experience that look similar to those of natives. On the other hand, those that finished schooling in a poor country and then migrated, have lower returns to U.S. experience. Of course, immigrants that arrived in the U.S. for schooling may be selected differently than those that arrived afterwards. Thus this evidence should be taken as supportive rather than definitive.

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 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 2005–2012 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

²There are important exceptions to this fact, in particular Mexico, as we discuss in Section 4 below.

data from earlier Censuses because their sample size were smaller (1 percent instead of 5 percent) and immigrants were a much smaller share of the population before 1980. The 2000 Census was the last to include a long form with detailed questionnaires sent to a subset of the population; the ACS, an annual 1 percent sample of the American population, is the successor to the Census long form. Most questions and responses were maintained in the transition, so that combining the data is straightforward.

Our basic sample includes wage workers who are age 16 or older. We exclude those who are self-employed or unpaid workers because their recorded incomes are unreliable or difficult to interpret. 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. We aggregate the provided codes slightly in a few cases to produce countries more in line with standard international data (for example we aggregate the Azores into Portugal, or the subregions of the United Kingdom). We exclude any codes that correspond to “other” or “not elsewhere classified” groups and focus on well-defined countries. 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 – primarily immigrants – reports very low levels of schooling. Following [Lagakos, Moll, Porzio, and Qian \(2013\)](#), we define experience as $\text{age} - 14$ for anyone with less than eight years of schooling, under the assumption that no one acquires significant useful experience before age 14. Given this variable, we focus our attention on the subsample with between 0 and 45 years of experience, inclusive. For immigrants we split their experience into foreign (birth country) and domestic (U.S.) experience. We exclude workers with more than 35 years of foreign experience because we have little data for immigrants with so much foreign experience.

For immigrants, we also distinguish between three different age at arrival groups. Our baseline results are for immigrants who enter after their expected date of graduation in the U.S. However, we also present results for two other groups: those who arrive to the U.S. during their education; and those

who arrive to the U.S. before age 6, or before the start of their education. Table 3 shows the ten countries with the most immigrants in our sample in total and breaks down the totals by age at arrival category. The table shows that even for the top ten source countries there is a reasonable mixture of rich and poor birth countries, with the income per capita range from roughly 3,200 to 43,000 dollars in 2010 (PPP GDP p.c., PWT 7.1, Vietnam to Canada).

We construct the hourly wage using information on annual wage and salary income for the prior year, usual hours worked per week, and weeks worked in the prior year.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 three Census-provided controls in our analysis. The first is state of residence, which is designed to help capture the large cross-state differences in cost of living that would otherwise bias our results. The second is English-language ability. The Census has included a self-reported measure of English language ability throughout this time, with five options ranging from “Does not speak English” to “Yes, speaks only English.” Given that we study immigrants this is a useful control. We further parse the data by creating a sixth category for U.S. born persons, so that the remaining categories all capture variation within the immigrant population. The last control is a gender dummy.

2.2. New Immigrants

In this section illustrate the main spirit of our exercise in the simplest possible way by focusing on new immigrants, which we define as immigrants as those who arrived in the United States during one of our census years. What makes new immigrants interesting is that they all have less than one year of work experience in the United States, and yet the amount of foreign experience they bring with them to the U.S. labor market varies across individuals. Thus we can look, for each country, at how variation in foreign experience is related to variation in wages in the United States, to learn about the value of foreign experience from that country.

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). Letting w_{it} be the earnings of worker i in time period t , S_{it} be their years of schooling, X_{it} be their years of foreign potential experience, 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 2008–2012 data.

estimate for each country:

$$\log(w_{it}) = \alpha + \theta S_{it} + \sum_{\ell=1}^4 \phi_{\ell} X_{it} + \mu_t + \varepsilon_{it} \quad (1)$$

where μ_t is a time dummy and ε_{it} is an error term for individual i in t . We choose a quartic polynomial in potential experience following [Murphy and Welch \(1990\)](#), who show that a quartic fits the data far better than Mincer’s original quadratic. The regression coefficients ($\alpha, \theta, \phi_1, \phi_2, \phi_3$ and ϕ_4) naturally differ across countries, but we suppress country indices for simplicity.

For each country we focus only on *new immigrants*, or those that have been in the United States less than one year. For illustrative purposes we begin by presenting the results for four countries only: the United Kingdom, Germany, Jamaica and the Philippines, each of which has more than 1,000 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. Notably, returns to foreign experience are high in the United Kingdom and Germany and much more modest in Jamaica and the Philippines. Relative to a new immigrant with zero years of foreign potential experience (i.e. one that never worked in her home country), an immigrant from the United Kingdom or Germany with twenty years of foreign experience earns more than twice as much, all else equal. In the Philippines and Jamaica, in contrast, immigrants with twenty years of potential experience earn roughly 20 percent and 35 percent premiums. These findings suggest that returns to experience can vary dramatically across immigrants from different countries.⁴

2.2.2. Richer Specification

We now consider a richer specification that allows for cohort-of-immigration effects, following the work of [Borjas \(1985\)](#), to capture the idea that immigrants who enter in different years may be drawn from different parts of the income or talent distribution in their birth country. We also pool all countries for which we have at least 1,000 new immigrants, include the natives, and add controls for state of residence, gender and english-language ability. We now estimate

$$\log(w_{it}) = \alpha + \beta X_{it} + \theta S_{it} + \sum_{\ell=1}^4 \phi_{\ell} X_{it} + \mu_t + \sum_c \omega_{ic} D_{ic} + \varepsilon_{it} \quad (2)$$

where X_{it} is a vector of controls for state, gender and english ability, and D_{ic} is a dummy for decadal cohort of immigration. As before, each of the estimated coefficients is country specific, but we sup-

⁴We have also estimated specifications of (1) where the returns to experience are captured more flexibly using one dummy variable for each year of experience. We find similar results to the ones presented here, and hence omit them for brevity.

press country indices for simplicity.

In Figure 2 we plot our estimated returns to experience, using (2), using one simple summary statistic: the returns to twenty years foreign experience. We plot this statistic for each country against the country's GDP per capita in 2010. One can see that the returns to foreign experience vary positively with GDP per capita. The simple linear regression line (drawn in solid blue) has a slope of 0.176, and is significant at the 1% level. We find a similar positive and significant slope for other experience levels, such as 5 and 10 years foreign experience, but do not report these for brevity. We conclude that among new immigrants, returns to foreign experience are higher for immigrants from richer countries than immigrants from poorer countries.

2.3. Full Set of Immigrants

We now consider returns to experience using the entire sample of immigrants in our data. In general, these immigrants have potential experience that accrued in their source country and potential experience that accrued in the United States. Similarly, their schooling may have been completed in their source country, or in the United States, or some combination of the two.

To motivate our estimation, we first group immigrants by the fraction of their experience that accumulated in their home countries. We group people by the years of experience they have acquired in their home country before coming to the United States. Specifically, we group them into quintiles based on the percent of their potential experience that is foreign. Denote by D_{it}^q the dummy for being in quintile q for individual i in time t , where D_{it}^1 is the quintile with 0 to 20 percent foreign experience, D_{it}^2 is the quintile with 20 to 40 percent foreign experience, and so forth. We then estimate:

$$\log(w_{it}) = \alpha_i + \theta_i s_{it} + \sum_{q=1}^5 D_{it}^q \sum_{\ell=1}^4 \phi_{\ell}^q x_{it}^{\ell} + \mu_t + \varepsilon_{it}.$$

The resulting experience profiles (implied by the estimated ϕ_{ℓ}^q terms) are displayed in Figure 3. What the figure shows is that the experience profiles for those with intermediate experience levels lie between those with either mostly foreign experience or mostly U.S. experience. This is particularly apparent in Jamaica and the Philippines for example, where the gap between the 80-to-100-percent-foreign profile and the 0-to-20-percent-foreign profiles is particularly wide, and the intermediate profiles lie in between, ordered by percent foreign.

These findings motivate a parsimonious *mixture* specification, in which m_{it} is defined as the percent of an immigrant's experience that is foreign, given by

$$\log(w_{it}) = \alpha_i + \theta_i s_{it} + \sum_{\ell=1}^4 \left[\phi_{f,\ell} m_{it} x_{it}^{\ell} + \phi_{u,\ell} (1 - m_{it}) x_{it}^{\ell} \right] + \mu_t + \varepsilon_{it}.$$

This semi-parametric specific allows us to estimate the returns to both foreign and U.S. experience with just eight parameters for each country (in addition to the country-specific returns to schooling and other controls): four $\phi_{f,\ell}$ terms and four $\phi_{u,\ell}$ terms.

Figure 4 presents the results. For each country of origin we present two estimates: first, the returns to twenty years foreign experience, and second, the returns to twenty years U.S. experience. The blue dots in the figure represent the returns to foreign experience. As the figure shows, these tend to be lower in the countries with lower GDP per capita than in the countries with higher GDP per capita. The slope coefficient from a regression of the return to twenty years foreign experience on log GDP per capita is 0.163, 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 twenty years of U.S. experience on log GDP per capita is 0.067, and significant at the one percent level.

2.4. Robustness

We now explore the robustness of our stylized results given above. One way to do so is to alter our sample selection criteria, control variables, and so on, and to reproduce Figure 4 to see if the same basic patterns prevail. A more parsimonious way to achieve the same ends is to focus on the regression lines plotted in that figure. They capture the extent to which returns to experience vary with (the log of) GDP per capita in the birth country. Table 1 displays the coefficient on log GDP per capita from these regressions. Focusing on the row “Baseline” shows that there is a statistically significant and positive relationship for immigrants who get both their schooling and experience abroad, but that this relationship weakens and eventually becomes only marginally significant as immigrants acquire more of their experience and schooling in the U.S.

The remaining rows show the same results for alternative robustness checks. For example, the second row shows the relationship we would estimate if we limited our initial immigrant sample to only those with at least a college degree. For this subsample, we can see that the relationship is even stronger for foreign-educated immigrants, but that it is significantly weaker for the remaining groups. Our basic pattern also applies looking at only those with more than a high school degree or those without advanced degrees.

The next five checks focus on subsamples defined by where they work. The pattern is weakened for those who work for the government, but is essentially the same for manufacturing and service industry workers; there are too few immigrants who work in agriculture to estimate a meaningful relationship. Likewise, the same pattern holds if we focus on immigrants who work in occupations that are commonly licensed or not; see [Schoellman \(2012\)](#) for the details of how this is constructed. The pattern also prevails if we focus on those who speak only English or speak English well; on heads

of household, on men; or on those living outside of ethnic enclaves, defined as a public use micro data area with more than 5 percent of the population sharing the same birth country or a metropolitan area with more than 2.5 percent of the population sharing the same birth country; these restrictions eliminate around one-third of immigrants from our sample. Our results go through, indicating that ethnic enclaves do not drive our results. We restrict our attention to household heads and find very similar results. We use only the year 2000+ data, which may be interesting because for these years the year of immigration is coded exactly rather than in ranges; the results are essentially identical to the baseline. Finally, we try excluding some of the control variables and find that does not matter. We also explored using Canadian census data, but found that there were too few immigrants from too few countries, particularly poor countries, to complete our exercises.

3. Model of Immigrant Returns to Experience

In this section we present a Ben-Porath model of human capital accumulation that captures three different theories of the immigrant returns to experience presented above. The first is differential selection, and states that immigrants from poor countries are less selected on learning ability on average than immigrants from rich countries. The second 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. The third is differential human capital accumulation in source country, and says that immigrants from poor countries accumulate less human capital over the lifecycle than immigrants from rich countries.

3.1. Benchmark Ben-Porath Model

We model the human capital accumulation decision of an individual from country c who may work either in his country of origin, acquiring foreign experience or in the United States, acquiring U.S. experience. We denote variables observed abroad with asterisk superscripts, and those observed in the U.S. without superscripts. For instance, the wage of an individual from country c who works in his country of origin is $w_c^*(t)$ and if he works in the U.S. it is $w_c(t)$. Individuals devote a fraction $\ell_c(t)$ of their time to human capital accumulation. If they work in their home country their human capital accumulates according to

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

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

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

We assume that $\phi(\ell) = \ell^\gamma$, $\gamma < 1$ and that the depreciation rate $\delta \geq 0$. The parameters B_c and B_c^* determine how quickly human capital accumulates for a given amount of time devoted to human

capital accumulation. B_c^* may vary across countries and may be different from B_c , capturing the idea that countries differ in the quality of their “learning environment.” We also allow the “learning environment” in the U.S. B_c to vary across countries so that it matters where an individual is born, even after he migrated to the U.S. In Section 3.2 we extend the model to feature individual-specific heterogeneity in the parameters B_c and B_c^* so as to explore the issue of selection of migrants with different learning abilities.

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

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

where the human capital at the beginning of the work life, $h_c(0)$, is given. Workers abroad solve an analogous problem. At some level of foreign experience x^* , workers from country c migrate to the U.S.. In our benchmark model we assume that when workers migrate, they take with them their entire human capital stock so that $h_c(x^*) = h_c^*(x^*)$. In Section 3.2 we extend the model to allow for “skill loss” upon migration, that is $h_c(x^*) < h_c^*(x^*)$. Finally, in our benchmark exercise we focus on parameter constellations such that there is an interior solution for the time allocation decision, $0 < \ell(t) < 1$ for all $t < T$ so that in particular individuals earn a strictly positive wage.

We make the following assumptions so as to capture the facts documented in Section 2.

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

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

Assumption 3 *Individuals do not anticipate migration.*

For future reference, denote by $w_c(t; x^*)$ the wage of a migrant to the U.S. who immigrates with x^* years of foreign experience. It is also useful to define the following objects of interest. First, we define the *return to foreign experience for non-migrants measured in their home country* as

$$R_c^*(x^*, y^*) = \log w_c^*(y^*) - \log w_c^*(x^*)$$

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

$$R_c(x^*, y^*) = \log w_c(y^*; y^*) - \log w_c(x^*; x^*)$$

for any levels of foreign experience $y^* > x^*$. Finally, define the *return to U.S. experience* of a migrant who immigrated with x^* years of foreign experience as

$$\rho_c(x, y; x^*) = \log w_c(y; x^*) - \log w_c(x; x^*)$$

for any levels of U.S. experience $y > x$ and foreign experience x^* .

Given Assumptions 1 to 2, we can prove the following results (all proofs are in the Appendix).

Result 1: Non-migrants that are born and work in poor countries accumulate less human capital $h_c^*(x^*)$ for any foreign experience level x^* . They also have smaller returns to experience $R_c^*(x^*, y^*)$ between any foreign experience levels x^* and y^* , the poorer is the country of origin c .

Result 2: Returns to foreign experience between any foreign experience levels x^* and y^* *measured in the U.S.* immediately after arrival $R_c(x^*, y^*)$ are lower the poorer is the country of origin c .

Result 3: Returns to U.S. experience $\rho_c(x, y; x^*)$ between any U.S. experience levels x and y and for any level of foreign experience x^* are lower the poorer is the country of origin c .

Assumption 4 *The gap B_c/B_c^* is larger for poorer countries.*

Given Assumption 4, we can also obtain an additional result.

Result 4: The gap between returns to U.S. experience and returns to foreign experience measured in the U.S. $\rho_c(x, y; x^*) - R_c(x^*, y^*)$ is larger for poorer countries.

3.2. Selection and Skill Loss in the Ben-Porath Model

4. Distinguishing Between Theories

In this section we draw on new data that compares characteristics of immigrants and non-migrants from a large set of countries. We draw on three basic facts that help us distinguish between the theories above. First, returns to foreign potential experience among immigrants are similar to returns to potential experience among non-migrants. Second, immigrants tend to have more schooling than non-migrants, and particularly so for the poorest countries. 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 potential experience among immigrants to the returns among *non-migrants* estimated by [Lagakos, Moll, Porzio, and Qian \(2013\)](#). As the theory of Section 3 shows, returns to experience are informative about the human capital production functions in the countries of origin. We can make these comparisons in the 32 countries for which we have an estimate of immigrant returns, and for which [Lagakos, Moll, Porzio, and Qian \(2013\)](#) calculate returns using a representative sample of non-migrants. We begin by plotting the estimated returns to twenty years experience against GDP per capita.

Figure 5 plots these estimates. 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 0.263 for the immigrants, with a P-value less than 0.001. For non-migrants the slope coefficient is 0.211 and the P-value is also less than 0.001. One can also see that for certain countries the correspondence between the two sets of estimates is better than for others.

Figure 6 plots the estimated returns to twenty years experience for immigrants against the same estimated return for non-migrants. The 45-degree line is also plotted for reference. As one can see, there is a positive relationship between the two sets of estimates, though the relationship is far from completely linear. The correlation coefficient between the two estimates is 0.637 with a P-value of 0.001. Countries like Germany and the Netherlands are high both among immigrants and non-migrants, and most of the developing countries have low returns in both groups. Prominent outliers include Canada, with substantially higher returns for non-migrants than immigrants, and Guatemala, Nicaragua, Mexico, Egypt and Uruguay, with substantially lower returns among immigrants.

The fact that estimated returns to experience from poor countries are low both for immigrants and migrants provides one piece of evidence against differential selection as a theory of the immigrant evidence. If low returns to experience among immigrants were driven solely by negative selection by immigrants from poor countries, one would expect that returns to experience among non-migrants

were similar in countries of all income levels. As Figures 5 and 6 show, this is not the case. Of course, the fact that returns differ so much between immigrants and non-migrants for particular countries, such as Guatemala, Mexico and Nicaragua, suggest that selection is indeed driving some of the discrepancies between the two estimates we will return to this in the following section.

The broad similarity between returns to experience among immigrants and non-migrants is also evidence against differential skill loss as a theory of the immigrant returns. If low returns among immigrants from poor countries were solely due to skill loss, one would again expect that returns to experience among non-migrants would be similar in countries of all income levels. This prediction is not borne out in the Figures. Instead, the figures suggest a world where workers in poor countries do not acquire much human capital while in their home countries.

4.2. Years of Schooling Among Immigrants and Non-Migrants

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

Figure 7 shows the two data sets plotted against one another, with the 45-degree line for reference. As can be seen in the figure, the average schooling level of immigrants is higher than that of non-migrants in every country in the world (except for the U.S., which lies on the line by construction.) The schooling gaps are particularly large for the poorest countries, where immigrants average more than 12 years of schooling, and non-migrants average far less, in many cases less than 6 years.

The fact that schooling levels are so much higher among immigrants than non-migrants in most countries suggests that immigrants are, in general, positively selected on learning ability. The fact that the schooling gaps between immigrants and non-migrants are generally higher among poor countries than rich countries suggests that, if anything, immigrants from poor countries are more positively selected than immigrants from rich countries. This figure provides little support for the view that immigrants from poor countries are negatively selected on learning ability, thus explaining their low returns to experience both in their home countries and in the United States after migrating.

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

for the low estimated returns to experience among immigrants from Mexico and Guatemala, whose immigrant returns were substantially below returns for non-migrants.

We note that our conclusions in this section are consistent with findings of previous studies. [Chiquiar and Hanson \(2005\)](#) use census data from Mexico and the United States to argue that there is "intermediate" selection of immigrants from Mexico. Their key piece of evidence is that years of schooling attained are a bit higher among Mexican immigrants than Mexican non-migrants. [Grogger and 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.

4.3. Occupations of Educated Immigrants and Non-Migrants

In this section we compare the occupations of college-educated immigrants and non-migrants. The goal is to understand skill loss after immigrants, which we get at by comparing the fraction of immigrants that work at high-skilled occupations to the same fraction calculated for non-migrants. We focus on college educated individuals since presumably they are the ones who have the most skills to lose.⁵ As above, we ask specifically whether our proxies for skill loss are correlated with GDP per capita of the origin country.

We define occupations as either "high-skilled" or "low-skilled" using the international standard code of occupations constructed by IPUMS ([Minnesota Population Center, 2011](#)). We defined high skilled to be professionals, technicians and associate professionals, and legislators, senior officials and managers. We define low skilled to be clerks, service workers and shop and market sales, skilled agricultural and fishery workers, crafts and related trades workers, plant and machine operators and assemblers and elementary occupations. We omit individuals in the armed forces or other unspecified or unreported occupations.

Figure 8 shows that fraction of college-educated immigrants and non-migrants that work at high-skilled occupations. Not surprisingly, most countries lie below the 45-degree line, meaning that a larger fraction of educated non-migrants work at high-skilled occupations than educated immigrants. This confirms that skill loss is likely to be an important reality for immigrants. What matters for the current study, however, is whether skill loss tends to be high for immigrants from rich countries or immigrants from poor countries, on average.

Figure 9 plots the ratio of the high-skilled employment rate for immigrants to the the high-skilled employment rate for non-migrants against GDP per capita. In other words, we look at the ratio of the y-value to x-value of Figure 8 against GDP per capita. Most countries have ratios between 0.5 and 1.0, meaning that on average a substantially smaller fraction of immigrants work at high-

⁵When we focus on high-school graduates or higher we research similar conclusions as the ones presented below.

skilled occupations than non-migrants. However, this ratio seems to be largely uncorrelated with GDP per capita. This lack of correlation suggests that skill loss is not disproportionately a phenomenon pertaining to immigrants from poor countries. Instead, skill loss after migrating seems to be present in immigrants from countries of all income level, and on average the magnitudes appear similar across the income distribution. Thus, it seems unlikely that the facts on immigrant returns to experience we document are explained simply by skill loss that mostly affects immigrants from poor countries.

5. Conclusion

This paper seeks to understand whether workers in richer countries acquire more human capital over the lifecycle 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 human capital accumulation over the lifetime 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, 2010; Lagakos, Moll, Porzio, and Qian, 2013).

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.

References

- BARRO, R. J., AND J.-W. LEE (2012): “A New Data Set of Educational Attainment in the World, 1950-2010,” *Journal of Development Economics*.
- 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.
- (2000): “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 *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf., 679-741. Elsevier.
- CHIQUEAR, 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.
- 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 Immigrations 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.
- 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 *NBER Macroeconomics Annual 1997*, ed. by B. S. Bernanke, and J. Rotemberg. MIT Press, Cambridge.
- LAGAKOS, D., B. MOLL, T. PORZIO, AND N. QIAN (2013): “Experience Matters: Human Capital and Development Accounting,” Unpublished Manuscript, Princeton University.
- MANUELLI, R. E., AND A. SESHADRI (2010): “Human Capital and the Wealth of Nations,” Unpublished Manuscript, Washington University in St. Louis.
- MINCER, J. (1974): *Schooling, Experience and Earnings*. Columbia University Press, New York.
- MINNESOTA POPULATION CENTER (2011): *Integrated Public Use Microdata Series, International: Version 6.1 [Machine-readable database]*. University of Minnesota.
- MURPHY, K. M., AND F. WELCH (1990): “Empirical Age-Earnings Profiles,” *Journal of Labor Economics*, 8(2), 202–229.
- PSACHAROPOULOS, G. (1994): “Returns to Investment in Education: A Global Update,” *World Development*, 22(9), 1325–43.
- RUGGLES, S. J., T. ALEXANDER, K. GENADEK, R. GOEKEN, M. B. SHROEDER, AND M. SOBEK (2010): “Integrated Public Use Microdata Series: Version 5.0 [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: Returns to Experience and GDP per capita: Robustness of Results

Sample	Experience, School Location			
	For, For	US, For	US, Mix	US, US
Baseline	0.163***	0.067***	0.029**	0.036*
≥College Graduate	0.255***	0.055*	0.029	-0.004
≤H.S. Graduate	0.157***	0.074***	0.013	0.068***
No Advanced Degrees	0.176***	0.073***	0.021	0.031
Public Employee	0.092	-0.015	0.006	X
Manufacturing	0.247***	0.109**	0.009	X
Service Industry	0.181***	0.075***	0.041**	0.021
Licensed Occupation	0.135**	0.000	-0.001	X
Unlicensed Occupation	0.237***	0.086**	0.058***	X
Excellent English	0.217***	0.094***	0.018	0.025
Head of Household	0.248***	0.093***	0.036	-0.032
Men	0.211***	0.088***	0.056***	0.050**
No Ethnic Enclaves	0.183***	0.074***	0.023	0.006
Year 2000+	0.184***	0.078***	0.029*	0.016
No State, Gender Controls	0.165***	0.068***	0.030**	0.032

Note: Each estimate in the table represents the slope coefficient from a regression of the estimated returns to twenty years of potential experience on GDP per capita. The column For, For is the return to foreign experience for those immigrants who got their schooling in their home country. The column US, For is the return U.S. experience for those getting their schooling in their home country. The column US, Mix is the return to US experience for those getting some schooling abroad and then some schooling in the United States. The column US, US is the returns to US experience for those immigrating and then getting their schooling in the United States.

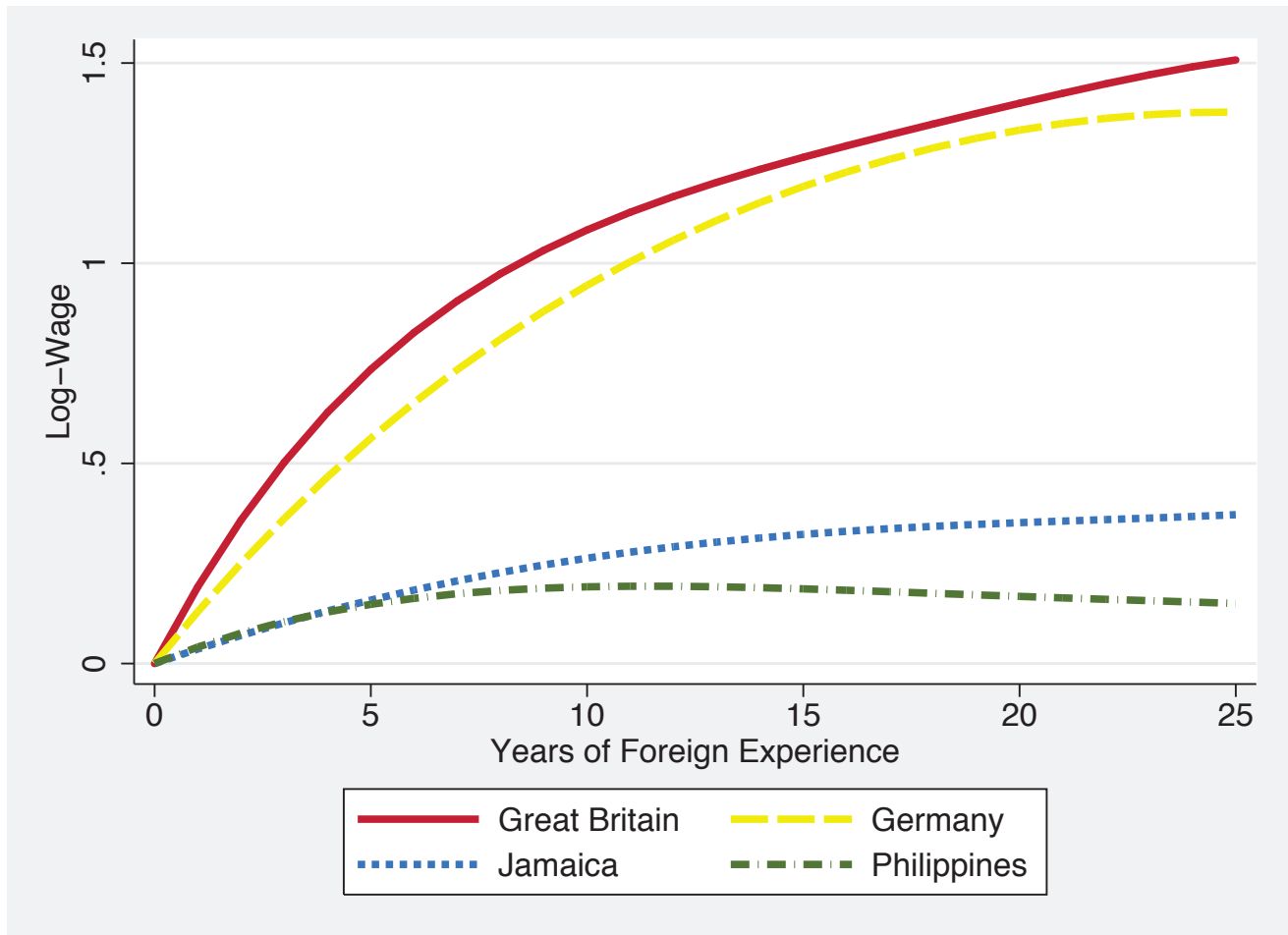


Figure 1: Returns to Foreign Experience Among New Immigrants, Select Countries

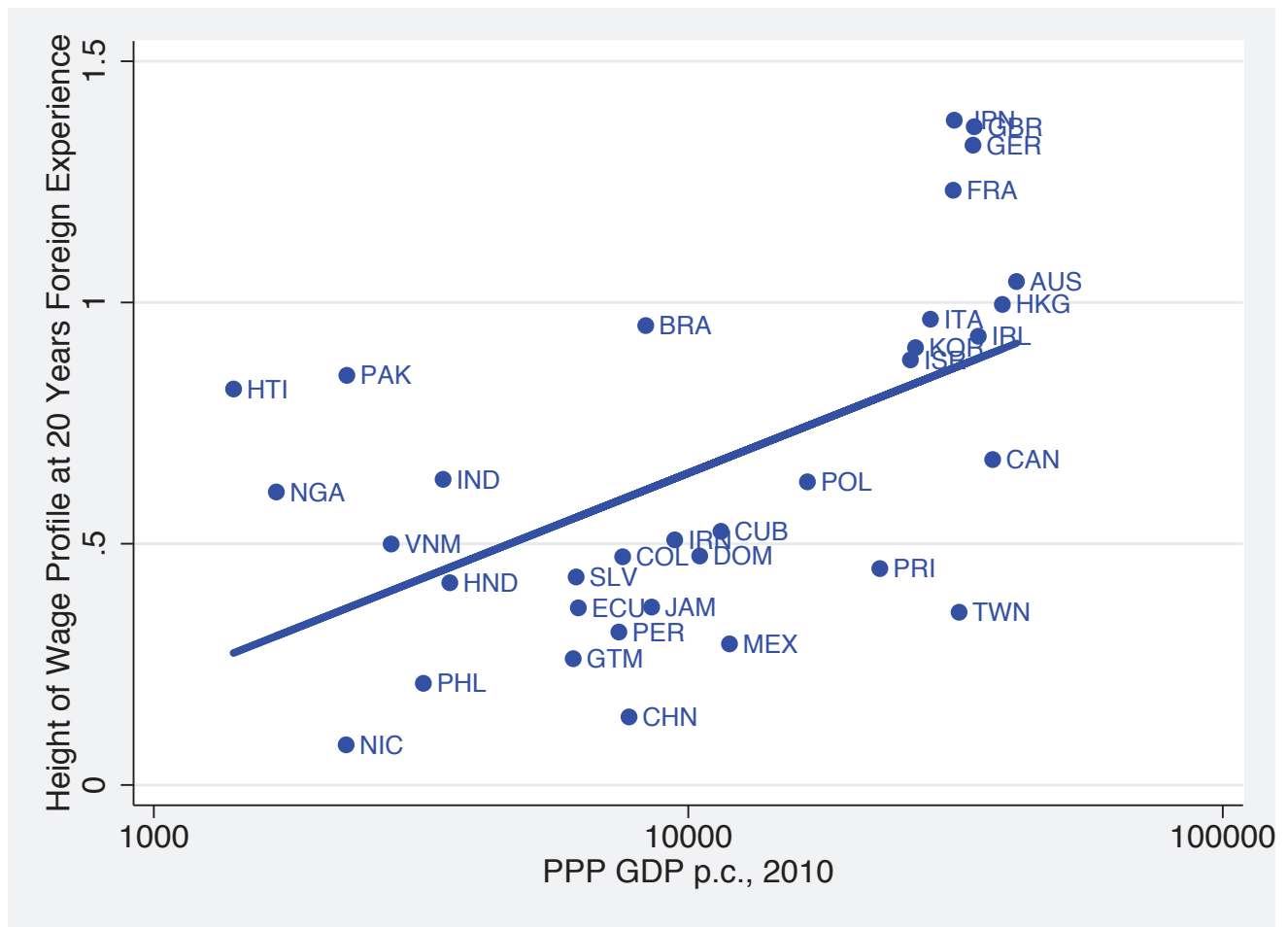


Figure 2: Returns to Twenty Years Foreign Experience Among New Immigrants, by GDP per capita of Origin Country

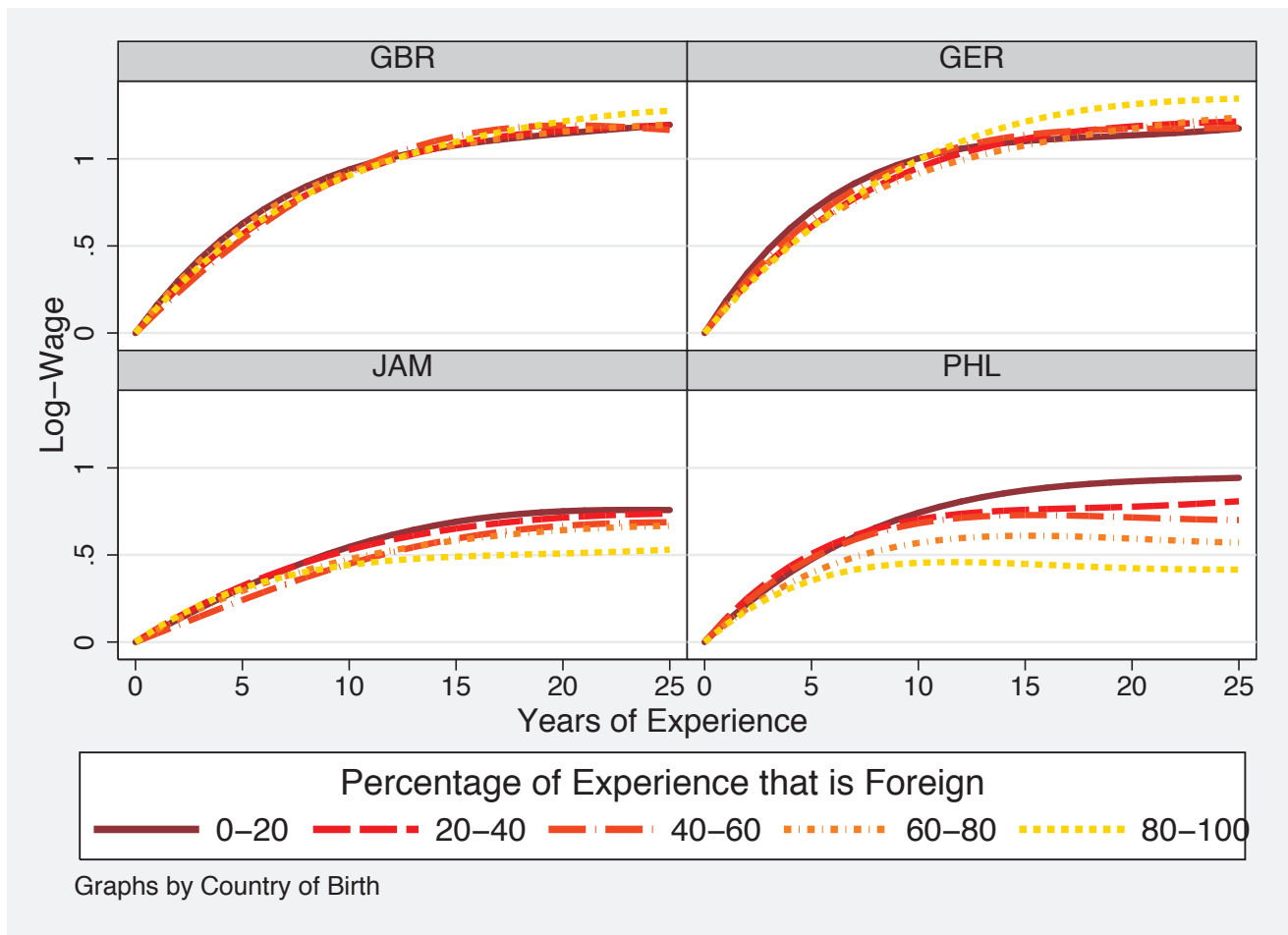


Figure 3: Returns to Experience by Percent Foreign, Select Countries

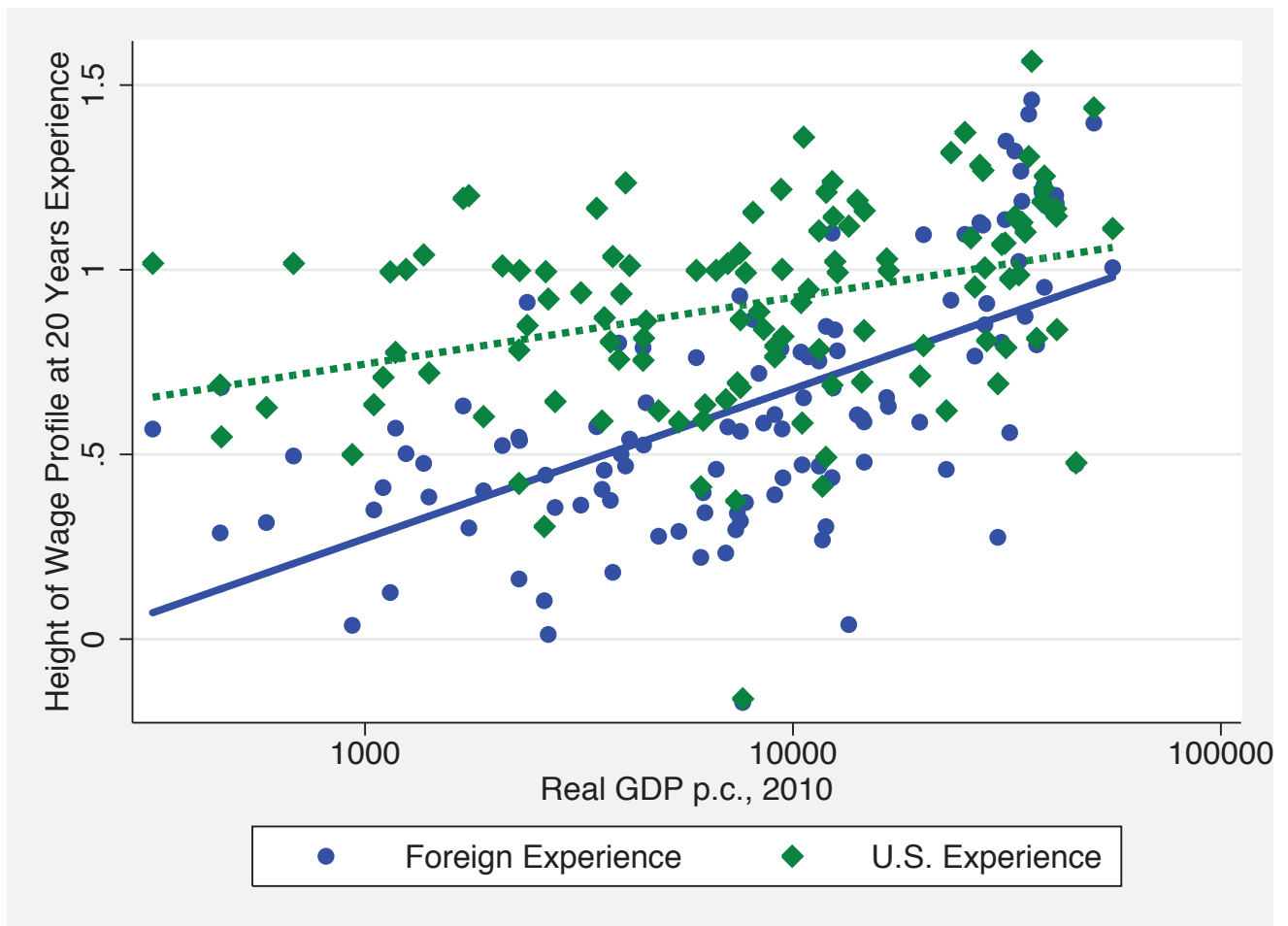


Figure 4: Returns to Twenty Years Foreign and U.S. Experience by GDP per capita

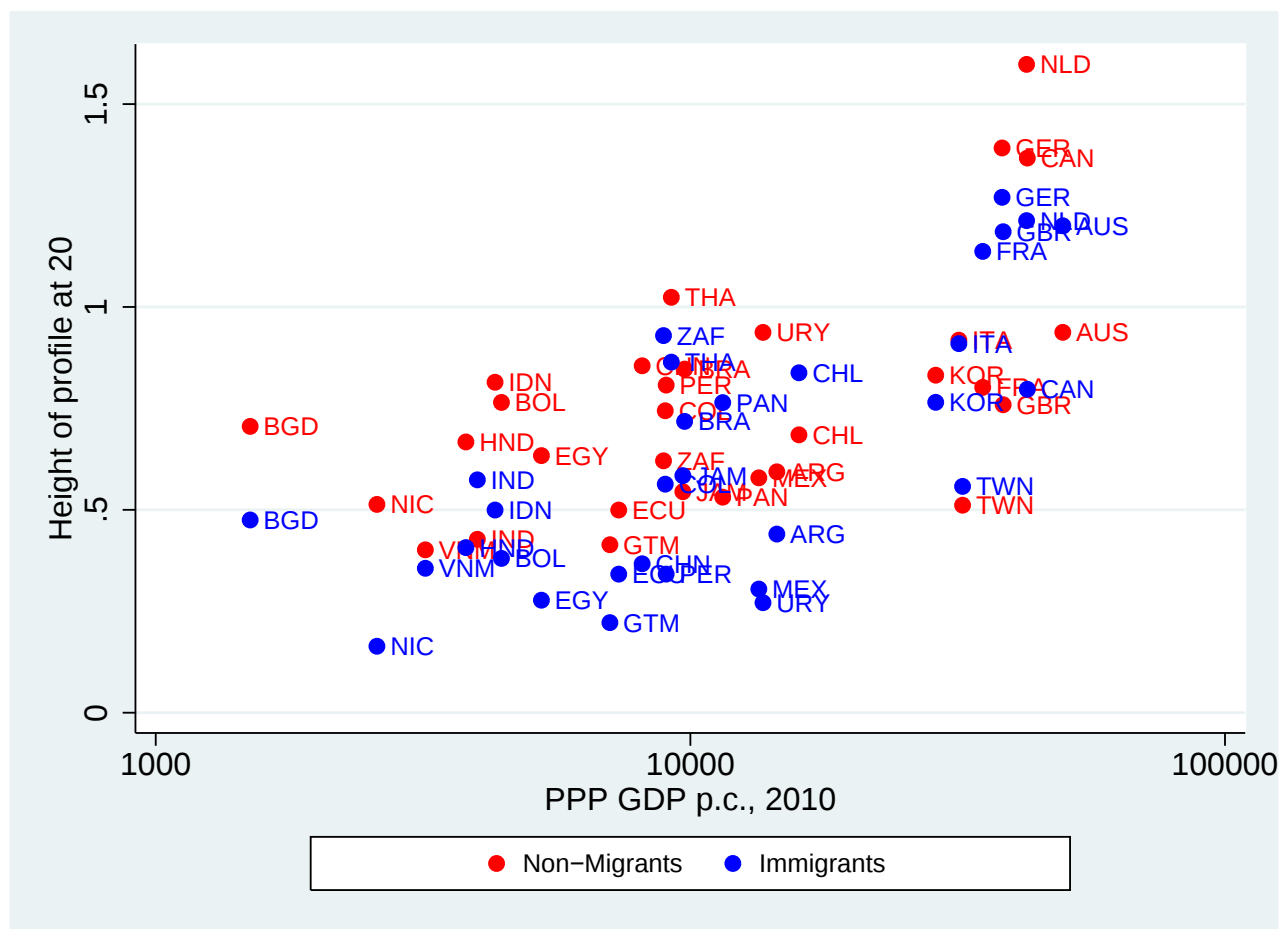


Figure 5: Returns to Twenty Years Experience, Immigrants and Non-migrants

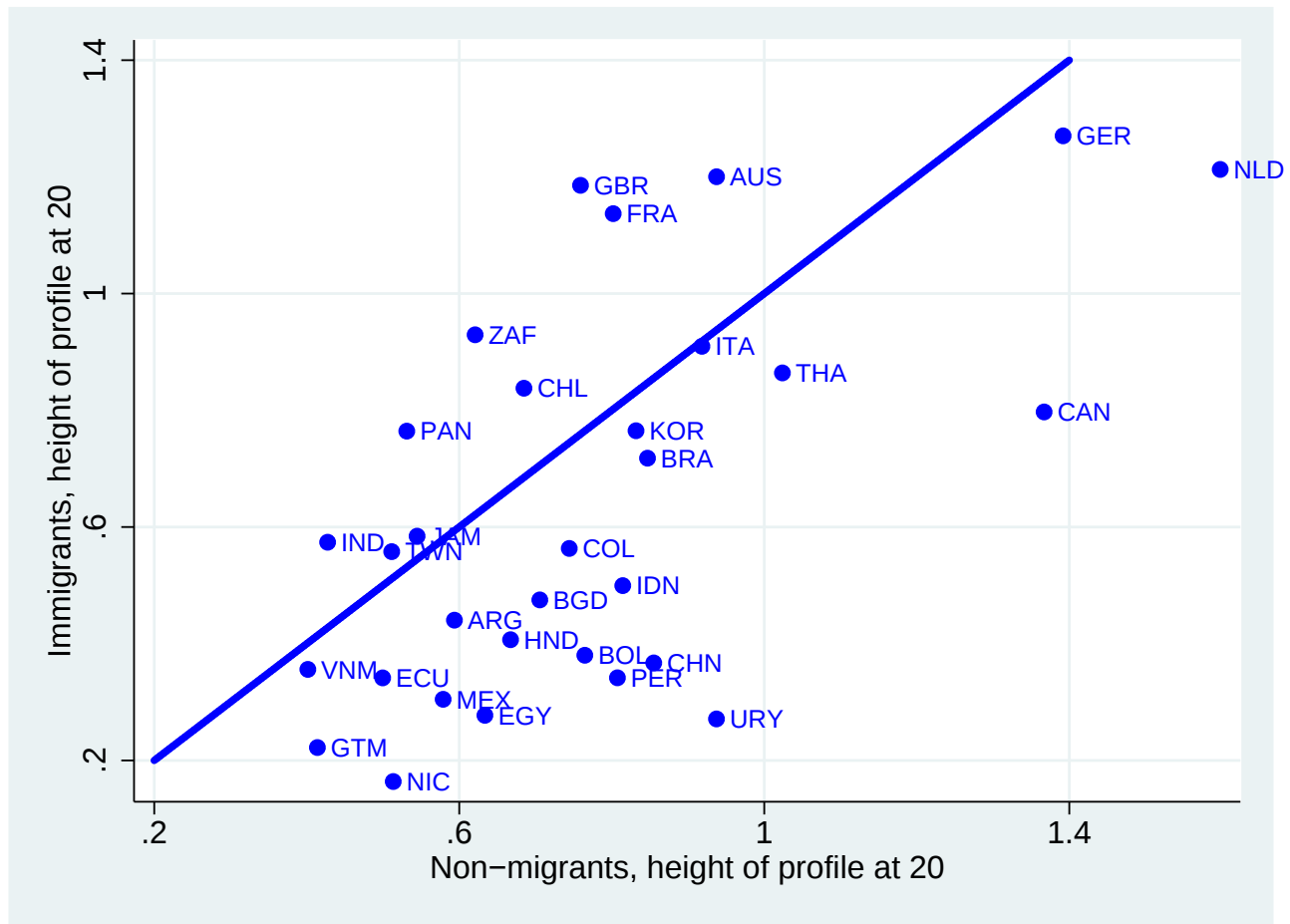


Figure 6: Returns to Twenty Years Experience, Immigrants vs Non-migrants

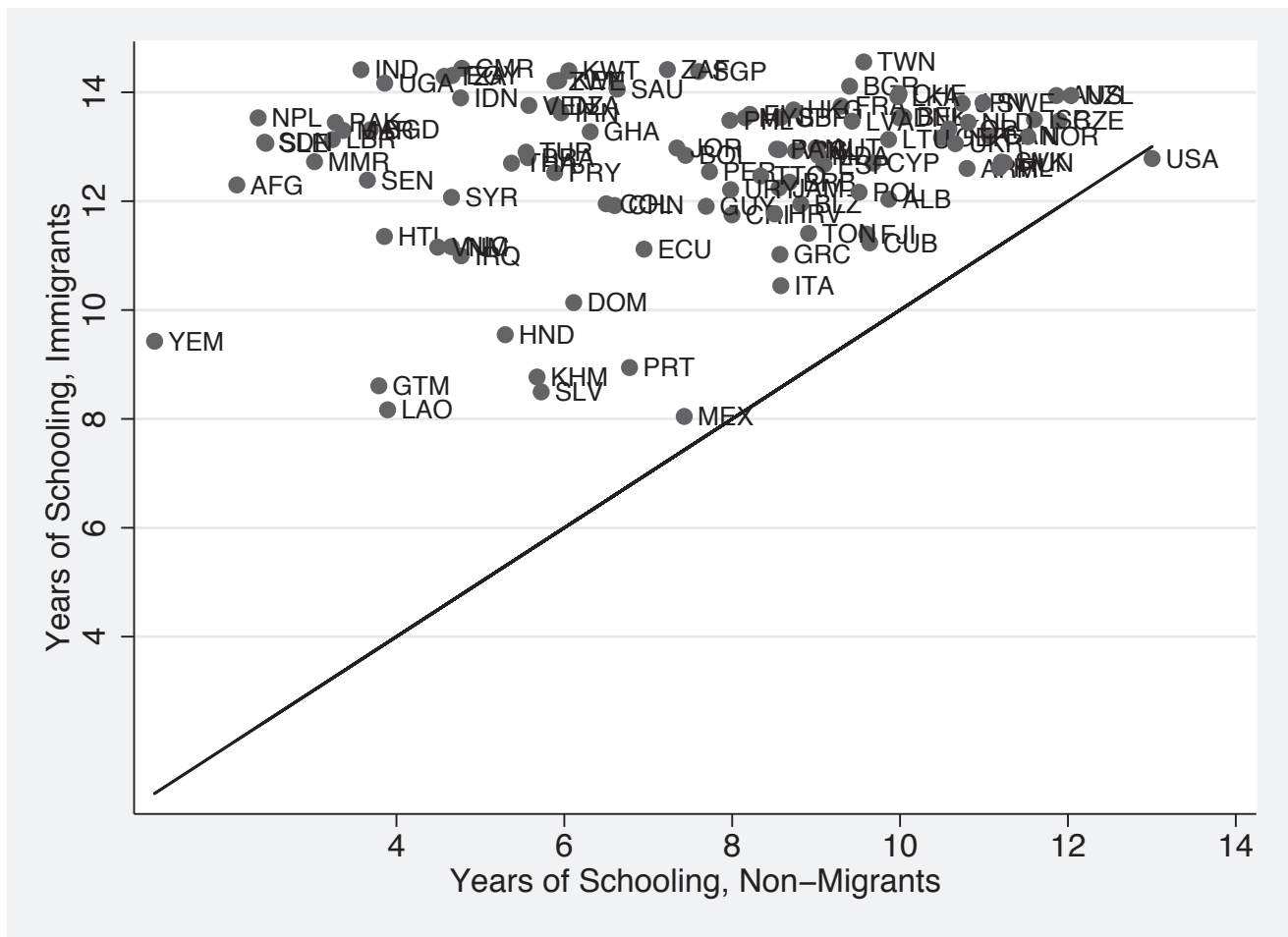


Figure 7: Years of Schooling Completed Among Immigrants and Non-migrants

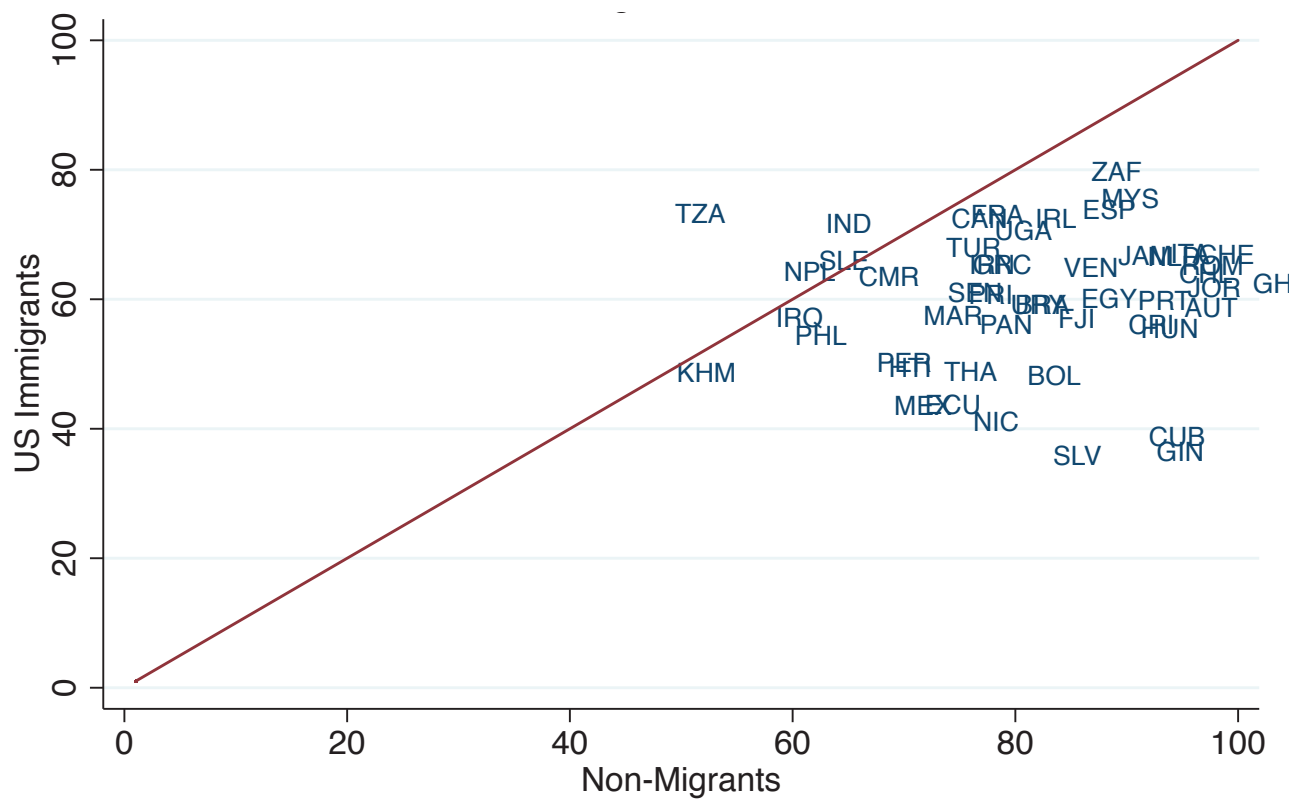


Figure 8: Percent of Educated Workers in High-Skilled Occupations, Immigrants and Non-migrants

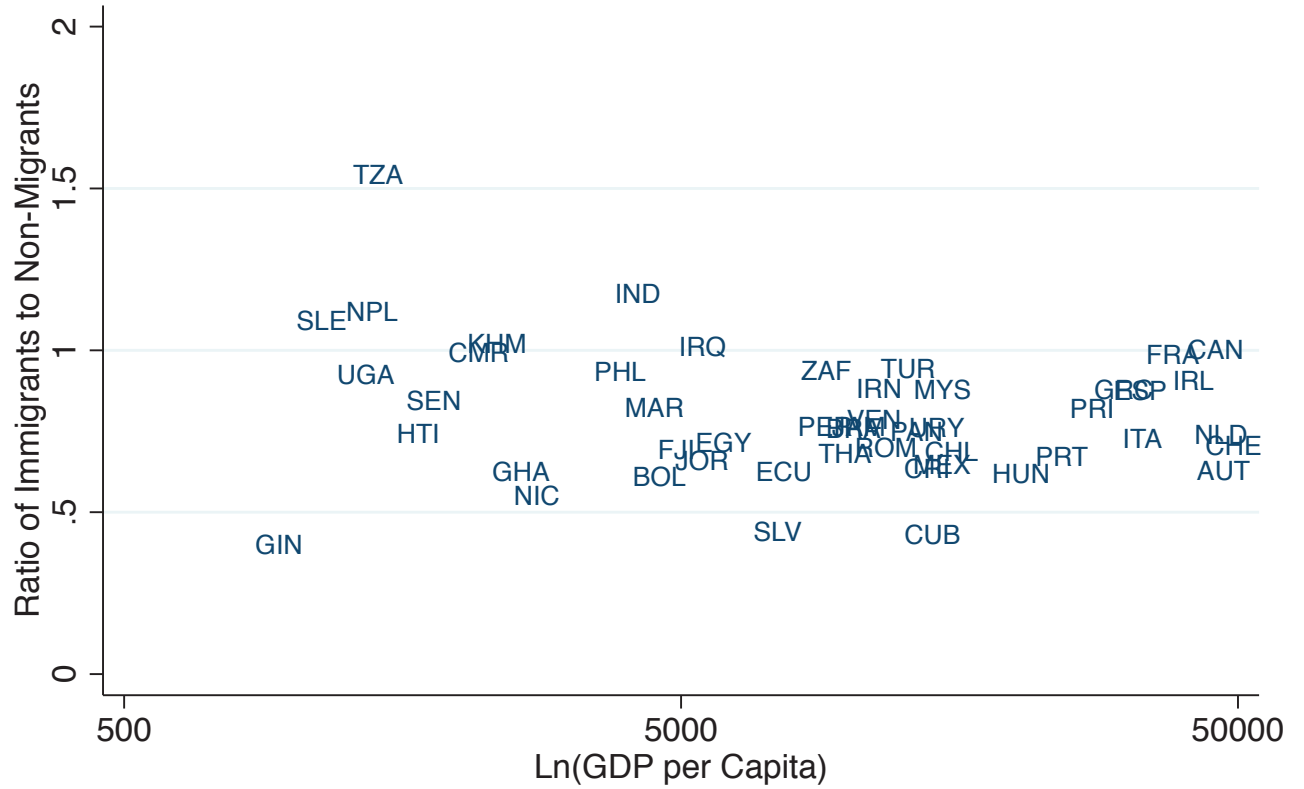


Figure 9: Percent of Educated in High-Skilled Occupations: Ratio of Immigrants to Non-migrants

Appendix

A. Estimating Returns to Experience Among Immigrants

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

$$\log(w^N) = \beta^N X^N + \phi^N E^N + \omega^N Y^N + \mu^N S^N + \varepsilon^N \quad (3)$$

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 domestically-acquired schooling. We also want to distinguish between the return to foreign (birth country) and domestic (U.S.) experience for immigrants FE and DE . We will also allow the return to domestic experience to be different for immigrants and natives.

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

$$\log(w^I) = \beta^I X^I + \phi_1^I FE^I + \phi_2^I DE^I + \omega^I Y^I + \mu^I S^I + \alpha^I AA^I + \gamma^I C^I + \delta^I YUS^I + \varepsilon^I \quad (4)$$

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 (4) 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 (4) 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.⁶ In this case the time effects can be

⁶A less restrictive assumption is to require that immigrants share the time effects of a particular subgroup of natives. I have never seen this idea actually implemented.

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 domestic 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 domestic 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 domestic 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 ? for non-migrants, it seems that we will be able to make a concrete contribution to the interpretation of these coefficients.

An alternative way to explain our identification is to compare the wages of hypothetical immigrants who are constructed to clarify the sources of identifying variation. Table 2 does this. Immigrant 1 is an arbitrarily constructed “baseline” immigrant with year, year of immigration, age, and years of schooling listed in Panel A. These are the primitive statistics available in the Census. Panel B then shows the statistics we would construct and use in the regression: foreign and domestic experience, schooling, and year of immigration cohort. Immigrants 2–5 are constructed to allow for the identification of a single parameter of interest. For example, immigrant 2 is identical to immigrant 1 in almost all characteristics and is observed in the same dataset year. The only difference is that this immigrant is a year older. Inspection reveals that immigrant 2 differs from immigrant 1 only in foreign experience, so comparing the wages of these two immigrants will allow us to identify the role of foreign experience. Likewise, immigrant 3 is a year older than immigrant 1, but also has a year more schooling. When we construct right-hand side variables this immigrant will be identical to immigrant 1, except with an additional year of schooling. Comparing these two immigrants allows us to identify the effects of foreign schooling.

The remaining two parameters are a bit trickier to identify. To capture domestic experience effects we want to study an otherwise identical immigrant who has one additional year in the U.S. To do so, we need to observe this immigrant a year later, in the 2011 data.⁷ If we assume that dataset year effects are the same for immigrants as for natives, then we can estimate this effect using natives and net it out. Then the comparison between immigrants 4 and 1 identifies the effect of domestic experience. Finally, a similar problem applies to cohort effects. The cleanest way to identify cohort effects is to consider immigrant 5, who is born and immigrates to the U.S. a year later. Again, we have to

⁷An alternative approach is to study an immigrant who is a year older and immigrates a year earlier; but this again involves changing two variables simultaneously (in this case domestic experience and cohort).

Table 2: Identification: Hypothetical Immigrants

	1	2	3	4	5
<i>Panel A: Primitive Measures</i>					
<i>Y</i>	2010	2010	2010	2011	2011
<i>YI</i>	2005	2005	2005	2005	2006
<i>A</i>	30	31	31	31	30
<i>S</i>	12	12	13	12	12
<i>Panel B: Derived Statistics</i>					
<i>FE</i>	7	8	7	7	7
<i>DE</i>	5	5	5	6	5
<i>C</i>	2005	2005	2005	2005	2006
<i>S</i>	12	12	13	12	12
<i>Panel C: Identification</i>					
		<i>FE</i>	<i>S</i>	<i>DE</i>	<i>C</i>

change two variables at once, but if we can estimate dataset year effects using natives then this is not a problem. Under the same assumption, the comparison between immigrants 5 and 1 identifies the effect of entering in a particular cohort. These latter two examples clarify the role of the assumption that year effects are shared between natives and immigrants.

A simple way to reduce these equations is as follows: we identify the effects of foreign experience using older immigrants; the effects of foreign schooling using older and better educated students; the effect of domestic experience using older immigrants observed in later years; and cohort effects using immigrants who enter later and are observed later. Now that issues of identification are clear, we turn to the data and estimation procedures.

B. Proofs

While in their home country, individuals solve

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

Note our assumption that individuals do not anticipate migrating to the United States so that they optimize assuming they will live in their country of origin over their entire time horizon $[0, T]$. If an

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

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

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

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

$$\frac{\dot{w}}{w} = \frac{\dot{h}}{h} - \frac{\dot{\ell}}{1 - \ell} = B\ell^\gamma - \delta - \frac{\dot{\ell}}{1 - \ell} \quad (5)$$

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

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

2.1. Proof of Lemma 1

The Hamiltonian of the model is given by

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

and the condition for optimality are

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

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

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

and we notice that⁸ h cancels out in the optimality equation and so the solution is fully characterized by

$$\begin{aligned} \ell &= \min \left\{ 1, (\gamma \lambda B)^{\frac{1}{1-\gamma}} \right\} \\ \dot{\lambda} &= (r + \delta - B \ell^\gamma) \lambda - (1 - \ell) \\ \lambda(T) &= 0 \end{aligned}$$

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

$$\dot{\ell} = \left(\frac{1}{1-\gamma} \right) [(r + \delta) \ell - ((1-\gamma) B) \ell^{\gamma+1} - \gamma B \ell^\gamma] \quad (6)$$

together with the terminal condition $\ell(T) = 0$. From (6), we have $\partial \ell(t) / \partial B < 0$ 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 6 that defines the path for $\ell(t)$ is independent of $h(t)$. $\square$

2.2. Proof of Result 1

That human capital accumulation is lower in poor countries follows immediately from Lemma 1 and Assumption 1. That $R_c^*(x^*, y^*)$ is lower follows from Lemma 1 and that

$$R_c^*(x, y) = \log w_c^*(y^*) - \log w_c^*(x^*) = \int_{x^*}^{y^*} \frac{w_c^*(t)}{w_c^*(t)} dt. \square$$

⁸This is due to the assumption of constant returns in the human capital accumulation technology.

2.3. Proof of Result 2

The goal is to compare $\log w_c(x_2^*; x_2^*) - \log w_c(x_1^*; x_1^*)$ for different levels of x_2^* and x_1^* where wage immediately after immigration satisfies

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

We have that

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

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

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

and from Assumption 1 richer countries have both higher B_c and B_c^* , we obtain the desired result. $\square$

2.4. Proof of Result 3

The result follows directly from Lemma 1, Assumption 1 and that

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

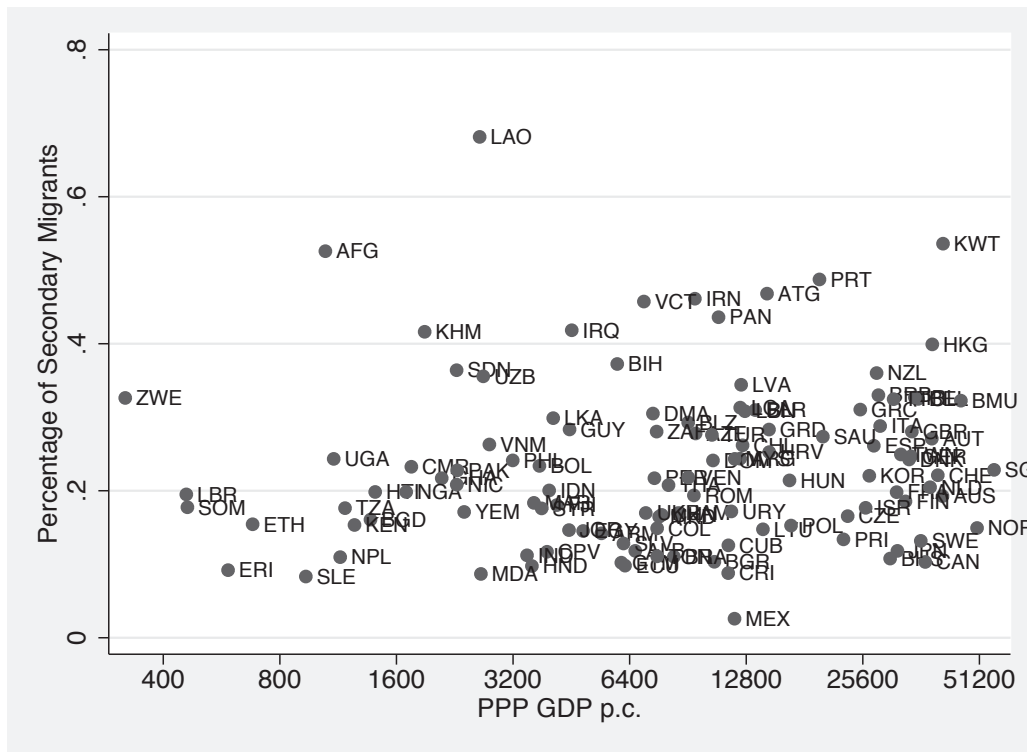
2.5. Proof of Result 4

TO BE DONE

Table 3: Ten Largest Countries by Number of Immigrants

Country	Total	By Age at Arrival		
		Early Childhood	School	Post-Graduation
Mexico	639,544	51,838	136,378	451,328
Philippines	146,518	11,622	34,905	99,991
Germany	103,118	52,342	20,121	30,655
India	96,225	3,415	27,222	65,588
Canada	76,331	18,283	20,910	37,138
Vietnam	74,635	5,644	23,321	45,670
Puerto Rico	74,612	17,240	23,469	33,903
United Kingdom	72,922	17,513	15,963	39,446
Cuba	68,293	9,212	22,677	36,404

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.

**Figure 10: Secondary Migration Rate by GDP per capita**

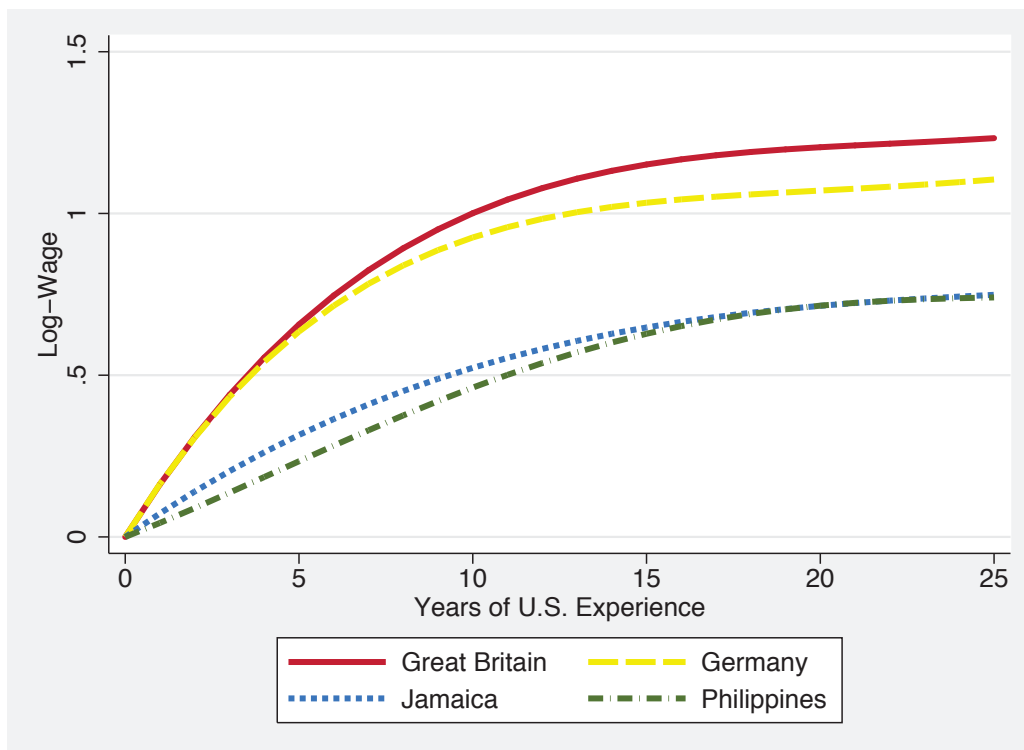


Figure 11: Returns to U.S. Experience, Select Countries

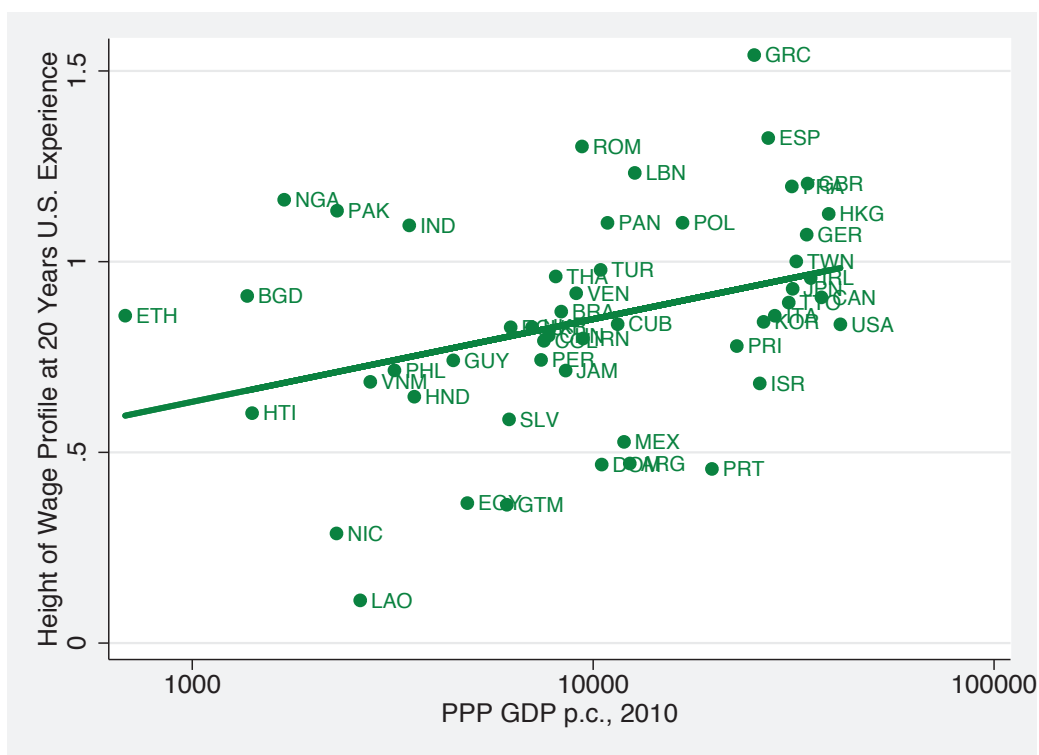


Figure 12: Returns to Twenty Years U.S. Experience Among Immigrants

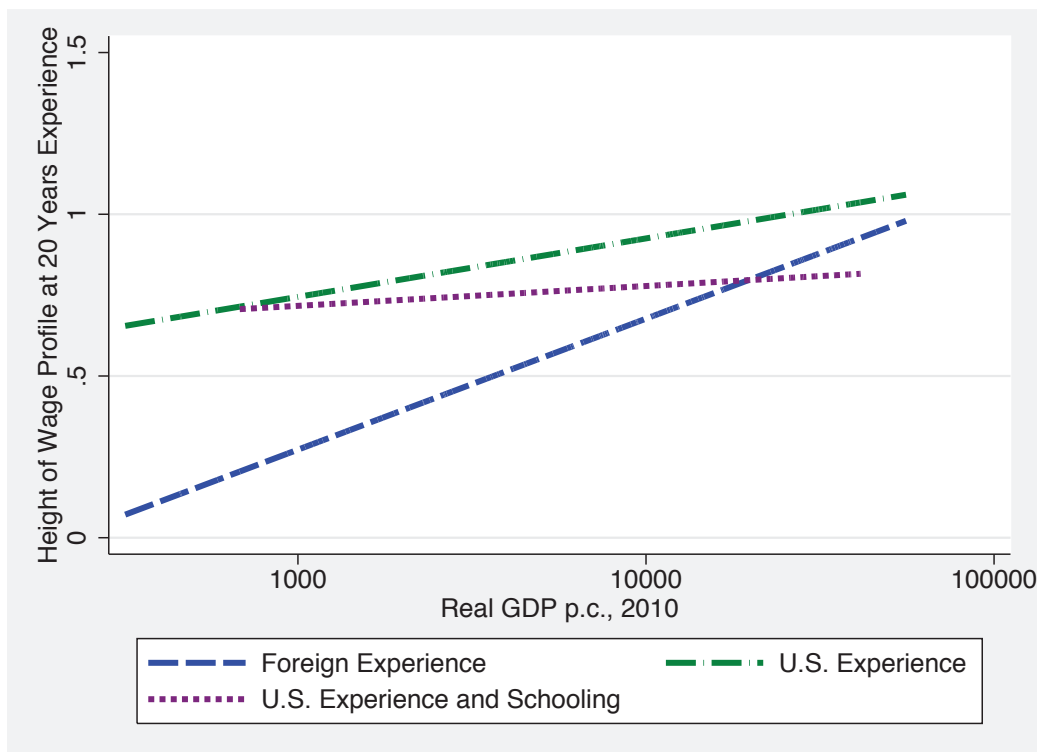


Figure 13: Returns to Twenty Years Experience by Location of Schooling and GDP per Capita (Regression Lines)