EXPLAINING U.S. IMMIGRATION, 1971–1998

Ximena Clark, Timothy J. Hatton, and Jeffrey G. Williamson*

Abstract—In this paper we develop and estimate a model to explain variations in immigration to the United States by source country since the early 1970s. The explanatory variables include ratios to the United States of source country income and education as well as relative inequality. In addition, we incorporate the stock of previous immigrants and a variety of variables representing different dimensions of the immigration quotas set by policy. We use the results to shed light on the impact of policy by simulating the effects of the key changes in immigration policy since the late 1970s. We also examine the factors that influenced the composition of U.S. immigration by source region over the entire period.

I. Introduction

SINCE 1950 more that 25 million immigrants have been admitted to the United States, about 20 million of whom arrived after 1970. This mass influx has stimulated a lively debate about the gains from immigration and the implications for immigration policy. Much of the literature has concentrated on the economic outcomes for the immigrants themselves and on the labor market impacts on native-born labor. These effects typically depend on how U.S. immigrants are selected—both within and between countries of origin—and models of this process are at the heart of the analysis. But while the literature is long on examining the outcomes of immigration, it is surprisingly short on estimating the determinants of immigration and on testing the models of immigrant selection that underpin our understanding of those outcomes. Our goal here is to develop and test just such a model.

This paper offers new estimates of the determinants of immigration rates by source from 1971 to 1998. It isolates the economic and demographic fundamentals that determine immigration rates across source countries and over time. These are real incomes, education, demographic composition, and inequality. We also allow for persistence in the flows arising from the stock of previous immigrants from the same source—accounting for the widely acknowledged but rarely estimated "friends and relatives effect." While existing studies typically include some of these variables, they often omit one or more of the key influences suggested by migration theory. More important is their neglect of immigration policy. Here we include policy variables that are derived directly from the quotas allocated to different visa categories. Finally, we examine more source countries over a longer period than does existing work on late twentieth century U.S. immigration.

We start in the next section by providing some background to U.S. immigration and immigration policy. We then set out a theoretical framework that is used to guide the choice of variables for regression analysis and to interpret the results. After presenting our econometric results, we evaluate the effects of major shifts in immigration policy on the total numbers and the effects of economic and demographic variables on the composition of immigration by source region.

II. Immigration and Immigration Policy

Changes in U.S. immigration over the last fifty years are well-known. As table 1 shows, the overall number legally admitted rose from a quarter of a million per year in the 1950s to nearly half a million in the 1970s and close to a million in the 1990s. The change in source composition has been even more dramatic. Europeans formed over half of the total in the 1950s, and the bulk of these were from Western Europe; by the 1990s, Western Europeans were a mere 5% of the total. The counterpart to this is the sharp rise in the proportion coming from Asia; the other notable feature is the ongoing rise in the share from Mexico. The sharpest change in the composition occurred between the 1950s and the 1970s and was associated with a major policy shift in 1965. Since then the composition of the flows has been more stable, although Western Europe has continued to decline and Mexico has continued to increase.

The most radical shift in postwar immigration policy was the 1965 Amendments to the Immigration and Nationality Act. Before this date country-of-origin quotas allocated the bulk of the visas to European countries, and two-thirds of these went to Germany and the United Kingdom. The 1965 legislation (effective 1968) abolished the quotas so that immigrants from all countries could compete more equally for the available visas. It established a maximum quota of 20,000 for each Eastern Hemisphere country, subject to an overall ceiling of 170,000. Within the quota, visas were allocated according to a seven-category preference system, which gave 64% of visas to relatives of U.S. citizens or residents, 6% to refugees, and 30% to employment-based categories. Children and spouses of U.S. citizens were exempt from the quota, reflecting a strong emphasis on family reunion.

Received for publication June 3, 2003. Revision accepted for publication February 27, 2006.

* World Bank; University of Essex and Australian National University; and Harvard University, respectively.

Earlier versions of this paper were presented at the Population Association of America Meetings, Atlanta, Georgia (May 9–11, 2002) and the LACEA meetings, Madrid (October 11–13, 2002). We are grateful for the helpful comments of Gary Chamberlain, Andy Mason, Sam Thompson, and participants at the PAA and LACEA meetings, as well as for the excellent research assistance of Rachael Wagner. We are also grateful to the editor and two anonymous referees for valuable comments. Hatton acknowledges with pleasure the financial support of the British Academy, and Williamson does the same for the National Science Foundation SES-0001362. The views expressed herein are those of the authors and not necessarily those of the World Bank.

The Review of Economics and Statistics, May 2007, 89(2): 359–373

TABLE 1.—SOURCE AREA COMPOSITION OF U.S. IMMIGRATION, 1951–2000
(PERCENT OF TOTAL FROM EACH SOURCE)

Region of Origin	1951–60	1961–70	1971–80	1981–90	1991–2000
Europe	52.7	33.8	17.8	10.4	14.9
Western	49.8	30.8	14.7	7.5	5.9
Eastern	2.9	3.0	3.1	2.9	9.0
Asia	6.1	12.9	35.3	37.3	30.7
Americas	39.6	51.7	44.1	49.3	49.3
Canada	15.0	12.4	3.8	2.1	2.1
Mexico	11.9	13.7	14.2	22.6	24.7
Caribbean	4.9	14.2	16.5	11.9	10.8
Central America	1.8	3.1	3.0	6.4	5.8
South America	3.6	7.8	6.6	6.3	5.9
Africa	0.6	0.9	1.8	2.4	3.9
Oceania	0.5	0.8	0.9	0.6	0.6
Total (000s)	2,515	3,322	4,493	7,338	9,095

Source: *Statistical: Yearbook of the Immigration and Naturalization Service for 2000,* table 2.

Notes: Immigrants classified by country of last residence. Percentages exclude the category "origin not specified." Western Europe is defined as the countries of the European Union (EU-15), excluding Finland but including Norway and Switzerland. Eastern Europe includes the category "other Europe."

In addition, a ceiling of 120,000 visas was set for the Western Hemisphere, but without country quotas or a preference system.[1]

Immigration legislation was amended again by an act of 1976 (effective 1977) when quotas of 20,000 per country, together with the system of preferences, was extended to Western Hemisphere countries, and an act of 1978 (effective 1979) when the hemispheric ceilings were combined into an overall quota of 290,000. In 1980 the preference category for refugees was removed and the worldwide ceiling was reduced to 270,000 (effective 1981). In 1986 the Immigration Reform and Control Act (IRCA) provided for the legalization of illegal immigrants who had resided in the United States since before 1982. It also expanded the H-2 program for temporary foreign workers and introduced temporary visas for agricultural workers with three years' residence in the United States.

The most important amendment to the post-1965 regulations came in the 1990 Immigration Act (effective 1992). This legislation introduced an overall quota of 675,000, divided into three classes. First, a total of 480,000 visas was allocated to family immigrants, with immediate relatives of U.S. citizens coming under the quota for the first time. Within this total, a minimum of 226,000, allocated according to a four-part preference system, was given to family-sponsored nonimmediate relatives of U.S. citizens and resident aliens.[2] Second, the 1990 act increased the number of employment-based visas to 140,000 (from 60,000 previously), under a five-part preference system.[3] Third, 55,000 visas were allocated on top of the overall quota for "diversity" immigrants—those from countries with relatively low immigration since 1965.[4]

The current (and past) legislation provides different routes into the United States. Differences among source regions in levels of economic development and immigration histories are reflected in the composition of entry routes. Table 2 illustrates these differences for 1998. Overall, just 12% entered on employment-based preference categories, but the figures are substantially higher for immigrants from Western Europe and Canada. Employment-based entry is particularly low for Eastern Europe and Africa, where refugee and asylee admissions are significant, and also from Mexico and the Caribbean. It is notable also that reunion with immediate family is the entry route for more than half of Western Hemisphere immigrants except for Canada. The data suggest that the persistence effects of past immigration has waned for Western Europe and Canada, as reflected in the small share of family-sponsored preferences (a fact partly represented in the diversity category). It is also small for Africa, a source country for whom American mass immigration has only just begun. It is *very* large for the remaining regions in transition: 34% for Asia (74% when immediate relatives are included) and the Americas (86% when immediate relatives are included), reaching an enormous 42% for Mexico (88% when immediate relatives are included).

There are two important indirect routes that have affected the sources of immigration. One is illegal immigration, which has increased over time and is currently running at about 300,000 per year. Mass legalization of 2.7 million illegal immigrants took place in the decade after the Immigration Reform and Control Act of 1986. This provided an additional route to legal immigration largely for Western Hemisphere immigrants, and especially those from Mexico. The other route is that of temporary workers or trainees with H, O, and P visas, the numbers of whom soared from 75,000 in 1985 to 430,000 in 1998. This rising number originated chiefly from Europe and Asia. Although they are not part of the overall immigration total, temporary visas clearly have been used as an intermediate step before adjusting to permanent status.

III. Modeling Immigration

Immigration is determined partly by individual incentives and constraints, and partly by policy. Immigration policy can be seen as a filter though which ex ante migration

[1] Further details of numbers allocated to different preference categories are given in Appendix B4.

[2] The maximum number of visas allocated to nonimmediate family members is the difference between 480,000 and the actual number of visas issued to immediate relatives in the previous year, subject to a minimum of 226,000. Thus under the "flexible cap" system the total number admitted under the quota can exceed the overall cap in a particular year.

[3] The quotas for different preferences in the employment-based category are detailed in Appendix B4.

[4] In the transitional period between 1992 and 1994, the overall quota was raised to 700,000 with 465,000 visas reserved for close family immigration, but the diversity program was limited to 40,000.

TABLE 2.—CLASS OF ADMISSION BY SOURCE AREA, 1998
(PERCENT OF TOTAL FOR EACH SOURCE)

Class of Admission	Family-Sponsored Preferences	Employment-Based Preferences	Immediate Relatives of U.S. Citizens	Refugee and Asylee Adjustments	Diversity Program
All immigrants	29.0	11.7	42.9	8.3	6.9
Europe	9.6	15.0	32.0	20.8	20.9
West	12.1	27.5	46.4	2.3	11.2
East	8.4	8.5	25.3	30.3	25.8
Asia	35.5	16.9	37.9	5.3	3.9
Americas	34.2	7.3	51.9	5.1	0.9
Canada	14.3	43.8	35.4	0.1	4.8
Mexico	42.1	2.8	45.6	0.0	0.0
Caribbean	33.8	3.1	42.9	18.8	1.3
Cnt. America	26.7	11.1	58.7	2.4	0.5
Sth. America	24.5	12.6	58.9	1.6	2.1
Africa	8.2	7.2	35.8	10.8	37.7
Oceania	30.0	14.1	42.5	0.6	12.4
Total (000s)	191.5	77.5	283.4	54.6	45.5

Source: *Statistical Yearbook of the Immigration and Naturalization Service for 1998,* table 9.
Notes: Immigrants classified by country of last residence. Rows do not add to 100 because they exclude certain other classes of admission. Western Europe is defined as the countries of the European Union (EU-15), excluding Finland but including Iceland, Norway, and Switzerland.

decisions are translated into ex post migration. The economics of the migration decision has been widely studied, most notably by Larry Sjaastad (1962), George Borjas (1987), and Barry Chiswick (2000). Here we set out a heuristic framework, which follows in that tradition. It emphasizes the effects of income differentials, skill differentials, migration costs, demographic at-risk sensitivity, and immigration policy on the probability that individuals will move from one country to another.

Individual i ($i = 1, \ldots, n$) residing in source country y receives the wage $w_y(s_i)$, where s_i is the individual's skill level. The wage the individual would receive in the destination country x is $w_x(s_i)$. Thus the gains to migration for individual i are represented by the difference $w_x(s_i) - w_y(s_i)$. Migration costs depend on four elements. First, there is an individual-specific migration cost, z_i. This may be interpreted as reflecting individual preferences for migration in terms of equivalent income. This compensating differential differs across individuals, but would be expected to be positive on average. Factors such as having relatives in the destination country are likely to lower the psychic cost component of z_i. It will also reflect the lower direct cost of immigration through family reunion or family-sponsored preference categories as compared with other routes, including illegal migration.

Second, there is a direct cost, c_1, which is the same for all migrants from source country y, but which may differ across source countries according to distance from the destination. It may also reflect immigration policy: tougher immigration policy raises the cost of migration for all immigrants by raising c_1. Third, there is the cost to migrants associated with quantitative restrictions: the greater is the total quota, the lower is the cost in terms of waiting time, or the cost and effort of moving to a higher-preference category. Thus the cost-equivalent effect of quotas is represented by $c_2(q)$, which applies to all potential migrants, given their status under the quota. Finally, skill-selective immigration policy is represented by a term $\gamma(\delta - s_i)$; the higher the individual's skill level, relative to benchmark level δ, the lower the costs of migration. A rise in δ increases the overall standard for admission, while an increase in the skill selectivity of immigration policy, for a given threshold, is represented by an increase in the parameter γ.

Putting these elements together, the probability that individual i will migrate from country y to country x is

$$m_i = \text{Prob}(v_i > 0), \text{ where } v_i = w_x(s_i) - w_y(s_i) - z_i - c_1 + c_2(q) - \gamma(\delta - s_i). \quad (1)$$

Across individuals in country y, $w_x(s_i)$, $w_y(s_i)$, z_i, and s_i are assumed to be normally distributed with means μ_x, μ_y, μ_z, and μ_s respectively. Summing over all n individuals in the source country y, the emigration rate to x is:

$$M = 1 - \Phi\left[\frac{-\mu_x + \mu_y + \mu_z + c_1 - c_2(q) + \gamma(\delta - \mu_s)}{\sigma_v}\right], \quad (2)$$

where Φ is the standard normal distribution function and σ_v is the standard deviation of the net benefit function v. This is simply a modified version of the Roy model advanced by Borjas (1987), among others.

Higher mean wage rates in the destination country or lower mean wage rates in the source country (for a given skill level) increase the migration rate, as does a fall in the mean of personal migration costs, μ_z, or a fall in the fixed migration cost, c_1. An increase in the average skill level in country y would increase the migration rate if there is

skill-selective immigration policy in country x ($\gamma > 0$) and could increase the migration rate through the wage differential, if the function w_x is steeper than w_y. The variances will also matter, and the effect of changing wage and skill distributions will depend on their effect on σ_v, and the sign of the mean of $-v_i$, that is $-[\mu_x - \mu_y - \mu_z - c_1 + c_2(q) - \gamma(\delta - \mu_s)]$. These effects are examined further in appendix A. To take one example, if the mean of v_i is positive (the destination is relatively rich), then the migration rate will be an inverse U-shaped function of the ratio of source to destination wage inequality (as a relative proxy for the return on skills).

Immigration policy will also influence the volume of migration through several different channels represented by the terms in equation (2). Widening of family reunification policies, by reducing z_i for some potential emigrants, will lower its mean μ_z and increase migration. A reduction in the overall quota, q, would raise direct migration costs through $c_2(q)$ and therefore reduce migration. An increase in skill selectivity through raising the threshold value, δ, would be expected to reduce the migration rate, while the effect of increasing the value of γ could raise or lower the migration rate (see Appendix A).

Since migration is a forward-looking decision, it is useful to think of the gains to migration in present-value terms. Thus $w_x(s_i)$ and $w_y(s_i)$ can be thought of as discounted income streams for individual i in the destination and source, respectively. For any individual, the present value of migration as represented by the difference between these income streams, net of costs, will depend on the length of working life remaining. Hence, the net gain represented by equation (2) will be greater the younger the potential migrant is in the source country. It follows that the source-country age structure should also matter: the larger the share of young adults, the greater will be the migration rate for a given positive wage gap, net of costs.[5]

IV. Explaining Immigration

Recent studies of U.S. immigration highlight some of the economic forces that determine immigration rates across source countries. The dependent variable is typically taken as the number of immigrants to the United States relative to the source-country population, representing the propensity to emigrate to the United States. Borjas (1987) found that, for a cross section of average emigration rates from 1951 to 1980, migration was negatively related to origin-country income per capita and to distance from the United States. In addition, the emigration rate was negatively related to inequality in the origin country, implying negative within-country selection.[6] Using a cross section of source-country immigration rates for 1982–1986, Philip Yang (1995) confirmed the income effects but found the stock of previous immigrants from each source country to be the single-most important determinant of the immigration flow.

David Karemera, Victor Oguledo, and Bobby Davis (2000) used panel data on emigration rates to the United States and to Canada for the decade 1976–1986, including a wide range of explanatory variables for both the United States and countries of origin. They found that emigration rates were related negatively to distance from the United States, negatively to origin-country income, positively to U.S. income, and negatively to the U.S. unemployment rate. In addition, they found that migration was positively related to measures of political rights and individual freedom in source countries, and negatively to political instability. Thus, their results confirm the importance of economic variables, migration costs, and civil rights in determining migration. Immigration policy in the United States was modeled as a dummy variable only.

More recently, two studies have explored the determinants of migration flows using panels for multiple origin and destination countries. Anna Maria Mayda (2004) analyzed migrants to fourteen OECD destinations for 1980–1995. One of the key findings was that income and education affect migration with opposite signs and that origin-country effects are smaller than destination-country effects. One interpretation of this is that for potential migrants in poor countries, an increase in income both reduces the incentive and increases the ability to migrate. Peder Pedersen, Mariola Pytlikova, and Nina Smith (2004) studied immigration to 27 OECD countries from 129 source countries in the 1990s, also finding source-country effects to be weaker than destination-country effects. Both studies provide evidence of positive effects on migration associated with migrant networks and of negative effects associated with the economic, cultural, and geographical distance between origin and destination.

Here we focus on the United States, which has the widest range of immigration sources over the longest period. We include in our analysis a more comprehensive set of variables than previously, all of which are suggested by migration theory, but some which are often omitted from existing studies. These variables include human capital, inequality, the demographic structure of the origin country, and the stock of previous migrants living at the destination. We also derive a set of variables to capture different dimensions of U.S. immigration policy, something that has been previously modeled only by dummy variables. And finally, we use our results to see how these variables account for the

[5] Let the wage difference (destination minus source country) per year of working life be a constant D. If the age range of potential working-age migrants, a, runs from 20 to 65 and the discount rate is r, then the present value of the gains will be $PV(a) = Dr[1 - (1 + r)^{-(46-a)}]$, which is a decreasing function of a.

[6] This is consistent with the Roy model summarized in equation (1) above, where source countries are relatively poor and relatively unequal compared with the United States (see Appendix A).

source-country composition of immigration and to simulate the effects over time of key changes in immigration policy.

We attempt to capture the determinants of the emigration rate to the United States with the following specification:

$$\begin{aligned}(mig/pop)_{j,t} &= \beta_0 + \beta_1(y_j/y_{\text{us}})_t + \beta_2(syr_j/syr_{\text{us}})_t \\ &+ \beta_3 age_{j,t} + \beta_4(ineq_j/ineq_{\text{us}})_t + \beta_5(ineq_j/ineq_{\text{us}})^2{}_t \\ &+ \beta_6 pov_{j,t} + \beta_7 dist_j + \beta_8 land_j + \beta_9 eng_j \\ &+ \beta_{10}(stock_{j,t-1}/pop_{j,t}) + \beta_{11}(stock_{j,t-1}/pop_{j,t})^2 \\ &+ \beta_{12}X_{r,j,t}(stock_{j,t-1}/pop_{j,t}) + \beta_{13}X_{e,j,t}(syr_j/syr_{\text{us}})_t \\ &+ \beta_{14}X_{d,j,t} + \beta_{15}X_{a,j,t}civ_{j,t} + \beta_{16}X_{irc,j,t} + \beta_{17}X_b.\end{aligned} \tag{3}$$

The left-hand-side variable is migration to the United States from country j in year t as proportion of the origin-country population.

Economic and demographic fundamentals are reflected by the first five terms, while the others represent costs. The first term, the ratio of the average (purchasing power parity adjusted) income in j relative to the United States is expected to have a negative effect; $\beta_1 < 0$. The second term is the ratio of average years of schooling (*syr*) in j relative to the United States. Since the income variable reflects both the amount of human capital and the average return on human capital, it must be "deflated" by human capital stocks in order to reflect the relative return alone. Thus, relative schooling years is expected to have a positive effect on immigration; $\beta_2 > 0$. The variable age in the origin country is the share of population aged 15–29. It reflects the fact that the present value of migration is higher, for a given wage incentive, at younger ages: thus, $\beta_3 > 0$. The ratio of inequality in the origin relative to the United States (*ineq*) is entered in quadratic form. According to the Roy model, when the destination country is richer than the source (adjusted for migration costs) the effects of inequality follow an inverse U-shape. When the origin country is relatively unequal, an increase in its relative inequality will reduce the migration rate; when the source country is relatively equal, an increase in its inequality will increase the migration rate (see appendix A). Hence $\beta_4 > 0$, $\beta_5 < 0$. Here inequality is represented by the Gini coefficient of household income.

Migration costs constrain the move. The proportion of the source-country population living in poverty is represented by the inverse of its income squared. As income increases, poverty quickly evaporates, releasing constraints on migration and hence $\beta_6 < 0$.[7] As in any gravity model, costs rise with distance from the United States; hence, $\beta_7 < 0$. Such costs are also associated with whether the source country is landlocked and whether it is predominantly English speaking; $\beta_8 < 0$, $\beta_9 > 0$. Current migration costs are also represented by the stock of previous immigrants from the sending country. This is defined as the ratio of the number born in country j residing in the United States at time $t - 1$ relative to the population of country j. Since relatives (and friends) abroad reduce migration costs, $\beta_{10} > 0$, although we expect this effect to diminish with size (over the relevant range, hence $\beta_{11} < 0$) if it is accounted for by immigrants' job search and settlement costs (diminishing returns to network externalities) rather than by remittances releasing the financial constraint.

The remaining variables represent the effects of immigration policies, through the different routes of entry. These are interacted with other variables to represent the ease of access to these channels for migrants from a given country. The variables X_r, X_e, X_d, and X_a represent the number of visas available by different entry routes, divided by the total population of the countries that qualify for them. These are derived separately for each major channel of entry, and are calculated for each country, as described in Appendix B4. This reflects the scarcity of visas and hence the cost of immigration. A fall in X as a result of a reduction in the quota will therefore reduce migration; thus β_{12} through β_{15} are expected to be positive.

The variable X_r represents the quota for nonimmediate relatives, and it is interacted with the immigrant stock divided by origin-country population. Thus, the higher the stock of foreign born from a given country, the lower the average cost of migration from that country and the more migrants choose that route. X_e represents the quota of employment visas and is interacted with the ratio of schooling years to capture the element of skill selectivity. X_d reflects the quota of diversity visas available since 1992, prior to which it takes the value of 0. Since diversity visas are awarded by lottery, it is not interacted with country characteristics. X_a represents the allocation of visas to refugees, which since 1980 has been set year by year rather than coming under the legislated quota. This variable is interacted with a dummy for civil war—the main cause of refugee flights (for example, Hatton & Williamson, 2003).

The final two variables represent somewhat special circumstances. X_{irc} is intended to capture the effects of the IRCA legalization program. It is the estimated number of illegal immigrants from a given country residing in the United States preceding the legalization program divided by that country's population. It is applied only to the years 1989–1991, when the bulk of legalizations took place, and β_{16} is therefore expected to be positive. Finally, X_b is a dummy for the years 1995–1998 when, due to administrative changes in the processing of visa applications, there was a progressive rise in the backlog. As a result, recorded

[7] Unfortunately there are insufficient data to construct a direct measure of poverty incidence for the country/period observations in our data. Using cross-country estimates for recent years, Ravallion (2004) finds that poverty head count ratio is inversely related to the square of per capita income.

immigration for these years was lower than it would otherwise have been, and the dummy is therefore expected to be negative; $\beta_{17} < 0$. Details of the derivation of these variables are given in Appendix B4.

V. Econometric Results

We estimate our migration model on panel data for immigration to the United States by place of birth for 81 source countries across the 28 years from 1971 to 1998 (see Appendix B, sections 1 and 5). These countries form 82.5% of all U.S. immigration over this entire period. For relative income we use purchasing power parity adjusted GDP per head from the Penn World Tables; years of education is based on the series derived by Barro and Lee. Total population and population aged 15–29 come from the UN demographic database; the Gini coefficient for household income (a crude measure for the return to skills) is calculated from data collected by the World Bank and the WIDER Institute. These sources are further detailed in Appendix B3. The stock of foreign born from each source country is calculated using census and CPS data and then interpolating using gross immigration flows in order to obtain annual series. The sources and methods of calculation are discussed in Appendix B2.

Our estimating equation is based on equation (3) but, because the gross immigration rate is bounded at zero, the left-hand-side variable is transformed by taking natural logs. This constrains the predicted values to be positive in the regression and in the counterfactual predictions examined below. The right-hand-side variables are as in equation (3). Our estimating method is fixed-effects regression. The variables distance, English speaking, and landlocked are thus eliminated since they have no time variation. These effects nevertheless influence the country means and we will return to them in what follows.

The results from estimating the equation on this panel data set appear in table 3. The first column excludes all the policy-related variables. The coefficients are significant at conventional levels with the exception of relative years of schooling and the share of source-country population aged 15–29. This probably reflects the fact that these variables are interpolated between five-year benchmarks. The full model in the second column includes the policy variables; adding these has little effect on the coefficients of the economic and demographic fundamentals. We include the uninteracted policy effects as well as the interactions with other variables. All these variables are significant with the exception of the interaction between the employment-based quota and relative education years and the processing backlog. The third column omits the variables that were insignificant in column 2.

One possible concern is that the per-country numerical limits within the worldwide immigration quota may be binding on some countries and that the coefficients might be affected by the inclusion of constrained countries. That concern is partially allayed by the facts that only about half of all immigrants come under the worldwide quota and that very few countries have been at or near the per-country limit for a sustained period.[8] The exceptions are Mexico, India, China, the Philippines, the Dominican Republic, and Korea.[9] It is reassuring also that the coefficients in column 3 are changed very little when the model is estimated excluding these countries.[10]

It is worth examining the quantitative implications of some of the estimated coefficients, focusing on the third column. The relative income term implies that a 10% increase in a source country's income per capita reduces the immigration rate by 4.4%. The inverse of source-country income squared takes a negative coefficient consistent with the idea that higher absolute income releases the poverty trap. This provides an offsetting effect of about 1% when evaluated at the mean income level. However, when evaluated at the mean for the poorest region, Africa, a 10% increase in income per capita reduces the immigration rate for an individual country by 1.2% through the relative income effect, but increases the immigration rate by 2.4% through the poverty trap effect.

The coefficients on relative inequality and its square, positive and negative respectively, support the predictions of the Roy model (see Appendix A), with the maximum occurring at an inequality ratio less than 1. Evaluated at the mean, a 10% increase in source-country inequality causes immigration to fall by 7.5%, but these effects will differ by country. At the mean for South America (1.20), a 10% increase in relative inequality will cause a fall of 11.9% in immigration as the incentives for migration decline for those higher up in the income distribution. At the mean for Western Europe (0.82), a 10% increase in inequality reduces migration by only 1.3% as incentives decrease for those higher up but increase for those lower down.

[8] Some evidence of substitution of quota and nonquota categories is offered by Jasso, Rosenzweig, and Smith (2000). They examined immigration of husbands of U.S. citizens, finding that the numbers entering under this nonquota category are sensitive to changes in the rules governing that route of entry, but also to changes in the rules for entry under the quota.

[9] The worldwide quota excludes immediate relatives of U.S. citizens, refugees, and legalizations under IRCA. Per-country limits were 20,000 until 1991 and subsequently 7% of the worldwide quota (25,620 in 1998). In 1998 the countries with the highest number of immigrants admitted as part of the worldwide quota were Mexico (58,758), India (24,988), China (23,231), the Philippines (17,496), the Dominican Republic (9,595), and Korea (9,105). The figure for Mexico is explained by the fact that it shares with the Dominican Republic an exemption from the per-country limit for immigrants entering under category 2 of the family-sponsored preferences. The number of Mexicans admitted under the family-sponsored and employment-based preferences was 18,795. Although India and China are currently close to the limit, this was not the case before the 1990s.

[10] A Wald test for the full set of interactions between each of the variables and a dummy for the six countries yields a value 19.4, which is not significant for $\chi_{(15)}^2$ at the 5% level.

TABLE 3.— GROSS IMMIGRATION RATE REGRESSIONS
(81 COUNTRIES, 28 YEARS; DEPENDENT VARIABLE: LOG IMMIGRANTS ADMITTED/SOURCE-COUNTRY POPULATION)

	(1)	(2)	(3)
Constant	−8.74	−8.84	−9.09
	(13.4)	(26.2)	(38.8)
GDP per capita ratio (foreign/U.S.)	−1.49	−1.42	−1.41
	(7.1)	(7.0)	(7.1)
Schooling years ratio (popn. > 14) (foreign/U.S.)	−0.02	0.04	
	(0.1)	(0.2)	
Share of population aged 15–29 (foreign)	0.49	−0.56	
	(0.7)	(0.8)	
Inequality ratio (foreign/U.S.)	1.50	1.45	1.64
	(3.5)	(5.1)	(4.1)
Inequality ratio (foreign/U.S.) squared	−1.18	−0.95	−1.01
	(6.2)	(5.1)	(5.7)
Inverse of income squared (foreign)	−0.20	−0.18	−0.18
	(4.2)	(4.1)	(4.2)
Immigrant stock ($t - 1$)/foreign population	8.63	6.91	6.71
	(4.2)	(3.3)	(3.2)
(Immigrant stock ($t - 1$)/foreign population)^2	−47.34	−40.78	−40.38
	(4.9)	(4.3)	(4.3)
Quota X_r		−0.60	−0.59
		(3.3)	(3.0)
Quota $X_r \times$ (immigrant stock ($t - 1$)/foreign population)		32.58	33.37
		(3.1)	(3.3)
Quota X_e		9.16	4.61
		(3.3)	(4.5)
Quota $X_e \times$ schooling years ratio		−8.73	
		(1.6)	
Diversity quota X_d		0.40	0.40
		(7.6)	(7.6)
Refugee quota X_a		1.27	1.27
		(6.6)	(6.7)
Refugee quota $X_a \times$ civil war		1.45	1.49
		(3.6)	(3.7)
IRCA legalization X_{irc}		0.07	0.07
		(7.9)	(8.1)
Processing backlog X_b		−0.03	
		(0.9)	
R^2 (within)	0.15	0.23	0.23
No. of observations	2,268	2,268	2,268

The coefficients on the migrant stock by itself (that is, ignoring its interaction with X_r) reflect the nonpolicy component of the "friends and relatives effect." While the linear term is positive as expected, the squared term is strongly negative, implying that the marginal effect is large when the stock is small but diminishes as the stock increases. That marginal effect eventually falls to 0 when the migrant stock in the United States reaches 8.3% of the source-country population. At the mean (1.3% of source-country population), the coefficients imply that if the immigrant stock from a given source is raised by 1,000, the annual flow from that source would be increased by 4.7 immigrants. This direct effect is augmented by an indirect effect working through the policy variable representing the quota on nonimmediate relatives (X_r). This adds a further 1.6, yielding 6.3 more immigrants per year for every 1,000 added to the existing immigrant stock. Thus, the overall friends and relatives effect is equivalent to compounding the immigrant stock by about 0.6% per year, which is somewhat less than the "depreciation" of about 1% in the immigrant stock through deaths and return migration.

The effects of immigration policy are discernible and have the expected effects. The quota for family members works through the uninteracted term, which is negative, and the positive interaction term. An increase of 10% in the family quota raises immigration from a country by 0.3%. The same proportionate increase in employment visas raises it by 1.4%. A 10% increase in the refugee allowance raises immigration by 0.5%, while the effect of the diversity quota is minimal. By contrast, the effects of the Immigration Reform and Control Act were relatively large; these are discussed further below.

VI. The Effects of Policy Changes

The impact of immigration policy on immigrant numbers is assessed by means of counterfactual simulations. These simulations provide an important check on the model as well as a gauge of the effects of policy. Dynamic simulations are made for each of the 81 countries in the data set, using the estimated equation in the third column of table 3. A counterfactual change in one of

TABLE 4.— THE EFFECTS OF IMMIGRATION POLICY
(ACTUAL/NO-POLICY-CHANGE COUNTERFACTUAL)

			Merging Hemispheres			
Year	1976	1977	1978	1979	1980	1981
Eastern Hemisphere	100.0	99.9	101.9	101.6	101.4	101.3
Western Hemisphere	100.0	132.5	131.2	92.6	89.1	84.7
World	100.0	113.2	111.5	97.8	96.4	93.8
			Immigration Control and Reform Act			
Year	1988	1989	1990	1991	1992	1993
Eastern Hemisphere	100.0	102.3	102.4	101.9	100.1	100.0
Western Hemisphere	100.0	207.4	213.6	233.0	105.5	103.9
World	100.0	148.1	158.1	169.1	102.4	101.5
			1990 Immigration Act			
Year	1991	1992	1993	1994	1995	1996
Eastern Hemisphere	100.0	107.5	107.3	107.2	107.4	107.2
Western Hemisphere	100.0	110.0	109.7	107.3	107.0	106.7
World	100.0	108.6	108.2	108.1	107.4	107.1

the explanatory variables (in this case policy-related variables) serves to change the level of gross immigration, which in turn alters the immigrant stock at the end of that year. The updated immigrant stock then influences the counterfactual level the following year, and so on. The effects of changes in policy can be assessed by comparing the counterfactual level of immigration with the actual level.[11]

The first case is the period in the late 1970s when the separate quotas for the Eastern and Western Hemispheres were merged into a worldwide quota. This affected the total number of visas for both nonimmediate family members and employment-based immigration. As noted earlier, the Western Hemisphere quota for nonimmediate relatives was cut by 26%, and then in 1979 the Eastern and Western Hemispheres were merged, cutting the total numbers under the quota by a further 7%. The quota for Western Hemisphere employment visas was raised from zero to 24,000 in 1977, and then in 1979, it was merged with the Eastern Hemisphere quota (of 34,000), with reductions in the total taking place in 1980 and 1981.

In the counterfactual simulation the quotas are held constant at the 1976 levels from 1977 onward, retaining the distinction between Eastern and Western Hemisphere countries. The results are displayed in the first panel of table 4. These figures are calculated as the ratio of the actual immigrant numbers to the counterfactual simulation and hence they reflect the effect of policy change in relative terms. In the years 1977–1978 the effect of the increase in employment visas outweighs that of the decline in family-based visas for the Western Hemisphere. The subsequent sharp decline in the Western Hemisphere total reflects the crowding out of Western Hemisphere immigration when the two sectors were merged. The overall decline in immigration between 1978 and the early 1980s is just a shade larger than the 5% cut in the overall quota.

The second change is the Immigration Reform and Control Act of 1986. As is well-known, the effects of IRCA were very large; this is reflected in the ratios in the second panel of table 4. The IRCA effects are especially marked for Western Hemisphere countries but only marginal for the Eastern Hemisphere. The figures can be compared with the ratio of IRCA legalizations to all other classes of admissions recorded in the INS immigration statistics. Over the years 1989 to 1991, IRCA legalizations were 126% of non-IRCA admissions, somewhat more than the figures estimated here. This suggests that the legalization program may have substituted for other admissions to some degree, particularly by reducing nonquota immigrants.

The third panel of table 4 simulates the effects of the Immigration Act of 1990, which took effect in 1992. The 1990 act increased the number of visas available to nonimmediate relatives by about a third between 1991 and 1992, a figure that was cut by 20% in 1995. In addition, the number of employment visas was more than doubled and a new category of diversity visas was introduced. Overall, these policy changes amounted to approximately a 75% increase in the number of available visas between 1991 and 1992–1994. However, the net effect on admissions would have been much less than this because some previously non quota categories, such as immediate relatives and certain employment-based immigrants, were absorbed under the quota for the first time. Our estimated effects of these changes, taken together, suggest that between 1991 and 1992–1994 the overall effect was to raise immigration by a little over 8%. This is smaller than the increase in the numbers for non-IRCA immigrants, which rose by 18% over the same period.

[11] These simulations include the equation errors and use the same depreciation parameter that was generated for each country/decade when calculating the immigrant stock (see appendix B2). Thus, a simulation using the actual values of the explanatory variables would exactly replicate the data.

TABLE 5.— GROSS IMMIGRATION RATE REGRESSIONS (81 COUNTRIES, 28 YEARS; DEPENDENT VARIABLE: LOG IMMIGRANTS ADMITTED/SOURCE-COUNTRY POPULATION)

	(1)	(2)	(3)
Constant	−18.31	−16.89	−16.65
	(7.3)	(8.4)	(38.8)
GDP per capita ratio (foreign/U.S.)	−2.47	−1.77	−2.76
	(3.0)	(2.6)	(4.0)
Schooling years ratio (popn. > 14) (foreign/U.S.)	4.00	3.08	3.79
	(4.2)	(4.0)	(4.2)
Share of population aged 15–29 (foreign)	12.46	10.32	
	(1.6)	(1.6)	
Inequality ratio (foreign/U.S.)	13.30	7.51	14.92
	(3.0)	(2.0)	(4.0)
Inequality ratio (foreign/U.S.) squared	−5.79	−3.07	−6.44
	(2.9)	(1.9)	(2.8)
Inverse of income squared (foreign)	−0.61	−0.33	−0.70
	(2.2)	(4.2)	(2.8)
Distance from U.S.	−0.28	−0.09	−0.20
	(7.7)	(2.1)	(5.0)
Landlocked	−0.36	−0.33	
	(1.0)	(1.1)	
English-speaking origin	1.19	0.31	1.04
	(3.8)	(1.0)	(3.5)
Immigrant stock ($t - 1$)/foreign population		89.90	
		(5.9)	
(Immigrant stock ($t - 1$)/foreign population)^2		−418.74	
		(8.4)	
IRCA legalization X_{irc}			1.37
			(3.4)
R^2 (between)	0.68	0.80	0.71
No. of observations	2,268	2,268	2,268

VII. The Regional Composition of U.S. Immigration

The source-country composition of U.S. immigration changed dramatically from the 1960s; this change has been widely ascribed to the effects of the 1965 amendments to the Immigration Act. Abolishing the quotas favoring European countries widened the opportunities for immigrants from a wider range of countries to compete on a more or less equal basis for the available visas. So what influence did source-country variables have on the regional composition of immigration that emerged under the post-1965 regime?

To get a feel for these magnitudes, we reestimate the model using the "between" estimator for the country means. This allows us to include the variables distance, English speaking, and landlocked, which may be important determinants of the source-country composition, but that do not vary over time. The first column of table 5 shows that the signs on the variables for economic fundamentals are similar to those in the fixed-effects regressions of table 3, although the magnitudes are larger. As expected, distance from the United States has a strong negative effect, while English speaking is strongly positive. As column 2 shows, these effects diminish in size when the immigrant stock and its square are added. Since the immigrant stock reflects past immigration, it captures much of the effect of slow-moving fundamentals over the longer term.

In order to capture the full effect of the economic fundamentals, the third equation in table 5 omits the immigrant stock terms as well as the insignificant variables for source-country population aged 15–29 and landlocked. Among the policy variables, only the IRCA legalization has a significant effect on the source-country composition, and so this is also included in column 3. We use this equation to pose the following question: What would the regional composition of immigration look like if a given variable took the same value for all source countries? To do this we first set each country's value of a given variable to the mean across all source countries for that year. We then predict a counterfactual immigration level for each country/year in the data set.[12]

Table 6 shows the differences, in percentage points, between the counterfactual composition of total immigration between 1971 and 1998 and the actual. The actual shares are given in the top line. Thus for the relative income counterfactual, the first entry in line 1 indicates that, had the income of Western European countries (relative to the United States) been the same as the mean for all sending countries, then this lower income level would have increased Western Europe's share of total immigration by 9.7 percentage points, from 7.3% to 17%. It would also raise the

[12] Setting each variable in turn to the annual mean ensures that total predicted immigration remains approximately constant—keeping the counterfactual in line with the overall immigration policy constraint. Predictions for the individual countries by year are summed to regional aggregates, adjusting to allow for different degrees of underrepresentation of the regional totals by the countries in our data set.

TABLE 6.—EFFECTS OF VARIABLES ON THE REGIONAL COMPOSITION OF U.S. IMMIGRATION, 1971–1998

	Western Europe	Eastern Europe	Middle East	East Asia	Africa	Oceania	Canada	Mexico	Caribbean	Central America	South America
	Baseline Shares: Percent of Total Immigration 1971–1998										
	7.3	6.0	3.9	31.5	3.0	0.6	1.8	21.7	12.8	5.2	6.2
Variable adjusted	Deviation from Baseline Due to Changing a Variable to the All-Country Mean for Each Year										
(1) Relative income	9.7	−1.9	−0.4	−6.9	−1.4	1.0	7.1	0.2	−3.2	−2.3	−1.9
(2) Education years	−3.0	−4.1	1.6	−1.1	3.2	−0.4	−1.5	−3.2	4.4	3.9	0.3
(3) Rel. inc. + education	3.2	−4.2	1.8	−7.4	1.1	−0.2	0.2	2.8	1.9	1.3	−0.5
(4) Inequality	3.2	6.5	−1.0	0.3	1.1	−0.1	0.8	−4.4	−3.9	−1.2	−1.2
(5) Poverty	−1.4	−0.6	−0.6	4.6	3.5	−0.1	−0.3	−3.6	−1.1	0.1	−0.4
(6) Inc. + edu. + ineq. + pov.	4.5	−2.7	−0.5	0.8	6.5	−0.3	0.6	−3.7	−3.6	0.1	−1.8
(7) Distance	−2.2	−0.1	1.2	29.4	2.0	1.0	−1.4	−15.3	−8.6	−3.5	−2.6
(8) English speaking	−0.3	1.4	0.9	−5.1	−0.7	−0.3	−1.0	5.2	−2.0	1.2	0.6
(9) Dist. + Eng.	−1.5	2.3	3.3	22.4	0.9	0.2	−1.6	−12.7	−8.8	−2.8	−1.7
(10) IRCA	0.4	2.8	0.0	13.0	1.4	0.0	−0.1	−12.4	−3.6	−2.4	0.9

shares of Oceania, Canada, and Mexico while reducing those of other regions, especially East Asia. As line 2 in the table shows, education effects often work in the opposite direction. Thus the combined effect of relative income and relative education in line 3 tends to be smaller than for income alone. The exceptions are Eastern Europe and East Asia, both of which have low income relative to their education levels, and where the combined effect boosted immigration substantially compared with other regions.

The effects of inequality are also important for some regions, as line 4 shows. Average levels of inequality slightly below the mean tended to increase immigration from Europe relative to Mexico, the Caribbean, and Central and South America, where inequality is much higher. Line 5 shows that the effect of the inverse of income squared, reflecting the poverty trap, is most important for the poorest regions, East Asia and Africa. For Africa, the effect of higher absolute income dominates the relative income effect, indicating that the poverty was a key constraining factor there. When all the economic fundamentals are adjusted to the mean values, there are offsetting influences (line 6). Thus, for Western Europe immigration was lower than otherwise because of high relative income and low inequality, whereas for Africa it was chiefly the poverty constraint that kept immigration lower than otherwise. For Mexico, the Caribbean, and South America, economic fundamentals boosted immigration compared with the counterfactual, but only modestly.

Line 7 shows that distance effects massively reduced the share of immigrants from East Asia while dramatically increasing the shares from Latin America.[13] For these two major source regions, distance effects dominate the effects of fundamentals. Those effects are somewhat offset by the English-speaking effect (line 8). Finally, the effect of IRCA (line 10) gave a major boost to (legal) immigration from Mexico, the Caribbean, and Central America compared with all other regions.

These results shed some light on the issue of differences by source region in the composition of immigration to Canada and the United States. One argument is that the Canadian points system has the effect of reducing the shares from regions that generate low-skilled immigrants (Borjas, 1993). Another view stresses the proximity of the United States to Latin America, and especially Mexico, in accounting for the lower average skill levels of its immigrants (Antecol, Cobb-Clark, & Trejo, 2003). Just as an illustration, increasing the distance from the United States of all Latin American countries by 1,500 miles (while preserving the overall mean) reduces the share from these sources by 10 percentage points. This is almost a third of the difference between the U.S. and Canadian shares.

VIII. Conclusion

Our results offer strong support for a model of U.S. immigration that stresses both individual incentives and policy constraints. Relative incomes and absolute incomes matter in a manner predicted by the theory, and the nonlinear effects of inequality support the predictions of the Roy model. But other variables matter, too—variables that are widely acknowledged to be important but that are often omitted in empirical work. The stock of previous immigrants from a given source country has substantial effect—drawing about six more immigrants annually per thousand of the stock. A part of this reflects the stance of immigration policy that encourages family reunion, but it is mostly the traditional friends and relatives effect that has been identified even in the absence of such policies (Hatton & Williamson, 1998). The effects of other policy changes are also discernible in the data, particularly changes in the size and structure of the immigration quota and the IRCA legalization program.

The effects of changes in immigration policy can be clearly discerned in the data, and we have made an effort to

[13] It should be noted, however, that such a massive expansion in immigration from Asia would likely be constrained by the per-country quota noted above, particularly in the cases of China and India.

incorporate not only the overall quota level, but also key elements of its structure. Our evidence confirms that different components of the quota also interact with variables like the immigrant stock in ways that are plausible but rarely implemented in studies of immigration flows. Major policy shifts such as the merging of hemispheres into a worldwide quota, the IRCA legislation, and the 1990 Act affected both the level and the source-region composition of immigration in a manner that is consistent with other evidence. This provides further support for our model as a realistic account of the factors that drive U.S. immigration.

Finally, the effects of the economic and demographic fundamentals on the composition of U.S. immigration by source region are mixed. While the effects of differences in source-country per capita income shifted the composition away from developed regions toward poorer regions, education effects generally work in the opposite direction. It is important to recognize that, both theoretically and empirically, what drives migration is income relative to education. Since these are strongly correlated across countries, the migration incentives are not as large as income gaps alone would suggest. Another key finding is the effects of variables like distance and English speaking, which do have decisive effects on the composition of U.S. immigration. As in models of international trade, gravity effects are important, even in the presence of a wide range of other variables.

REFERENCES

Antecol, Heather, Deborah A. Cobb-Clark, and Stephen J. Trejo, "Immigration Policy and the Skills of Immigrants to Australia, Canada and the United States," *Journal of Human Resources* 38 (Winter 2003), 192–218.

Borjas, George, J., "Self Selection and the Earnings of Immigrants," *American Economic Review* 77 (Sept. 1987), 531–553.

——— "Immigration Policy, National Origin, and Immigrant Skills: A Comparison of Canada and the United States" (pp. 21–43), in David Card and Richard B. Freeman (Eds.), *Small Differences that Matter: Labor Markets and Income Maintenance in Canada and the United States* (Chicago: University of Chicago Press, 1993).

Chiswick, Barry R., "Are Immigrants Favorably Self-Selected? An Economic Analysis" (pp. 61–76), in Caroline B. Brettell and James F. Hollifield (Eds.), *Migration Theory: Talking Across Disciplines* (New York: Routledge, 2000).

De Laet, Debra L., *U.S. Immigration and Politics in an Age of Rights* (Westport, CT: Praeger, 2000).

Hatton, Timothy J., and Jeffrey G. Williamson, *The Age of Mass Migration: Causes and Economic Effects* (New York: Oxford University Press, 1998).

——— "Demographic and Economic Pressure on Emigration Out of Africa," *Scandinavian Journal of Economics* 105 (September 2003), 465–486.

Jasso, Guillermina, Mark R. Rosenzweig, and James P. Smith, "The Changing Skill of New Immigrants to the United States: Recent Trends and Their Determinants" (pp. 185–225), in George J. Borjas (Ed.), *Issues in the Economics of Immigration* (Chicago: National Bureau of Economic Research, 2000).

Karemera, David, Victor I. Oguledo, and Bobby Davis, "A Gravity Model Analysis of International Migration to North America," *Applied Economics* 32 (October 2000), 1745–1755.

Mayda, Anna Maria, "International Migration: A Panel Data Analysis of Economic and Noneconomic Determinants," Georgetown University, unpublished paper (2004).

Pedersen, Peder J., Mariola Pytlikova, and Nina Smith, "Selection or Network Effects? Migration Flows into 27 OECD Countries, 1990–2000," IZA discussion paper 1104 (2004).

Ravallion, Martin, "Growth, Inequality and Poverty: Looking Beyond Averages" (pp. 62–80), in Anthony F. Shorrocks and Rolph van der Hoeven (Eds.), *Growth, Inequality and Poverty: Prospects for Pro-poor Economic Development* (Oxford: Oxford University Press, 2004).

Sjaastad, Larry, "The Costs and Returns of Human Migration," *Journal of Political Economy* 70 (October 1962), 80–93.

U.S. Immigration and Naturalization Service (INS), *Statistical Yearbook of the Immigration and Naturalization Service, 1998* (Washington, DC: U.S. Government Printing Office, 2000).

Warren, Robert, and Jeffrey S. Passel, "A Count of the Uncountable: Estimates of Undocumented Aliens Counted in the 1980 United States Census," *Demography* 24 (August 1987), 375–393.

Yang, Philip Q., *Post-1965 Immigration to the United States* (Westport, CT: Praeger, 1995).

APPENDIX A

Migration and Selection

This appendix provides a fuller derivation of equation (2) in the text, and it illustrates the effects on migration flows of changes in relative inequality between source and destination countries. Here we ignore the effect of age on the net present value of migration and examine the migration decision for individuals at a given age.

In the source country, *y,* skill endowments follow a normal distribution: $s \sim N(\mu_s, \sigma_s^2)$. The incomes that individual i ($i = 1, \ldots, n$) receives at home in country *y,* and would receive if he/she were to migrate to country *x,* are the following:

$$\text{Income in destination: } w_{xi} = \alpha_x + \beta_x s_i, \text{ distributed as } w_x \sim N(\mu_x, \sigma_x^2). \tag{A1}$$

$$\text{Income in origin: } w_{yi} = \alpha_y + \beta_y s_i; \text{ distributed as } w_y \sim N(\mu_y, \sigma_y^2).$$

Thus incomes, and income inequality, differ in origin and destination but incomes in x are perfectly correlated with those in y across individuals in the origin country. This simplifying assumption could be relaxed without qualitatively altering the results, provided that cov (w_x, w_y) is sufficiently positive (see Borjas, 1987, p. 533).

As discussed in the text, the cost elements are the following. Individual-specific migration costs, z_i, follow a normal distribution, $z \sim N(\mu_z, \sigma_z^2)$, with mean, μ_z, and variance σ_z^2, where z is independent of s (Cov $(s,z) = 0$). The constant cost elements, $c_1 - c_2(q)$, are the same for all potential immigrants. The cost associated with the skill-selective element of immigration policy is $\gamma(\delta - s_i)$, where δ is a threshold or benchmark skill level.

As shown in the text, the probability that an individual, *i,* will migrate from country y to x, m_i, is:

$$m_i = Prob(v_i > 0), \text{ where } v_i = w_{xi} - w_{yi} - z_i - c_1 + c_2(q) - \gamma(\delta - s_i). \tag{A2}$$

Summing over all n individuals in source country *y,* the emigration rate to x is:

$$M = 1 - \Phi\left[\frac{-\mu_x + \mu_y + \mu_z + c_1 - c_2(q) + \gamma(\delta - \mu_s)}{\sigma_v}\right], \tag{A3}$$

where Φ is the standard normal cumulative distribution function.

The standard deviation of v can be written as

$$\sigma_v = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \gamma^2 - 2\sigma_x\sigma_y + 2\sigma_x\gamma\sigma_s - 2\sigma_y\gamma\sigma_s}. \tag{A4}$$

TABLE A1.—EFFECTS OF INCOME DISTRIBUTION AND IMMIGRATION POLICY ON MIGRATION

Effect on migration rate of:	Destination is "relatively rich": $\mu_x > \mu_y + \mu_z + c_1 - c_2(q) + \gamma(\delta - \mu_s)$	Destination is "relatively poor": $\mu_x < \mu_y + \mu_z + c_1 - c_2(q) + \gamma(\delta - \mu_s)$
Income distribution in destination country	$dM/d\sigma_x > 0$ if: $\sigma_x < \sigma_y - \gamma\sigma_s$	$dM/d\sigma_x > 0$ if: $\sigma_x > \sigma_y - \gamma\sigma_s$
Income distribution in source country	$dM/d\sigma_y > 0$ if: $\sigma_y < \sigma_x + \gamma\sigma_s$	$dM/d\sigma_y > 0$ if: $\sigma_y > \sigma_x + \gamma\sigma_s$
Selective immigration policy	$dM/d\gamma > 0$ if: $\gamma > (\sigma_y - \sigma_x)\sigma_s + (\delta - \mu_s)(\sigma_v/v)$	$dM/d\gamma > 0$ if: $\gamma > (\sigma_y - \sigma_x)\sigma_s + (\delta - \mu_s)(\sigma_v/v)$

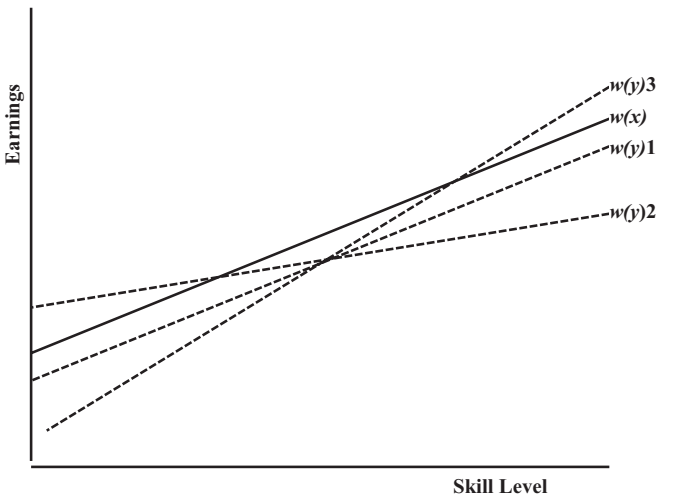
The effects of changes in income distribution and in the selectivity of immigration policy depend on the sign of the numerator in the bracketed term in (A3) as well as on the sign of the derivative of σ_v with respect to σ_x, σ_y, and γ. The following table gives the conditions for these effects to be positive on total migration, holding the underlying skill distribution constant.

We examine the case where destination-country income exceeds source-country income adjusted for migration costs ($\mu_x > \mu_y + \mu_z + c_1 - c_2(q) + \gamma(\delta - \mu_s)$, and assume γ is small. For a source country that is initially relatively equal ($\sigma_y < \sigma_x - \gamma\sigma_s$), rising inequality will increase immigration up to the point where $\frac{\sigma_y}{\sigma_x} = 1 + \frac{\gamma\sigma_s}{\sigma_x}$, beyond which immigration will decline. The effect of changing inequality in the destination is the exact opposite. Thus the immigration rate is an inverse U-shaped function of the ratio of source to destination inequality. Note also that, in the presence of skill-selective immigration policy ($\gamma > 0$), the peak immigration rate will occur at a point where the inequality ratio exceeds 1.

These effects are illustrated in Figure A1. The figure shows wage earning profile, $w(x)$, for the destination and three alternative profiles, $w(y)$, for the source country. The source country profiles are net of migration costs and they intersect at a mean income level that is lower than the mean of $w(x)$. When source and destination profiles are parallel, as in $w(x)$ and $w(y)1$, then all individuals in the source country (with sufficiently low z) have an incentive to migrate. If the source country has a more equal income distribution, as in $w(y)2$, then low-skill individuals for whom $w(y)2 > w(x)$ will not migrate and total migration will be lower than previously. In the case where the source country is more unequal than the destination, as in profile $w(y)3$, migration will also be lower than in the case of parallel profiles, and migrants will be negatively selected.

These relationships will be shifted by skill-selective immigration policy. This is equivalent to steepening the slope of $w(y)$ in Figure A1, at the same time as shifting the profile down at the median skill level. Increasingly selective policy always increases the positive selection of immigrants, and could increase migration, an effect that is more likely the lower inequality is in the source country and if $\frac{\sigma_y}{\sigma_x} > 1 + \frac{\gamma\sigma_s}{\sigma_x}$. In this case the shift effect dominates the slope effect.

FIGURE A1.— EARNINGS BY SKILL PROFILES IN DESTINATION AND SOURCE COUNTRIES



APPENDIX B

Data Used in Estimation: Sources and Methods

1. The INS Gross Immigration Data

The data for the number of immigrants to the United States by country is taken from the U.S. Immigration and Naturalization Service (INS) *Statistical Yearbooks*.[14] The data cover all legal immigration, including refugees, and include those who applied from abroad and those who are already in the United States and are adjusting to permanent status. The country-of-origin classification used here is by country of birth rather than by country of last residence. Choosing country of birth rather than country of last residence allows us to gain consistency between the immigrant flow and the stock of resident immigrants, which is available only by place of birth.

Before 1976, the INS defined a fiscal year as July 1 through June 30. For example, fiscal 1974 began on July 1, 1973, and ended on June 30, 1974. In 1976, however, the INS changed its definition of a fiscal year to October 1 through September 30. Because this change occurred within our data period, the pre-1976 annual observations have been adjusted to conform to the 1976 definition of a fiscal year. The INS did not report monthly totals of immigrants admitted by country of birth, so we used data that the INS labeled as "Immigrants Admitted by Region and Country of Birth" for the third quarter (July 1 through September 30) of 1976. To convert the 1976 "June" fiscal year into a "September" fiscal year, we added the 1976 third quarter totals to the June fiscal year 1976 totals for each country. These sums represent the total immigration from each country to the United States during the fifteen-month period from July 1, 1975, to September 30, 1976. To estimate the immigration for the twelve months of the new September fiscal year 1976, we multiplied the fifteen-month totals by 0.8. This operation gives four-fifths of the fifteen-month totals, results that should be roughly equivalent to the amount of immigration that occurred during four of the five quarters represented from July 1, 1975, to September 30, 1976. This process was then repeated for the previous years. Thus, to convert the June fiscal year 1975 into a September fiscal year, we added one-fifth of the fifteen-month totals that we used to adjust fiscal year 1976 to the June fiscal year 1975 figures. We then took four-fifths of these sums as the data for the new September fiscal year 1975. Thus, all of the annual gross immigration figures reported in this adjusted INS database now represent October to September totals.

2. Annual U.S. Foreign-Born Stock Values

Benchmark Estimates: Foreign-born population stock data for census years 1970, 1980, and 1990 are taken from the Census Bureau, Population Division, technical working paper no. 29, *Historical Census Statistics on the Foreign-Born Population of the United States: 1850–1990* (1999). This paper by Campbell J. Gibson and Emily Lennon is available online

[14] Since 2002 this has become the *Yearbook of Immigration Statistics* of the Department of Homeland Security.

at http://www.census.gov/population/www/documentation/twps0029/twps0029.html. Since the 2000 Census figures were not yet available at the time of writing, the only source of post-1990 foreign-born stock values is the Census Bureau's annual Current Population Survey (CPS) March demographic supplement. These data were obtained from the online data extraction service at http://ferret.bls.census.gov/cgi-bin/ferret. A description of the survey's methodology is available online at http://www.bls.census.gov/cgi-bin/dms?Folder=657. The CPS uses a system of supplemental weights to estimate nationwide foreign-born stock values from the information it collects from its sample. Although the CPS data are useful for displaying demographic trends, the small sample size makes the estimates highly variable. Furthermore, CPS data is available only after 1994 (and up to 1998). To fill out our data set, we used the 1990 Census values and the 1994–1998 CPS data to estimate a simple source-country-specific regression against time. The regression was then used to generate predicted foreign-born by source country for 1998.

Interpolating Between-Census Years: In order to obtain annual estimates of the foreign-born stock by country, we interpolate between the benchmarks established obtained from the Census or calculated from the CPS, using the following stock adjustment equation:

$$S_{t+1} = M_t + dS_t,$$

where S_t is the stock at the beginning of year t and M_t is the flow during that year. We use the gross flow series by birthplace (as defined above) in order to update the stock. The stock observed midway through a year is updated with the flow beginning in that year but carrying through to the next year.

As noted in the text, the parameter d reflects deaths, return migration, and illegal immigration, which subtract or add to the stock independently of the additions through gross immigration and hence $1 - d$ is the rate at which the stock depreciates. This depreciation rate is calculated for each interval between Census or CPS benchmarks using an iterative procedure beginning with S_t, such that the value of S_{t+10} obtained by cumulating forward is reconciled with that of the next Census benchmark. Thus there is a different value of d for each country for each interval between benchmarks. However, in some cases no Census estimate was available for 1970; in that case the value of d calculated for the 1980–1990 interval was used, together with the gross migration series, to extrapolate back to 1970. Similarly, where it was not possible to construct a benchmark figure for 1998 using the CPS data, we use the 1980–1990 value of d to extrapolate forward to 1998.

3. *Economic and Demographic Variables*

The relative income variable is real GDP per capita at 1985 international prices from the Penn World Tables version 5.6 updated by the World Bank, available at http://www.worldbank.org/research/growth/GDNdata. Average years of education for the population age fifteen and over are from the database of Barro and Lee, available at http://www2.cid.harvard.edu. Since the frequency is five years, the data for each country were linearly interpolated. The share of population aged 15–19 are taken from the annual data (available on CD) underlying the United Nations, *World Population Prospects: The 2000 Revision* (Geneva: UN Population Division). Household income inequality is based on the data originally assembled by Deininger and Squire at the World Bank, now augmented and available from the WIDER Institute at http://www.wider.unu.edu/wiid/wiid.htm. The observations selected are (almost) exclusively those labeled as "high quality" with linear interpolations between these observations. Certain adjustments were made according to whether the observations were for income/expenditure, gross/net income, or individuals/households.

4. *Immigration Policy Variables*

Immigration policy is characterized in equation (3) in the text by a series of variables denoted by X. The Xs are variables reflecting the quota limits, which are interacted where appropriate, with different variables representing country characteristics. The derivation of the Xs for each category is detailed below:

Nonimmediate Relatives: (X_r): Nonimmediate relatives enter under the following preference categories in the post-1990 legislation (with total numbers in parentheses): (a) adult married children of U.S. citizens (23,400); (b) spouses and unmarried children of U.S. residents, 75% of whom must be minors (114,200); (c) married children of U.S. citizens (23,400); and (d) siblings of adult U.S. citizens (65,000). Before 1992 the preference categories were broadly similar (with percentages of total quota in parentheses): (a) unmarried children of U.S. citizens (20%); (b) spouses and unmarried children of resident aliens (20%); (c) married children of U.S. citizens (10%); and (d) siblings of U.S. citizens (24%).

The total number of visas available for these categories is calculated as follows:

Eastern Hemisphere 1968–78: 170,000	World 1979–81: 214,600
Western Hemisphere 1968–76: 120,000	World 1981–91: 210,000
Western Hemisphere 1977–78: 88,800	World 1992–94: 281,000
	World 1995–98: 226,000

Note that until 1976 there were no preference categories for the Western Hemisphere, and so the entire quota is included under this heading. For 1977–1978, when a preference system was in force, the number is the total quota net of employment and refugee categories. From 1992 the figure is calculated as the total quota net of employment, diversity, and immediate family categories *plus* the floor of 226,000 for nonimmediate relatives.

The variable X_r is the total number of visas divided by world population (excluding the United States), and that value is applied to each country. Before 1978 it is calculated to produce a separate value for each hemisphere by using the respective hemispheric populations.

Employment Visas (X_e): From 1992 employment-related visas were given under the following categories (with total numbers in parentheses): (a) individuals of outstanding ability (40,000); (b) professionals with advanced degrees or with exceptional abilities (40,000); (c) skilled workers or unskilled shortage workers (40,000); (d) special occupations including religious workers (10,000); and (e) investors (10,000). Before 1992 there were just two employment categories (with percentage of quota in parentheses): (a) exceptional professionals, scientists, and artists (10%); and (b) skilled and unskilled workers in shortage occupations (10%).

The total number of visas for these categories is calculated as follows:

Eastern Hemisphere 1968–78: 34,000	World 1979: 58,000
Western Hemisphere 1968–76: 0	World 1980: 56,000
Western Hemisphere 1977–78: 24,000	World 1981–91: 54,000
	World 1992–98: 140,000

The variable X_e is the total number of visas divided by the world population. Before 1979, it is calculated to produce a separate value for each hemisphere by using respective hemispheric populations.

Diversity Immigrants (X_d): The diversity category was introduced for the first time in the 1990 Immigration Act. Diversity visas are a special category applied to countries that were under-represented in U.S. immigration following the 1965 amendments. Countries eligible for diversity visas are those with less than 50,000 immigrants in the preceding five years. In the period 1992–1994, 40,000 (AA-1) visas were available and these were awarded among the applicants by lottery. For those years, the list of eligible countries comprised mainly Europe (excluding the former Soviet Union), Canada, and a few other countries. Within this list there was a quota specific to Ireland, with the rest distributed among the other eligible countries. From 1995, 55,000 (DV) visas were available, and the list of eligible countries includes most of the world, with a few specific exceptions. For these years the total allocation was divided into quotas by continent, with no specific country quotas and a per-country ceiling of 7% of the worldwide total.

The variable X_d is defined only for 1992–1998 and only for those countries eligible to participate; otherwise it takes the value of 0. For 1992–1994 it is defined for each participating country as the total number of non-Irish visas available divided by the total population of countries eligible to participate, excluding Ireland. The variable for Ireland is the Irish quota divided by Irish population. For 1995–1998 it is calculated by continent and applied to each eligible country within that continent.

Refugees and Asylees (X_a): Refugees and asylees were integrated in the total quota until the 1980 Refugee Act. Since then the number, which is not part of the overall ceiling, is determined annually. The "quotas" for refugees are as follows:

Eastern Hemisphere	1968–78: 10,200		
Western Hemisphere	1968–76: 0		
Western Hemisphere	1977–78: 7,200		
World	1979: 50,000	1986: 67,000	1993: 116,000
	1980: 213,700	1987: 70,000	1994: 117,500
	1981: 217,000	1988: 87,500	1995: 111,000
	1982: 140,000	1989: 116,500	1996: 90,000
	1983: 90,000	1990: 110,000	1997: 78,000
	1984: 72,000	1991: 116,000	1998: 83,000
	1985: 70,000	1992: 123,500	

The variable X_a is defined as the refugee quota divided by the country population. Before 1979 it is calculated to produce a separate value for each hemisphere by using respective hemispheric populations. From 1980 the overall allocation was divided into regional totals. A separate value was therefore calculated for each region and applied to all countries in that region.

Immigration Reform and Control Act (X_{irc}): As regards permanent admissions, IRCA made two major provisions. The first was legalization of illegal immigrants who had resided in the United States continuously since before 1982. After first applying for temporary status (during a window in 1987–1988), these immigrants could gain permanent status after eighteen months. The second granted temporary visas to seasonal agricultural workers (SAWs), previously working illegally, with the right to become permanent immigrants after one year. Further temporary visas were made available for new agricultural workers, with the right to become permanent after two years. The IRCA provisions are relevant here only insofar as they offered a new channel for permanent immigration. Most of the illegal immigrants eligible for adjustment under IRCA were from Mexico and Central America (especially the former), and the bulk of these adjustments took place in 1989–1991.

Our variable X_{irc} is derived from the number of illegal immigrants living in the United States in 1980 estimated by Warren and Passell (1987, pp. 380–381). Estimates for 1980 are appropriate given that legalizations applied to those living in the United States since before 1982. The estimates are based on a comparison of Census data for 1980 and measures of the stock of legal immigrants based on INS data. The total number of just over two million is considered as a lower bound. Figures are given for specific countries and for continental remainders; the latter were distributed across countries using 1980 population weights. The variable X_{irc} was obtained by dividing the number of illegals, thus calculated, by the origin country population in 1990. It is applied only to the years 1989–1991.

Backlog (X_b): In 1995 the burden of dealing with adjustments shifted from consular offices to the INS, as a result of abolishing the requirement that eligible immigrants present in the United States had to leave the country and apply for immigrant visas through consular offices abroad. As a result, between the end of fiscal 1994 and fiscal 1998 the backlog of applications pending a decision increased from 121,000 to 811,000. The INS estimated that, in the absence of the increase in the pending caseload, legal immigration would have been 110,000 to 140,000 higher for each of the years 1995 to 1998 (INS, 2000, p. 15).

Our variable X_b is simply a dummy for the years 1995–1998.

5. *The Balanced Panel*

In our econometric work and in the simulations that follow, we use a balanced panel of 81 countries across 28 years. Although there are about twice this number of source countries separately identified in the INS immigration series, the remainder were dropped from the sample because one or more of the explanatory variables was not available for some or all of the period. In cases where countries have split or amalgamated during the period, they have been reaggregated to the combined total throughout. Thus for immigration and the foreign-born stocks, Czechoslovakia, Yugoslavia, and the Soviet Union have been reassembled. East and West Germany are together throughout as are China and Taiwan. In these cases the economic and demographic variables used to explain immigration are aggregated using current population weights.

Panel A of Table B1 lists all the countries in the data set by region. As panel B shows, these account for 82.5% of all immigration over the period. But, as reflected in panel C, under-representation is greater for some regions than others. This is especially important for Africa, the Caribbean, and the Middle East. Important countries that are omitted include Vietnam, Iraq, and Lebanon in Asia; Ethiopia, Somalia, and Nigeria in Africa; and Cuba and Haiti in the Caribbean.

TABLE B1.—THE BALANCED PANEL FOR U.S. IMMIGRATION, 1971–1998

A: Countries in the Balanced Panel

Western Europe	Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom (16)
Eastern Europe	Czechoslovakia (frmr.), Hungary, Poland, Romania, Soviet Union (frmr.), Yugoslavia (frmr.) (6)
East Asia	Bangladesh, China (inc. Taiwan), Hong Kong, India, Indonesia, Japan, Korea (South), Malaysia, Nepal, Pakistan, Philippines, Singapore, Sri Lanka, Thailand (14)
Middle East	Cyprus, Iran, Israel, Jordan, Turkey (5)
North America	Canada, Mexico (2)
Caribbean	Barbados, Dominican Republic, Jamaica, Trinidad and Tobago (4)
Central America	Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, Panama (6)
South America	Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Guyana, Paraguay, Peru, Uruguay, Venezuela (11)
Africa	Algeria, Cameroon, Egypt, Ghana, Kenya, Senegal, Sierra Leone, South Africa, Sudan, Tanzania, Tunisia, Uganda, Zambia, Zimbabwe (14)
Oceania	Australia, Fiji, New Zealand (3)

B: Numbers in Balanced Panel and in Total Immigration 1971–1998, by Period

Period	Immigrants in Sample	Total Immigration	Percent in Data Set
1971–1980	3,656,107	4,389,630	83.3
1981–1990	5,913,094	7,337,806	80.6
1991–1998	6,374,841	7,597,762	83.9
1971–1998	15,944,042	19,325,630	82.5

C: Numbers in the Balanced Panel and Total Immigration, 1990–1998, by Region

Region	Immigrants in Data Set	Total Immigration	Percent in Data Set
Europe	2,507,796	2,575,018	97.4
Asia	4,959,606	6,839,410	72.5
Africa and Oceania	379,085	700,070	54.1
North America	6,923,475	8,034,314	86.2
South America	1,174,080	1,176,386	99.8