WPS3252

EXPLAINING U.S. IMMIGRATION 1971-98

by

Ximena Clark, World Bank
Timothy J. Hatton, University of Essex
Jeffrey G. Williamson, Harvard University

World Bank Policy Research Working Paper 3252, March 2004

Revision of a paper 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, the LACEA, and the WIDER Conference on World Migration (September 26-28, 2002), as well as for the excellent research assistance of Rachael Wagner. Hatton acknowledges with pleasure the financial support of the British Academy and Williamson does the same for the National Science Foundation SES-0001362.

Abstract

This paper develops and estimates a model explaining the level and country-source composition of United States immigration since the early 1970s. The model incorporates ratios to the U.S. of source country income, education, and demographic structure, as well as relative inequality. Our model also incorporates both network effects, as reflected in the stock of previous immigrants, and various controls for immigration quota policy. The model is estimated on a panel of 81 source countries for 1971-98. The results strongly support the influence of economic, demographic and geographic variables as well as policy. The regression results are used to identify those factors that most influenced the changing composition of U.S. immigration by source.

Ximena Clark
The World Bank
181 H Street, NW
Washington DC 20433
xclark@worldbank.org

Timothy J. Hatton
Department of Economics
University of Essex
Colchester, UK
hatton@essex.ac.uk

Jeffrey G. Williamson
Department of Economics
Harvard University
Cambridge, Massachusetts 02138
and NBER
jwilliam@kuznets.fas.harvard.edu

JEL No. F22, J1, O15

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 US 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 income, 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" network 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 countries over a longer period than does existing work on late 20th century US immigration.

We start in the next section by providing some background to US 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 economic and demographic variables on the composition of immigration by source region and the effects of major shifts in immigration policy on the total numbers.

Immigration and Immigration Policy

Changes in US immigration over the last 50 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 percent of the total. The counterpart to this is the sharp rise in the proportion coming from Asia; the other notable feature being 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 percent of visas to relatives of US citizens or residents, 6 percent to refugees, and 30 percent to employment-based categories. Children and

spouses of US citizens were exempt from the quota, underpinning the strong emphasis on family reunification. 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 provided for the legalization of illegal immigrants who had resided in the US 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 US 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, were given to family-sponsored non-immediate relatives of US 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]

---

[1] Further details of numbers allocated to different preference categories are given in Appendix 2D.
[2] The maximum number of visas allocated to non-immediate 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 2D.
[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.

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 percent 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 percent for Asia (74 percent when "immediate relatives" are included) and the Americas (86 percent when "immediate relatives" are included), reaching an enormous 42 percent for Mexico (88 percent 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. These provided an additional route to legal immigration largely for Western Hemisphere immigrants, and particularly from Mexico. The other source is by those entering as temporary workers and trainees with H, O and P visas, the numbers of which soared from 75,000 in 1985 to 430,000 in 1998. This rising source originated chiefly from Europe and Asia. These are not part of the overall immigration total, but temporary visas clearly have been used as an intermediate step before adjusting to permanent status.

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 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 that follows in this tradition. It emphasizes the roles 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,....,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 are 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 1. 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 an inverse 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 1).

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 is the potential migrant 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]

Explaining Immigration

Recent studies of US 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 US relative to the source country population,

[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: $PV(a) = \frac{D}{r}\left[1-(1+r)^{-(46-a)}\right]$, which is a decreasing function of a.

representing the propensity to emigrate to the United States. Borjas (1987) found that, for a cross section of average emigration rates 1951-80, 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 1982-6 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.

More recently David Karemera, Victor Oguledo and Bobby Davis (2000) used panel data on emigration rates 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 US income, and negatively to the US 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 US was modeled as a dummy variable only.

In order to study the effects of policy change, Guillerma Jasso, Mark Rosenzweig and James Smith (2000) modeled male immigrants admitted as husbands of US citizens over the period 1972-90. They argued that this category, which was not subject to the quota, was nevertheless influenced by immigration policy, both directly, due to tightening eligibility conditions, and indirectly, due to substitution across entry categories in response to changes in the conditions for entering through other routes. In addition to income and education, policy dummies were found to matter. In particular, application of the preference system to

[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 1).

the Western Hemisphere raised the numbers arriving as male spouses from that region, while the Immigration Marriage Fraud Amendments of 1986 reduced the numbers.

Previous studies have contributed much, but they suffer a number of shortcomings. First, they either use country cross-sections, or cover a limited number of years in time series, or only explore a subset of all immigration. We think there is an advantage to being more comprehensive: by covering emigration regions in decline, ascension, and transition we are more likely to identify the economic and demographic fundamentals driving changing immigrant source. Second, a number of key variables stressed by theory are often omitted. Among these are the age structure of sending population and measures of human capital and/or the return to skills. The omission of variables that are central to migration theory make it impossible to assess the role of sending country demographic and human capital attributes. Third, despite the obvious importance of "chain migration" effects which have been greatly reinforced by family reunification policies, proxies for these effects -- like the resident immigrant stock -- are often omitted from the analysis. Finally, shifts in immigration policy are typically reflected by dummies rather than by variables that take full account of changes in the size and structure of quotas, and to whom they apply.

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

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

The left-hand side variable is migration to the US 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 US. 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 US (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 1). Hence $\beta_4 > 0$, $\beta_5 < 0$. Here inequality is represented by the gini coefficient of household income.

Migration costs constrain the move. As in any gravity model, these costs rise with distance from the US; hence, $\beta_6 < 0$. Such costs are also associated with whether the source country is landlocked and whether it is predominantly English-speaking; $\beta_7 < 0$, $\beta_8 > 0$. 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 US at time t-1 relative to the population of country j. Relatives (and friends) abroad reduce migration costs, $\beta_9 > 0$. We expect this effect to diminish with size (over the relevant range, hence $\beta_{10} < 0$) if it is being driven by immigrant job search and settlement costs (diminishing returns to network externalities). We would not expect the effect to diminish with size if instead the forces reflected 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 2D. 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_{11} through β_{14} are expected to be positive.

The variable X_r represents the quota for non-immediate 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 choosing 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 zero. 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 (e.g. Hatton and 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 1989-91, when the bulk of legalizations took place, and β_{15} is therefore expected to be positive. Finally, X_b is a dummy for the years 1995-98 when, due to administrative changes in the processing of visa applications, there was a progressive rise in the backlog. As a result, recorded immigration for

these years was lower than it would otherwise have been, and the dummy is therefore expected to be negative; $\beta_{16} < 0$. Details of the derivation of these variables are given in Appendix 2D.

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 1990 (see Appendix 2A and E). These countries form 82.5 percent of all US 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 2C. 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 2B.

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. The right hand side variables are as in equation (3). We also include fixed effects for nine geographical regions (not reported in Table 3). These are assumed to capture, among other things, the availability of alternative migrant destinations, since third country effects are not included in the model. We also include separate dummies for the border states, Canada and Mexico.

The results from estimating the equation on this pooled cross section/time series dataset appear in Table 3. The first column excludes the immigrant stock variable and all the policy related variables. All the explanatory variables are significant with the expected signs

and they account for nearly three quarters of the variation in the dependent variable. When, in the second column, the (lagged) immigrant stock is added the coefficients of the other variables are somewhat attenuated, but the overall pattern remains the same. The attenuation has an obvious explanation: the lagged immigrant stock variable has embedded in it the influence of previous economic and demographic fundamentals that are correlated with present fundamentals. The full model in the third column includes the policy variables; adding these has little effect on the coefficients of the economic and demographic fundamentals. All these variables take the expected signs although those representing diversity, refugees and the processing backlog have low levels of significance. Other variables suggested by the literature, such as an index of source country civil rights, and the US unemployment rate, failed to provide significant coefficients and thus were excluded throughout.

It is worth examining the quantitative implications of some of these estimated coefficients, focusing on the third column. The relative income term implies that a 10 percent increase in a country's income per capita reduces the immigration rate by around 5 percent. A more dramatic implication of the coefficient on relative income is this: moving from an income level typical of Western Europe to one typical of South America raises the immigration rate by about 70 percent. A 10 percent increase in a country's years of schooling (equivalent to 0.55 years for the average source country) increases the immigration rate by 13 percent. Again, a more dramatic implication of the coefficient on relative schooling is this: moving from an education level typical of Western Europe to one typical of South America reduces the immigration rate by about 50 percent. Relative schooling is also interacted with the variable X_e, and the positive coefficient supports the interpretation that this element of immigration policy is (mildly) skill-selective. Through this channel, increasing a country's relative schooling level by 10 percent raises its immigration rate by a further one percentage point.

Raising a source country's share aged 15-29 by 10 per thousand of population increases immigration rate by 3.7 percent or by 2.6 per thousand. Thus, demographic effects have a significant direct influence by changing the number of potential immigrants, although their impact would be bigger if we added the effect of demography on the sending country's per capita income. The influence of relative inequality is more complex because the variable enters non-linearly. The quadratic peaks at a ratio of the foreign/US gini coefficient of 1.3, consistent with the prediction of the Roy model (a peak greater than one) in the presence of selective immigration policy (see Appendix 1). Thus, moving from a relative inequality ratio typical of South America (1.20, higher inequality than the US) to one typical of Western Europe (0.82, lower inequality than the US) reduces the immigration rate by 25 percent, a very sizeable effect. This is because, even if mean incomes were the same in both regions, South Americans would have an incentive to migrate across a wide range of skill levels whereas, for Europeans, that incentive would be concentrated only among those at the higher skill levels.

The coefficient on the migrant stock by itself (i.e. ignoring its interaction with Xr) is of particular interest because it reflects the non-policy 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 zero when the migrant stock in the US reaches 11.6 percent of the source country population. At the mean (1.3 percent 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 27.7 immigrants. This direct effect is augmented by an indirect effect working through the policy variable representing the quota on non-immediate relatives (Xr). This adds a further 2.2 immigrants, yielding 29.9 more immigrants per year for every 1,000 added to the existing immigrant stock. Thus, the overall 'friends and relatives effect' is powerful. It is equivalent to compounding the immigrant stock by 3 percent per year and it is more than enough to

compensate for the 'depreciation' of the immigrant stock through deaths and return migration (about 1 percent on average).

The effects of immigration policy are discernible although not as strong as source country economic, demographic and geographical characteristics. An increase of 10 percent in the quota for family members raises immigration from a country by 0.4 percent; the same proportionate increase in employment visas raises it by 1 percent. Proportionate changes in the diversity quota and in the refugee allowances have relatively small effects since these are small components of the overall immigration program. By contrast, the effects of the Immigration Reform and Control Act were relatively large and these are discussed further below.

The Regional Composition of US Immigration

The source-country composition of US immigration changed dramatically from the 1960s and this 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 pose the 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 use equation (3) in Table 3 to predict a counterfactual immigration level for each country/year in the dataset. 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 total immigration by region are adjusted to allow for different degrees of under-representation of the regional totals by the countries in our dataset.

Table 4 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 per capita income counterfactual, the first entry in line (1) indicates that, had the per capita incomes of Western European countries 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 5.2 percentage points, from 7.3 percent to 12.5 percent. It would also raise the shares of Canada and Mexico while reducing those of other regions, especially East Asia. On the whole, the effects of education work in the opposite direction but not for Eastern Europe and East Asia, both of which have low income relative to their education levels. Line (3) shows that the combined effect of income and education boosted immigration from these regions compared with all others while the opposite is true for Mexico and the Caribbean.

The effects of inequality are smaller than income and education; levels of inequality below the mean tended to suppress immigration from Europe and Asia relative to Mexico, the Caribbean, Central and South America. Similarly, the effects of demographic structure have relatively small overall effects on the composition of immigration, with ageing populations tending to suppress the European share and youthful populations boosting slightly the shares from Latin America. Because the effects of economic and demographic fundamentals often cut in opposite directions, their combined effects shown in line (6) are often modest. The most important joint effect of these fundamentals was to reduce the share of immigrants coming from Europe and to raise the share coming from East Asia.

Line (7) shows that giving each country the mean for the landlocked variable (0.12) makes very little difference to regional composition of immigration. By contrast, adjusting each of the countries to the mean of the English-speaking dummy (0.28) has much larger

effects--effects that particularly favored East Asia and disadvantaged Mexico. More important still are the effects of distance shown in line (9). Distance effects massively reduced the share of immigrants from East Asia while dramatically increasing the shares from Latin America. They also dominate the combined effects of landlocked, English speaking and distance shown in line (10).

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 et. al., 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 11 percentage points. This is a one-third of the difference between the US and Canadian shares. However, this will be an underestimate since it does not account for border effects, or for the long-run endogeneity of the immigrant stock.

The final line in Table 4 shows the effects of the immigrant stock. As we have seen, the 'friends and relatives effect', resulting from chain migration underpinned by the family reunification element in immigration policy, has potentially large consequences. Setting the stock to source country population equal across all countries increases the share of immigrants coming from East Asia (low stock/population ratio) by nearly 16 percentage points while it reduces the share from Mexico and the Caribbean combined (high stock/population ratio) by 23 percentage points. Clearly such effects are endogenous in the long run and they largely reflect the cumulative effect of the other variables driving immigration. Taking them into account would magnify the effects of the fundamentals on the immigration shares reported in Table 4.

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 dataset, again using the estimated equation in the third column of Table 3. A counterfactual change in one of 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.[7]

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 non-immediate family members and employment-based immigration. As noted earlier, the Western Hemisphere quota for non-immediate relatives was cut by 26 percent, and then in 1979 the Eastern and Western Hemispheres were merged, cutting the total numbers under the quota by a further 7 percent. 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 onwards, retaining the distinction between Eastern and Western Hemisphere countries. The results are displayed in the first panel of Table 5. 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 1977-8 the effect of the increase in employment visas outweighs that of the decline in family-based visas for the Western

[7] These simulations include the equation errors and use the same "depreciation" parameter that was generated for each year when calculating the immigrant stock (see Appendix 2B). Thus a simulation

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 reflects the cut in the overall quota, although, here again, the effects are much larger than the change in the quota.

The second change is the Immigration Reform and Control Act of 1986. As is well known, the effects of IRCA were very large and this is reflected in the ratios in the second panel of Table 5. The IRCA effects are especially marked for Western Hemisphere countries and 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 percent of non-IRCA admissions, somewhat less than the figures estimated here. This suggests that the legalization program added a further, indirect, twist to total immigration by also increasing the number of non-quota immigrants.

The third panel of Table 5 simulates the effects of the Immigration Act of 1990, which took effect in 1992. The 1990 Act increased the number of visas available to non-immediate relatives by about a third between 1991 and 1992, a figure that was cut by 20 percent in 1995. In addition, the number of employment visas was more than doubled and the new category of diversity visas was introduced. Overall, these policy changes amounted to approximately a 75 percent increase in the number of available visas between 1991 and 1992-94. 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 1992-4 the overall effect was to raise immigration by 17 percent. This is highly consistent with the trend in the INS statistics for non-IRCA immigrants, which rose by 18 percent over the same period.

---

using the actual values of the explanatory variables would exactly replicate the data.

Conclusion

Our results offer strong support for a model of US immigration that stresses both individual incentives and policy constraints. Relative incomes and education, and source country demography all matter in a manner predicted by the theory. In addition, the non-linear 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 a substantial effect--drawing 30 more immigrants annually per thousand of the stock. In part this reflects the stance of immigration policy that encourages family reunion, and it provides a powerful cumulative effect. 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 the economic and demographic fundamentals on the composition of US 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 US immigration. As in models of international trade, gravity effects are important, even in the presence of a wide range of other variables.

Finally the effects of changes in immigration policy can be clearly discerned in the data and we have made an effort to 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 and relative education levels 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 US immigration.

References

Antecol, H., Cobb-Clark, D. A. and Trejo S. J. (2003), "Immigration Policy and the Skills of Immigrants to Australia, Canada and the United States," *Journal of Human Resources 38* (1), pp. 192-218.

Borjas, George J. (1987), "Self Selection and the Earnings of Immigrants," *American Economic Review,* 77, pp. 531-53.

Borjas, George J. (1993), "Immigration Policy, National Origin, and Immigrant Skills: A Comparison of Canada and the United States," in D. Card and R. B. Freeman (eds.), *Small Differences that Matter: Labor Markets and Income Maintenance in Canada and the United States* (Chicago: University of Chicago Press).

Chiswick, Barry R. (2000), "Are Immigrants Favorably Self-Selected? An Economic Analysis," in C. D. Brettell and J. F. Hollifield, eds., *Migration Theory: Talking Across Disciplines* (New York: Routledge).

Hatton, T. J. and Williamson, J. G. (2003), "Demographic and Economic Pressure on African Emigration," *Scandinavian Journal of Economics* (forthcoming).

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

Karemera, D., Oguledo, V. I., and Davis, B. (2000), "A Gravity Model Analysis of International Migration to North America," *Applied Economics,* 32 (13), pp. 1745-55.

Sjaastad, L. (1962), "The Costs and Returns of Human Migration," *Journal of Political Economy,* 70, pp. 80-93.

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

APPENDIX 1

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 for 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, …, n) receives at home in country y, and would receive if he/she were to migrate to country x, are:

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

Income in origin: $w_{yi} = \alpha_y + \beta_y s_i$; distributed as $w_y \sim N(\mu_y, \sigma_y^2)$. (A1)

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:

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

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

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

where Φ is the standard normal distribution function.

The standard deviation of v, can be written as:

$$\sigma_v = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \gamma^2 - 2\sigma_x\sigma_y + 2\sigma_x\gamma\sigma_s - 2\sigma_y\gamma\sigma_s} \quad \text{(A4)}$$

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 (3) 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.

TableA1.1
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: $\gamma > (\sigma_y - \sigma_x)/\sigma_s + (\delta - \mu_s)(\sigma_v/v)$

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 A1.1

Figure A1.1

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 A1.1, 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 is inequality 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.

APPENDIX 2
Data Used in Estimation: Sources and Methods

A: The INS Gross Immigration Data

The data for the number of immigrants to the United States by country is taken from the US Immigration and Naturalization Service (INS) *Statistical Yearbooks.* The data covers all legal immigration, including refugees, and it includes both those who applied from abroad and those who are already in the US 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 only available by place of birth.

Before 1976, the Immigration and Naturalization Service (INS) defined a fiscal year as July 1 through June 30. For example, FY1974 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. Under the new definition of a fiscal year, FY1981 began on October 1, 1980, and ended on September 30, 1981. Because this change occurred during the time series with which we are working, the original data collected from the INS *Annual Reports* and *Statistical Yearbooks* have now been adjusted. The pre-1976 annual observations now conform to the 1976 definition of a fiscal year, one which begins in October 1 and ends in September 30.

The INS does not report monthly totals of immigrants admitted by country of birth, so some assumptions were invoked to make the adjustment. To do so, we used data that the INS labeled as "Immigrants Admitted by Region and Country of Birth" for the Third Quarter (July 1 – 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" FY1976 totals for each country. These sums represent the total immigration from each country to the United States during the 15-month period from July 1, 1975 to September 30, 1976. To estimate the immigration for the twelve months of the new "September" FY1976, we multiplied the 15-month totals by 0.8. This operation gives four-fifths of the 15-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 year. Thus, to convert the "June" FY1975 into a "September" fiscal year, we added one fifth of the 15-month totals that we

used to adjust FY1976 to the "June" FY1975 figures. We then took four-fifths of these sums as the data for the new "September" FY1975. This process was carried back to FY1960, the first year in the data set. Thus, all of the annual gross immigration figures reported in this adjusted INS database now represent October to September totals.

B: Annual US 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 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 only available 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 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 though 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 1980-90 value of d to extrapolate forward to 1998.

C: 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 15 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.

D: Immigration Policy Variables

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

Non-immediate relatives: (X_r)

Non-immediate relatives enter under the following preference categories in the post 1990 legislation (with total numbers in parentheses): (1) adult married children of US citizens (23,400); (2) spouses and unmarried children of US residents, 75 percent of whom must be minors (114,200); (3) married children of US citizens (23,400); and (4) siblings of adult US citizens (65,000). Before 1992 the preference categories were broadly similar (with percentages of total quota in parentheses): (1) unmarried children of US citizens (20%); (2) spouses and unmarried children of resident aliens (20%); (3) married children of US citizens (10%); and (4) siblings of US citizens (24%).

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

Eastern Hemisphere	1966-78: 170,000	World	1979-81: 214,600
Western Hemisphere	1966-76: 120,000	"	1981-91: 210,000
" "	1977-78: 88,800	"	1992-94: 281,000
		"	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 1977-8, 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 non-immediate relatives.

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

Employment visas: (X_e)

From 1992 the employment-related visas are given under the following categories (with total numbers in parentheses): (1) individuals of outstanding ability (40,000); (2) professionals with advanced degrees or with exceptional abilities (40, 000); (3) skilled workers or unskilled shortage workers (40,000); and (5) special occupations including religious workers (10,000); (6) investors (10,000). Before 1992 there were just two employment categories (with percentage of quota in parentheses): (1) exceptional professional, scientists and artists (10%); (2) skilled and unskilled workers in shortage

occupations (10%).

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

Eastern Hemisphere	1966-78:	34,000	World	1979:	58,000
Western Hemisphere	1966-76:	0	"	1980:	56,000
Western Hemisphere	1977-78:	24,000	"	1981-91:	54,000
			"	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 introduced to apply to countries that were under-represented in US 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 1992-4, 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 percent of the worldwide total.

The variable X_d is defined only for 1992-8 and only for those countries eligible to participate; otherwise it takes the value of zero. For 1992-4 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 1995-8 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	1966-78:	10,200			
Western Hemisphere	1966-76:	0			
" "	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 US continuously since before 1982. After first applying for temporary status (during a window in 1987-8) these immigrants could

gain permanent status after 18 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 1989-91.

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. 380-1. Estimates for 1980 are appropriate given that legalizations applied to those living in the US 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 1989-91.

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 US 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 estimates 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 1995-8.

E: The Balanced Panel

In our econometric work and in the simulations, 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 re-aggregated to the combined total throughout. Thus for immigration and the foreign-born stocks, Czechoslovakia, Yugoslavia, and the Soviet Union have been re-assembled. 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 A1 lists all the countries in the dataset by region. As panel B shows, these account for 82.5 percent 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 and the Middle East; Ethiopia, Somalia and Nigeria in Africa; Cuba and Haiti in the Caribbean.

Table A1

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)
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 1971-98, by period

Period	Immigrants in Sample	Total Immigration	Percent in Dataset
1971-80	3,656,107	4,389,630	83.3
1981-90	5,913,094	7,337,806	80.6
1991-98	6,374,841	7,597,762	83.9
1971-98	15,944,042	19,325,630	82.5

C: Numbers in the Balanced Panel and Total Immigration by Region 1990-8

Region	Immigrants in Dataset	Total Immigration	Percent in Dataset
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

Table 1

Source Area Composition of US 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 (000's)	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, excluding Finland but including Norway and Switzerland. Eastern Europe includes the category "Other Europe".

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 US 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
Western	12.1	27.5	46.4	2.3	11.2
Eastern	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 (000's)	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, excluding Finland but including Iceland, Norway, and Switzerland.

Table 3

Gross Immigration Rate Regressions
(81 countries, 28 years; dependent variable: log immigrants admitted/source country population)

	(1)	(2)	(3)
Constant	−12.11	−11.48	−11.77
	(30.9)	(31.1)	(31.2)
GDP per capita ratio (foreign/US)	−2.23	−1.71	−1.64
	(13.3)	(10.7)	(10.2)
Schooling years ratio (popn. >14) (foreign/US)	3.31	2.68	2.42
	(19.0)	(16.1)	(13.9)
Share of population aged 15-29 (foreign)	5.22	3.31	3.70
	(5.0)	(3.4)	(3.7)
Inequality ratio (foreign/US)	4.74	2.72	3.01
	(7.9)	(4.7)	(5.1)
Inequality ratio (foreign/US) Squared	−2.03	−1.05	−1.15
	(7.9)	(4.3)	(4.6)
Distance from US	−0.24	−0.18	−0.18
	(17.2)	(13.4)	(13.2)
Landlocked	−0.35	−0.36	−0.34
	(5.6)	(6.0)	(5.8)
English speaking origin	1.49	1.01	0.97
	(25.1)	(17.1)	(15.2)
Immigrant stock(t-1)/foreign population		46.65	44.04
		(17.1)	(15.9)
(Immigrant stock(t-1)/foreign population)^2		−196.95	−190.24
		(14.3)	(13.8)
Quota Xr × immigrant stock			44.42
			(3.4)
Quota Xe × schooling years ratio			14.80
			(4.1)
Diversity quota			0.15
			(1.6)
Refugee quota × civil war			1.14
			(1.8)
IRCA legalization			0.05
			(3.1)
Processing backlog			−0.06
			(0.09)
Adj R^2	0.73	0.76	0.77
Hetero ($\chi^2_{(1)}$)	6.11	1.31	2.46
No. of observations	2268	2268	2268

Table 4
Effects of Variables on the Regional Composition of US Immigration, 1971-98

	Western Europe	Eastern Europe	Middle 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) Per capita income	5.2	–0.9	–0.1	–4.3	–0.8	0.5	3.1	1.4	–1.7	–1.3	–1.0
(2) Education years	–2.4	–3.2	1.1	–0.9	2.1	–0.4	–1.3	–1.6	3.4	2.7	0.4
(3) Income+education	1.3	–3.4	1.2	–4.9	1.0	–0.2	–0.2	1.9	2.0	1.3	–0.1
(4) Inequality	1.2	1.5	–0.1	1.5	0.1	0.0	0.3	–2.5	–0.9	–0.6	–0.5
(5) Share aged 15-29	1.5	1.1	0.1	–0.8	0.0	0.0	0.2	–1.4	–0.4	–0.1	–0.2
(6) Inc+edu+ineq+age	4.6	–2.1	1.1	–3.3	1.2	–0.1	0.3	–1.7	0.5	0.5	–0.8
(7) Landlocked	0.0	0.1	0.0	–0.1	0.1	0.0	0.0	–0.1	0.0	0.0	0.1
(8) English-speaking	–0.3	1.4	0.9	–4.9	–0.6	–0.3	–1.0	4.9	–1.9	1.2	0.6
(9) Distance	–1.9	0.1	1.2	27.0	1.8	0.9	–1.3	–14.2	–8.0	–3.3	–2.3
(10) Land+eng+dist	–1.3	2.4	3.1	20.0	0.9	0.2	–1.5	–11.5	–8.3	–2.6	–1.4
(11) Immigrant stock	0.2	3.6	1.9	15.9	2.6	0.3	–0.9	–13.8	–9.2	2.4	1.7

Table 5
The Effects of Immigration Policy
(actual/no-policy-change counterfactual)

	Merging Hemispheres					
	1976	1977	1978	1979	1980	1981
Eastern Hemisphere	100.0	99.8	99.7	103.7	103.1	102.6
Western Hemisphere	100.0	141.2	137.2	73.6	64.1	53.5
World	100.0	116.0	113.2	88.5	83.5	74.4
	Immigration Control and Reform Act					
	1988	1989	1990	1991	1992	1993
Eastern Hemisphere	100.0	101.6	101.7	101.4	100.1	100.1
Western Hemisphere	100.0	170.1	180.7	205.2	120.1	115.5
World	100.0	134.7	144.7	157.8	108.3	105.7
	1990 Immigration Act					
	1991	1992	1993	1994	1995	1996
Eastern Hemisphere	100.0	117.3	117.1	117.3	116.1	115.5
Western Hemisphere	100.0	117.7	117.4	117.6	113.6	113.6
World	100.0	117.5	117.2	117.4	115.1	114.7