NBER WORKING PAPER SERIES

WHY ARE SOME IMMIGRANT GROUPS MORE SUCCESSFUL THAN OTHERS?

Edward P. Lazear

Working Paper 23548
http://www.nber.org/papers/w23548

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2017

Why Are Some Immigrant Groups More Successful than Others?
Edward P. Lazear
NBER Working Paper No. 23548
June 2017
JEL No. F22,J01,J15,J61,M5

ABSTRACT

Success, measured by earnings or education, of immigrants in the US varies dramatically by country of origin. For example, average educational attainment among immigrants ranges from 9 to 16 years, depending on source country. Perhaps surprisingly, immigrants from Algeria have higher educational attainment than those from Israel or Japan. Also true is that there is a strong inverse relation of attainment to number of immigrants from that country. These patterns result because in the US, immigrant slots are rationed. Selection from the top of the source country's ability distribution is assumed and modeled. The main implications are that average immigrant attainment is inversely related to the number admitted from a source country and positively related to the population of that source country. The results are unequivocally supported by results from the American Community Survey. Additionally, a structural model that is more explicit in the assumptions and predictions fits the data well.

Edward P. Lazear
Graduate School of Business
Stanford University
Stanford, CA 94305
and Hoover Institution
and also NBER
lazear@stanford.edu

Algeria, Israel, and Japan, along with over one hundred other countries, are sources of immigrants in the United States. Try the following thought experiment: Rank those three countries' immigrants, highest to lowest, by educational attainment. The ranking is Algeria, Israel, and Japan. Surprised? Consider an additional fact: Algerians make up .0004 of immigrants, Israelis comprise .003 of our immigrants whereas about 1% of immigrants are from Japan. The largest source country of immigrants is Mexico, accounting for 27% of the immigrant population in the US. Mexican immigrants rank 134th out of 136 in educational attainment, as compared with Algerian immigrants who rank 25th . Yet, average education attainment in Mexico is 8.5 years, whereas in Algeria it is only 7.6 years. The group with the highest educational attainment are those from the former Soviet Union, who make up .001 of all immigrants.

The attainment of immigrants in the US varies greatly by country of origin. Average educational attainment by country of origin ranges from a low of 9 years of schooling to a high of 16 years. Similarly, average annual earnings by country of origin ranges from about $16,000 to $64,000. Not surprisingly, the correlation between these two measures of attainment across 134 source countries is .7 . Additionally, two of the largest sources of US immigrants are at both extremes, with Mexico's migrants to the US having a mean educational attainment of 9 years, whereas India's migrants to the US have a mean of 16 years. Contrast that with the fact that the average educational level in Mexico is almost twice that of the average educational attainment in India. What explains these immigrant attainment differences across origin countries and the counterintuitive patterns?

Because the US admits as immigrants only a small fraction of most origin countries' populations, almost every country has a large enough group of highly educated people from which the US could draw its immigrants. Many of these individuals might be willing to move to the US if admitted, which implies that the average educational attainment of immigrants from any source country depends in large part on whom the US is willing to admit. The more selective is immigration policy, the higher is the educational attainment of the group.

Few Algerians ever obtain permission to come to the United States and those who do tend to be highly educated. In contrast, most Mexican immigration is based on family reunification. Although family reunification is a worthy goal, a selection system based primarily on that factor will not generate a population of immigrants with the highest levels of education.

Rather than modeling US immigration as driven by supply considerations such as the wage that an individual can earn in the US relative to what he or she earns at home, the focus here is on the rationing mechanism. Most models of immigration treat migrants as if they are mobile labor, moving from one sector to another freely, without any constraints on the migrant imposed by policy.[1] These models have served analysts well in considering who chooses to come to the US, but they are less well-suited to describing the composition of the immigrant population and its educational attainment and wage levels.

Supply factors that would determine who would move from one occupation or industry to

[1]Models, such as the one by Roy (1951) are frequently used to describe the flow of immigrants to the US. See, for example, the seminal work by Borjas (1987).

another in an open economy with free mobility are de-emphasized (although not ignored) here because would-be immigrants do not have free choice over whether they are admitted to a country. Models that assume that individuals are free to choose the country to which they migrate omit an important consideration, namely, that immigration slots are rationed in many countries and in the United States, in particular. For example, between 2009 and 2014, approximately 1 million individuals per year were granted permanent resident status. In each of those years, there was a large number of applicants who were in the queue for resident status, equaling about four times as many as the number granted permanent residency.[2] There is excess supply of immigrants, even by measures of those who apply. There are surely many more who would apply if they thought they would be admitted, but decline to do so because the likelihood that their application will succeed is too low.

To explain the attainment of immigrants, another extreme approach is adopted, albeit a caricature of the true situation. The analysis here emphasizes rationing. It is assumed that anyone who is offered admission to the US from another country accepts that offer and migrates to the US. The shape of the education distribution in the origin country, the country's population, and importantly, the target immigrant number from that country determine the composition and attainment of immigrants in the US. Supply considerations enter primarily by affecting the educational distribution that prevails in the immigrant's country, but are also relevant in providing a rationale for the assumption that the most talented individuals from each origin country are selected by the US.[3]

Although the point of this analysis is not unique to the United States, the focus is the US because the assumption that the selection filter, rather than supply considerations, are paramount is more applicable to the US than to most other countries. The excess supply of immigrants to the US means that the rationing rule is a key factor in determining the nature of immigrants in the US.

---

[2]U.S. Department of State, Bureau of Consular Affairs, "Annual Immigrant Visa Waiting List Report as of November 1, 2015," 2015, and Department of Homeland Security, "Yearbook of Immigration Statistics 2014," 2016.

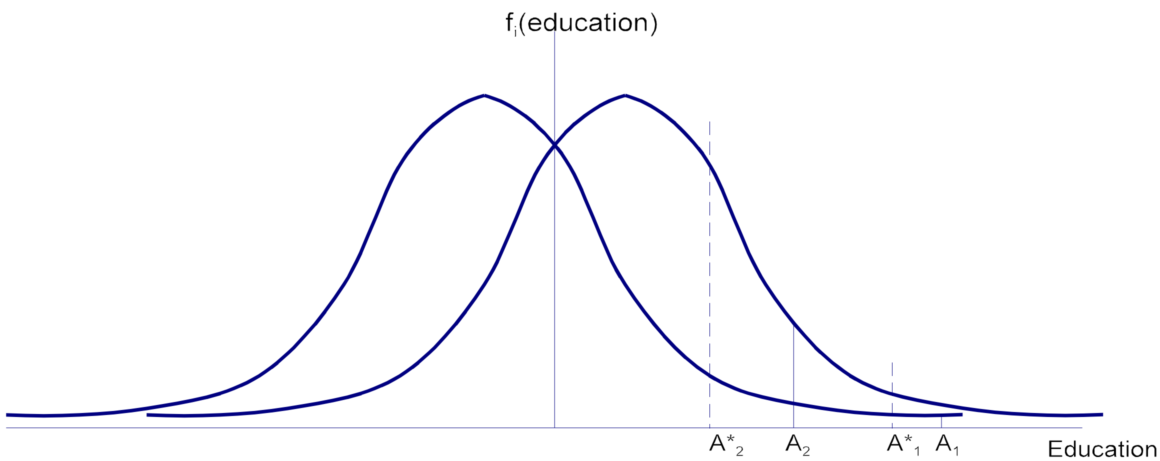
[3]Borjas and Friedberg (2009) argue that selection rules explain the rise in wages of immigrants that occurred between 1990 and 2000. Additionally, Kerr and Lincoln (2010) examine the effect of the rise in H-1B visas on innovation and document its importance. Kerr, Kerr, and Lincoln (2015) use the LEHD data and find that a skilled immigrant labor is an important determinant of the overall increase in the skilled workforce of American firms. H-1Bs are important because the H-1B program is always over-subscribed and the number of H-1Bs permitted is a policy choice that results in selection from the top, made by the US government.

Two Hunt papers, Hunt (2011) and Hunt (2015), examine the performance of highly skilled immigrants in the US. Another measure of ability to perform in the US relates to English fluency, which was explored in Lazear (1999). Lewis (2013) picks up on that theme and discusses the ability of immigrants to substitute for native-born US labor and how that relates to language skills. Similarly, Peri, Shih and Sparber (2015) estimate the effect of STEM immigration on productivity and find it is substantial, using city differences. None of these papers speaks to the empirical validity of the assumption that selection is from the top of an origin country's distribution, but the fact that many of the origin countries have low educational levels implies that the high skilled ones are selected from the top.

Figure 1 illustrates the main argument. Consider two countries, 1 and 2, with equal population sizes and with educational distributions as shown. Country 2's distribution is a rightward displacement of country 1's distribution.

Suppose that the US were to decide that 3% of its immigrants will come from country 1, but that 30% will come from country 2. If the US also allows only the most educated immigrants in first from each country, or alternatively, if the most educated in each country are most attracted to the US or most able to negotiate their way through the immigration process, then the upper tail of each distribution will end up migrating to the US. In this case, because the US targets 3% from country 1 and 30% from country 2, the minimum cutoff level of education in each country is A_1* and A_2*, respectively. Note that the educational cutoff level for country 1, the lower education country, is considerably above that for country 2, the higher education

Figure 1



country, because so many more are being admitted from country 2. Given the cutoffs and the underlying distributions, the average level of education among immigrants from country 1 and country 2 are A_1 and A_2. The educational attainment of immigrants from 1 exceeds that from 2, i.e., $A_1 > A_2$, even though country 1's education level at home is below that of country 2 at home.

Of course, this is not a necessary outcome. It depends on the amount by which country 2's education level dominates country 1 and in particular on the number of immigrants that the US admits from each of the two countries. But figure 1 illustrates that other things equal, the smaller the proportion of immigrants in the US who come from a country, the higher is the expected level of education of immigrants in the US who are supplied by that country.

Model

The general model captures the intuition of the figure and discussion. Suppose that the

US chooses a selection rule such that I_i of the immigrants have origins in country i. The selection rule is taken to be exogenous, determined by policy, politics, or considerations outside the model.

Assume that anyone outside the US offered immigrant status in the US accepts it. Assume additionally that rationing is such that the top of the educational (or any other dimension of immigrant ability) is admitted first. This can either be a result of explicit US policy or a consequence of the supply side, where the people most likely to come to the US from any other country are at the top of the ability distribution, the latter resulting either because they are best able to navigate the immigrate hurdles or because they have the highest return from migrating. This policy-determined selection of immigrants fits those who come in legally with official documentation. Those who enter the country without documentation are more likely to fit a strict supply-determined mechanism because the rationing rule chosen by the US does not bind.[4]

Let N_i be the population of country i and let $f_i(A)$ be the density of education or some other measure of ability or attainment, A, in country i. Then, A_i^* is the cutoff ability level of immigrants from country i determined such that

$$N_i \int_{A_i^*}^{\infty} f_i(A)dA = I_i$$

or

$$N_i[1 - F_i(A_i^*)] - I_i = 0 \tag{1}$$

Equation (1) determines the cutoff level A_i^*. Given this cutoff for immigrants, the expected level of education among those from country i in the US is simply the conditional expectation or

$$\overline{A}_i = \frac{1}{1 - F_i(A_i^*)} \int_{A_i^*}^{\infty} A f_i(A)dA \tag{2}$$

The goal is to predict the effect of the key variables on the average educational level or other measure of attainment of a country's migrants to the US. The key variables are the number of immigrant slots allocated to country i, I_i, the population of country i, N_i, and the level of education in country i, μ_i, where μ_i is defined as the average level of education in source country i .

To do this, differentiate (2) with respect to I_i, N_i, and μ_i. In general, from (2), for any variable x,

[4]It is possible that the policy on internal enforcement and border control may have an effect even on those who do not apply through legal channels.

$$\frac{\partial \overline{A}_i}{\partial x} = \frac{f_i(A_i^*)\dfrac{\partial A_i^*}{\partial x}}{[1-F_i(A_i^*)]^2}\int_{A_i^*}^{\infty} Af_i(A)dA$$
$$+\frac{1}{1-F_i(A_i^*)}\int_{A_i^*}^{\infty} A\frac{\partial f_i(A)}{\partial x}dA$$
$$-\frac{A_i^* f_i(A_i^*)}{1-F_i(A_i^*)}\frac{\partial A_i^*}{\partial x}$$

or

$$\frac{\partial \overline{A}_i}{\partial x} = \frac{f_i(A_i^*)\dfrac{\partial A_i^*}{\partial x}}{[1-F_i(A_i^*)]}\left[\frac{\int_{A_i^*}^{\infty} Af_i(A)dA}{1-F_i(A_i^*)} - A_i^*\right]$$
$$+\frac{1}{1-F_i(A_i^*)}\int_{A_i^*}^{\infty} A\frac{\partial f_i(A)}{\partial x}dA \qquad (3)$$

The derivation in (3) allows the basic theoretical predictions to be stated. This is done in the form of propositions, the proofs of which are contained in the appendix.

Proposition 1:

$$\frac{\partial \overline{A}_i}{\partial I_i} < 0 \ .$$

Increasing the number of immigrants admitted from country i, I_i, lowers their expected level of attainment, $\overline{A}_i$.

Proposition 2:

$$\frac{\partial \overline{A}_i}{\partial N_i} > 0 \ .$$

For any given number of immigrants, I_i, the larger is the population in country i, the higher is the expected level of attainment of immigrants $\overline{A}_i$ from that country.

Finally, let $F_i(A) = F(A-\mu_i)$ so that every country's ability distribution is of the same form, but merely displaced by country-specific parameter μ_i. Countries with higher μ_i values have higher ability distributions. Assume further that f '(A) is negative for all $A \geq A_i^*$. This is likely to hold empirically under the assumption that the most able are taken first because there is no country that provides so many immigrants that the cutoff ability, A_i^*, would not be in the upper tail of the ability distribution, which is expected to be negatively sloped. Then,

Proposition 3:

$$\frac{\partial \overline{A}_i}{\partial \mu_i} > 0$$

As educational or other attainment in the source country rises, expected attainment among immigrants from that country also rises.

It is also possible to express the concepts in propositions 1-3 in terms of the "representation ratio" defined as

$$R_i \;/\; \frac{I_i / \sum I_i}{N_i / \sum N_i} \quad .$$

R_i should be interpreted as the over-representation of country i among immigrants, given the country's relative population importance. If R_i equals 1, the proportion of immigrants in the US from country i reflects its weight in the overall population of the world. If R_i exceeds 1, that country is over-represented among US immigrants. If R_i is less than 1, then country i is under-represented among immigrants in the US.

India is under-represented among immigrants, despite the fact that Indians are the third largest group of immigrants (behind Mexicans and Filipinos). In contrast, Jamaicans are over-represented, making up over forty times the number of immigrants as would be expected given Jamaica's population even though there are only one-third as many immigrants from Jamaica as there are from India.

It is then possible to state a corollary to propositions 1 and 2 in terms of the representation ratio:

Corollary 1:

$$\frac{\partial \overline{A}_i}{\partial R_i} < 0$$

Increasing a country's representation ratio, R_i, lowers the expected level of attainment $\overline{A}_i$.

Additional theoretical predictions can be derived. The model provides implications not only for immigrants and their relative standing in the recipient country, in this case the US, but also for their situation vis à vis the general population of the source countries. For these purposes, define A as referring to and only to levels of attained education. Define

$\Delta \;/\; E_i(A \mid A>A_i^*) - E_i(A)$

where E_i is the expected level of education within country i, given the distribution of education $f_i(A)$ in country i. Then Δ is interpreted as the difference between the attained education of immigrants in the US from country i and the average level of education of the overall population in country i.

Recall that $\bar{A}_i \equiv E_i(A \mid A>A_i^*)$ so

$$\Delta \equiv \bar{A}_i - \int_{-\infty}^{\infty} A f_i(A)\,dA$$

The empirically verifiable implications are stated as corollaries here and proved in the appendix.

Corollary 2: $\partial\Delta/\partial I_i < 0$.

The difference between the mean education of immigrants from source country i and the mean education of the population of that country falls in I_i.

Corollary 3: $\partial\Delta/\partial N_i > 0$.

The difference between the mean education of immigrants from source country i and the mean education of the population of that country rises in N_i.

Corollary 4: $\partial\Delta/\partial\mu_i = 0$.

Under the assumptions above, a shift in the mean of the source country's education distribution is neutral, having no effect on the difference between the mean education of immigrants from that country and the mean education in the source country's overall population.

Equivalently, if $I_i = I_j$ and $N_i = N_j$, then the difference between the mean level of education among immigrants from country i and country j equals the difference between the mean level of education in the origin country's overall population.

Discussion

The logic behind the propositions and corollaries fits the example in the introduction and, as will be shown below, more generally the data on immigrant background and educational attainment and earnings in the US.

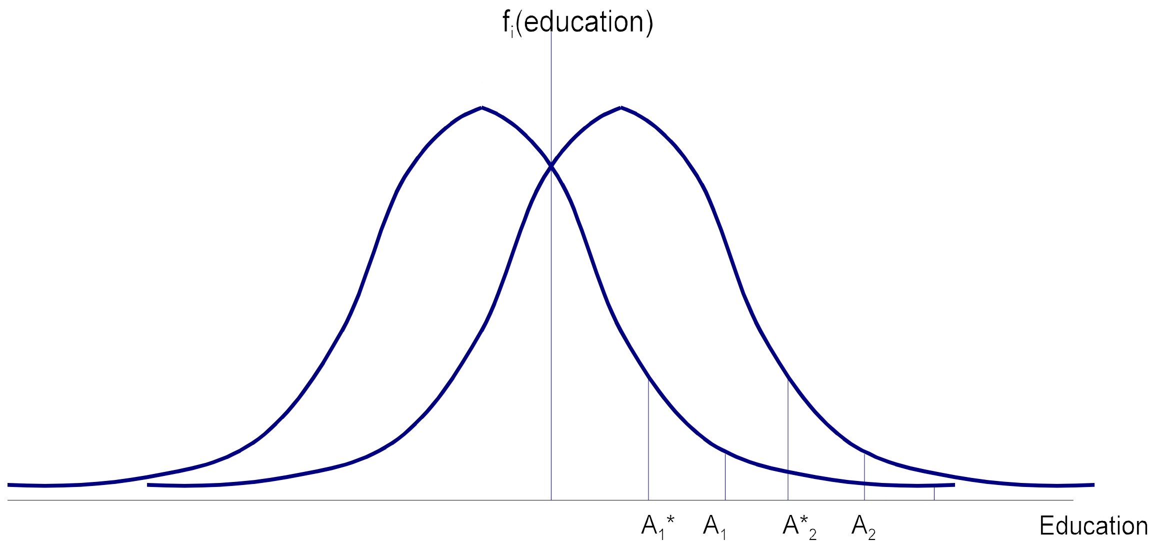
Proposition 1 predicts the basic point made at the outset. When a country is permitted to send only a small number of immigrants and when selection is from the top down, those who enter the US will be the most talented. The group with the highest level of educational attainment are those who came from the USSR. Note that this is "USSR," not Russia, which means that they entered before the breakup of the Soviet Union. Among the pool of immigrants in the sample, there were only 400 from the former Soviet Union as compared with over 92,000 from Mexico, the largest source of US immigrants. Those who entered from the USSR were a rare group, needing to obtain both exit permission from the USSR, as well as entry permission from US. A large proportion were highly-educated political dissidents, many of whom were elite academics. The same is true perhaps to a lesser extent of other countries, like Algeria. The average level of education in Algeria is well below the mean for source countries, but those who succeed in moving to the US were not typical Algerians. Instead, they were more educated than their compatriots, so much so that as a origin country, Algeria is in the top 20% of immigrant groups to the US in educational attainment. Algerians make up less than .0005 of our immigrants

- a tiny fraction - and those who have been admitted to the US have been selected for reasons that correlate well with education. The same educational attainment would not likely be found of Algerian migrants to France, where they make up a much larger fraction of the immigrant population.

Proposition 2 is slightly more subtle, but almost equally intuitive. Consider selecting the most highly educated 100,000 people from a tiny country like Laos with 7 million people versus 100,000 from India with 1.3 billion. If the distribution of underlying ability were the same in both countries, A* for India would be higher than A* for Laos because 100,000 people comprise a much smaller fraction of the top tail of the distribution when there are 1.3 billion than when there are 7 million. In fact, that is what is seen in the data. A substantial fraction of our immigrants come from India, comprising almost 5% of the sample of immigrants. But because India is so large, India is very much under-represented (by a factor of four), given its importance in world population. As such, those who come from India are from the top part of India's educational distribution. India itself is not a country with a high average level of education. Only 14 countries in the world have lower average levels of education than India has. However, because US immigration policy does not select randomly from origin countries nor from individuals within each country, Indians in the US rank second among immigrants in educational attainment.

Proposition 3 makes the more intuitive point that the more highly educated are people in the home country, the more educated are those who immigrate from them. Consider two countries with exactly the same populations and that make up exactly the same proportion of our immigrant pool. If the ability distribution from country 2 lies to the right of that from country 1, then selecting the same fraction of the upper tail from each results in a higher average level of education among those selected. This is shown in Figure 2. Two countries, 1 and 2, have similarly shaped distributions of talent, but country 2's distribution is a rightward shift of country 1's distribution. Consequently, A_2* exceeds A_1* by the difference in their means. Furthermore, as is the subject of corollary 4, $A_{2,}$ defined as the expectation of ability among immigrants from 2, exceeds A_1, defined as the expectation of ability among immigrants from 1, by the difference in their means. This is equivalent to the statement that A_2 exceeds the mean of distribution 2 by the same amount that A_1 exceeds the mean of distribution 1.

Figure 2



Corollary 1 simply states propositions 1 and 2 in another intuitive form. A country can be over- or under-represented among immigrants. When a country is over-represented, the cutoff level of ability from that population must be lower than it would be were that same country under-represented among immigrants.

Corollaries 2, 3, and 4 follow directly from the model, but empirically, they are independent tests of the model's logic. The dependent variable in the corollaries is Δ_i, which is not the same as $\overline{A}_i$. The average attainment among immigrants from country i, $\overline{A}_i$ could be high relative to that from country j, $\overline{A}_j$, but that does not imply mechanically that $\overline{A}_i$ is high relative to the average level of education in country i. For example, immigrants from the United Kingdom have average schooling attainment of 15 years, above the immigrant mean. But the average level of education in the UK is well above the mean education for other countries in the world, so the value of Δ for the UK is low. Conversely, immigrants in the US from Yemen have only 9 years of schooling, while the average level of education in Yemen is only 2.5 years, resulting in a high value of Δ. Obviously, as a statistical matter, Δ_i and $\overline{A}_i$ are related because the former contains the latter, but they are not the same. Corollaries 2-4 provide additional predictions that can be tested and could be rejected, even were propositions 1-3 to hold empirically. Furthermore, note that the predicted relation of Δ to μ_i is zero, not negative as would result from pure statistical bias.

In the analysis below, attainment among immigrants is measured by three different variables, namely average education, hourly wages among working immigrants, and average earnings among all immigrants from a particular country.

Without exception, all predictions of propositions 1 through 3 and corollary 1 are borne out for all three measures. Corollaries 2-4 are found to hold with respect to educational attainment, but cannot be tested for the two income variables because the data do not contain information on income for origin countries.

Specific Distributions of Ability and Implied Structural Estimates

Under more specific assumptions, the propositions and corollaries stated above can be parameterized. This approach has the advantage that it allows for interpretation of estimates, but more important, it provides additional checks on the credibility of the model and its assumptions.[5]

To begin, suppose that each origin country has a normal distribution of A with mean μ_i and variance σ^2. Recall that I_i is the exogenously determined policy variable. Let F_i* denote the cumulative normal distribution with mean μ_i and and variance σ^2. Further, let f_i* be the normal density with mean μ_i and and variance σ^2. Then, from (1), A_i* can be determined as

(1') $N_i\,[1-F_i*(A_i*)] - I_i = 0$

and similarly

(2') $$\bar{A}_i = \frac{1}{1-F_i*(A_i*)}\int_{A_i*}^{\infty} Af_i*(A)dA$$

If all the μ_i and σ were known, the exact $\bar{A}_i$ for each country i would be determined, given μ_i . That is, once the country's ability distribution is known precisely and once A_i* is given, it is simple work to compute the conditional expectation of those who are above A_i*. Then a goodness-of-fit statistic can be computed by checking to see how well the actual $\bar{A}_i$ compare to those predicted based on the observed μ_i of the origin country and σ.

The data below provide an average attained level of education for the 129 countries studied so that μ_i is observable. The only missing parameter is σ but this can be obtained by an iterative approach where a σ is selected, the $\bar{A}_i$ are estimated and then are compared to the actual values of attained education among immigrants in the US. The σ is chosen so as to

[5]I ask the reader to excuse my self-indulgence. One of (if not the) first structural estimation approaches in labor and applied microeconomics appeared in my 1977 paper on a different topic. See Lazear (1977).

maximize goodness-of-fit.

Data

The data are taken from the American Community Survey from years 2011 through 2015. This is a series of five consecutive cross-sectional data sets. By combining years, a larger sample is created so that more precision can be obtained. It is straightforward to adjust standard errors for population weights and to correct wage data for inflation to turn nominal wages and earnings into real values.

The variables of interest are those that measure attainment of the immigrants once in the US. They include educational attainment, wages, and income. The independent variables are internally constructed and also drawn from other data sets, the latter providing information on average education within the origin country, the percent with tertiary education in the origin country, and the population of the origin country. Additionally, information on GDP per capita and share in agriculture can be used to determine stage of development of a country. Appendix B contains the sources for each of the variables used.

Results

The results are presented in tables 1 through 3. Table 1 tests the propositions using the average educational attainment of immigrants from country i as the dependent variable. There are 129 countries with the required information. For each of those countries, observations on each individual in the ACS sample is used. The 129 observations consist of origin-country averages among those immigrants who are in the ACS sample.

Columns 1 through 4 of table 1 use different weighting schemes and subsamples. Column 1 is a full sample analysis where countries are weighted by the number of observations that are used to compute the mean of the dependent variable. Column 2 presents the same analysis on the same sample but weights every country equally. The results are similar, both in terms of sign and statistical significance. All the propositions are supported by these results. Specifically, the average educational attainment of immigrants from country i decline in $I_{i,}$ as predicted by proposition 1, increase in the N_i, as predicted by proposition 2, and increase in μ_i as predicted by proposition 3.

Note that in the empirical analysis, I_{i*} is used, where I_{i*} is the number of immigrants from country i in the ACS sample. The relation of I_{i*} to I_i is that the latter is a scalar times the former, where the scalar is the ratio of immigrants in the country to immigrants used to compute the averages in the sample. That scalar, as elaborated below in Appendix C, is 20.7. That is, there are about 1/20th as many immigrants in the ACS sample used as in the actual US population. This has no effect on the estimates, however, because all calculations below are merely expressed either in unit-free numbers or in units that relate to the sample. It is important to remember these scale issues in comparing coefficients on I_{i*} and N_i because N_i is the actual population of the origin country in billions, whereas I_{i*} is 1/20th of the immigrants in the US. The mean of I_i* is around 11,000 whereas the mean of N_i is .05 so the coefficient on N_i is much larger than that on I_i

in the table.

In accordance with proposition 1, a one standard deviation increase in a country's number of immigrants decreases the predicted level of educational attainment in the US by about .4 years on a mean level of education of 13.3 years. The predicted difference in education among immigrants from the highest immigrant provider, Mexico, and the lowest immigrant provider, Estonia is 4.3 years. The actual difference between immigrants from the two countries is not far off at 5.6 years in favor of those from Estonia.

Proposition 2 predicts that the larger the population of the origin country, the higher is the attained level of education among immigrants. A one standard deviation increase in the population of the origin country implies about .4 of a year increase in the education levels of the immigrants from that country. Coincidentally, this is the same number as a one standard deviation change in I_i produces. Compare a tiny country like Cape Verde with ½ million people to a large country like Nigeria with almost 200 million people. Both have similar average levels of education, but the average level of education among immigrants in the US is 9.8 years for Cape Verde immigrants versus over 15 years for those from Nigeria.[6]

The more intuitive prediction of proposition 3, that the education of immigrants from country i is positively associated with education among natives in country i, is borne out by the positive and significant coefficients on μ_i in table 1. A one standard deviation increase in the mean level of education in the origin country implies about a one year increase in education among immigrants. This is not completely consistent with the theory, however. The proof of proposition 3, derived in the appendix, yields the result that $\partial \bar{A}_i / \partial \mu_i = 1$. The coefficient in table 1 on μ_i is always less than one. Some of this may be attributed to an errors-in-variables issue, where μ_i is mis-measured, but the coefficient is substantially below 1, seeming to deviate from the literal prediction of the model.

Finally, table 1, columns 7 and 8, speak to corollary 1, where the representation ratio is used in place of I_i and N_i . Corollary 1 is supported by the results, which yield negative and statistically significant coefficients. Recall that India is very much under-represented among US immigrants, despite being our third largest supplier of immigrants. Because India is the second most populous country in the world, it is under-represented by a factor of 4 among US immigrants. The implied difference between the education of immigrants from India and that from, say, El Salvador, which is highly over-represented is about 3.5 years based on this factor alone. The actual difference is about 6 years.

Columns 3 - 6 repeat the analysis, but with different sub-samples. In columns, 3 and 4, the four largest source countries of immigrants are omitted to ensure that the results are not driven by a few countries. In columns 5 and 6, the smallest countries are omitted. Qualitatively, the results are unchanged, although the magnitudes do change some, especially and not surprisingly for the coefficient on N_i, because the omission of very large population countries changes the scale of that factor. Additionally, the conclusions are robust to weighting and sample

---

[6]The actual difference is much greater than that predicted from the coefficients in table 1, but the direction and nature of the effect is as predicted. Also, there are about 6 times as many immigrants from Nigeria than there are from Cape Verde, but there are 400 times more people in Nigeria than in Cape Verde. Indeed, Cape Verde is over-represented among immigrants by a factor of 12, whereas Nigeria is under-represented by a factor of 5.

choice.

Table 2 performs the same analysis, but defines attainment in terms of the hourly wage received among those working rather than educational attainment. Columns 1 through 7 of table 2 mirror columns 1 through 7 of table 1, with all signs and statistical precision being similar across dependent variable definitions. A one standard deviation decrease in I_i implies about a $1.20 increase in the wage on a mean wage of $28 and a one standard deviation decrease in the representation ratio implies an over $5 increase. Just as was the case for education, the larger the number of immigrants from an origin country, the lower is the average wage, and the larger is the population of the country, the higher is the average wage among those immigrants. Additionally, the higher is the level of education in the origin country, which serves as a proxy for the average wage in the origin country, the higher is the immigrant wage. Indeed, substituting GDP per capita for education as a proxy for wage in the origin country gives essentially the same results. Also, as was the case in table 1, all robustness checks on weightings and sub-samples confirm the initial results of the first column.

Table 3 is analogous to table 2, but the dependent variable is average earnings among all immigrants from country i. This takes into account wage conditional on working, as in table 2, but also is affected by hours of work and employment rates among the immigrant population. Again, results are qualitatively identical. The one slight difference is that μ_i does not have as important nor as precisely estimated an effect on the dependent variable as it did in table 2.

The effects estimated in table 3 can be quite large. For example, the predicted difference between average earnings of those from the Philippines, which supplies the second largest fraction of US immigrants and Mexico, which supplies the most, is $12,495 in favor of Philippines. Although Filipinos are the second largest group of US immigrants, there are almost five times as many Mexican immigrants in the US as Filipino. That difference accounts for the much higher predicted income for Philippine immigrants. The actual difference between the two groups is about $17,000 in favor of the Philippines. Note that the both origin countries have average levels of education of around 9 years, so the source countries are comparable at least in this respect. Furthermore, Mexico's per capital GDP is much higher than that of the Philippines so the result among US immigrants from these two countries reflects the implicit rationing rule.

Corollary 1 has already been discussed in the context of tables 1-3. Corollary 1 simply condenses propositions 1 and 2 into the representation ratio, which is analyzed in columns 7 and 8 of the three tables. As discussed earlier, the data strongly support the contention of corollary 1, namely that the lower the representation ratio, the higher the achievement of a given country's immigrants in the US.

Corollaries 2-4 are testable in the same way that Propositions 1-3 were tested, with two exceptions. First, because there is no direct measure of average wage or earnings for all the countries in the dataset, the analysis is restricted to estimating only Δ defined in terms of education, namely the education of immigrants from country i minus the average level of education in country i.[7]

[7]It is important to use measures that are on the same scale, particularly with respect to proposition 4. The closest to wages or earnings in the US would be purchasing power parity GDP per capita, but this would not be a good proxy of comparable earnings at home if for no reason other than scaling.

Second, there is a standard statistical problem introduced by regressing Δ on independent variables that include μ because μ is part of Δ. Recall that Δ is defined as $\bar{A}_i - \mu_i$, so any errors in measurement of μ_i bias the coefficient on μ_i downward. The standard solution for this problem is to instrument the independent variable, in this case, using something that is correlated with μ_i but does not have the same measurement error associated with it. Fortunately, there is another measure of education at home that is correlated with average education (as is evidenced by the first stage), but is not the same variable. The proportion of the population with tertiary education from Barro and Lee (2010) is another measure of a country's education level that is different from μ_i but related to it. That variable is used as an instrument.

Table 4 reports the results. Column 1 provides the estimates from the instrumental variables approach. As corollary 2 predicts, the sign on I_i is negative and as corollary 3 predicts, the sign on N_i is positive. The larger is the group of immigrants in the US, the smaller is the difference between the immigrants' educational attainment in the US and the average educational attainment of the population in the origin country. The larger is the population of the home country, the larger is the difference between the immigrants' educational attainment in the US and the average educational attainment of the population in the origin country. Finally, as predicted, the effect of μ is not significantly different from zero. Of course, the failure to find a significant effect of μ does not imply that corollary 4 is proved, but merely that it is not refuted. The r-squared from the first stage is .54 with an $F(1, 66) = 28.1$, which provides some additional evidence on the validity of the estimates.

Column 2 reports the OLS results. They are similar, with the exception of the coefficient on μ, which, not surprisingly, is negative. Entering the same variable on both sides of the equation results in bias, in this case negative.

Structural Estimates

Under the assumption of normality, a predicted attainment of immigrants in the US can be determined for each of the 129 countries. The details of the estimation are described in Appendix C. Briefly, a starting value of σ is selected. Given that value and the actual values of I_i and N_i, it is possible, using the structure in (1') and (2'), to estimate A_i^* implied by truncating the top tail of the education distribution of the origin country and assuming that all immigrants to the US from that country come from that tail. Again, this makes the extreme and obviously incorrect assumption that every immigrant in the US is from that truncated upper tail in accordance with (1') and that everyone in that tail is offered and accepts the offer to come to the United States. Given A_i^* and the distribution of underlying education in each origin country, the predicted average attained level of education, $\bar{A}_i$,among immigrants from each origin country i can be obtained. Goodness-of-fit can be calculated by regressing the actual $\bar{A}_i$ on the estimated one.

The result is that a value of $\sigma = 4.7$ maximizes r-squared at .53.[8] Thus, slightly more than

[8]This weights observations by the number of immigrants from each country, as does column 1 of table 1. The unweighted version yields an r-squared of .503.

half the variation in average attained level of education is explained by the structural model that builds in explicitly selection from the top of a normal distribution. The functional form of the normal distribution, coupled with the assumption that A* is chosen so as to take only those from the top in accordance with (1), gives an exact predicted value of $\overline{A}_i$. This can be compared to the unconstrained version of table 1, column 2, which yields an r-squared of .72. The structural version of the model, which relies heavily on the assumption that immigrants are taken exclusively from the top of each countries distribution, is about 74% as effective in explaining the variation in educational attainment among immigrants as the unconstrained reduced form.

Other Factors

Table 5 allows other variables that are not explicitly modeled to affect the educational attainment of immigrants in the US. In particular, variables that measure growth and stage of development are included. They are GDP per capita, the last five years' growth rate, and the share of GDP that is comprised of agricultural output. Only agricultural share maintains its sign through all specifications and is significant in three of four, but the magnitude varies greatly by specification. Still, in the preferred specification, column 1, the coefficient is large and significant. One possibility is that the larger is agriculture's share, the greater incentive for high education individuals to seek residency in the US. This would be more consistent with a supply side explanation than a policy selection one. Most important, however, is that the inclusion of these variables does not affect in any substantial way the coefficients on the three key variables, I_i, N_i, and μ_i.

The conclusion from the empirical analysis is that the model works well in predicting who ends up in the US. Although supply considerations may matter, particularly in determining who is successful in being admitted to the US, a structure that assumes all who are admitted come and that the US admits from the top of the attainment distribution of each country after determining how many to admit from each county explains the data well. All predictions are borne out and the structural model provides a good fit with the data.

Other Factors and Explanations

Selection from the Top

The most obvious issue is whether the assumptions of the model are valid. Even though the view is admittedly stylized, it would be useful to find some evidence that lends some credence to the assumptions. Significantly, the assumption is that potential migrants are selected from the top of the origin country's attainment distribution. US immigration policy does not do that explicitly. Indeed, an important part of the policy, namely family reunification, seems to run counter to that assumption.

There is at least some support for the view that immigrants are selected from the top of the distribution. Recall that Δ is defined as the average educational attainment among immigrants and the average attainment in the origin country. Furthermore, as was argued earlier, to the extent that immigration slots to the US are scarce and desirable, ability may come into play in finding ways to make it into the US. Some of this is explicit. A number of skills-based green cards are issued to highly educated foreign citizens who eventually become residents or citizens of the US. But given the number who enter through other legal channels, not to mention those who come in illegally, it is worth exploring the validity of the assumption that the highly able from any given country are selected into the United States.

There is an internal check on the nature of immigrants relative to the population of the origin country. Specifically, the educational attainment of immigrants and that of the origin country population are available in the data. This is simply $\Delta_{i,}$ defined above. In 129 out of 129 cases, Δ_i is positive, meaning that the average educational attainment of immigrants from country i exceeds the average level of education in country i. The difference averages 4.8 years, with a low of ½ year and a high of 11 ½ years. Still, this may not be conclusive because educational attainment of those who are in the US might be higher than that of those who remain in the origin country if for no other reason than the US has higher average education than the world as a whole. After arriving in the US, immigrants might invest in education, consistent with being in a more highly educated society. So while the evidence on Δ is consistent and supportive of the theory, it is is not sufficient to prove that the assumptions of the model hold.[9]

Reverse Causation

The primary implication of the model is that those countries that provide a large number of migrants to the US have lower educational attainment. The assumption is that I_i, the number selected from country i, is the policy variable, not the educational attainment level itself. But suppose the US had an explicit policy of letting in only low-skilled individuals, say to prevent competition with the higher skilled native-born Americans, and let in more from low-skilled countries intentionally.

First, $\mu_{i,}$ which measures education in the origin country, should control for this factor to some extent, but it is still worth exploring the reverse causation more directly. The best evidence is provided by the simple correlation of I_i and μ_i, which is essentially zero at -.02.[10] There does not seem to be any pattern of selecting more immigrants from countries with low levels of educational attainment.

Another possibility is that immigrants with low levels of education are more likely to be admitted to the United States on an individual basis. Although possible, there is no reason why this would result in a negative correlation between number from a particular country and average skill level. Low-skill immigrants might be selected, but there are plenty of low-skilled immigrants in the world, and there is nothing that implies that the lowest skilled would all be

---

[9]Chiswick and Miller (2011) argue that there is some negative assimilation that occurs after migration. This pertains particularly when the origin country is the same as the destination country so may have less validity in this context. Chiswick and Miller (2012) investigate both negative and positive assimilation.

[10]It is -.004 when weighted by the number of immigrants from the country.

from one country. For example, India, which contributes a large number of migrants to the US, has one of the lowest levels of education of the 134 countries. Yet Indians in the US are highly educated. Given that the US allows in many Indians and given that there are many low-educated Indians, if the policy were to select negatively on the basis of skill, one would expect that the Indians who are here would be of low skill. The model implies the opposite, namely that selection is from the top of the distribution, not the bottom. Indians in the US are highly educated because a small number of Indians are admitted relative to the Indian population. That selection rule is consistent with the results of tables 1-5. and that is borne out by the data.

Additionally, a policy that seeks low-skilled immigrants does not appear to be a true description of the data. The average educational attainment among immigrants in the sample as a whole is a little over 12 years, which is not much below that for the native-born American population.

The Explicit Policy Deviates from that Assumed

In a typical year, over 60% of those issued permanent resident status are family sponsored. Although this does not rule out that those individuals are from the top of the educational attainment distribution at home, it does not appear to be a criterion that is closely related to selecting the most able from each country.

There are a number of factors to consider. First , the countries most likely to have immigrants who come in on family sponsorship are those that have the largest number already in the US. They may bring the average level of education down, not because there are so many of them, but because they are selected on family basis, rather than skill. Although possible and probably true to some extent, the relationships predicted by the propositions and corollaries hold irrespective of the inclusion and heavy weighting of the high immigrant countries (Mexico, Philippines, India, and China).

Second, as already discussed, it is likely that in addition to the rationing rule, supply side considerations enter. Obtaining permission to reside in the United States may be more easily acquired by those who are the most educated, given that slots are rationed. This pushes in the direction of getting the top immigrants from any given country, as assumed.

Third, recall that the average educational attainment among immigrants is significantly above that of the educational attainment in the origin country. Typically, immigrants are more educated than those who stay behind.

Fourth, Hanson, Liu, and McIntosh (2017) find that even for Mexico, which supplies the largest number of migrants to the US and which also is the largest source of unskilled labor, those who come to the United States are drawn from middle-income Mexicans, not from the lowest part of the income distribution. Grogger and Hanson (2015) find selective preference to stay in the US among foreign students in the US who have more educated parents and merit-based financial support. Again, this is a supply-side justification for the assumption that selection ends up being from the top of the distribution. Additionally, Docquier, Lohest, and Marfouk (2007) pay explicit attention to “brain drain” from developing countries to more advanced places and document this, noting in particular that it has increased over time. Selection from the top is the recipient country's description of what the source country calls brain drain.

Random Selection of Immigrants

The implications on I_i and N_i that form the basis of propositions 1 and 2 and corollaries 2 and 3, are a result of a model that assumes immigration slots are rationed on the basis of ability, from the top down. Were the selection process random, there would be no reason to expect that countries that were sources of a larger number of immigrants, I_i, would have lower levels of attainment. The same is true of population of the origin country. Were immigrants selected randomly, then the distribution of talent in the US would mimic that of the origin country. Proposition 3 and corollary 4 would result from random selection, however, because countries with higher levels of educational attainment would also send migrants with higher levels of education under random selection. The most important implications, however, are violated by the random selection model.

Supply Determined by Comparative Advantage

Consider a pure supply theory of migration, where those who get the most out of migrating move to the United States. There are a variety of versions of this, with the earliest being that of Sjaastad (1962). Already mentioned is the well-known work of Borjas (1987), formalizing and applying the Roy (1951) model to migration.[11]

Supply considerations may be important in providing a rationale for the rationing of slots according to ability or expected attainment. A modified supply model could give the implication that the larger the number of immigrants, the lower expected attainment, and the larger the population of the origin country, the higher attainment of immigrants.

One logical possibility is simply to allow the costs of migrating to the US to differ by country. This would be analogous to the gravity models used in trade economics[12], where individuals from nearby countries face lower migration costs and are therefore more likely to migrate. If this were the case, then the cutoff in gain from migrating to the US would be lower for nearby countries.

It is surely true that a migration version of gravity models explains some of the migration pattern. When countries are ranked by their representation ratios, nine of the top ten are from the Western Hemisphere. Without exception, these are countries with low populations, the largest being El Salvador with about 6 million people.

There still remain two issue that are not easily resolved by a supply model. It is the gain

[11]Grogger and Hanson (2011) find some support for the wage differential model driving selection of immigrants into a country. In particular, because there is a large difference between wages of high and low-skilled immigrants in the US, skilled immigrants tend to prefer the US as a destination country. What seems most relevant is the difference between the wages of the skilled in the destination country and the origin country, but this is likely to be correlated with the destination country's skill premium.

[12]See, for example, Bergstrand (1985) and Lewer, and Karemera, Oguledo and Davis (2000) and Van den Berg (2008), the latter two of which are direct applications to immigration.

from migration, not the final attainment, that should be related to the desire to migrate. As a consequence, it is necessary to argue that those who attain the highest levels of education in the US also have the most to gain from migration, given the costs. This is not an obvious proposition, but it is not unreasonable, given that the value of education in a highly developed economy may well be higher than that in a less developed one.

The second issue is that four of the five most important origin countries as sources of US immigrants are in thousands of miles away from the US. Mexico is the largest source of immigrants, but Philippines, India, China and Vietnam are very large contributors. Only four (Mexico, El Salvador, Canada, and Cuba) out of the top ten origin countries are in the Western Hemisphere. Each country may have an idiosyncratic reason for being a large source country (e.g., dislocation caused by the Vietnam war or Cuban migration associated with the ascension of Castro), but a model that integrates all factors is preferable.

There is no doubt that supply considerations matter. Economists believe that market equilibrium is determined by supply and demand. But in an environment where slots are rationed, a pure supply approach is only part of the story. Understanding the rationing rule is essential to predicting the attainment of US immigrants.

Conclusion

The larger the number of immigrants from a given country, the lower is the educational attainment and wages of that group. The pattern is a result of a selective immigration process that rations slots. Because of variations in the way the various origin countries are treated by the immigration system, a particular distribution of immigrants results, and this gives rise to differences in educational attainment and earnings by country of origin. Those countries that are given the largest number of slots tend to supply lower average ability immigrants.

For the same reason, the larger the population of the origin country, the higher the attainment number of immigrants from that country. It is easier to select one million highly educated people from India with 1.3 billion people than it is from Laos with 7 million people. Consequently, immigrants of Indian origin have higher levels of educational attainment than do immigrants of Laotian origin.

A model of selection is constructed that yields seven specific empirical implications, all of which are borne out by data from the American Communities Survey, 2011-2015. The larger the number of immigrants from an origin country, the lower the level of educational attainment, of wages, and of earnings in the US. The larger the population of the origin country, the higher the educational attainment, the higher the wages, and the higher the earnings of those immigrants in the US. A more parsimonious approach expresses predictions in terms of a representation ratio, which is a measure of how under- or over-represented a country is in the US immigrant stock. Countries that are more over-represented are predicted and found to have lower attainment in education, wages and earnings.

The theory also has implications for the difference between attainment of immigrants and that of the population in the origin country. This provides a separate test of the model, and all implications are borne out. In particular, the larger is the stock of immigrants from any given country, the smaller is the difference between the attainment of immigrants from that country and the origin population. Additionally, the larger is the population of the origin country, the larger is the difference between the attainment of immigrants from that country and the origin population.

A structural approach that assumes a particular functional form and specific selection rule performs well in explaining the data, yielding a goodness-of-fit statistic that is 74% as high as the unconstrained reduced form version. Overall, the model that postulates selection from the top of origin countries' ability distribution does well in describing the actual data.

References

Barro, Robert and Jong-Wha Lee. 2010. "A New Data Set of Educational Attainment in the World. 1950-2010." *Journal of Development Economics* 104(C): 184-198.

J.H. Bergstrand. 1985. "The gravity equation in international trade: Some microeconomic foundations and empirical evidence." *The Review of Economics and Statistics* 67(3): 474-481.

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." *The American Economic Review* 77(4): 531-553.

Borjas, George J. And Rachel M. Friedberg. 2009. "Recent Trends in the Earnings of New Immigrants to the United States." Working Paper no. 15406, NBER, Cambridge, MA.

Chiswick, Barry R. and Paul Miller. 2011. "The 'Negative' Assimilation of Immigrants: A Special Case." *Industrial and Labor Relations Review* 64(3): 502-525.

Chiswick, Barry R. and Paul Miller. 2012. "Negative and Positive Assimilation, Skill Transferability, and Linguistic Distance." *Journal of Human Capital* 6(1): 35-55.

Docquier, Frederic, Olivier Lohest, and Abdeslam Marfouk. "Brain Drain in Developing Countries." *The World Bank Economic Review* 21(2): 193-218.

Grogger, Jeffrey and Gordon Hanson. 2011. "Income Maximization and the Selection and Sorting of International Migrants." *Journal of Development Economics* 95(1): 42-57.

Grogger, Jeffrey and Gordon H. Hanson. 2015. "Attracting Talent: Location Choices of Foreign-Born PhDs in the United States." *Journal of Labor Economics* 33(S1): S5-S38.

Hanson, Gordon, Chen Liu, and Craig McIntosh. 2017. "Along the Watchtower: The Rise and Fall of U.S. Low-Skilled Immigration." Brookings Papers on Economic Activity, Washington, D.C.

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

Hunt, Jennifer. 2011. "Which Immigrants are Most Innovative and Entrepreneurial? Distinctions by Entry Visa" *Journal of Labor Economics* 29(3): 417-457.

Hunt, Jennifer. 2015. "Are Immigrants the Most Skilled US Computer and Engineering Workers?" *Journal of Labor Economics* 33(S1): S39-S77.

Karemera, David, Victor Iwuagwu Oguledo, and Bobby Davis. 2000. "A gravity model analysis of international migration to North America." *Journal of Applied Economics* 32(13): 1745-55.

Kerr, William R. and William F. Lincoln. 2010. "The Supply Side of Innovation: H1-B Visa Reforms and U.S. Ethnic Invention." *Journal of Labor Economics* 28(3): 473-508.

Kerr, William R., Sari Pekkala Kerr, and William F. Lincoln. 2015. "Skilled Immigration and the Employment Structure of US Firms." *Journal of Labor Economics* 33(S1): S147-S186.

Lazear, Edward P. 1999. "Culture and Language." *Journal of Political Economy* 107(S6): S95-S126.

Lazear, Edward P. 1977. "Education: Consumption or Production?" *Journal of Political Economy* 85:3 (June 1977): 569-97.

Lewis, Ethan. 2013. "Immigrant-Native Substitutability and the Role of Language" in *Immigration, Policy, and Socioeconomic Inequality*, ed. David Card and Stephen Raphael. New York: Russell Sage Foundation, 2013, pp. 60-97.

Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. "STEM Workers, H-1B Visas, and Productivity in US Cities." *Journal of Labor Economics* 33(S1): S225-S255.

Roy, Arthur. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3(2): 135-146.

Sjaastad, Larry A. 1962. "The Costs and Returns of Human Migration." *Journal of Political Economy* 70(5): 80-93.

U.S. Department of Homeland Security, Office of Immigration Statistics. *Yearbook of Immigration Statistics: 2014*. Washington, D.C., August 2016. <https://www.dhs.gov/sites/default/files/publications/DHS%202014%20Yearbook.pdf>

U.S. Department of Homeland Security. *U.S. Lawful Permanent Residents: 2015*. by Ryan Baugh and Katherine Witsman, Washington D.C., March 2017. <https://www.dhs.gov/sites/default/files/publications/Lawful_Permanent_Residents_2015.pdf>

U.S. Department of Homeland Security. *U.S. Lawful Permanent Residents: 2014.* by Nadwa Mossaad, Washington D.C., April 2016. <https://www.dhs.gov/sites/default/files/publications/Lawful_Permanent_Residents_2014.pdf>

U.S. Department of State, Bureau of Consular Affairs. Annual Immigrant Visa Waiting List Report as of November 1, 2015. 2015.

Van den Berg, Hendrik and Joshua L. Lewer. 2008. “A gravity model of immigration.” *Economic Letters* 99(1):164-167.

Appendix A

Proofs

Proposition 1:
From (1), using the implicit function theorem,

$$\frac{\partial A_i *}{\partial I_i} = -\frac{\partial / \partial I_i}{\partial / \partial A_i *}$$
$$= \frac{-1}{N_i f_i(A_i *)}$$

which is negative.

Note also that in (3), $\left[\frac{\int_{A_i*}^{\infty} A f_i(A)dA}{1-F_i(A_i*)} - A_i*\right]$ is just the conditional expectation of A, given $A>A_i*$, minus A_i*, which is necessarily positive since the conditional expectation must exceed its lower limit. This implies that the sign of the first term in (3) is just the sign of MA_i*/Mx . Additionally, when underlying distribution of ability, $f_i(A_i*)$, is independent of x as it is for $x=I_i$, the sign of the derivative in (3) is the same as that of MA_i*/Mx because the second term is zero.

Thus, since $MA_i*/MI_i < 0$, $\overline{A}_i$ decreases in I_i . |||

Proposition 2:
Analogously,

$$\frac{\partial A_i *}{\partial N_i} = -\frac{\partial / \partial N_i}{\partial / \partial A_i *}$$
$$= \frac{1-F_i(A_i *)}{N_i f_i(A_i *)}$$

which is positive. By the same logic is in the proof above, and since $f_i(A)$ does not depend on N_i,

$$\frac{\partial \overline{A}_i}{\partial N_i} > 0 . \qquad |||$$

Proposition 3:

First note that when $F_i(A) = F(A-\mu_i)$,

$$\frac{\partial A_i *}{\partial \mu_i} = -\frac{\partial / \partial \mu_i}{\partial / \partial A_i *}$$
$$= \frac{N_i f(A_i * - \mu_i)}{N_i f(A_i * - \mu_i)}$$
$$= 1$$

from (1).

Substituting into (3) yields

$$\frac{\partial \overline{A}_i}{\partial \mu_i} = \frac{f_i(A_i*-\mu_i)}{1-F(A_i*-\mu_i)}\left[\frac{\int_{A_i*}^{\infty} Af_i(A)dA}{1-F_i(A_i*)} - A_i*\right]$$
$$+\frac{1}{1-F(A_i*-\mu_i)}\int_{A_i*}^{\infty} A\frac{\partial f(A-\mu_i)}{\partial \mu_i}dA$$

The first term is positive because, as stated before, the conditional expectation exceeds its lower limit. The second term is positive because

$$\frac{\partial f(A-\mu_i)}{\partial \mu_i} = -f'(A-\mu_i)$$

and because f '(A) <0 for $A>A_i*$. |||

Corollaries

The proofs of the corollaries follow.

Corollary 1:

Rewrite (1) as

$$[1-F_i(A_i*)]-R_i = 0$$

Then

$$\frac{\partial A_i*}{\partial R_i} = \frac{\frac{-\partial}{\partial R_i}}{\frac{-\partial}{\partial A_i*}}$$
$$= \frac{-1}{f_i(A_i*)}$$

which is negative. Using the same logic as in the proof of proposition 1 and noting that $F_i(A)$ does not depend on R_i,

$$\frac{\partial \overline{A}_i}{\partial R_i} < 0$$ |||

To prove the corollaries that relate to Δ, note that for any variable x,

$$\frac{\partial \Delta}{\partial x} = \frac{f_i(A_i^*)\dfrac{\partial A_i^*}{\partial x}}{[1-F_i(A_i^*)]}\left[\frac{\int_{A_i^*}^{\infty} A f_i(A)dA}{1-F_i(A_i^*)} - A_i^*\right]$$

$$+\frac{1}{1-F_i(A_i^*)}\int_{A_i^*}^{\infty} A\frac{\partial f_i(A)}{\partial x}dA$$

$$-\int_{-\infty}^{\infty} A\frac{\partial f_i(A)}{\partial x}dA$$

Corollary 2:

Since $F_i(A)$ does not depend on I_i and since $\dfrac{\partial A_i^*}{\partial I_i} = \dfrac{-1}{N_i f_i(A_i^*)}$,

$$\frac{\partial \bar{A}_i}{\partial I_i} = \frac{f_i(A_i^*)\left[\dfrac{-1}{N_i f_i(A_i^*)}\right]}{[1-F_i(A_i^*)]}\left[\frac{\int_{A_i^*}^{\infty} A f_i(A)dA}{1-F_i(A_i^*)} - A_i^*\right] < 0$$

|||

Corollary 3: Since $F_i(A)$ does not depend on N_i and since $\dfrac{\partial A_i^*}{\partial N_i} = \dfrac{1-F_i(A_i^*)}{N_i f_i(A_i^*)}$, which is positive,

$$\frac{\partial \bar{A}_i}{\partial Ni} = \frac{f_i(A_i^*)\left[\dfrac{1-F_i(A_i^*)}{N_i f_i(A_i^*)}\right]}{[1-F_i(A_i^*)]}\left[\frac{\int_{A_i^*}^{\infty} A f_i(A)dA}{1-F_i(A_i^*)} - A_i^*\right] > 0$$

|||

Corollary 4:

Consider two countries. Define the base country as having an ability distribution given by F(A) and arbitrary country i as having an ability distributions $F_i(A)$, where as in the text, $F_i(A)$ is a displacement of F(A) by $\mu_{i.}$

The cutoff level for the base country is A* to satisfy (1) such that

$$N_0 [1-F(A^*)] = I_0$$

where I_0 is the policy determined number of immigrants from the base country. The expectation of A among those who immigrate.

Recall that

$$\frac{\partial A_i^*}{\partial \mu_i} = 1$$

so that $A_i^* = A^* + \mu_i$. The goal is to show that Δ is invariant with respect to μ. This is equivalent to showing that average ability among immigrants, $\overline{A}_i$, is greater than $\overline{A}$ by μ_i because

$$\begin{aligned}
&\Delta_i = \overline{A}_i - E(A_i) \\
&\Delta = \overline{A} - E(A) \\
&and \\
&E(Ai) - E(A) = \mu_i \\
&so\, if \;\; \Delta_i = \Delta, \\
&then \\
&\overline{A}_i - \overline{A} = \mu_i
\end{aligned}$$

s

Definitionally,

$$\overline{A}_i = \frac{1}{1-F_i(A_i^*)} \int_{A_i^*}^{\infty} A f_i(A) dA$$

or

$$\overline{A}_i = \frac{1}{1-F(A^*)} \int_{A^*+\mu_i}^{\infty} A f(A-\mu_i) dA$$

because $F_i(A) = F(A-\mu_i)$, $f_i(A) = f(A-\mu_i)$ and $A_i^* = A^* + \mu_i$.

A change of variables allows this to be rewritten as

$$
\begin{aligned}
\overline{A}_i &= \frac{1}{1-F(A^*)}\int_{A^*}^{\infty}(A+\mu_i)f(A)dA \\
&= \frac{1}{1-F(A^*)}\int_{A^*}^{\infty}Af(A)dA + \frac{\mu_i}{1-F(A^*)}\int_{A^*}^{\infty}f(A)dA \\
&= \frac{1}{1-F(A^*)}\int_{A^*}^{\infty}Af(A)dA + \frac{\mu_i[1-F(A^*)]}{1-F(A^*)} \\
&= \overline{A}+\mu_i
\end{aligned}
$$

|||

Table 1. Attainment of Immigrants in US

Dependent Variable: A_i measured as average years of schooling completed among immigrants from country i

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I_{i*}	-0.0000109***	-0.0000146***	-0.0000384***	-0.0000543***	-0.0000103***	-0.0000108***		
	(0.000000694)	(0.00000319)	(0.00000598)	(0.0000103)	(0.00000145)	(0.00000384)		
N_i	2.530***	2.851***	15.27***	12.00***	2.702***	2.899***		
	(0.309)	(0.688)	(1.933)	(2.050)	(0.631)	(0.829)		
μ_i	0.324***	0.267***	0.405***	0.293***	0.364***	0.401***	0.192**	0.252***
	(0.0532)	(0.0412)	(0.0365)	(0.0367)	(0.124)	(0.109)	(0.0799)	(0.0419)
R_i							-0.112***	-0.0458***
							(0.0168)	(0.0101)
Weight	I_i	None	I_i	None	I_i	None	I_i	None
Sample	Full	Full	Without Mexico, Philipp., India, China	Without Mexico, Philipp., India, China	$I_i > 10,000$	$I_i > 10,000$	Full	Full
Constant	10.19***	11.09***	9.373***	10.84***	9.631***	9.348***	11.41***	11.47***
	(0.496)	(0.380)	(0.373)	(0.351)	(1.179)	(1.037)	(0.703)	(0.381)
Observations	129	129	125	125	29	29	129	129
Adjusted R^2	0.720	0.331	0.624	0.447	0.731	0.457	0.267	0.293

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2. Attainment of Immigrants in the US
Dependent Variable: A_i measured as average hourly wage among working immigrants from country i

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8
I_{i*}	-0.0000332^{***}	-0.0000432^{***}	-0.000120^{***}	-0.000145^{***}	-0.0000306^{***}	-0.0000325^{*}		
	(0.00000288)	(0.0000155)	(0.0000278)	(0.0000547)	(0.00000606)	(0.0000172)		
N_i	12.08^{***}	13.18^{***}	48.90^{***}	38.32^{***}	12.53^{***}	13.39^{***}		
	(1.282)	(3.338)	(8.969)	(10.86)	(2.639)	(3.718)		
μ_i	1.484^{***}	1.521^{***}	1.868^{***}	1.607^{***}	1.446^{***}	1.752^{***}	0.819^{***}	1.440^{***}
	(0.221)	(0.200)	(0.170)	(0.195)	(0.518)	(0.488)	(0.295)	(0.198)
R_i							-0.414^{***}	-0.196^{***}
							(0.0621)	(0.0476)
Weight	I_i	None	I_i	None	I_i	None	I_i	None
Sample	Full	Full	Without Mexico, Philipp., India, China	Without Mexico, Philipp., India, China	$I_i > 10{,}000$	$I_i > 10{,}000$	Full	Full
Constant	13.89^{***}	14.63^{***}	10.47^{***}	13.77^{***}	13.22^{**}	10.75^{**}	21.33^{***}	16.65^{***}
	(2.060)	(1.843)	(1.731)	(1.859)	(4.934)	(4.650)	(2.599)	(1.805)
Observations	129	129	125	125	29	29	129	129
Adjusted R^2	0.649	0.343	0.556	0.375	0.653	0.413	0.275	0.339

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3. Attainment of Immigrants in the US

Dependent Variable: A_i measured as average earnings among immigrants from country i

	(1) Variable 1	(2) Variable 2	(3) Variable 3	(4) Variable 4	(5) Variable 5	(6) Variable 6	(7) Variable 7	(8) Variable 8
I_{i*}	-0.0401***	-0.0557**	-0.196***	-0.216**	-0.0364***	-0.0425		
	(0.00449)	(0.0239)	(0.0434)	(0.0847)	(0.00940)	(0.0275)		
N_i	15811.9***	17487.7***	60725.9***	52094.2***	16086.6***	17307.1***		
	(1996.2)	(5156.7)	(14029.1)	(16802.8)	(4090.0)	(5942.8)		
μ_i	985.4***	1206.6***	1585.7***	1334.4***	753.9	1446.6*	104.3	1093.7***
	(343.9)	(308.7)	(265.2)	(301.2)	(802.6)	(780.1)	(418.1)	(309.1)
R_i							-501.7***	-226.1***
							(87.98)	(74.20)
Weight	I_i	None	I_i	None	I_i	None	I_i	None
Sample	Full	Full	Without Mexico, Philipp., India, China	Without Mexico, Philipp., India, China	$I_i > 10,000$	$I_i > 10,000$	Full	Full
Constant	22766.2***	22583.9***	18438.9***	21453.6***	23413.6***	18274.8**	32686.0***	25141.0***
	(3208.5)	(2846.8)	(2708.3)	(2878.3)	(7646.9)	(7432.9)	(3683.2)	(2813.7)
Observations	129	129	125	125	29	29	129	129
Adjusted R^2	0.528	0.151	0.348	0.177	0.538	0.225	0.193	0.130

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4

	(1) Δ_i IV	(2) Δ_i OLS
I_{i*}	-.0000109***	-.0000109***
	(.0000006)	(.0000006)
N_i	3.40***	2.53***
	(0.61)	(0.31)
μ_i	-.279	-.675***
	(.183)	(.053)
Observations	70	70
R-squared	0.79	.83

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Attainment of Immigrants in U.S.
Dependent Variable: A_i measured as average years of schooling completed among immigrants from country i

	(1)	(2)	(3)	(4)
I_i	-0.00000801***	-0.0000118***	-0.0000349***	-0.0000510***
	(0.00000108)	(0.00000313)	(0.0000112)	(0.0000145)
N_i	1.965***	2.110***	10.89***	12.00***
	(0.414)	(0.693)	(3.275)	(3.014)
μ_i	0.611***	0.384***	0.577***	0.401***
	(0.131)	(0.107)	(0.114)	(0.0948)
GDP per capita	0.0000245	0.0000161	-0.000000847	0.0000144
	(0.0000241)	(0.0000176)	(0.0000200)	(0.0000156)
GDP growth rate (last 5 years)	-0.0425	0.0776*	0.0932*	0.0882**
	(0.0592)	(0.0451)	(0.0501)	(0.0401)
Share of GDP from agriculture	0.241***	0.0721**	0.0811*	0.0421
	(0.0513)	(0.0358)	(0.0451)	(0.0319)
Constant	5.783***	8.988***	7.072***	8.984***
	(1.399)	(1.123)	(1.142)	(0.999)
Weight	I_i	None	I_i	None
Sample	Full	Full	Without Mexico, Philipp., India, China	Without Mexico, Philipp., India, China
Observations	69	69	65	65
Adjusted R^2	0.835	0.367	0.530	0.452

Standard errors clustered at the area level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix B

Additional Tables

Table B-1

Variable	Description	Source	Mean	Std Dev
A_i (Education)	Mean years of schooling among immigrants from country i	ACS	13.32207	1.641941
A_i (Hourly Wage)	Mean wage among immigrants from country i (condition on working)	ACS	27.79516	7.738721
A_i (Annual Earnings)	Mean earnings among immigrants from country i (unconditional)	ACS	33062.38	10567.58
I_i	Number of immigrants in US from country i	ACS	11291.1	36397.28
N_i	Population of country i	World Bank Database 2015	.0527341	.170682
μ_i	Mean schooling in country i	UN Development Reports 2016	8.611628	2.796416
R_i	Representation ratio	created	6.279349	12.64804
$Tertiary_i$	Percentage with tertiary education in country i	Barro and Lee	10.6199	6.762148
$GDP/person_i$	Per capita GDP in purchasing power parity dollars	Heston and Summers	18071.52	13331.35
$GDP\ growth_i$	5 year growth rate of GDP	Heston and Summers	3.774583	3.437578
Agr share in GDP_i	Agricultural output / total output in country i	World Bank	6.976657	6.589

Appendix C

Recall that I_i refers to the actual number of immigrants from country i. For most of the analysis, the number is the sample was used because the scaling of I_i relative to N_i is of no consequence. But for this analysis, which relies on the actual number of immigrants and of the population of the native country, that scalar is relevant, but easy to obtain.

Immigrants make up .0993 of the ACS sample (of those 25 years old and r above). In the middle sample year of 2013, the population of the US was 316 million meaning that about 31 million were immigrants. Thus, to estimate the actual number of immigrants from country i, it is simply necessary to blow up the ACS by the ratio of observations relative to the population. There are 1.513 million immigrants used to compile the averages used in tables 1-4. Since the actual number of immigrants is 31 million,

$$I_i = \text{(sample number of immigrants)} \times (31/1.513)$$
$$= \text{(sample number of immigrants)} \times 20.7$$

The number of immigrants from country i is given by (1) as $N_i\,[1\text{-}F_i(A_i^*)\,] = I_i$.

Given assumptions, $F_i(A_i^*)$ is a normal, where μ_i displaces the distribution as specific to country i. Define the cumulative standard normal as $G(z_i)$ with $z_i = (A_i^* - \mu_i) / \sigma$. Then

$$N_i\,[\,1\text{-}\,G(z_i)\,] = I_i$$

or

$$G(z_i) = 1\text{-}\,I_i / N_i$$

Thus,

$$z_i = G^{-1}(1\text{-}\,I_i / N_i)$$

so

$$A_i^* = \sigma\, G^{-1}(1\text{-}\,I_i / N_i) + \mu_i \qquad \text{(C1)}$$

which is easily obtainable.

Given A_i^*, it is possible to estimate $A_{i.}$ That is done as follows. Because A_i^* is in the upper tail of the normal for all countries (even Mexicans at about 7 million equal only about 5% of Mexico's population), the tail is distribution is flat, close to linear, and negatively sloped.

Because A_i^* is so high relative to μ_i, it is assumed that the upper tail has value $f_i(A_i^*)$ at A_i^*, but tapers down to zero at A_i^*+1. Then the distribution of the part above A_i^* is a triangle with a height of 2 at A_i* and 0 at A_i* + 1. (The results are highly insensitive to this assumption because variation in A_i* is the major predictor of average education among immigrants.) Thus

$$E(A_i \mid A_i > A_i^*) = \int x(2A_i^* + 2 - 2x)dx$$

$$= A_i^* + 1/3 .$$

Iterating over σ yields σ=4.7 which maximizes the r-squared for the actual average attained level of education on the estimated $E(A_i \mid A_i > A_i^*)$. The r-squared is .53 (weighted by number of immigrants from each country). Thus, the model that assumes selection from above explains about half the actual variation in average attained education among immigrants, but does about 3/4 as well as the unconstrained version.