Life-Cycle Human Capital Accumulation across Countries: Lessons from US Immigrants

David Lagakos
University of California, San Diego, and National Bureau of Economic Research

Benjamin Moll
Princeton University and National Bureau of Economic Research

Tommaso Porzio
University of California, San Diego, and Centre for Economic Policy Research

Nancy Qian
Northwestern University, Centre for Economic Policy Research, National Bureau of Economic Research, and Bureau for Research in Economic Analysis of Development

Todd Schoellman
Federal Reserve Bank of Minneapolis

This paper assesses cross-country variation in life-cycle human capital accumulation, using new evidence from US immigrants. The returns to experience accumulated in an immigrant's birth country before migrating are positively correlated with birth-country GDP per capita. To understand this fact, we build a model of life-cycle human capital accumulation that features three potential theories: differential human capital accumulation, differential selection, and differential skill loss. We use new data on the characteristics of immigrants and nonmigrants from a large set of countries to distinguish between these theories. The most likely theory is that immigrants from poor countries accumulate less human capital in their birth countries before migrating. Our findings imply that life-cycle human capital stocks are much larger in rich countries.

For helpful comments we thank Rachel Friedberg, Gordon Hanson, Pete Klenow, Gary Becker, Bob Lucas, Christos Makridis, two anonymous reviewers, 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. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. All potential errors are our own.

[*Journal of Human Capital*, 2018, vol. 12, no. 2]
© 2018 by The University of Chicago. All rights reserved. 1932-8575/2018/1202-0006$10.00
https://www.journals.uchicago.edu/doi/10.1086/697245

I. 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 2 between the richest and poorest countries, whereas actual output per worker varies by a factor of more than 20.

One reason the existing literature has found such a modest role for human capital is that it has focused largely on human capital accumulated in schooling. Several previous studies have included human capital accumulated over the life cycle, that is, 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 1998, 2000). 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 from 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 et al. (2018) 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 US immigrants to help measure and understand differences in life-cycle human capital accumulation across countries. Studying US 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. For example, this fact can help us isolate differences in the quantity of experience human capital generated in different countries from possible differ-

[1] This conclusion has been arrived at by others as well, including Caselli (2005) and Erosa, Koreshkova, and Restuccia (2010). See the summary of Hsieh and Klenow (2010) for a clear overview of the development accounting literature.

ences in the price of experience human capital, because we observe all immigrants in a common country with a common set of prices.[2] Second, data for all workers come from a common data source, the US census, thus minimizing worries about international data comparability. Finally, the data span more than 3 decades in time and cover US natives, allowing us to isolate cohort-of-migration and time effects consistently for workers from a large set of countries. The insight of using immigrants to study human capital accumulation across countries is based on the work of Hanushek and Kimko (2000), Hendricks (2002), and Schoellman (2012), although this paper is the first to measure and explain stocks of human capital from experience using US immigrants.

We begin by documenting a key fact about immigrant returns to experience: returns to experience are lower among immigrants from poor countries than among immigrants from rich countries. We find that this is true both for returns to "foreign experience," acquired before migrating, and returns to "US 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 1 year, and considers returns only to foreign experience (which is essentially all they have). The second version considers all US immigrants and estimates the return to foreign and US 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 US experience are increasing in GDP per capita of the birth country, but not as sharply as for foreign experience. These facts are consistent with earlier work by Chiswick (1978) and Coulombe, Grenier, and Nadeau (2014) that shows similar patterns in other data sets. The main contribution 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 nonmigrants.

To understand the cross-country differences in returns to experience, 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 human capital accumulation, which says that workers in poor countries accumulate less human capital. The second 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 third theory is differential skill loss, which says that immigrants from poor countries tend to lose a larger fraction of their skills after migrating. All

[2] Note that this point is logically distinct from the concern that the selective nature of immigration may itself be changing the relative wage structure of the United States, as in Ehrlich and Kim (2015). Our approach relies on the fact that the relative prices are the same for all workers in the United States, regardless of how those prices are determined.

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 nonmigrants in a large set of countries. The data contain the average years of school completed by immigrants and nonmigrants and the fraction of both groups working at "high-skilled" occupations, both of which are taken from national census data from around the world. The data also contain the returns to experience for immigrants and nonmigrants, taken from this study and Lagakos et al. (2018), respectively.

The data on immigrants and nonmigrants 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 US immigrants. The reasons are as follows. First, returns to experience among nonmigrants look quite similar to returns to foreign experience among immigrants for most countries. This is consistent with individuals in poor countries accumulating less human capital, while it 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 completed and premigration wages suggests that immigrants from poorer countries are typically 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 differ only in these capital stocks. We depart from the literature in that we use our estimated returns to experience among US 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 in 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 II, we describe the facts that we document about returns to experience among US immigrants. In Section III, we present a model capturing the three different

theories of the facts described above, and in Section IV, we draw on evidence comparing immigrants and nonmigrants to help distinguish between the theories. In Section V, we illustrate what our empirical findings imply for development accounting. In Section VI, we conclude.

II. Immigrant Returns to Experience: The Facts

A. *Sample and Data*

Our data on immigrants draw on the 1980–2000 US population censuses as well as the 2005–2013 American Community Surveys (ACSs) from the Integrated Public Use Microdata Series (IPUMS; Ruggles et al. 2015). Each of these data sets includes a large, representative cross section of the US population in a particular year. We choose not to use data from earlier censuses because their sample sizes 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 as close as possible to that of Lagakos et al. (2018), since one goal is to compare our estimated results for immigrants to those for nonmigrants. We focus on men aged 18 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. (2018) 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 hours or weeks worked, labor income, and education; such people are relatively rare in the census.

We identify immigrants by 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 US immigrants.[3] Our data sets 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 cod-

[3] We find that most immigrants report being in their country of birth right before migrating: 87 percent report being in their birth country 5 years before migrating, and 83 percent report being there 1 year before migrating. There also appears to be no systematic relationship between this secondary migration and GDP per capita.

ing 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 also use this variable to exclude immigrants who likely immigrated before completing their schooling for our baseline analysis, although we return to this group below.

We construct potential experience (henceforth "experience") by using information on age and educational attainment. In the 1980 census, the raw data were years of schooling, while from 1990 onward they were recorded as educational attainment (e.g., high school graduate). We recode educational attainment into years in the standard fashion. We then define experience as age minus schooling minus 6. A small subset of our sample reports very low levels of schooling. Following Lagakos et al. (2018), we define experience as age minus 18 for anyone with less than 12 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 (US) experience.

We construct the hourly wage by using information on annual wage and salary income for the prior year, usual hours worked per week, and weeks worked in the prior year.[4] 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 US natives, so that the remaining categories all capture variation within the immigrant population.

Table 1 shows summary statistics for our baseline sample (means and standard deviations) as well as means and sample sizes for immigrants from the 10 most common birth countries in our sample. The country means convey the heterogeneity in our sample, as there is a reasonable mixture of rich and poor birth countries, with the income per capita range from roughly $3,200 to $43,000 in 2010 (real GDP per capita; Penn World Table 7.1, Vietnam to Canada). Immigrants from poorer countries, especially those in the Americas, have lower mean education, are less likely to speak

[4] Weeks 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.

TABLE 1
SUMMARY STATISTICS FOR SAMPLE AND SELECT COUNTRIES

	Wage	US Experience	Foreign Experience	Schooling	English	Number
Summary statistics, all immigrants:						
Mean	.96	13.0	7.6	11.1	3.0	771,952
Standard deviation	1.00	8.8	7.4	4.7	1.2	
Sample means, select countries:						
Mexico	.63	13.3	5.8	8.5	2.4	280,955
India	1.82	10.0	7.0	16.3	3.8	38,208
Philippines	1.14	14.4	9.8	14.4	3.6	29,905
El Salvador	.67	13.7	6.1	8.6	2.6	26,476
China	1.26	12.7	9.8	14.1	2.8	24,075
Vietnam	.90	16.2	9.2	11.3	2.8	20,058
Guatemala	.59	10.7	6.1	8.0	2.4	18,465
United Kingdom	2.40	12.4	10.2	15.4	4.9	18,231
Cuba	.78	11.7	12.5	12.4	2.6	16,266
Canada	2.16	12.3	9.9	15.1	4.7	15,567

Note.—Summary statistics for the baseline sample for all immigrants (first row) and the 10 countries with the most immigrants in the baseline sample. Wage is hourly wage, taken relative to natives in the given year and averaged across years. (Potential) experience and schooling are measured in years. "English" refers to the respondent's self-rated English ability, ranking from 1 to 5 where 1 corresponds to "does not speak English" and 5 corresponds to "speaks only English." "Number" is the sample size. All data are weighted by within-sample weights and summed across years.

English, and earn lower wages than natives, whereas immigrants from rich countries are generally more educated, speak English well, and earn higher wages than natives.

B. *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 who arrived in the United States in the year before a census. The advantage of looking at new immigrants is that they have a negligible amount of US work experience. Thus, we can estimate the returns to foreign experience for each country using the standard approach, without having to worry yet about US experience and the interaction between foreign and US experience. We also start with the simplest possible specification, with the minimal number of control variables. We then build toward our baseline specification, which includes all immigrants and a richer set of control variables.

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

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

where D_{it}^x is a dummy variable that takes the value of 1 if a worker is in experience group $x \in X = \{5\text{–}9, 10\text{–}14, \ldots\}$; the omitted category is less than 5 years. This specification allows us to capture nonlinearities 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 that of workers with less than 5 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 suppress country indices for simplicity.

For each country we focus only on new immigrants, who arrived in the United States in the year before 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 next 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 who 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.[5]

2. Richer Specification

We now consider a richer specification that allows for controls for state of residence and English-language ability, pools all countries for which we have at least 500 new immigrants, and includes native-born workers. We estimate

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

[5] We have also estimated equation (1) with immigrants who arrived within 2 years of a census. We find similar results, available upon request.

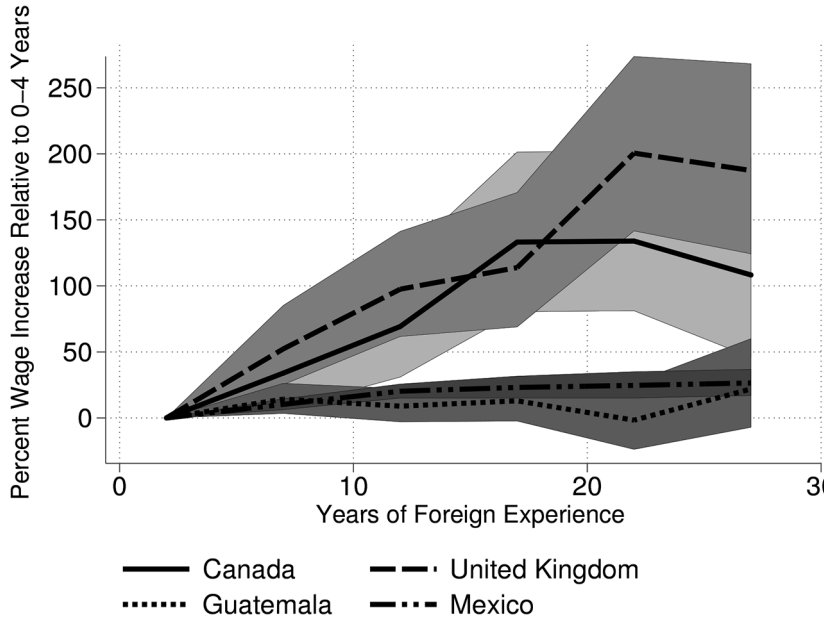

Figure 1.—Returns to foreign experience among new immigrants, select countries. Shaded regions are the 95 percent heteroskedasticity-robust confidence intervals.

where α is a country fixed effect, z_{it} is a vector of controls for state and English ability, θ is country-specific return to schooling, and the ϕ_x are the country-specific returns to experience group x. As above, 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.

We find similar results. In particular, if we focus again on the returns to 20–24 years of foreign experience, we find a strong relationship between the returns to foreign experience and (log) GDP per capita; the estimated coefficient from a simple regression of returns to experience on GDP per capita is 63.7 and is significant at the 1 percent level.[6] We conclude that among new immigrants, returns to foreign experience are higher for immigrants from richer countries than for immigrants from poorer countries.[7]

[6] Throughout the rest of the paper, we use heteroskedasticity-consistent standard errors in second-stage regressions where estimated heights of experience profiles are treated as data, consistent with Lewis and Linzer (2005) and Caron, Fally, and Markusen (2014).

[7] One potential source of bias in our calculations comes from selection on which types of immigrants obtain jobs within a year of migrating. This would drive our results if the selection is such that those with low ability from poor countries are more likely to land jobs when they first arrive, while those with high ability from rich countries are more likely to land jobs when they first arrive. In fact, we find that virtually all immigrants are employed within a year of migrating, casting doubt on this possible bias. Another possibility is that the types of jobs that

C. *Baseline Specification: All 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 US 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 US and foreign experience.[8]

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

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

This semiparametric specification allows us to estimate the returns to foreign and US experience as above. Now $D_{it}^{\mathrm{f},x}$ is a dummy variable that takes the value of 1 if a worker is in foreign-experience group $x \in X = \{5\text{–}9, 10\text{–}14, \ldots\}$, and $D_{it}^{x,\mathrm{u}}$ is a similar dummy variable for US experience; $g(x_{\mathrm{f}}, x_{\mathrm{u}})$ is the polynomial that controls for interactions between foreign and US experience; and D_{ic} is a dummy for decadal cohort of immigration to capture the possibility that immigrants who entered during different decades may differ in unobserved ability; while the remaining controls are similar to those in equation (2).

Figure 2 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 US experience. The first thing to note is that our sample size is much larger than that in figure 1; we now have estimated returns to experience for 70 countries. The circles in the figure

are taken by new immigrants are better reflective of their skills for immigrants from rich countries than for immigrants from poor countries. In Section IV, we compare the occupations of immigrants and nonmigrants country by country and find little support for this possibility.

[8] By this we mean controls for the product of US and foreign experience, the product of US experience and the square of foreign experience, and the product of foreign experience and the square of US experience. This approach of allowing for interaction terms has been discussed and used elsewhere in the literature (Chiswick 1978; Coulombe, Grenier, and Nadeau 2014). 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.

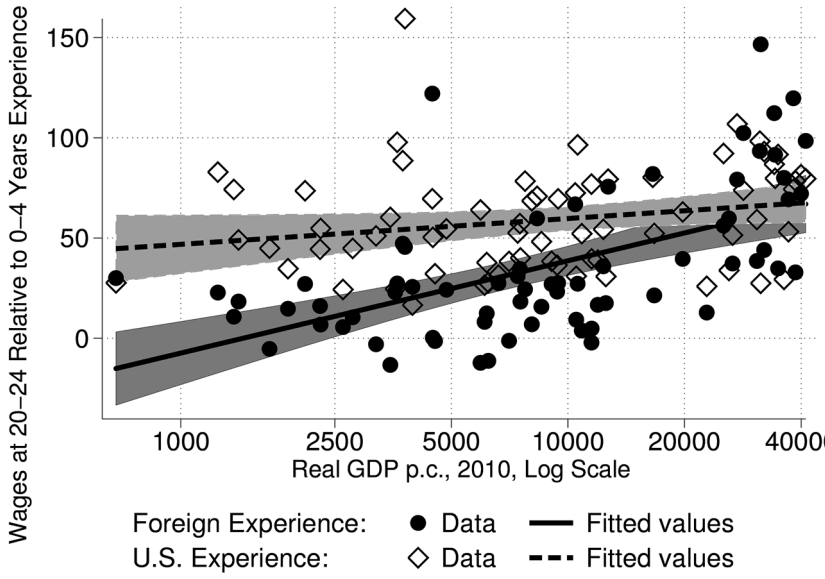

Figure 2.—Returns to foreign and US experience by real GDP per capita (p.c.) of birth country. Shaded regions are the 95 percent heteroskedasticity-robust confidence intervals.

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 1 percent level. The diamonds show the returns to US experience. As can be seen, these are also higher in countries with higher GDP per capita, yet the relationship is weaker than that for foreign experience. The slope coefficient from a regression of 20–24 years of US experience on log GDP per capita is 5.6, which is significant only at the 10 percent level. Finally, we find that the two slopes are significantly different from each other, also at the 1 percent level.

D. 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 US experience were weakly related. Recent research has suggested an important complementary relationship between education and the returns to experience (Lemieux 2006; Lagakos et al. 2018). 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 reestimate the returns to experience for country and education level. We focus on the returns to US experience, since this holds fixed the country of experience and isolates the effect of quantity of schooling. The result is shown in figure 3, which plots the estimated heights of the profiles at 20–24 years of experience against GDP per capita. We include the regression lines and confidence intervals, which show that immigrants with more education have higher returns to US experience (consistent with the literature) but that the difference is not significantly related to GDP per capita.[9]

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 study an alternative sample of immigrants who have a largely but not entirely US education, defined as those who arrive in the United States after age 12 but at least 2 years before their expected age of graduation. 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 examines the returns to US experience, since immigrants who move to the United States before graduation have only US experience. The main finding is shown in figure 4, which plots the estimated heights of the profiles at 20–24 years of US experience against GDP per capita for immigrants who migrated before and those who migrated after their expected age of graduation. The main finding is that immigrants with US education have higher returns to US 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 the 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. Overall, it seems that, in addition to supporting the growing finding that quantity of education and returns to experience are correlated, we can also offer evidence that the location of education matters.

E. Sensitivity

We now explore the sensitivity of our results along four 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 re-

[9] Note that focusing on subsamples with different levels of education also allows us to test whether our results are sensitive to allowing the returns to school to vary by education level. This may be important, given that wages are increasingly nonlinear functions of schooling, as noted by Heckman, Lochner, and Todd (2008) and others.

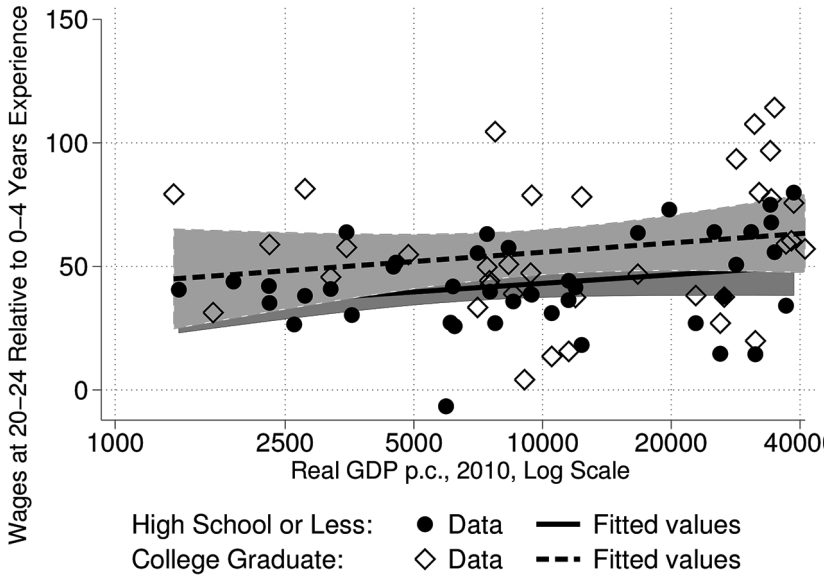

Figure 3.—Returns to US experience by education level by real GDP per capita (p.c.) of birth country. Shaded regions are the 95 percent heteroskedasticity-robust confidence intervals.

sults are robust to different ways of thinking about measurement error in the construction of experience. Finally, we explore whether the results are robust to controlling for possible confounding influences relevant for immigrants. Throughout, we focus on the relationship between the

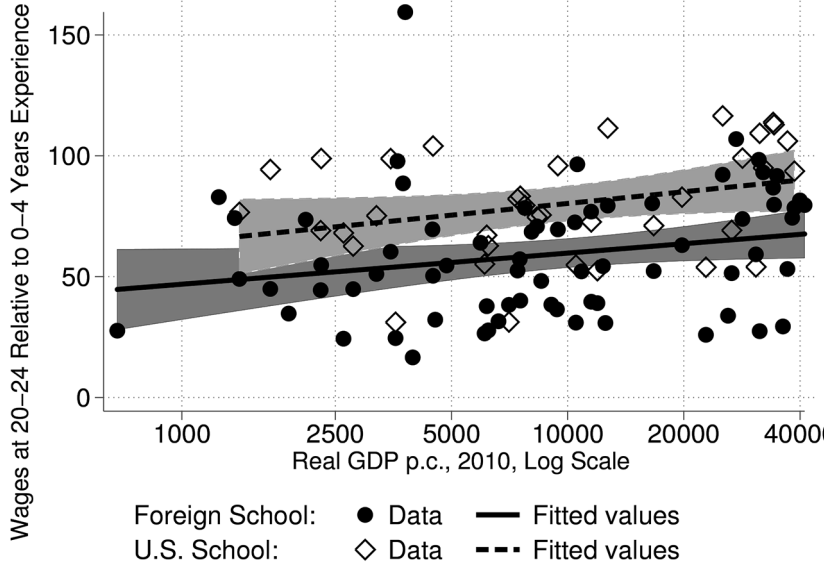

Figure 4.—Returns to US experience by education location by real GDP per capita (p.c.) of birth country. Shaded regions are the 95 percent heteroskedasticity-robust confidence intervals.

life-cycle wage growth and birth-country real GDP per capita, in line with figure 2. 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 US experience of foreign-educated workers, and the US experience of US-educated workers. As discussed above, the returns to experience are much more strongly related to development (birth-country real GDP per capita) for foreign than for US experience. For US experience, it seems to matter little whether the immigrant was entirely educated abroad or was partially educated in the United States.

TABLE 2
Returns to Experience and GDP per Capita: Robustness of Results

	Foreign Schooling, Foreign Experience	Foreign Schooling, US Experience	US Schooling, US Experience
Baseline	20.0***	5.6*	7.0*
Alternative metrics:			
15–19 years experience	15.2***	3.3	1.5
Average height of profile	16.4***	5.5**	6.1
Discounted average height of profile	8.4***	2.6**	3.3**
Weighted by sample size	23.2***	3.7	1.9
Weighted by population	31.3***	6.0	−.3
Broadening the sample:			
Include women	20.2***	3.6	2.6
Include some part-time	20.5***	5.3*	7.1*
Include all part-time	23.7***	7.8***	3.9
Include public sector	21.4***	6.3**	10.9***
Sensitivity to experience measurement:			
Year 2000 onward	21.9***	6.3*	7.0*
Country-specific school start date	18.7***	1.9	NA
Experience at 16	20.9***	3.3	1.5
Ages 18–65	19.9***	5.6*	10.3
Ages 22–55	23.1***	6.1*	7.4**
Sensitivity to immigrant sample:			
College graduate or more	38.6***	5.4	10.2*
High school graduate or less	13.8***	5.0*	NA
Variable return to schooling	20.9***	.6	NA
Manufacturing	15.7***	5.4	NA
Service industry	23.2***	11.5***	7.0
Excellent English	26.4***	6.6*	6.2*
English-speaking country	14.3***	2.6	NA
No ethnic enclaves	22.5***	6.2*	7.3*
Fewer unauthorized immigrants	20.9***	6.3**	5.3

Note.—Estimates from a second-stage regression of the estimated height of experience profiles on log Real GDP per capita. Each row corresponds to a sample for the first-stage regression, a specification for the first-stage regression, or a measure of the height of experience profiles; see the text for details. The columns correspond to the estimated value of foreign experience, US experience for foreign-educated immigrants, and US experience for immigrants with some US education. "NA" denotes a check that is not available because it is logically inconsistent (country-specific school start date) or there are too few countries in the sample (<15).

* Significant at the 10 percent level.

** Significant at the 5 percent level.

*** Significant at the 1 percent level.

The next five rows explore alternative metrics for the height or steepness of experience profiles. We see that the same results prevail if we focus on the height of profiles measured at different ages; we show here the result for 15–19 years of experience. 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 lifetime earnings calculation in the spirit of what is often done in the education literature. Finally, we explore the possibility of weighting observations, rather than treating each country equally, in the second-stage regression. We explore weighting by the sample size from the first-stage regression, which gives more precise estimates more weight, and weighting by the country's population, which gives more populous countries more weight. These alternative weights imply, if anything, a larger gap between the returns to foreign and US experience.

The next four rows explore the possibility of broadening the sample. Our baseline sample excludes several categories of workers because we are worried that wages may not be set competitively or that potential experience may be a poor measure of actual experience. Here, we explore including women, part-time workers, and public-sector workers in the sample; in the case of including women, we also include a gender dummy as a control variable. The exact results vary somewhat, but throughout we find that the value of foreign experience is closely tied to development, while the value of US experience is much less so.

The next five rows explore sensitivity to the measurement of experience. A particular concern is measurement error in experience, which would attenuate the measured experience-wage profiles. In our companion paper, we bound the possible role of measurement error in age and years of schooling and show that it is unlikely to explain the sizable differences in returns to experience between poor and rich countries; interested readers should see Lagakos et al. (2018) for details. Here, we present a few new exercises more relevant to immigrants. First, we check whether our results apply for the year 2000 onward, when we measure year of immigration exactly and so can partition experience into foreign and US experience more precisely. Second, we explore accounting for cross-country variation in the age children start school. We construct experience for every worker as age minus schooling minus age of first school, where the age of first school is matched from the World Development Indicators by country and birth cohort (World Bank 2015).[10] This slightly flattens the returns to both foreign and domestic experience. Third, we explore including 16–17-year-olds in the sample and allowing experience to start at age 16 rather than age 18, which may be more appropriate for immigrants from some poorer countries. Finally, we narrow the age range of immi-

[10] We thank an anonymous reviewer for pointing us to these data.

grants we consider, in case we may not fully capture experience for very young or very old workers. We find broadly similar results.

The final nine rows explore sensitivity to possible concerns that might apply to an immigrant sample. First, we test sensitivity to educational background, in case we are confounding returns to experience and those to education. We find that the results apply broadly from high school graduates to college graduates. We also explore a more relaxed functional form that allows the returns to schooling to vary by country and birth decade rather than just by country. This captures the possibility that educational quality may have improved at different rates in different countries. We find systematic variation in the returns to schooling over time, but this has little effect on the relationship between the returns to experience and development. We also test sensitivity to sector of employment, in case we are confounding returns to experience and those to job type, but find that the patterns are similar in services and manufacturing. We explore the role of language, by including only workers who self-report that they speak English well or very well or by including only workers born in English-speaking countries, and find broadly similar results. The results are not sensitive to excluding immigrants who live in ethnic enclaves, defined as a public-use-microdata area where more than 5 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, we explore excluding the 10 countries with the highest estimated rates of unauthorized immigrants, taken from Office of Policy and Planning, US Immigration and Naturalization Service (2003). Unauthorized immigrants mostly enter the United States from a small set of countries. Thus, these 10 countries include all countries for which more than one-quarter of immigrants are unauthorized immigrants. Also, more than 80 percent of all unauthorized immigrants to the United States come from these 10 countries. Unauthorized immigrants are a concern for two reasons. First, unauthorized immigrants may be selected very differently from other immigrants. Second, unauthorized immigrants may be much less likely to respond to data collection efforts or to respond accurately, which would make inferences of the type we want challenging. Fortunately, we find that our results are robust to excluding them.

Overall, our main stylized fact is that the value of foreign experience in the US labor market is strongly correlated with development, as measured by birth-country real GDP per capita. The estimated effect is consistently large and statistically significant throughout table 2. Our second stylized fact is that the value of US labor market experience is much less strongly correlated with development, whether immigrants were educated abroad or in the United States. The estimated effects are consistently smaller than those for foreign experience and are less often statistically significant. We now turn to a model to help think through why these facts might hold.

III. 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 questions of 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, which says that immigrants from poor countries accumulate less human capital over the life cycle than immigrants from rich countries. The second is differential selection, which 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, which 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 IV, we then bring additional evidence to the table to distinguish between the three different theories laid out in this section.

In our model, there are 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 US experience. We denote variables observed in immigrants' birth countries without superscripts and those observed in the United States with asterisks. 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 United States 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.[11] But, as discussed below, migrants may come from a selected part of the population. We further denote by $x_{ic}(t)$ and $x_{ic}^*(t)$, respectively, the amount of birth-country and US experience an individual has accumulated up to time t. If an individual works in his birth country, his human capital accumulates passively according to

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

[11] These assumptions are not necessary for the analysis that follows; we make them to simplify and focus attention on the selection of immigrants. Allowing $\mathbb{E}[\eta_{ic}]$ to vary by country has no effect on our results, given that we compare log wages of experienced and inexperienced workers; see equation (7). Allowing $\mathbb{E}[z_{ic}]$ to vary by country provides an alternative possible interpretation of ϕ_c; see below.

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

$$\dot{h}^*_{ic}(t) = z_{ic}\phi^*_c(x^*_{ic}(t))h^*_{ic}(t). \tag{5}$$

In particular, we allow for the possibility that, upon arrival in the United States, 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 United States. 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 United States 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 nonmigrants. More precisely, we denote the set of individuals with experience level x who migrate from country c to the United States $M_c(x)$ and allow for the possibility that

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

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 US experience is

$$w^*_{ic}(x, x^*) = \omega^*_c h^*_{ic}(x, x^*)e^{\varepsilon_{ic}}, \tag{6}$$

where ω_c^* is the skill price earned by immigrants from country c in the United States and ε_{ic} is an error term. Given our assumptions, the immigrant's human capital can be solved for in closed form and satisfies[12]

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

Combining with equation (6), the wage of a "new" immigrant, that is, one with zero years of US 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 we run using data on new immigrants only is therefore

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

The measured return to foreign experience $R_c(x)$ may be low for one of three reasons. First, the true returns to experience $f_c(x)$ may be low. Second, there may be selection. This can take the form of either experience-dependent selection on initial ability (i.e., $\mathbb{E}[\log \eta_{ic} | i \in M_c(x)]$ decreases with x) or 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 1 or decreasing). Third, there may be experience-dependent skill loss, $\theta_c < 1$. Estimates from regression (9) by themselves do not allow us to distinguish between these three 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.

IV. Distinguishing between Theories

In this section, we draw on new data to compare the characteristics of immigrants and nonmigrants from a large set of countries. We draw on three

[12] To see this, note that, for example, equation (4) can be integrated to yield

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

Integrating equation (5) in the same way and applying the assumed functional form for skill loss yields equation (7).

basic facts that help us distinguish between the theories above. First, returns to foreign experience among immigrants are similar to returns to experience among nonmigrants. Second, immigrants from poor countries tend to be more selected on premigration characteristics, such as years of schooling. Third, educated immigrants tend to work in high-skilled occupations at a lower frequency than nonmigrants, though at a similar rate in rich and poor countries alike.

A. *Returns to Experience among Immigrants and Nonmigrants*

We begin by comparing our returns to foreign experience among immigrants to the returns among nonmigrants estimated by Lagakos et al. (2018). We can make these comparisons in the 16 countries for which we have an estimate of immigrant returns and for which Lagakos et al. (2018) calculate returns using a representative sample of nonmigrants. Since we have followed the sample selection and variable construction of Lagakos et al. (2018) closely, the comparability of the results is informative about the extent to which life-cycle wage growth differs between immigrants and nonmigrants. Figure 5 plots the estimated returns to 20–24 years of experience for immigrants against the same estimated return for nonmigrants. 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 0.806, with a p-value of .001, although there are some outliers, such as Indonesia and South Korea. Countries such as Germany, the United Kingdom, and Australia have high returns among both immigrants and nonmigrants, and most of the developing countries have low returns in both groups.[13]

Perhaps the most striking fact is that these trends do not simply reflect the general influence of being more developed. To make this point, we first regress the estimated returns to experience of both immigrants and nonmigrants on log real GDP per capita. We then study the residuals from these regressions, which remove the general effect of income. These residuals are also quite strongly correlated: the correlation coefficient is 0.4760, with a p-value of .06. This estimate shows that if a country's residents have unusually steep wage profiles for its income level, then its emigrants to the United States are likely to have unusually steep wage profiles too. This provides further evidence in favor of an explanation that stresses a common factor between immigrants and nonmigrants, such as human capital accumulation. One additional type of evidence favoring the theory that experience human capital accumulation is higher in the United States than in developing countries comes from return migrants. Reinhold and Thom (2013) find

[13] The estimated relationship between the height of profiles and GDP per capita is also similar for immigrants and nonmigrants. Among immigrants, the slope coefficient in a regression of GDP per capita is 29.6, with a p-value less than .001. Among nonmigrants, the slope coefficient is 25.2 and the p-value is less than .001.

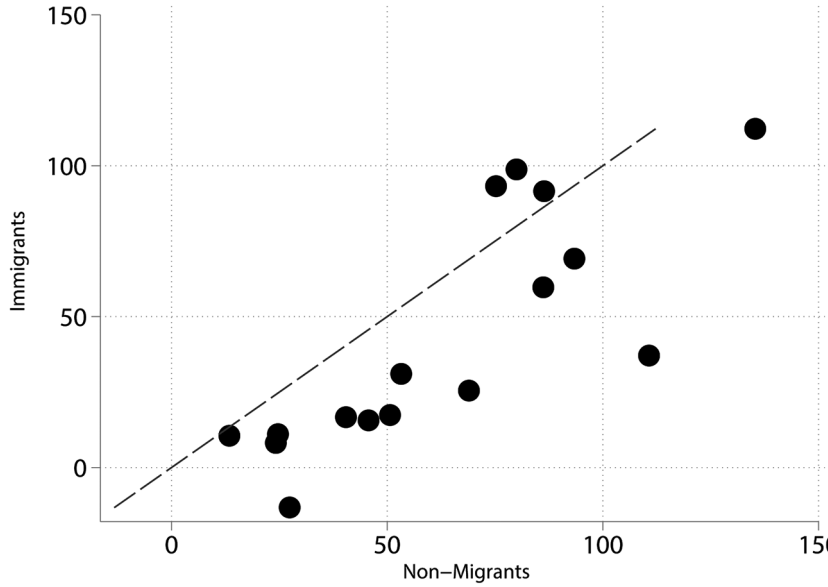

Figure 5.—Returns to foreign experience, immigrants versus nonmigrants.

that Mexican immigrants to the United States earn a large premium on their US experience when returning to Mexico.

The fact that estimated returns to experience from poor countries are low both for immigrants and nonmigrants 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 nonmigrants to be similar in countries of all income levels. As figure 5 shows, this is not the case. The broad similarity between returns to experience among immigrants and nonmigrants 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 nonmigrants 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.

B. Comparing Other Characteristics of Immigrants and Nonmigrants

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 nonmigrants 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 2015). 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 nonmigrants. 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 6 shows the results. The left-hand panel shows that average years of schooling among nonmigrants is strongly correlated with log GDP per capita, with less than 5 years of schooling on average in the poorest countries and more than 12 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 12 years

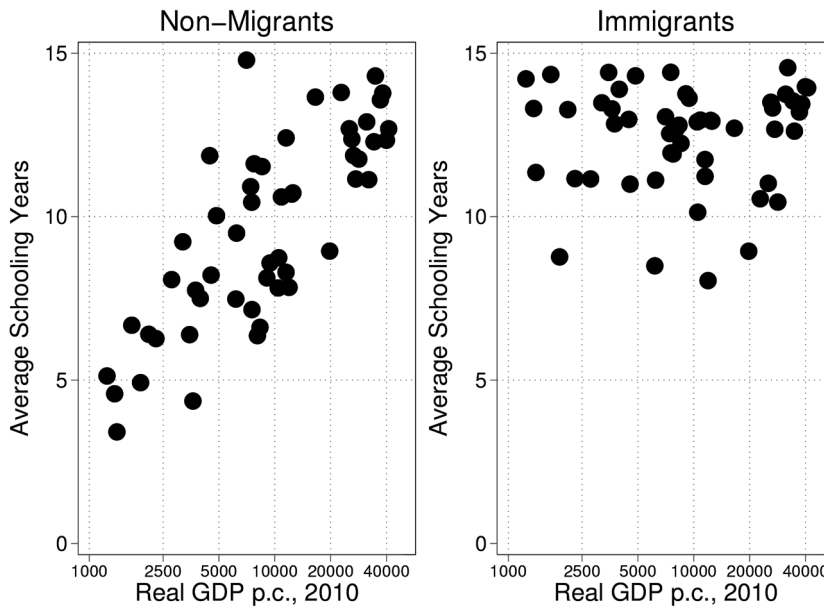

Figure 6.—Years of schooling completed among immigrants and nonmigrants; p.c.: per capita.

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 data suggest that some alternative force, such as differences in education quality, or the type of work performed, or the incentives to invest in human capital accumulation on the job, is a more likely explanation of flat life-cycle wage profiles for immigrants from poor countries.

These findings are consistent with most evidence from previous studies. 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) show a similar result for migrants to a wide range of countries. Hendricks and Schoellman (2018) show that immigrants to the United States are strongly selected on a large range of characteristics, including premigration wages, occupation, and education, and that immigrants from poorer countries are more selected on these dimensions. There are perhaps some key exceptions, mostly the countries near the United States where unauthorized immigration plays a large role. For example, Moraga (2011) suggests that Mexican immigrants may be negatively selected. Fortunately, as we documented above, our results are robust to focusing on countries with low rates of unauthorized immigration.

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 United States than can immigrants from poor countries. To test this hypothesis, we compare rates of skill loss for immigrants and nonmigrants. 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.[14] For each country, we calculate the fraction of all college-educated immigrants and nonmigrants who work in high-skilled occupations.

Figure 7 plots the results. The left-hand panel shows that among nonmigrants, 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

[14] IPUMS has standardized occupation codes across all our data sources. We define high-skilled occupations to be "professionals," "technicians and associate professionals," and "legislators, senior officials and managers," and low-skilled ocupations 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.

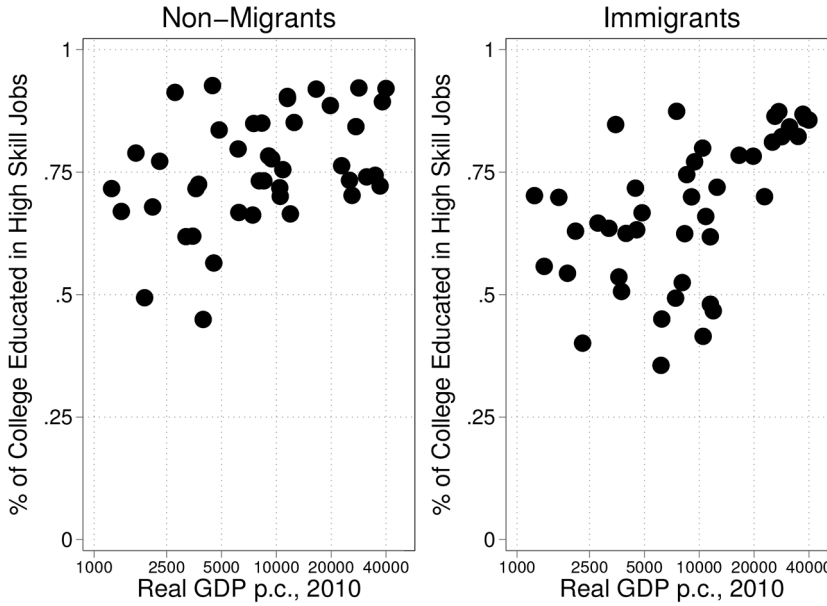

Figure 7.—Percent of educated workers in high-skilled occupations, immigrants and nonmigrants; p.c.: per capita.

likely to work at high-skilled jobs than nonmigrants in countries of any income level. This is consistent with the presence of skill loss. However, the slopes for immigrants and nonmigrants 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 (less than 10 years) and those with high (10 or more years) experience. We find that the relationship between education or skill loss and GDP per capita shown in figures 6 and 7 is similar for workers with low and those with high experience from a particular country (results available upon request). This fact suggests that selection or skill loss that is correlated with experience is not driving our results.

The results on selection and skill loss, when combined with the fact that returns-to-experience patterns are strongly correlated with the same patterns for nonmigrants, suggest that cross-country differences in human capital accumulation are the most plausible interpretation of the data.

V. Development Accounting

In this section, we use development accounting to quantify the economic importance of the empirical results shown in Section II. 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

$$\text{success}_1 = \frac{\text{Var}(\log Y_{KH,c})}{\text{Var}(\log Y_c)},$$

where $Y_{KH,c} = K_c^{\alpha} H_c^{1-\alpha}$ is the component of output due to factors of production. Values of success_1 close to 1 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 $\bar{x}_c$, respectively (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 II 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 II.

We consider two assumptions on the life-cycle labor market productivity profile, corresponding to whether we view human capital as the result of passive investment (simple learning by doing, as in Sec. III) or active accumulation (Ben-Porath, as in app. 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, respectively.[15] 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 United States, $R_c^*(x, 0)$ 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 toward 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 a 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 $\log[(1 - \ell_c^*(x))/(1 - \ell_c(0))]$ 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 8. This figure plots the human capital stocks implied by our upper bound and the slope from a regression of human capital stocks measured using only schooling on log GDP per capita. As the figure shows, our estimated human capital stocks are substantially larger in rich countries than in poor countries once experience is included.

[15] The 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 $\log h_c(x) = \int_0^x (\phi_c(\ell_c(\tilde{x})) - \delta)d\tilde{x}$.

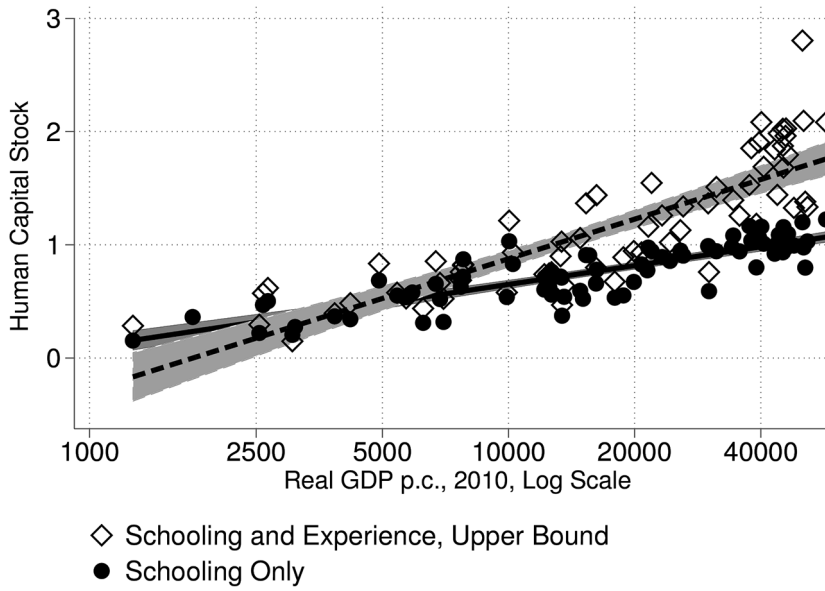

Figure 8.—Human capital stocks from both experience and schooling (computed with the upper-bound method) and human capital stocks from schooling only by real GDP per capita (p.c.). Shaded regions are the 95 percent heteroskedasticity-robust confidence intervals.

The accounting under this upper bound is presented in panel A 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.

TABLE 3
Development Accounting

Human Capital Measure	Success_1 (1)	Slope ($\log(Y_{KH})$, log (GDP)) (2)
	A. Upper Bound	
Schooling	.44	.63
Experience	.48	.65
Schooling + experience	.74	.81
	B. Lower Bound	
Schooling	.44	.63
Experience	.42	.60
Schooling + experience	.66	.76

Note.—Success_1 is defined as the ratio of the variance of $\log(Y_{KH})$ to the variance of $\log(Y)$. Slope ($\log(Y_{KH})$, $\log(\text{GDP})$) is defined as the slope coefficient from a regression of $\log(Y_{KH})$ on $\log(\text{GDP})$ and a constant. Each row of each panel describes how the human capital stock to compute Y_{KH} is computed. The first row includes only human capital from schooling, the second one only human capital from experience, and the third one includes human capital from both schooling and experience.

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 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 $\log(\text{GDP})$ is 0.63 or 0.65, respectively. 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 with 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 United States ($T - t$ and B_c^* in the notation of app. 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 US experience 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 US 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 US-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 US-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 panel B of table 3. This time, when human capital from both schooling and experience is 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 $\log(\text{GDP})$ 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 cap-

ital for development accounting: between 0.66 and 0.74 by the first criterion and between 0.76 and 0.81 by the second criterion. 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.

VI. 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 1998, 2000; Caselli 2005). Yet more recent work claims that human capital plays a much more central role (Manuelli and Seshadri 2015; Lagakos et al. 2018).

To address this question, this paper draws on evidence from US 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 who did not migrate, whereas immigrants from richer countries are only modestly more educated than nonmigrants 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 those of nonmigrants, which suggests similar skill loss across countries. Finally, the fact that returns to experience are similar between immigrants and nonmigrants, 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) does not 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 US 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 in rich countries. One possible explanation is that the quantity and type of schooling result 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 US experience among foreign-educated workers are lower than the returns to US experience for natives. At the same time, we have also documented that the returns to US experience among US-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.

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

$$\log(w^{\mathrm{N}}) = \beta^{\mathrm{N}} X^{\mathrm{N}} + \phi^{\mathrm{N}} E^{\mathrm{N}} + \omega^{\mathrm{N}} Y^{\mathrm{N}} + \mu^{\mathrm{N}} S^{\mathrm{N}} + \varepsilon^{\mathrm{N}}, \tag{A1}$$

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. Like Chiswick (1978) and Schoellman (2012), we want to allow the return to foreign-acquired schooling to differ from that to US-acquired schooling. We also want to distinguish between the return to foreign (birth-country) and domestic (US) experience for immigrants, FE and DE, respectively. We 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 has 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 United States. Finally, some authors have suggested allowing a role for years in the United States, YUS, to capture the assimilation of immigrants. Combining all of these potential factors would suggest a Mincerian regression equation for immigrants:

$$\begin{aligned}\log(w^{\mathrm{I}}) = {} & \beta^{\mathrm{I}} X^{\mathrm{I}} + \phi_1^{\mathrm{I}} \mathrm{FE}^{\mathrm{I}} + \phi_2^{\mathrm{I}} \mathrm{DE}^{\mathrm{I}} + \omega^{\mathrm{I}} Y^{\mathrm{I}} + \mu^{\mathrm{I}} S^{\mathrm{I}} \\ & + \alpha^{\mathrm{I}} \mathrm{AA}^{\mathrm{I}} + \gamma^{\mathrm{I}} C^{\mathrm{I}} + \delta^{\mathrm{I}} \mathrm{YUS}^{\mathrm{I}} + \varepsilon^{\mathrm{I}},\end{aligned} \tag{A2}$$

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 (A2) 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 with linear combinations of four survey questions: age, years of schooling, data set year, and year of immigration YI^{I}. Years of schooling and data set year enter the regression equation directly; five other variables in that equation are linear combinations of these four survey questions:

1. $\mathrm{FE}^{\mathrm{I}} = A^{\mathrm{I}} - S^{\mathrm{I}} - 6 - (Y^{\mathrm{I}} - \mathrm{YI}^{\mathrm{I}})$;
2. $\mathrm{DE}^{\mathrm{I}} = Y^{\mathrm{I}} - \mathrm{YI}^{\mathrm{I}}$;
3. $\mathrm{AA}^{\mathrm{I}} = A^{\mathrm{I}} - (Y^{\mathrm{I}} - \mathrm{YI}^{\mathrm{I}})$;
4. $C^{\mathrm{I}} = \mathrm{YI}^{\mathrm{I}}$;
5. $\mathrm{YUS}^{\mathrm{I}} = Y^{\mathrm{I}} - \mathrm{YI}^{\mathrm{I}}$.

Equation (A2) 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 $\omega^{\mathrm{N}} = \omega^{\mathrm{I}}$. 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 US experience and years in United States are, in fact, defined in the same manner. In this case, it is impossible to identify separately the effect of US potential experience from any other, more general effects of spending time in the United States, 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 US 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, $S^{\mathrm{I}} + 6$. 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 as a result of, say, adaptability. However, given that our estimated experience effects for immigrants look strikingly similar to those estimated in Lagakos et al. (2018) for nonmigrants, we believe that our results support assigning a dominant role to experience human capital.

Appendix B

Ben-Porath Model of Human Capital Accumulation

In this appendix, we extend the simple learning-by-doing model of Section III 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 V.

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 US experience. We denote variables observed in the United States with asterisks 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 United States 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 United States, 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. The value of 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 United States, B_c^*, to vary across countries so that it matters where an individual is born, even after migration to the United States.

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 US wage. At some level of home-country experience x, workers from country c migrate to the United States. 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.[16] 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 III, 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, one can also extend the model to feature "skill loss" upon migration. All the insights discussed there would go through.

An individual who lives in his home country solves $\max_{\{\ell_c(t)\}}\int_0^T e^{-rt}w_c(t)dt$ such that

$$\begin{aligned} w_c(t) &= \omega(1 - \ell_c(t))h_c(t), \\ h_c(t) &= \phi_c(\ell_c(t))h_c(t) - \delta h_c(t), \\ 0 &\leq \ell_c(t) \leq 1, \end{aligned} \tag{P}$$

[16] This assumption can be relaxed at the expense of some extra notation.

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 that 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 United States, he thereafter solves $\max_{\{\ell_c^*(t)\}} \int_x^T e^{-r(t-x)} w_c^*(t, x)\, dt$ such that

$$\begin{aligned} w_c^*(t, x) &= \omega\big(1 - \ell_c^*(t)\big) h_c^*(t), \\ h_c^*(t) &= \phi_c^*\big(\ell_c^*(t)\big) h_c^*(t) - \delta h_c^*(t), \\ h_c^*(x) &= h_c(x), \\ 0 \le \ell_c^*(t) &\le 1. \end{aligned} \tag{P^*}$$

One useful feature of our Ben-Porath model is that it can be mapped very directly to the empirical model in Section II, in particular Mincer type regressions such as equation (1). To see this, note that the wage of a worker right before migration satisfies

$$\begin{aligned} \log w_c(x) &= \log \omega_c + \log(1 - \ell_c(x)) + \log h_c(x), \\ \log h_c(x) &= \int_0^x (\phi_c(\ell_c(t)) - \delta)\, dt. \end{aligned}$$

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

$$\begin{aligned} \log w_c(x) &= \alpha_c + R_c(x), \\ R_c(x) &= L_c(x) + \log h_c(x), \end{aligned}$$

where $L_c(x) = \log[(1 - \ell_c(x))/(1 - \ell_c(0))]$ reflects the nonmigrant's time allocation decision. Hence, measured nonmigrant returns to experience, $R_c(x)$, reflect both the nonmigrant'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 United States who immigrates with x years of home-country experience and accumulates x^* years of US experience (and so has $x + x^*$ years of total experience). Given our assumptions, this wage can be written as

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

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

$$\begin{aligned} \log w_c^*(x,0) &= \alpha_c^* + R_c^*(x,0), \\ R_c^*(x,0) &= L_c^*(x) + \log h_c^*(x,0), \end{aligned} \tag{B1}$$

where $L_c^*(x) = \log[(1 - \ell_c^*(x))/(1 - \ell_c^*(0))]$. Similarly to those for a nonmigrant 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 in Sec. B.3) will be useful below.

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

$$\ell_c^*(t) = \min\left\{1, \left(\sigma\lambda_c^*(t)B_c^*\right)^{1/(1-\sigma)}\right\}, \tag{B2}$$

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

$$\begin{aligned} \dot{\lambda}_c^*(t) &= \left(r + \delta - B_c^*\ell_c^*(t)^\sigma\right)\lambda_c^*(t) - \left(1 - \ell_c^*(t)\right), \\ \lambda_c^*(T) &= 0. \end{aligned} \tag{B3}$$

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 United States, her time allocation decision and wage path are defined uniquely from her productivity of human capital production while in the United States, 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 V 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 $H_c = h_c(\bar{x}_c, \bar{s}_x)$ where $\bar{x}_c$ and $\bar{s}_c$ are a country's average experience and schooling attainments, respectively.

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 regression (B1) 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 fig. 8):

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

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 $h_c^*(\bar{x}_c)$, that is, to assume that the experience-wage profile is not affected by the time allocation changing over the life cycle. From equation (B1), since time allocation is declining over the life cycle, $\ell_c^*(x) \leq \ell_c(0)$ for $x > 0$, 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 $\mathrm{Cov}(\log H_c, \log Y_c) \leq \mathrm{Cov}(R_c^*(\bar{x}_c, 0), \log Y_c)$, and hence,

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

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 US experience; the second are 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

$$\begin{aligned} \log w_c^*(e, 0) &= \alpha_c^* + R_c^*(e, 0), \\ R_c^*(e, 0) &= L_c^*(e) + \log h_c^*(e, 0), \end{aligned} \tag{B4}$$

$$\begin{aligned} \log w_c^*(0, e) &= \alpha_c^* + R_c^*(0, e), \\ R_c^*(0, e) &= L_c^*(e) + \log h_c^*(0, e). \end{aligned} \tag{B5}$$

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, as a result of lemma 1, the term reflecting the time allocation, $L_c^*(e)$, is identical for the two types of individuals.

Lemma 2. We have that

$$\beta = \frac{\mathrm{Cov}(\log H_c, \log Y_c)}{\mathrm{Var}(\log Y_c)} \geq \frac{\mathrm{Cov}\big(R_c^*(\bar{x}_c, 0) - R_c^*(0, \bar{x}_c), \log Y_c\big)}{\mathrm{Var}(\log Y_c)}, \tag{B6}$$

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. From the two regressions (B4) and (B5), 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 nonmigrants and those who migrated straight after high school.

Next, we explain how this can be used to obtain the lower bound (eq. [B6]). To fix ideas, consider first the case $B_c^* = B^*$ for all c, that is, immigrants from all countries learn at the same pace once they arrive in the United States. In that case, the human capital stock of individuals who spent their entire working life in the United States, $h_c^*(0,e)$, does not depend on their country of origin c. Therefore, $\mathrm{Cov}(\log h_c^*(0,e), \log Y_c) = 0$ for all e, and hence $\mathrm{Cov}(R_c^*(e,0) - R_c^*(0,e), \log Y_c) = \mathrm{Cov}(\Delta \log h_c(e), \log Y_c) = \mathrm{Cov}(\log h_c(e), \log Y_c)$. 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, equation (B6) would not be only a lower bound on the regression coefficient but also 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 that $\mathrm{Cov}(\log h_c^*(0,e), \log Y_c) \geq 0$. In that case, one can obtain only a lower bound:

$$\begin{aligned}\mathrm{Cov}\big(R_c^*(e,0) - R_c^*(0,e), \log Y_c\big) &= \mathrm{Cov}\big(\log h_c(e) - \log h_c^*(0,e), \log Y_c\big)\\ &\leq \mathrm{Cov}(\log h_c(e), \log Y_c).\end{aligned}$$

Using that $H_c = h_c(\bar{x}_c)$, we obtain equation (B6). QED

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 conditions for optimality are

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

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

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

and we note that h cancels out in the optimality equation and so the solution is fully characterized by equations (B2) and (B3).[17]

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

$$\dot{\ell} = \left(\frac{1}{1-\sigma}\right)\{(r+\delta)\ell - [(1-\sigma)B]\ell^{\sigma+1} - \sigma B\ell^{\sigma}\}, \tag{B7}$$

together with the terminal condition $\ell(T) = 0$. From equation (B7), we have $\partial\dot{\ell}(t)/\partial B < 0$ for all t. Given the terminal condition $\ell(T) = 0$, therefore, $\partial\ell(t)/\partial B > 0$ 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, individuals 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 (B7) that defines the path for $\ell(t)$ is independent of $h(t)$. QED

References

Bils, Mark, and Peter J. Klenow. 1998. "Does Schooling Cause Growth or the Other Way Around?" Working Paper no. 6393 (February), NBER, Cambridge, MA.

———. 2000. "Does Schooling Cause Growth?" *A.E.R.* 90 (5): 1160–83.

Borjas, George J. 1985. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." *J. Labor Econ.* 3 (4): 463–89.

Caron, Justin, Thibault Fally, and James R. Markusen. 2014. "International Trade Puzzles: A Solution Linking Production and Preferences." *Q.J.E.* 129 (3): 1501–52.

Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." In *Handbook of Economic Growth*, vol. 1A, edited by Philippe Aghion and Steven N. Durlauf, 679–741. Amsterdam: Elsevier.

Chiswick, Barry R. 1978. "The Effect of Americanization on the Earnings of Foreign-Born Men." *J.P.E.* 86 (5): 897–921.

Coulombe, Serge, Gilles Grenier, and Serge Nadeau. 2014. "Quality of Work Experience and Economic Development: Estimates Using Canadian Immigrant Data." *J. Human Capital* 8 (3): 199–234.

Ehrlich, Isaac, and Jinyoung Kim. 2015. "Immigration, Human Capital Formation, and Endogenous Economic Growth." *J. Human Capital* 9 (4): 518–63.

Erosa, Andrés, Tatyana Koreshkova, and Diego Restuccia. 2010. "How Important Is Human Capital? A Quantitative Theory Assessment of World Income Inequality." *Rev. Econ. Studies* 77 (4): 1421–49.

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

Grogger, Jeffrey, and Gordon H. Hanson. 2011. "Income Maximization and the Selection and Sorting of International Migrants." *J. Development Econ.* 95 (1): 42–57.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Q.J.E.* 114 (1): 83–116.

Hanushek, Eric A., and Dennis D. Kimko. 2000. "Schooling, Labor-Force Quality, and the Growth of Nations." *A.E.R.* 90 (5): 1184–1208.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2008. "Earnings Functions and Rates of Return." *J. Human Capital* 2 (1): 1–31.

[17] This is due to the assumption of constant returns in the human capital accumulation technology.

Hendricks, Lutz. 2002. "How Important Is Human Capital for Development? Evidence from Immigrant Earnings." *A.E.R.* 92 (1): 198–219.

Hendricks, Lutz, and Todd Schoellman. 2016. "Human Capital and Development Accounting: New Evidence from Wage Gains at Migration." *Q.J.E.* 133 (2): 665–700.

Hsieh, Chang-Tai, and Peter J. Klenow. 2010. "Development Accounting." *American Econ. J.: Macroeconomics* 2 (1): 207–23.

Klenow, Peter J., and Andrés Rodríguez-Clare. 1997. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" In *NBER Macroeconomics Annual 1997*, vol. 12, edited by Ben S. Bernanke and Julio Rotemberg. Cambridge: MIT Press.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2018. "Life Cycle Wage Growth across Countries." *J.P.E.* 126 (2): 797–849.

Lemieux, Thomas. 2006. "The 'Mincer Equation' Thirty Years after *Schooling, Experience, and Earnings*." In *Jacob Mincer: A Pioneer of Modern Labor Economics*, edited by Shoshana Grossbard, 127–45. New York: Springer.

Lewis, Jeffrey B., and Drew A. Linzer. 2005. "Estimating Regression Models in Which the Dependent Variable is Based on Estimates." *Polit. Analysis* 13 (4): 345–64.

Manuelli, Rodolfo E., and Ananth Seshadri. 2015. "Human Capital and the Wealth of Nations." *A.E.R.* 104 (9): 2736–62.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings*. New York: Columbia Univ. Press.

Minnesota Population Center. 2015. "Integrated Public Use Microdata Series–International," version 6.2. Minneapolis: Minnesota Population Center, Univ. Minnesota. https://international.ipums.org/international.

Moraga, Jesús F.-H. 2011. "New Evidence on Emigrant Selection." *Rev. Econ. and Statis.* 93 (1): 72–96.

Office of Policy and Planning, US Immigration and Naturalization Service. 2003. "Estimates of the Unauthorized Immigrant Population Residing in the United States: 1990 to 2000." https://www.dhs.gov/sites/default/files/publications/Unauthorized%20Immigrant%20Population%20Estimates%20in%20the%20US%201990%20to%202000.pdf.

Psacharopoulos, George. 1994. "Returns to Investment in Education: A Global Update." *World Development 22*(9), 1325–43.

Reinhold, Steffen, and Kevin Thom. 2013. "Migration Experience and Earnings in the Mexican Labor Market." *J. Human Resources* 48 (3): 768–820.

Ruggles, Steven, Katie Genadek, Josiah Grover, and Matthew Sobek. 2015. "Integrated Public Use Microdata Series," version 6.0. Minneapolis: Univ. Minnesota.

Schoellman, Todd. 2012. "Education Quality and Development Accounting." *Rev. Econ. Studies* 79 (1): 388–417.

World Bank. 2015. "World Development Indicators." http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators.