NBER WORKING PAPER SERIES

WHERE DO U.S. IMMIGRANTS COME FROM, AND WHY?

Ximena Clark
Timothy J. Hatton
Jeffrey G. Williamson

Working Paper 8998
http://www.nber.org/papers/w8998

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2002

Paper presented at the Population Association of America Meetings, Atlanta, Georgia (May 9-11, 2002). We are grateful for the helpful comments of Andy Mason and Sam Thompson, as well as for the excellent research assistance of Rachael Wagner. Williamson acknowledges with pleasure the financial support of the National Science Foundation SES-0001362. The views expressed herein are those of the authors and not necessarily those of the World Bank. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Where Do U.S. Immigrants Come From, and Why?
Ximena Clark, Timothy J. Hatton and Jeffrey G. Williamson
NBER Working Paper No. 8998
June 2002
JEL No. F22, J1, O15

ABSTRACT

The United States has experienced rising immigration levels and changing source since the 1950s. The changes in source have been attributed to the 1965 Amendments to the Immigration Act that abolished country-quotas and replaced them with a system that emphasized family reunification. Some believed that the Amendments would not change the "traditional" sources of US immigrants. Given this view, it seems all the more remarkable that the sources of immigration changed so dramatically. This paper isolates the economic and demographic fundamentals that determined immigration rates by source from 1971 to 1998 -- income, education, demographic composition and inequality. The paper also allows for persistence - big US foreign-born stocks implying a strong 'friends and neighbors' pull on current immigrant flows. Specific policy variables are included which are derived directly from the quotas allocated to different visa categories. Parameter estimates from the panel data are then used to implement counterfactual simulations that serve to isolate the effects of immigration policy as well as source-country economic and demographic conditions.

Ximena Clark
The World Bank
1818 H Street, N.W.
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, MA 02138
and NBER
jwilliam@harvard.edu

Introduction

The United States has experienced rising immigration since the 1950s and this has been accompanied by a growing debate about its economy-wide impact. Observers have stressed the decline in the human capital content of recently arrived immigrants and have associated this with shifts in source country composition.[1] These have been dramatic. Between the 1950s and the 1990s, the proportion of immigrants arriving from Europe fell from over half to just 15 percent, while those from Asia rose from 6 percent to 30 percent, and those from Mexico alone rose from 12 percent to 25 percent. A decline in the skills and schooling of successive immigrant cohorts relative to native-born Americans has been closely associated with the changing immigrant origin. Given this fact, it is surprising that there has been so little analysis of *why* immigrant source has changed so dramatically in such a short period of time. So, where have US immigrants come from, and why?

These changes in source have taken place mostly since the 1960s and they have been attributed to the 1965 Amendments to the Immigration Act. Before 1965, quotas were set for the maximum number of immigrants from a given country. Since quota allocations were based on the 1921 immigrant stock, they strongly favored Europeans, especially those from Western Europe. Immigrants from the Western Hemisphere were restricted under a separate quota. The 1965 Amendments abolished quotas and replaced them with a new non-discriminatory system that strongly emphasized family reunification as a criterion for admission. It was believed by some that this would still be consistent with the result that the bulk of immigrants would come from the "traditional" sources. Given this policy view, it seems all the more remarkable that the sources of immigration changed so dramatically. The interesting question is: why?

[1] These issues are surveyed in Borjas (1994). Studies that debate the long-term decline in immigrant 'quality', and its apparent reversal sometime in the 1980s include Borjas (1995), Barrett (1996) and Jasso, Rosenzweig and Smith (2000)

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. The paper also allows for persistence in these patterns as they arise from the impact of the existing immigrant stock – big foreign-born stocks implying strong 'friends and neighbors' effects. Specific policy variables are included which are derived directly from the quotas allocated to different visa categories. Finally, the paper examines far more countries over a longer period than has been true of previous work on late 20th century US immigration.

After outlining the course of US policy, we set out a model of immigration, which is then estimated on a panel of 81 countries for the years 1971 to 1998. Economic and demographic variables, the immigrant stock, and a series of policy-related variables all emerge as significant determinants of migration rates as predicted by the theory. These estimates are then used to conduct counterfactual simulations so as to isolate the effects of immigration policy as well as the role of 'convergence' and 'divergence' in source country economic and demographic variables. The results indicate that variables like relative income and education had substantial effects on the composition of US immigration, while persistence wore off quickly

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 quarter of a million per year in the 1950s to nearly half a million in the 1970s and reaching close to a million in the early 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. Europeans had been an even bigger share in earlier decades: 62 percent in the 1920s and 92 percent in the 1900s. Canada and other so-called New World countries shared this long-term decline in relative and absolute numbers.

The counterpart to Europe's decline as an immigrant source has been the rise in less developed parts of the world. All less developed regions increased their US immigration shares between the 1950s and 1970s, but the trends since have varied considerably. While the Caribbean share reached its peak in the 1970s, and while the Asian share peaked in the 1980s, the shares from Mexico, Africa and Eastern Europe continued to increase. The recent rise in immigrants from Eastern Europe clearly reflects non-economic and non-demographic changes as that region undergoes political transition. Furthermore, even though the share of some regions has stabilized, the absolute numbers from those areas has increased.

These differences across countries and regions suggest that economic and demographic forces have been important. The golden age of Western European economic growth coupled with an ageing population has been associated with the collapse of European immigration. Declining trends from the 1970s in parts of Asia -- first in East Asia and subsequently in South East Asia -- also seem to point to economic and demographic fundamentals at work. Similarly, demographic pressure and lagging economic growth in Africa is associated with an upward trend, although from a low initial level.

The level and the composition of immigration have both been mediated, of course, by policy.[2] Country of origin quotas were first introduced in an emergency Act of 1921, made permanent by the 1924 Immigration Act (effective 1929),[3] and further modified by the Immigration and Nationality Act of 1952. Under the 1952 Act, children and spouses of US

[2] Useful summaries of US immigration legislation can be found in De Laet (2000, Appendix A), and Jasso, Rosenzweig and Smith (2000).

[3] The 1924 Act set quotas based on the proportion of each nationality in the 1920, subject to an overall Eastern Hemisphere ceiling of 150,000. Wives and minor children of US citizens were exempt from the quota. Western Hemisphere natives were not subject to the quota, but the 'Asiatic Barred Zone' introduced by the 1917 Immigration Act was retained.

citizens were exempt from the quota. Within the quota limit, visas were allocated according to a system of preferences that gave up to 50 percent to those with special skills, up to 30 percent to parents of adult US citizens, and up to 20 percent to spouses and children of legal aliens. Twenty-five percent of unused visas under the previous categories were given to siblings and married children of US citizens and the remainder to immigrants outside the preference categories. Admissions of Western Hemisphere immigrants were not subject to a quota, but small quotas were allotted to countries in the Asia-Pacific Triangle. Under this system, the bulk of the visas subject to quota were allocated to European countries, and among these two thirds went to just two countries: Germany and the UK.

The 1965 Amendments to the Immigration and Nationality Act contained a radical shift in previous policy. The 1965 legislation 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. As before, children and spouses of US citizens were exempt from the quota. In addition, a ceiling of 120,000 visas was set for the Western Hemisphere, but without country quotas or a preference system.[4]

This new system strongly favored family reunification over employment-based immigration. Indeed, family reunification had been the main entry mechanism for Eastern Hemisphere immigrants even before 1965. Some lobby groups and their congressional sympathisers believed that while the new policy could be seen as non-discriminatory, the composition of the existing immigrant stock would nevertheless ensure that immigrants would largely come from the traditional European sources (Briggs, 1984, p. 69; Daniels and Otis, 2001, p. 43-4). On the other hand, the abolition of the national origins system did represent an increase in the opportunities for immigration from non-European countries.

[4] Further details of numbers allocated to different preference categories are given in Appendix 2D.

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.[5] Second, the 1990 Act increased the number of employment based visas to 140,000 (from 54,000 previously), under a five-part preference system.[6] Third, 55,000 visas were allocated on top of the overall quota for "diversity" immigrants -- those from countries with relatively low immigration since 1965.[7]

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 of visas were issued under employment based preference

[5] 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.

[6] The quotas for different preferences in the employment-based category are detailed in Appendix 2D.

[7] In the transitional period between 1992 and 1994, the overall quota was raised to 700,000 with

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 reason for the establishment of the diversity category). It is also small for Africa, a source country for whom American mass immigration has only just begun (Hatton and Williamson 2001). 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 route is represented 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. They 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

465,000 visas reserved for close family immigration, but the diversity program was limited to 40,000.

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), as well as by Hatton and Williamson (1998) for the European Mass migrations before the 1920s. Here we set out a heuristic framework which follows in that 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. Third there is the cost to migration associated with quantitative restrictions on immigration: the greater is the total quota, q, 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 γ, 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, m_i, is:

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

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 $-v$, the numerator in equation (2). Even for a given value of v, migration will be a non-monotonic function of the relative return on skills in the source and the destination. These effects are examined further in Appendix 1.

Immigration policy will also influence the volume of migration through 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.[8]

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, 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

[8] 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.

rate was negatively related to inequality in the origin country. 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 to be the single most important determinant.

More recently, David Kamemera, 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 negatively related to distance from the United States and 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 modelled as a dummy variable only.

In order to study the effects of policy change, Guillerma Jasso and 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, as the result of substitution across visa categories. In addition to income and education, policy dummies were found to matter. In particular, application of the preference system raised the numbers arriving as male spouses from the Western Hemisphere, 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 population and measures of human capital and/or the return to skills. Their omission makes it impossible to assess the role of sending country demographic and human capital attributes, variables that theory suggests should matter. Foremost among these might well be the sending country's position in the demographic transition. 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. We believe this is a mistake since only by doing so can we isolate the role of persistence in the immigration flows. Finally, shifts in immigration policy are typically reflected by time 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 in the following specification:

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

The left-hand side variable is migration to the US from country j in year t as proportion of the source 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 will be non-monotonic. When the source country has a relatively unequal income distribution, an increase in its relative inequality will reduce the migration rate. When the source country has a relatively equal distribution, an increase in its relative inequality will increase the migration rate (see Appendix 1). Thus the effect of relative inequality on migration will be an inverse 'U' shape; 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. Since relatives (and friends) abroad reduce migration costs, $\beta_9 > 0$.

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_{10} through β_{13} 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 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 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 2001).

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_{14} is therefore expected to be positive. Finally, X_b is a dummy for the years 1995-8 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; $\beta_{15} < 0$. 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 1998 (see Appendix 2A and E). These countries form 82.5 percent of all US immigration over the entire period. For relative income we use purchasing power parity adjusted GDP per head, from the Penn World Tables9 while 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 benchmark figures from 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,

9 Later versions of this paper hope to use the ILO ppp-adjusted and occupation-specific wage data base currently being made ready for public use (Freeman and Oostendorp 2000).

in the second column, the (lagged) immigrant stock is added the coefficients of the other variables are somewhat attenuated, as might have been expected, but not by much. The full model appears in the third column and, while the coefficients of the other variables are little altered, most of the policy-related variables also enter strongly and with the expected signs. The only exceptions are the variables representing refugees and the processing backlog, which although taking the expected signs, are not significant. Additional variables such as an index of source country civil rights or the US unemployment rate failed to provide significant coefficients and so these 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 an increase of 10 percent in a country's income per capita (e.g. five years of catching up growth where per capita GDP grows 2% faster than in the US) reduces immigration to the US by around 6 percent. More dramatically, moving from an income level typical of Western Europe to one typical of South America would raise a country's immigration rate by 82 percent. Raising a country's years of schooling by 10 percent (equivalent to 0.55 years for the average source country) would increase the immigration rate by 15 percent. More significantly, moving from an education level typical of Western Europe to one typical of South America would reduce the immigration rate from a country by about 60 percent.

Raising the share aged 15-29 by ten per thousand of a source country's population increases immigration to the US by 4.5 percent or by 0.3 per thousand of the source country population. Thus demographic effects are quite significant. Inequality effects are more complex because the variable enters non-linearly. The quadratic peaks at a ratio of the foreign/US gini coefficient of 1.25. A peak greater than one would be predicted by the Roy model in the presence of selective immigration policy (see Appendix 1). Thus moving from an inequality ratio typical of South America to one typical of Western Europe (from 1.20 to 0.82) reduces a country's immigration rate by 34 percent--a sizeable effect. This is because,

for a given mean income, the lower is inequality in the source country (and therefore the lower is the return to skills) the less likely the low-skilled will have an incentive to migrate.

The variables reflecting fixed country characteristics are very powerful. The effect of distance is to reduce a country's migration rate to the US by about 21 percent for every additional thousand miles from the United States. The effect of being landlocked reduces a country's immigration rate to the US by 32 percent while the effect of being a predominantly English speaking country raises it by a massive 120 percent. While these fixed characteristics will always have an influence on the composition of US immigration, they can not have played a role in accounting for *changes* in that composition over time.

The coefficient on the migrant stock is of particular interest because it reflects the non-policy component of the 'friends and relatives effect.' This direct effect induces about 6 additional immigrants per year for every thousand of the foreign-born immigrant stock. To this direct effect, must be added an indirect effect: there is the additional effect working through the interaction with the policy variable Xr representing the quota on non-immediate relatives. This indirect effect adds, on average, a further 1.8 immigrants per year for every thousand of the foreign-born immigrant stock. These combined direct and indirect effects strike us as surprisingly modest,[10] although they do cumulate over time. In any case, ignoring deaths and return migration, these combined effects increase the immigrant stock of the typical country by 1.1 percent per year or about 12 percent per decade.

The policy-related effects are also important in the regression, but these are best treated by means of simulations in the next section.

The Impact of Immigration Policy

[10] Modest especially given that the stock effects do not fully account for immigrants entering on visas either under family preferences or not subject to world-wide numerical limits. In 1998 immigrants in those visa categories amounted to about 20 per thousand of the total foreign-born stock.

The impact of immigration policy is assessed by means of counterfactual simulations relative to a baseline simulation. 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 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 of immigration in the following year and so on. The effects of changing policy can be assessed by comparing the counterfactual level of immigration generated this way with a baseline simulation (where the explanatory variables take their actual values).[11]

The first case is the period in the late 1970s when the separate quotas for the Eastern and Western Hemispheres were merged in to a worldwide quota. This affected the total number of visas for both non-immediate family members and employment-based immigration. And it had differential effects on Eastern and Western Hemisphere countries. 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 employment visas was raised from zero to 24,000 in 1977, and then in 1979, this and the Eastern Hemisphere quota (of 34,000) were merged, with reductions in the total occurring 1980 and 1981.

In the counterfactual simulation the quotas are held constant at the 1976 levels from 1977 onwards, retaining the distinction between the 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 baseline simulation to the counterfactual simulation, and hence they reflect the effect of policy change in proportionate terms. In the years 1977-8 the effect of the

[11] The baseline simulation exactly replicates the data because (a) the immigration equation includes the equation errors, and (b) the equation used to update the immigrant stock uses the same "depreciation" parameter that was generated for each year when calculating the immigrant stock (see further below and Appendix 2B).

increase in employment visas massively outweighs that of the decline in family-based visas for the Western Hemisphere. The subsequent decline 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 policy 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 4. The IRCA effects are especially marked for Western Hemisphere countries and only marginal for the Eastern Hemisphere. These 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 4 simulates the effects of the Immigration Act of 1990, which took effect from 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 an about a 75 percent increase in the number of available visas between 1991 and 1992-4. Our estimated effects of these changes, taken together, are very much smaller than that. But they are broadly consistent with the trend in the INS statistics for total non-IRCA immigrants. Between 1991 and 1992-4 that total rose by 18 percent, a figure that is just a little under our estimate of around 21 percent for 1992-4.

One reason that the increase in predicted (and actual) immigration was less than in proportion to the increase in the overall quota is that some previously exempt categories were

absorbed into the quota for the first time. Specifically, these were immediate relatives and certain categories of employment-based immigrants. A second possible reason is that, in some years, the employment and diversity categories under-fulfilled their quotas.

Sources of Changing Country Composition

As we have seen, one of the major features in the evolution of US immigration in the last thirty or forty years has been the change in the source-country composition. That immigrant composition has also altered the composition of the stock of foreign-born. The 1965 Amendments to the Immigration Act are often seen as a critical vehicle of this change to the extent that it opened the door to immigrants from poor parts of the world where the incentive to migrate to the US is much greater than it is for Europeans. We cannot test the direct effects of the 1965 amendments since they fall outside our sample period. But we can pose some counterfactual questions that should shed considerable light on the issue. These counterfactuals ask: what if country economic and demographic characteristics had been the same across all sources?. As before, the counterfactuals are assessed by means of simulations, starting in 1971.[12]

Counterfactual regional compositions of immigration for the whole period 1971-98 and for the immigrant stock in 1997 are listed in Table 5. These counterfactuals examine the effects of 'convergence,' a term that has a precise meaning here: a given variable is set at the mean for each country for each year. This counterfactual deals only with the demographic and economic variation between source countries and makes no change in the mean value of the explanatory variables in any year. The total volume of immigration is thus kept approximately constant, keeping the counterfactual in line with the overall immigration policy constraint

[12] The simulation is complicated by the presence of persistence and the friends and neighbors effect. That is, for each country, a change in one of the explanatory variables must be allowed to feed into next year's foreign-born stock, which then influences the following year's immigration.

The baseline simulation is the prediction (appropriately adjusted) when the explanatory variables take their actual values. The second column of the table shows the effects of income convergence, and they are substantial. Western Europe had, of course, far higher income than the average sending region. Thus, had Western European incomes in each year been the same as the sending country average, the share of Western European-born in total immigration 1971-1998 would have been 13.5 percent rather than 7.3 percent, equivalent to an additional 1.2 million European migrants. By contrast, the numbers from Eastern Europe fall as their per-capita incomes are lifted to the average. These differences are closely reflected in the stock of foreign-born that emerges as the cumulative effect of the altered immigration flows. As a result, the share of Europeans would rise to 22.7 percent in 1997. But that would still be substantially below their 70 percent share in 1970 (and even below the 1980 share).

Other relatively rich areas like Canada and Oceania also would have undergone increased shares as their income gaps in the counterfactual rose relative to other sending regions: they would have had dramatically increased immigration shares as their incomes were forced to the world average. In contrast, the shares from Africa and Central America fall as their incomes rise in the counterfactual. Mexico's share increases slightly while that of East Asia falls and thus these two sources would still account for more than half of all immigration. Because these are middle income regions the proportionate change in immigration is modest. Nevertheless in absolute terms the counterfactual implies about a million fewer east Asians and 200,000 more Mexican immigrants over the whole period from 1971 to 1998.

The third column of Table 5 shows the effects of assuming that each country had the world average years of education and the effects are even more dramatic.[13] European migrants fall from 13.4 to just 5.9 percent of all immigrants and their share of the foreign-

[13] The counterfactual effects of education work through both the uninteracted term in the regression

born stock declines from 19.4 to 12.9 percent in 1997. Immigration from East Asia and Mexico also fall in the counterfactual since both undergo a fall in education to the lower sending-country average. The most dramatic increases come from Africa, the Caribbean, South America, and especially from Central America, where a counterfactual rise in education to the average generates a rise in immigration.

The fourth column shows what happens if each country in the counterfactual were forced to take on the sending country average proportion of population aged 15-29.[14] These effects are more modest than those of income and education, partly because there is less demographic variance across regions compared with per capita income and education. Still, larger young adult cohorts would have boosted the European share in US immigration from 13.4 to 16.8 percent. And smaller cohorts of young Mexicans would have reduced their share by 2 percent.

The last variable is inequality, and, surprisingly, its impact is as big as demographic influences. The increase in inequality implied for Europe raises immigration to the US because, as previously noted, European income distributions are relatively equal and they are therefore on the upward-sloping part of the inverted 'U' in the relationship between immigration and inequality. By contrast South and Central American immigration falls slightly as their high inequality is reduced, shifting them from the right-hand side of the peak to the left-hand side.

Overall these counterfactuals reinforce the point that the changing composition of immigration over the last three decades has been driven by a combination of economic, demographic and policy forces. In Europe, relatively high incomes, small youth cohorts and relatively equal income distributions have restrained immigration to the US. The effects of high relative education have worked in the opposite direction. In South and Central America the reverse has generally occurred.

---

equation and through the interaction with the employment quota.

[14] This clearly has implications for the relative rates of population growth that are not considered here.

What about the persistence effects of the immigrant stock? What was the legacy of the national origins system? Europeans formed 70 percent of the foreign-born stock in 1970; that figure had fallen to 22 percent by 1990. Had the 1965 Amendments to the Immigration Act been enacted earlier the European share would clearly have been lower in 1970. In turn, that would have led to even greater flows of immigration from non-European sources in subsequent years. Table 6 shows the result of a simulation, again starting in 1971, where the stock of Europeans in 1970 is cut by half and that of all other source countries is doubled. As compared to the baseline simulation in Table 5, regional shares of the total immigration flow over the three decades had changed—but not by very much. Thus the friends and neighbors effect appears to have played a minor role in influencing immigrant composition across the decades.

One reason why these persistence effects seem relatively weak is the difference in the rate at which the foreign-born stock "depreciates" across sending countries. The immigrant stock for each country is constructed as $S_{t+1} = M_t + dS_t$, where S_t is the stock at the beginning of year t and M_t is the flow during that year. The parameter d (see Appendix 2B), reflects the balance of deaths, return migration and illegal immigration. This is much lower for Europe where the average value is 0.97, than it is for Central America and Mexico where the average value exceeds one. The difference is explained partly by the fact that European-born populations are older and partly because there is more return migration among Europeans. Most importantly, however, the difference reflects a much higher incidence of illegal immigration from countries south of the United States border.

The effects of applying the mean depreciation rate, d, for each year to all countries is shown in the third and fourth columns of Table 6. The most dramatic effects are on Mexico where the share of immigration and the 1997 immigrant stock fall dramatically as a result of the much higher return migration rate (or less illegal immigration) implied by the counterfactual. The immigration share of Central America also declines but that of Europe

alters very little. Nevertheless, the European share of the immigrant stock increases dramatically to more than a quarter in 1998. However, this counterfactual should be treated with caution since patterns of return migration and patterns of illegal immigration are clearly responses to economic incentives. Thus the depreciation parameter itself reflects underlying economic and demographic forces.

Conclusion

This paper offers strong support for a model of US immigration that stresses both individual incentives and policy constraints. Income, education, and demography all matter in the manner predicted by 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 are almost always omitted in empirical work: the stock of previous immigrants and variables representing different dimensions of immigration policy.

What conclusions emerge from the simulations that were performed using the coefficients from the estimated equation? The first is that between 1971 and 1998 immigration policy had powerful effects on the volume and composition of US immigration. This conclusion is hardly surprising, but it is reassuring confirmation. The second is that experiments with economic and demographic fundamentals suggest that all of these worked in the expected direction. Taken together, the evolution of those fundamentals has had a profound impact on the source country composition of immigration filtered as it is through policies and quotas.[15] Furthermore, this paper has found that the effects of the migrant stock, though highly significant, are too small to have imposed a high degree of persistence in immigration patterns across the decades. Observers in the 1960s, who thought that a policy

[15] In future research, we hope to identify the extent to which US immigration policy has been determined by expected migrations and perceived impact, but the political economy of immigration policy has already been explored for the pre-1930 decades (Goldin 1994; Timmer and Williamson 1998).

emphasis on family reunification would serve to replicate the structure of immigration, have been proved wrong. We can see why: the persistence effects working through the foreign-born stocks simply have not been large enough to matter all that much.

Strong US policy effects and powerful economic and demographic influences in sending regions – as well as weak persistence -- are only part of the story reported here. These forces have changed in the past and will change in the future. But *fixed* effects were also very influential in determining the composition of US immigration, such as distance and proximity — forces that have been manifested by illegal immigration pressure from south of the border. Thus, it seems likely that immigration from Central and South America would have been sizeable under almost any plausible set economic and demographic trends in those countries.

The national origins systems introduced in the 1920s was important in determining subsequent sources of US immigration. While it may seem surprising that the origins system lasted until as late as 1965, perhaps the explanation is that the underlying economic and demographic fundamentals had changed very little in sending regions until the 1960s.

References

Barrett, A. (1996) “Did the Decline Continue? Comparing the labor Market Quality of United Sates Immigrants from the late 1970s to the late 1980s,” *Journal of Population Economics*, 9, pp. 57-63.

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

Borjas, George J. (1994), “the Economics of Immigration,” *Journal of Economic Literature*, 32, pp. 531-553.

Borjas, George J.(1995), “Assimilation and Changes in Cohort Quality Revisited. What Happened to Immigrant Earnings in the 1980s?” *Journal of Labor Economics*, 13, pp. 210-245.

Briggs, Vernon M. (1984), *Immigration Policy and the American Labor Market*, Baltimore Md: Johns Hopkins University 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).

Daniels, Roger and Graham, Otis L. (2001), *Debating American Immigration 1882-Present*, Lanham Md: Rowman and Littlefield.

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

Goldin, Claudia. (1994), "The Political Economy of Immigration Restriction in the United States, 1890 to 1921," in C. Goldin and G. D. Libecap, eds., *The Regulated Economy: A Historical Approach to Political Economy* (Chicago: University of Chicago Press).

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

Hatton, Timothy J. and Williamson, Jeffrey G. (2001), “Demographic and Economic Pressure on Emigration Out of Africa,” *NBER Working Paper 8124*, National Bureau of Economic Research, Cambridge, Mass. ((February).

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

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

Timmer, Ashley and Williamson, Jeffrey G. (1998), “Immigration Policy Prior to the Thirties: Labor Markets, Policy Interaction, and Globalization Backlash, *”Population and Development Review*, 24 (December), pp. 739-771.

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

Yang, Philip Q. (1995), *Post-1965 Immigration oi 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 $w_x \sim N(\mu_x, \sigma_x^2)$.

(A1)

Income in origin: $w_{yi} = \alpha_y + \beta_y s_i$; 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 > 0), where $v = w_{xi} - w_{yi} - z_i - c_1 + c_2(q) - \gamma(\delta - s_i)$ 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.

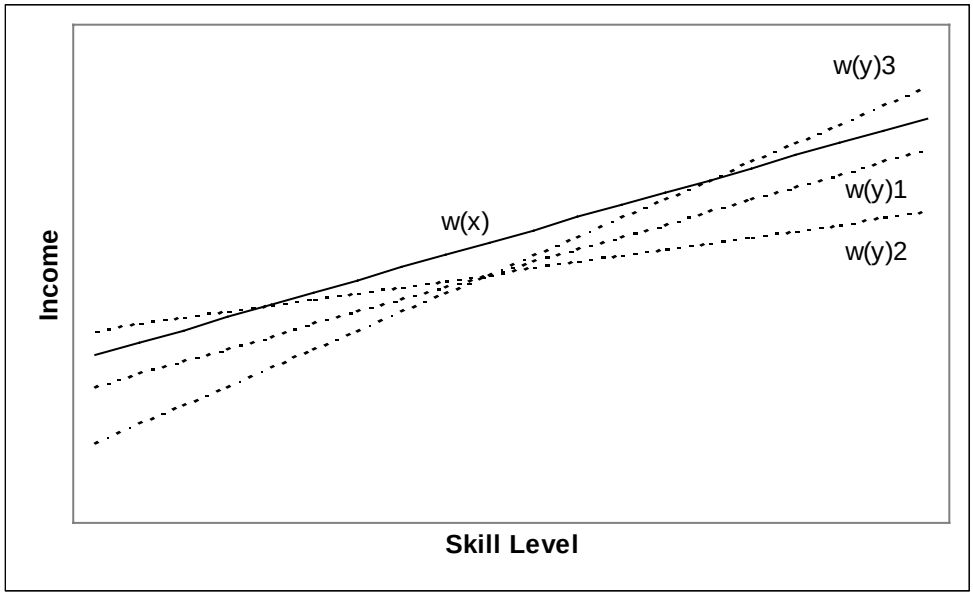
TableA1.1
Effects of Income Distribution and Immigration Policy on Migration

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 selective immigration, 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.

Data estimating years after 1990 (by sampling) come from the Census Bureau's Current Population Survey online data extraction service at

http://ferret.bls.census.gov/cgi-bin/ferret.

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. 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 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. Thus, for example, 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

[to be added]

D: Immigration Policy Variables

Immigration policy is characterized in equation (4) 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), p. 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 that follow, we use a balance 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 ommited include; Vietnam, Iraq and Lebanon in Asia; Ethipoia, Somalia and Nigeria in Africa; Cuba and Haiti in the Caribbbean.

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)
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 1971-98, 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, 1950-98
(percent of total from each source)

Region of origin	1951-60	1961-70	1971-80	1981-90	1991-8
Europe	52.7	33.8	17.8	10.3	14.9
West	47.1	30.2	14.5	7.2	4.9
East	5.6	3.6	3.3	3.1	10.0
Asia	6.1	12.9	35.3	37.3	30.8
Americas	39.6	51.7	44.1	49.3	49.7
Canada	15.0	12.4	3.8	2.1	2.1
Mexico	11.9	13.7	14.2	22.6	25.4
Caribbean	4.9	14.2	16.5	11.9	10.8
Central America	1.8	3.1	3.0	6.4	5.6
South America	3.6	7.8	6.6	6.3	5.8
Africa	0.6	0.9	1.8	2.4	3.7
Oceania	0.5	0.8	0.9	0.6	0.6
Total (000's)	2,515	3,322	4,493	7,338	7,605

Source*: Statistical Yearbook of the Immigration and Naturalization Service for 1998*, 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. East Europe includes the category "Other Europe".

Table 2

Class of Admission by Source Area, 1998
(percent of total for each source region)

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
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 (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	−12.01	−12.41
	(30.9)	(31.5)	(31.7)
GDP per capita ratio	−2.23	−1.95	−1.87
(foreign/US)	(13.3)	(11.6)	(11.1)
Schooling years ratio (popn. >14)	3.31	3.10	2.79
(foreign/US)	(19.0)	(18.0)	(15.5)
Share of population aged 15-29	5.22	4.16	4.54
(foreign)	(5.0)	(4.1)	(4.4)
Inequality ratio	4.74	4.49	4.77
(foreign/US)	(7.9)	(7.6)	(8.01)
Inequality ratio (foreign/US)	−2.03	−1.83	−1.91
squared	(7.9)	(7.3)	(7.6)
Distance from US	−0.24	−0.21	0.21
	(17.2)	(15.5)	(15.3)
Landlocked	−0.35	0.34	−0.32
	(5.6)	(5.4)	(5.2)
English speaking origin	1.49	1.25	1.20
	(25.1)	(19.8)	(18.9)
Immigrant stock(t-1)/foreign population		10.03	8.57
		(9.6)	(7.6)
Quota Xr × immigrant stock			36.60
			(2.6)
Quota Xe × schooling years ratio			17.23
			(4.55)
Diversity quota			0.26
			(2.64)
Refugee quota × civil war			1.00
			(1.5)
IRCA legalisation			0.07
			(4.17)
Processing backlog			−0.06
			(0.09)
Adj R^2	0.73	0.74	0.75
Hetero ($\chi^2_{(1)}$)	6.11	1.71	3.01
No. of observations	2268	2268	2268

Notes: 't' statistics in parentheses. Fixed effects, included but not reported are: Western Europe (excluded group), Eastern Europe, Middle East, East Asia, Africa, Oceania, Caribbean, Central America, South America. For the countries included in each region see Table A1, Appendix 2. In addition there are country dummies for Canada and Mexico.

Table 4

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

	Merging Hemispheres					
	1976	1977	1978	1979	1980	1981
Eastern Hemisphere	100	99.8	99.7	104.3	103.6	102.9
Western Hemisphere	100	155.3	149.6	81.4	75.3	66.2
World	100	120.3	116.6	93.3	90.5	83.8
	Immigration Control and Reform Act					
	1988	1989	1990	1991	1992	1993
Eastern Hemisphere	100	102.3	102.4	101.9	100.1	100.1
Western Hemisphere	100	207.5	215.5	237.6	111.9	110.7
World	100	148.2	159.1	170.8	105.1	104.0
	1990 Immigration Act					
	1991	1992	1993	1994	1995	1996
Eastern Hemisphere	100	122.9	122.5	124.0	118.7	117.9
Western Hemisphere	100	119.4	119.2	119.4	116.6	116.7
World	100	121.3	121.2	122.1	117.9	117.3

Table 5

Counterfactual Immigrant Composition
(percent of total)

	Shares of total gross immigration, 1971-1998				
	Baseline simulation	Income convergence	Education convergence	Demographic convergence	Inequality convergence
West Europe	7.3	13.5	3.8	9.3	9.0
East Europe	6.1	4.9	2.1	7.5	8.3
Middle East	3.9	3.8	4.0	4.1	3.7
East Asia	31.5	26.6	23.7	30.8	33.2
Africa	3.0	2.0	4.1	3.1	3.2
Oceania	0.6	1.2	0.2	0.7	0.6
Canada	1.8	6.3	0.4	2.0	2.3
Mexico	21.8	23.2	16.6	19.8	18.9
Caribbean	12.6	10.5	17.5	12.0	11.0
Cent America	5.2	3.3	19.2	5.1	4.4
South America	6.2	4.6	8.5	5.7	5.5
	Shares of foreign-born stock, 1997				
	Baseline simulation	Income convergence	Education convergence	Demographic convergence	Inequality convergence
West Europe	13.3	17.3	9.5	14.9	14.2
East Europe	6.1	5.4	3.4	7.1	7.3
Middle East	3.0	3.0	3.1	3.1	2.9
East Asia	22.6	19.7	19.1	22.0	24.2
Africa	1.9	1.4	2.7	1.9	2.0
Oceania	0.5	0.9	0.2	0.5	0.5
Canada	3.0	5.7	1.8	3.2	3.2
Mexico	27.0	28.6	23.1	25.8	25.2
Caribbean	10.3	8.6	15.7	9.5	9.1
Cent America	6.8	5.0	15.3	6.8	6.1
South America	5.6	4.6	6.1	5.3	5.3

Table 6
Immigrant Stock Counterfactual
(percent of total)

	Adjusted Initial (1970) values of the immigrant stock		Common rate of stock "depreciation"	
	Immigration 1971-98	Foreign-born stock, 1997	Immigration 1971-98	Foreign-born stock, 1997
West Europe	6.6	8.0	6.7	17.0
East Europe	5.7	4.5	5.5	9.3
Middle East	3.9	3.4	3.5	4.0
East Asia	30.9	24.9	27.8	21.1
Africa	2.9	1.9	2.7	2.2
Oceania	0.6	0.6	0.5	0.6
Canada	2.6	5.1	1.6	3.7
Mexico	21.0	25.5	15.3	13.8
Caribbean	13.5	10.8	25.4	19.6
Cent America	6.2	8.7	3.6	3.1
South America	6.1	6.5	7.3	5.7