Southern Economic Journal 2013, 80(2), 414–438
DOI: 10.4284/0038-4038-2011.377

The Short- and Long-Run Determinants of Less-Educated Immigrant Flows into U.S. States

Nicole B. Simpson* and Chad Sparber†

We use a gravity model of migration and alternative estimation strategies to analyze how income differentials affect the flow of immigrants into U.S. states using annual data from the American Community Survey. We add to existing literature by decomposing income differentials into short- and long-term components and by focusing on newly arrived less-educated immigrants between 2000 and 2009. Our sample is unique in that the vast majority of our observations take zero values. Models that include observations with zero-flow values find that recent male immigrants respond to differences in (short-term) GDP fluctuations between origin countries and U.S. states, and perhaps to (long-term) trend GDP differences as well. More specifically, GDP fluctuations pull less-educated male immigrants into certain U.S. states, whereas GDP trends push less-educated male immigrants out of their countries of origin. Effects for less-educated women are less robust, as GDP coefficients tend to be much smaller than for men.

JEL Classification: J61, E01

1. Introduction

Income is often cited as an important determinant of immigration, and some measure of income in the origin and/or destination country is included in almost every model explaining international migration. Recently, Clark, Hatton, and Williamson (2007), Lewer and Van den Berg (2008), Lewer, Pacheco, and Rossouw (2013), Ortega and Peri (2009), and Mayda (2010) all find evidence that per capita GDP (in the origin and/or destination country) is a significant predictor of cross-country immigrant flows. We add to this literature in three ways: (i) by analyzing recent inflows of less-educated immigrants into U.S. states between 2000 and 2009, (ii) by decomposing GDP into short- and long-run components, and (iii) by employing three distinct estimation methodologies: scaled ordinary least squares (SOLS), Eaton and Tamura (1994) threshold tobit, and two-part models.

First, we analyze the flow of new immigrants into U.S. states between 2000 and 2009 using U.S. Census and American Community Survey (ACS) data. Our work complements literature that focuses on the locational choice of new immigrants based on state-specific factors (e.g.,

* Department of Economics, 13 Oak Drive, Hamilton, NY 13346, USA; E-mail nsimpson@colgate.edu; corresponding author.

† Department of Economics, 13 Oak Drive, Hamilton, NY 13346, USA; E-mail csparber@colgate.edu.

The authors thank Piyush Chandra, Giovanni Peri, Cong S Pham, Todd Sorensen, seminar participants at Colgate University, Union College, and the 2010 Western Economic Association International in Seattle, and two anonymous referees for detailed comments. All errors are our own.

Received December 2011; accepted November 2012.

Bartel 1989; Zavodny 1997; Borjas 1999; Dodson 2001). These articles often explore the demographic characteristics of immigrants as potential determinants of their selected destination in the United States. Instead of analyzing individual decisions, we take a macro approach to estimate how U.S. immigrant flows respond to state-level economic conditions.

Our focus is on the flow of newly arrived male immigrants with a high school degree or less education who legally or illegally arrived to the United States. Our attention to men is driven by past evidence arguing that male migration decisions are more likely to be motivated by economic factors, whereas women more likely migrate for tied or associational reasons (Taylor 2006). The male labor market is especially interesting to study in the wake of the 2007–2009 recession when male unemployment rates were particularly high (Şahin, Song, and Hobijn 2010). We concentrate on flows of immigrants with little educational attainment (which account for 60% of immigrant labor flows during this period) because such individuals exhibit more volatility in employment than both their native counterparts and well-educated immigrants (Orrenius and Zavodny 2009). Interest in the determinants of less-educated immigrant flows is further driven by the group's relatively low level of popular support in the United States. A survey by Hainmueller and Hiscox (2010, p. 67), for example, argues that "although more than 60% of respondents state that they strongly disagree or somewhat disagree with an increase in low-skilled immigration, only 40% of respondents are opposed to an increase in highly skilled immigration."[1] Recent bipartisan immigration reform efforts in the U.S. Senate reflect this sentiment.[2]

Second, variation across countries and U.S. states allows us to consider whether differences in short-run GDP (i.e., fluctuations) and long-run GDP (i.e., trends) have distinct effects on gross immigrant flows. Surprisingly, there is little work that analyzes the response of immigrant *flows* to macroeconomic cycles (exceptions include Davis and Haltiwanger 1992; Hanson and Spilimbergo 1999; Borger 2008; Mandelman and Zlate 2012). Additionally, we further disentangle GDP differentials to separately identify push and pull factors, adding to recent work by Pedersen, Pytlikovà, and Smith (2008), Warin and Svaton (2008), Zaiceva and Zimmermann (2008), and Mayda (2010). This allows us to assess whether less-educated immigrants leave countries that are experiencing short-run downturns (i.e., recessions) or are attracted by states experiencing short-run booms. Similarly, we ask whether U.S. immigrants are pulled into U.S. states with higher income or are instead being pushed out by persistent poverty in their origin country.

Third, we estimate gravity models of immigration in the spirit of Karemera, Oguledo, and Davis (2000), Lewer and Van den Berg (2008), Ortega and Peri (2009), Mayda (2010), and Beine, Docquier, and Ozden (2011). However, we employ a number of techniques, including the two-part and Eaton and Tamura (1994) threshold Tobit models—methods that, to our knowledge, have not yet been used to analyze the determinants of immigration.[3] The use of these models is necessitated by unique features of our data. Specifically, we observe annual bilateral gross flows of less-educated immigrants in the labor force from 112 different source countries into each of the contiguous 48 U.S. states, but approximately 95% of our sample has an immigration flow value of zero. This presents estimation challenges since the standard gravity model adopts log-flows as the dependent variable. We first estimate our gravity model

[1] Also see Mayda (2006).

[2] For example, proposals by Lindsey Graham and Charles Schumer in February 2011 (Budoff Brown 2011) and Chris Coons, Jerry Moran, Marco Rubio, and Mark Warner in May 2012 (Weisman 2012) favored high-skilled immigration.

[3] A growing trade literature has provided ample support for the Eaton and Tamura technique, including Head and Ries (1998), Rauch and Trindade (2002), and Martin and Pham (2008).

using a scaled ordinary least squares (SOLS) regression in which we add 1 to each observed immigrant flow. Next, we follow the trade literature and apply a threshold tobit model in the spirit of Eaton and Tamura (1994) to account for the zero flows. Last, we employ the two-part model that estimates a probit regression, followed by an ordinary least squares (OLS) specification that drops all observations with zero flows.

Our results indicate that fluctuations in GDP positively affect the immigration of less-educated men, but only when the entire sample of immigrant flows is considered. If the observations with zero immigrant flows are dropped, we find no relationship between short-run GDP differentials and immigration. Effects from long-run GDP differentials follow a similar pattern, although baseline Eaton and Tamura estimation fails to find a significant relationship. In subsequent push and pull analysis, however, models that include zero-flow values robustly find that long-run GDP trends push less-educated men out of their origin countries, and that recent booms in U.S. states attract less-educated men from abroad. Conversely, there are no pull effects from long-run state GDP trends, nor do short-run origin-country GDP fluctuations spur men to emigrate.

Further analysis considers alternative subsamples of U.S. immigrants to see if certain groups respond similarly to GDP differentials. We briefly discuss how Mexican immigrants, who represent the vast majority of new less-educated male immigrants in the United States, affect the analysis. Results are robust to the exclusion of Mexico, however Mexican immigrants themselves are more responsive to short-run GDP differentials. Next, we find that the flows of less-educated female immigrants are much less responsive to short-run GDP fluctuations than their male counterparts. We also perform a number of additional robustness checks that are omitted from the article but available in an online appendix,[4] including alternative estimation procedures, regressions using a shorter panel, and specifications for male immigrants with a college degree. Most of our results remain robust to the various empirical specifications.

The article is organized as follows. First, we motivate our empirical specification with a simple model and provide a thorough explanation of the estimation techniques. Next, we summarize some of the important trends regarding recent U.S. immigration and describe our data in detail. We then present the results and consider various robustness checks. Finally, we discuss how our results add to existing literature.

2. Empirical Strategy

Theoretical Motivation

The canonical theoretical model of migration consists of an income maximization problem in which a potential immigrant from origin country o chooses destination d based on the relative returns to migrating after factoring out migration costs. Assume there is a discrete number of origin countries $o = \{1, 2, \dots, O\}$ and a discrete number of destinations $d = \{1, 2, \dots, D\}$.

Following the work of Ortega and Peri (2009), Beine, Docquier, and Ozden (2011), and Grogger and Hanson (2011), we assume a linear utility function. The utility of an agent from country o who remains in country o is therefore

$$u_{o,o} = Y_o + A_o + \varepsilon_o, \tag{1}$$

[4] The web appendix is available at www.colgate.edu/simpsonsparber_SEJwebappendix

where Y_o represents income in the origin country, A_o represents country-specific factors (such as amenities, etc.), and ε_o is the extreme-value distribution error term. Immigration researchers use either aggregate measures of income (i.e., GDP) or micro-level measures of income (i.e., wages) to model Y_o. We choose the former because we are assessing how less-educated immigrant flows respond to macroeconomic differences across a large set of destinations; that is, we are not trying to measure the response of immigrants to variations in the return to skill, for example.[5]

The utility of an agent from country o who decides to migrate to destination d is

$$u_{d,o} = Y_d + A_d - C_{d,o} + \varepsilon_d, \tag{2}$$

where migration costs are denoted by $C_{d,o}$ and can include costs that are specific to the destination (i.e., immigration restrictions), bilateral costs between the destination and origin country (i.e., language differences), or costs that are individual-specific (i.e., family members left back home).

The agent chooses the destination k that maximizes his/her utility:

$$\max_{k=\{1,\ldots,D\}} \{u_{k,o}\}. \tag{3}$$

Using this simplified model, the probability that an individual born in country o will move to destination d is then

$$\mathrm{pr}\left(u_{d,o} = \max_k \{u_{k,o}\}\right) = \frac{M_{d,o}}{M_o}, \tag{4}$$

where $M_{d,o}$ is the number of immigrants from origin country o in destination d and M_o is the native population of the origin country o. When the random term follows an independent and identically distributed extreme-value distribution, we can apply the results in McFadden (1984) to deliver

$$\frac{M_{d,o}}{M_o} = \frac{\exp(u_{d,o})}{\exp(u_{o,o})} \tag{5}$$

or, equivalently,

$$\frac{M_{d,o}}{M_o} = \frac{\exp(Y_d + A_d - C_{d,o})}{\exp(Y_o + A_o)}. \tag{6}$$

Taking natural logarithms of both sides yields

$$\ln\left(\frac{M_{d,o}}{M_o}\right) = Y_d - Y_o + A_d - A_o - C_{d,o} \tag{7}$$

or, equivalently,

$$\ln(M_{d,o}) = Y_d - Y_o - C_{d,o} + A_d - A_o + \ln(M_o). \tag{8}$$

---

[5] For a recent discussion of this issue, we refer the reader to Rosenzweig (2007).

Thus, immigrant flows depend on the aggregate income differential between the destination and origin ($Y_d - Y_o$), moving costs that depend on the destination and origin of the immigrant ($C_{d,o}$), origin- and destination-specific factors (A_d, A_o), and the population of the origin country (M_o).

Empirical Specification

Equation 8 motivates the basic empirical specification in Equation 9—a gravity model of immigration similar to Karemera, Oguledo, and Davis (2000), Lewer and Van den Berg (2008), Ortega and Peri (2009), Mayda (2010), and Beine, Docquier, and Ozden (2011).

$$\begin{aligned}\ln(M_{t+1,d,o}) = {} & \alpha + \beta(Y_{t,d} - Y_{t,o}) \\ & + \delta \ln(\text{Dist}_{d,o}) + \text{FE}_d + \text{FE}_o \\ & + \gamma \ln(\text{Stock}_{t,d,o}) + \eta \ln(\text{Pop}_{t,d}) + \mu \ln(\text{Pop}_{t,o}) + \text{FE}_t + \varepsilon_{t+1,d,o}\end{aligned} \tag{9}$$

The dependent variable $M_{t+1,d,o}$ measures the flow of immigrants from origin country o to destination state d at time $t + 1$. The income differential is measured using time t per capita GDP differentials, $Y_{t,d} - Y_{t,o}$. Notice that we lag the independent variables (by one year) to mitigate endogeneity issues. This lagged specification is also more appropriate if migration decisions are more likely to be based on past, as opposed to current, economic conditions.

We follow the literature in identifying control variables that proxy for migration costs. We include the natural log of the distance between the origin country's capital city and the state's geographic center ($\text{Dist}_{d,o}$). Time-variant factors include the natural log of a measure of the immigrant stock from country o residing in state d ($\text{Stock}_{t,d,o}$), the natural log of the state's population ($\text{Pop}_{t,d}$), and the natural log of the origin country population ($\text{Pop}_{t,o}$). Year fixed effects (FE_t) account for time trends as well as U.S. immigration policy decided at the national level (see Clark, Hatton, and Williamson [2007] for further discussion of the importance of policy). Destination and origin fixed effects (FE_d and FE_o) account for region-specific factors that do not change over time and imply that all coefficients of interest will be identified by variation within regions over time. The error term is represented by $\varepsilon_{t+1,d,o}$, and $\{\alpha, \beta, \delta, \gamma, \eta, \mu\}$ are the coefficients to be estimated.

We modify this framework by further decomposing GDP into two components. First, we consider a long-run country-specific GDP trend, $\hat{Y}_{t,c} = \hat{a}_c + \hat{b}_c T$ for $c = \{o, d\}$ with time trend T. The coefficients $\hat{a}_c$ and $\hat{b}_c$ are obtained by estimating the following country-specific regressions, where $e_{c,t}$ is an error term:

$$Y_{t,c} = a_c + b_c T + e_{c,t}. \tag{10}$$

We compute short-run fluctuations in GDP from its long-term trend, such that $\Delta Y_{t,c} = Y_{t,c} - \hat{Y}_{t,c}$. Thus, Equation 9 can be rewritten as Equation 11. Migration flow effects determined by differences between destination state and origin country trend GDP are measured by β_1, whereas β_2 represents the effect from differences in short-term GDP fluctuations.

$$\begin{aligned}\ln(M_{t+1,d,o}) = {} & \alpha + \beta_1 \left(\hat{Y}_{t,d} - \hat{Y}_{t,o}\right) + \beta_2(\Delta Y_{t,d} - \Delta Y_{t,o}) \\ & + \delta \ln(\text{Dist}_{d,o}) + \text{FE}_d + \text{FE}_o \\ & + \gamma \ln(\text{Stock}_{t,d,o}) + \eta \ln(\text{Pop}_{t,d}) + \mu \ln(\text{Pop}_{t,o}) + \text{FE}_t + \varepsilon_{t+1,d,o}\end{aligned} \tag{11}$$

Estimation Techniques

Anderson (2011) reports that gravity models were initially introduced to study immigration flows by Ravenstein (1889). However, they have been used most widely by trade economists to analyze bilateral export and import flows. The characteristics and limitations of the gravity model are therefore shared by these two fields, so knowledge from the trade literature is informative for our estimation technique.

Gravity models of international trade regress log bilateral trade flows (either exports or imports) on the economic mass of each trading partner, the geographic distance between them, and other covariates. Our procedure simply replaces trade flows with gross immigrant flows. Estimation problems arise, however, when country pairs experience zero flows because log values are undefined. This is a nontrivial issue both in trade and in our analysis. For example, half of the observations used in recent important work by Santos Silva and Tenreyro (2006) and Helpman, Melitz, and Rubinstein (2008) equaled zero. Summarizing trade data on the 10-digit harmonized system of goods classification (HS10), Baldwin and Harrigan (2011, p. 23) report that "The U.S. imports nearly 17,000 different HS10 categories from 228 countries, for a total of over 3.8 million potential trade flows [but] over 90 percent of these potential trade flows are zeros." In our data set of immigrant flows from origin countries to U.S. states, we encounter values of zero in roughly 95% of the observations. Thus, the proportion of zero values in our data set is quite similar to that confronted in trade, which motivates us to consider a variety of estimation techniques from the trade literature.

Martin and Pham (2008) thoroughly evaluate the efficacy of alternative estimation strategies when many zero values are present. SOLS offers a common method for overcoming this limitation and adds a scalar (usually 1) to each flow value before taking natural logs. Analysts can augment this approach by performing tobit estimation and censoring log values less than zero. Others estimate a truncated model (i.e., drop observations of zero flows). The two-part model (explained in the next section and often employed in health economics) first estimates a probit model to identify the determinants of whether positive values exist and then performs OLS estimation of the truncated model. Less-common methods include the Eaton and Tamura tobit estimator, the Heckman two-step estimator, and the pseudo Poisson maximum likelihood (PPML) procedure advocated in a well-known article by Santos Silva and Tenreyro (2006). Ultimately, Martin and Pham (2008, p. 20) argue that truncated OLS models outperform censored regressions and that "just solving the 'zero problem' and adding the zero valued observations to the sample is quite an unhelpful strategy." The smallest biases arise when using Eaton and Tamura tobit estimators (after controlling for heteroscedasticity). The Heckman two-step estimator performs well only if the true underlying data are governed by a Heckman selection model data–generating process. Otherwise, the Heckman model commonly fails to converge or produces massive biases.[6] PPML performs well "for analysis of nonlinear relationships in models where zero values of the dependent variable are infrequent" (Martin and Pham 2008, p.2), but the authors go on to emphasize that it provides severely biased estimates and is inferior to the Eaton and Tamura procedure when many observations equal zero.

[6] Moreover, the Heckman model requires one variable used in the first (selection) stage of the model to be omitted from the second (quantity) stage. In the context of immigration, this would require a variable that is related to the probability of positive immigration flows but unrelated to the size of immigrant flows among observations with positive values.

Within the literature on the determinants of migration, most economists using the gravity approach address the problem of zero flows by adopting truncated, SOLS, or censored methodologies.[7] Some eschew the gravity model and instead measure flows or emigration rates in levels (not logs).[8] A few, however, are beginning to take the issue of zero immigration flows more seriously. For example, the Falck et al. (2012) analysis of linguistic determinants of German regional migration is robust across truncated and PPML methodologies. PPML seems appropriate in their setting because only about 4% of their flows equal zero. Alternatively, Beine, Docquier, and Ozden (2011) estimate the role of diasporas (i.e., the stock of current immigrants) in predicting the current flow of immigrants using bilateral data from the OECD. They too have a large number of observations with zero values, but they favor the Heckman selection and instrumental variable methods to help with endogeneity issues.

Gravity model limitations are not, of course, limited to the problem of zero flows. An important emerging literature on multilateral resistance has argued that by estimating bilateral flows without taking into account phenomena occurring outside of origin countries and destination states, regression results could be biased (Hanson 2010). Bertoli and Fernandez-Huertas Moraga (2011) and Pesaran (2006) advocate a common correlated effects (CCE) estimator as a solution, with the former article finding that the method reduces the estimated effect of origin GDP on migration flows to Spain to two-thirds of the effect identified by standard models (although the coefficient is still significant).

The estimation strategy in this article employs traditional SOLS and two-part estimation because of the popularity of those models in the literature. The frequency of zero flows in our data set, coupled with recent evidence in Martin and Pham (2008), motivates us to also perform the Eaton and Tamura procedure. Although we believe the CCE estimator may be relevant in future work, we do not explore it in this article, in part because we fear that it would distract from our focus on comparing results across SOLS, Eaton and Tamura, and two-part models. Instead, we control for worldwide macroeconomic factors simply by including year indicators in our specifications.

Eaton and Tamura

The SOLS method of adding 1 to the dependent variable before taking logs, though common, is inherently biased in the sense that there is no reason to prefer an added scalar of 1 as opposed to any other value. Eaton and Tamura (1994) introduced a threshold tobit model to overcome this limitation. When analyzing Japanese and American trade patterns with a sample of countries in the late 1980s, the authors were confronted with a data set in which many trade flows equaled zero. Rather than simply adding 1 to each value before taking logs, they added λ, a value to be statistically estimated.[9]

---

[7] See Dodson (2001), Lewer and Van den Berg (2008), Lewer, Pacheco, and Rossouw (2009), Ortega and Peri (2009), Falck et al. (2012), or Beine, Docquier, and Ozden (2011) for recent examples.

[8] See Zavodny (1997), Dodson (2001), Pedersen, Pytlikova, and Smith (2008), Mayda (2010), or Adsera and Pytlikova (2012).

[9] Head and Ries (1998) note that one problem with adding 1 to each observation is that results will be sensitive to the units of measurement, whereas the Eaton and Tamura method overcomes this limitation.

Let the flow of immigrants ($M_{t+1,d,o}$) to destination state d from origin country o in year $t+1$ be defined by

$$M_{t+1,d,o} = \max\{0, \tilde{M}_{t+1,d,o}\} \tag{12}$$

The latent variable $\tilde{M}_{t+1,d,o}$ is a function of several year t determinants of migration ($X_{t,d,o}$), a mean-zero normally distributed error term ($\varepsilon_{t+1,d,o}$), and a threshold value (λ) that the function of explanatory variables must achieve before positive migration flows occur.[10]

$$\tilde{M}_{t,d,o} = -\lambda + \exp(\alpha + \beta X_{t,d,o} + \varepsilon_{t+1,d,o}) \tag{13}$$

By substituting Equation 13 into Equation 12, rearranging, and taking natural logs, we derive Equation 14. Eaton and Tamura (1994) provide the density function for $\tilde{M}_{t+1,d,o}$ and the necessary log-likelihood function for maximum likelihood estimation[11]; thus,

$$\ln(\lambda + M_{t+1,d,o}) = \begin{cases} \alpha + \beta X_{t,d,o} + \varepsilon_{t+1,d,o} & \text{if} \quad \tilde{M}_{t+1,d,o} > 0 \\ \ln(\lambda) & \text{if} \quad \tilde{M}_{t+1,d,o} \leq 0. \end{cases} \tag{14}$$

The Eaton and Tamura model is not altogether unfamiliar to economists who have examined immigration issues; Head and Ries (1998) and Rauch and Trindade (2002) used the methodology in their influential analyses of immigration's role in promoting international trade. To our knowledge, however, we are the first to apply the technique to a gravity model of the determinants of immigration. The model presents two limitations, however. First, because it is a nonlinear model estimated by maximum likelihood, it is possible that it will fail to converge to a solution. We do not encounter this problem in our analysis. Second, it can be difficult to interpret coefficient estimates, as is the case with the common SOLS solution of adding 1 to zero values. Strictly speaking, coefficients do not represent percentage changes of the dependent variable, although we follow the often-used convention of interpreting them in this manner.

Two-Part Model

The two-part model consists of first estimating a probit model with a latent variable formulation.[12] If $M_{t+1,d,o}$ is the flow of immigrants to destination state d from origin country o in year $t+1$, then let the indicator $M^* = 1$ if $M_{t+1,d,o} > 0$ and $M^* = 0$ otherwise. As before, the regressors are $X_{t,d,o}$. The two-part model for $M_{t+1,d,o}$ is then

$$f(M_{t+1,d,o}|X_{t,d,o}) = \left\{ \begin{array}{ll} = \Pr(M^* = 0|X_{t,d,o}) & \text{if} \quad M_{t+1,d,o} = 0 \\ = \Pr(M^* = 1|X_{t,d,o}) f(M_{t+1,d,o}|M^* = 1, X_{t,d,o}) & \text{if} \quad M_{t+1,d,o} > 0 \end{array} \right\}. \tag{15}$$

The two-part model consists of: (i) estimating a probit on M^* and (ii) estimating the truncated OLS for $M_{t+1,d,o}$ specified in Equation 11 if $M_{t+1,d,o} > 0$.

[10] Head and Ries (1998) interpret λ as undermeasurement.

[11] We are indebted to Cong S. Pham for kindly providing Stata code for the procedure.

[12] When the dependent variable exceeds zero, the model is a hurdle (or threshold) model.

3. Data

We focus our analysis on immigrants with a high-school degree or less education. We consider only those who are in the U.S. labor force (both employed and unemployed) at the time of survey and are between 18 and 89 years of age.[13] We first analyze the flow of male immigrants, but then incorporate female immigrants into our analysis in the next section.

We limit our analysis to the 2000–2009 period. Although this is a relatively short time series for analyzing short- and long-run GDP differences, it was a decade of considerable volatility in GDP, both in the United States and abroad (particularly when compared with the Great Moderation of the 1990s). One advantage of this short time series is that U.S. immigration policy was relatively unchanged during the period (with a few notable exceptions, including changes in the number of H-1B visas for college-educated workers). However, this decade witnessed the largest inflow of new immigrants in U.S. history, with approximately 14 million new (legal and illegal) immigrants (Camarota 2011). Additionally, new immigrants were more dispersed across the United States, with fewer immigrants going to traditional U.S. destinations compared with previous decades (Camarota 2011). As a result, we think it is important to understand how GDP differentials between origin countries and U.S. states affected the flow of new immigrants during this period.

We obtain bilateral immigration data from the Integrated Public Use Microdata Series for the 2000–2009 Census and ACS surveys (Ruggles et al. 2010). The value of this data set is that it is relatively large, provides annual measures of the combined legal and illegal U.S. immigrants in the labor force by country of origin, and identifies immigrants by state of residence. We believe the number of immigrants in each state who arrived in the United States within the last year (from the survey date) represents a reasonable proxy for new immigrant labor flows.[14] This definition of gross inflows intentionally excludes two groups that nonetheless warrant explicit recognition. First, it does not recognize internal migration of immigrants—those who previously arrived in the United States but recently moved to a new state.[15] Second, it omits circular or repeat migrants—those who recently re-entered the United States but first arrived in earlier years. Both of these groups instead represent part of a state's immigrant stock control variable. The one-year lag structure of our regression models implies that internal migrants and circular migrants returning to the United States in time t comprise part of the immigrant stock in time t, which is a determinant of time $t + 1$ immigrant flows. Because our focus is on newly arrived U.S. immigrants, our data measurement is appropriate and will not bias the results.

The Census and ACS data do, however, present a few limitations. First, there is likely a lag between arrival in the United States and being enumerated in the survey, and this lag may lead to a downward bias in immigrant flows. This issue may be especially salient in the case of less-educated illegal immigrants. Also, the ACS is administered monthly, but information is available only at the aggregated annual level. An economic shock in period t might have a

[13] Note that 99% of the flow of new immigrants into the labor force (i.e., those arriving to the United States for the first time) are between the ages 18 and 65.

[14] Beginning in the 2000 Census, the *yrimmig* variable reports the year an immigrant first entered the United States. In earlier surveys, *yrimmig* only provided a range of years that included the year of arrival. This, coupled with the nonexistence of annual ACS surveys, prohibited previous research from using Census and ACS data to generate accurate measures of newly arrived U.S. immigrants.

[15] Our analysis is related, but not directly comparable, to the work of Borjas (2001) and others that analyze how newly arrived immigrants (those who have been in the United States fewer than five years) respond to wage differentials within the United States.

Table 1. Characteristics of Immigrant Population, 2000–2009

Type of Immigrant	Flow	% of Total
Male, less-educated, labor force	3,136,560	28.0
Male, less-educated, not in labor force	628,320	5.6
Male, well-educated, labor force	1,777,795	15.8
Male, well-educated, not in labor force	463,279	4.1
Female, less-educated, labor force	1,167,661	10.4
Female, less-educated, not in labor force	1,726,462	15.3
Female, well-educated, labor force	1,016,070	9.0
Female, well-educated, not in labor force	1,339,901	11.9
Total	11,256,048	
Total male	6,005,954	53.4
Total less-educated	6,659,003	59.2
Total in labor force	7,098,086	63.1

Data are from the 2000–2009 ACS.

larger effect on potential migrants at the beginning of period $t + 1$ than at the end, but the ACS will not allow us to identify such a distinction. However, we are (to our knowledge) first to use the Census and ACS to generate annual gross inflow data for the United States and measure its response to state-level economic conditions. While there is little we can do to directly address these issues, we believe that they do not significantly bias our analysis. If anything, our data is understating immigrant flows and smoothing out business cycle responses, leading to estimates that are lower bounds.

Recent Immigration Trends

Our data set records approximately 11.26 million new immigrants (both legal and illegal) having entered the United States between 2000 and 2009. Table 1 shows the breakdown of the sample by gender, education level, and employment status. Approximately 53% of the new immigrants are male, 59% have a high school degree or less education, and 63% are in the labor force at the time of the survey. Less-educated male immigrants represent the largest subgroup of new immigrants with 28% of the sample (3.14 million people), and they constitute the bulk of our analysis. Female immigrants in the labor force of all education levels represent just 19% of new immigrants.

The first column of Table 2 reports the primary regions of origin of all new immigrants between 2000 and 2009. Latin America has provided the clear majority (55.7%) of new

Table 2. Regions of Origin, 2000–2009

Region	Flow as % of Total	Flow as % of Baseline Sample[a]
Latin America	55.72	86.14
Asia	26.11	7.09
Europe/Russia	11.42	3.88
Africa	3.43	2.06
Canada	2.55	0.49
Oceania	0.78	0.34

Data are from the 2000–2009 ACS.
[a] The "Baseline Sample" represents new less-educated male immigrants in the labor force.

Table 3. Largest Countries of Origin, 2000–2009

Region	Flow as % of Total	Flow as % of Baseline Sample[a]
Mexico	36.7	66.81
India	7.2	1.02
Philippines	3.7	0.61
China	3.6	1.65
Canada	2.5	0.49
Korea	2.5	0.29
Guatemala	2.3	4.65
El Salvador	2.1	2.97
Japan	2.0	0.19

[a] The "Baseline Sample" represents new less-educated male immigrants in the labor force. Note that Brazil and Honduras are among the top sending countries for the sample, representing about 1.9% each (but are not listed in the table).

immigrants. The second largest sending region is Asia, representing 26.1% of all new U.S. immigrants. Approximately 11.4% of all new immigrants originate from Europe. Not surprising, the distribution of sending regions for male less-educated immigrants in the labor force (reported in column 2) is quite different, with 86% coming from Latin America. Only 7.1% and 3.9% of these types of immigrants originate from Asia and Europe, respectively.

As evident in Table 3, more than one-third of all new immigrants and two-thirds of all new less-educated male immigrants in the labor force are from Mexico. India is the next largest sender, representing 7.2% of all immigrants. Immigrants from the Philippines and China represent approximately 3.7% of all U.S. immigrants. Notice, however, that the distribution of less-educated male immigrants in the labor force is more skewed to Latin American countries, with a disproportionate share coming from Guatemala, El Salvador, Brazil, and Honduras (in the second column). Recall that our data does not include circular migrants, which may be especially relevant for close countries, namely Mexico and Central American countries, for which individuals might find it easier to cross the border repeatedly. If circular migrants were included, these countries would represent even a larger share of immigrant flows.[16]

Table 4 provides a snapshot of where new immigrants are locating within the United States. For brevity, we categorize the U.S. states into six regions that are consistent with U.S. Census regions. Table 4 shows that approximately one-quarter of all new immigrants live in the Pacific and Southeast regions each. Approximately 20% of new immigrants live in the Northeast, 13% in the Midwest and South Central, and 8% in the Mountain region.[17] The distribution of the locational choices of less-educated male immigrants in the labor force is very similar (in column 2) to that of all immigrants (in column 1). The final column reports mean trend GDP by region and finds no clear correlation between trend regional income and recent immigrant flows.

[16] The Mexican Migration Project reports that roughly 60% of Mexicans arriving to the United States within the past year (during the 2000–2009 period) had resided in the United States before their latest arrival.

[17] The South Central region includes Arkansas, Louisiana, Oklahoma, and Texas.

Table 4. U.S. Destination Regions, 2000–2009

Regions	Flow as % of Total	Flow as % of Baseline Sample[a]	Trend GDP[b]
Pacific	24.18	23.02	\$50,360
Southeast	22.82	24.77	\$43,064
Northeast	19.85	15.33	\$53,814
Midwest	12.91	11.24	\$45,572
South Central	12.47	15.59	\$46,811
Mountain	7.76	10.05	\$44,492

Data are from the 2000–2009 ACS.
[a] The "Baseline Sample" represents new less-educated male immigrants in the labor force.
[b] Trend GDP reported in 2010 dollars. The six regions are consistent with Census Bureau regions and divisions.

Summary Statistics

Recall from Equation 10 that the trend component is estimated based on regressions (using data from 2000 to 2009) of $\hat{Y}_{t,c} = \hat{a}_c + \hat{b}_c T$ for $c = \{o, d\}$ and time trend T. After obtaining the trend component $\hat{Y}_{t,c}$, we then compute the short-run component as $\Delta Y_{t,c} = Y_{t,c} - \hat{Y}_{t,c}$, where $Y_{t,c}$ is the current year per capita GDP. Our main explanatory variables are differentials in GDP trend components, $\hat{Y}_{t,d} - \hat{Y}_{t,o}$, and differentials in the short-run GDP component, $\Delta Y_{t,d} - \Delta Y_{t,o}$, both of which are measured in real 2010 dollars. Origin country per capita real GDP comes from the World Development Indicators, and per capita GDP by state is from the Bureau of Economic Analysis. We include the 48 contiguous states in the analysis and drop the District of Columbia (which has an exceptionally high GDP per capita). We have nine years of data because we lag all of the independent variables.[18] We have complete data on 112 different source countries for a total of 48,374 observations ($48 \times 9 \times 112$). However, only 2609 observations have nonzero immigrant flows.

Table 5 reports the mean and standard deviation for each variable. The first two columns are for the entire sample, and the last two columns represent the sample of nonzero flows. The average bilateral flow of less-educated males between a country and a state is 1100 per year among observations with positive flows, but just 58 when including the entire sample. We report unweighted means to be consistent with our regression analysis that follows. Both samples exhibit tremendous variation in migrant flows.

GDP differentials, both long term and short run, are the independent variables of interest. Average trend per capita GDP of U.S. states is \$44,508 for the entire sample (with very little variation), whereas the average GDP for origin countries is \$13,282 (with high variation). Average fluctuations in per capita GDP would equal zero by construction if we used the entire time series because fluctuations are defined as the difference between current and trend GDP. Because the independent variables are lagged, however, we lose GDP data in 2000. The resulting averages are \$65 for origin countries and \$179 for U.S. states, leading to a \$114 gap in GDP fluctuations between the destination and origin of immigrants on average. Variation in GDP fluctuations is very high, with more variation coming from state fluctuations than from origin countries because the absolute deviations of income from its trend are higher for U.S. states compared with low-income countries.

[18] Our regressions ultimately include immigration flows from 2001 to 2009 and explanatory variables from 2000 to 2008.

Table 5. Summary Statistics

Variable	All Observations Mean	All Observations Standard Deviation	Observations with Positive Flows Mean	Observations with Positive Flows Standard Deviation
Immigrant flows	58.06	931.71	1100.545	3913.226
$\hat{Y}_{t,d}$: state trend (\$, $\times 10^3$)	44.508	7.845	47.724	7.057
$\hat{Y}_{t,o}$: country trend (\$, $\times 10^3$)	13.282	12.851	11.182	10.856
$\hat{Y}_{t,d} - \hat{Y}_{t,o}$: GDP trend differential (\$, $\times 10^3$)	31.226	14.986	36.542	13.343
$\Delta Y_{t,d}$: state fluctuations (\$, $\times 10^3$)	0.179	0.877	0.379	0.913
$\Delta Y_{t,o}$: country fluctuations (\$, $\times 10^3$)	0.065	0.462	0.060	0.344
$\Delta Y_{t,d} - \Delta Y_{t,o}$: GDP fluctuation differential (\$, $\times 10^3$)	0.114	0.924	0.319	0.916
Distance between origin country and destination state (miles)	5345	2168	4187	2527
Immigrant stock	4766	56,050	56,266	233,859
Country population ($\times 10^6$)	49.720	159.472	122.159	285.170
State population ($\times 10^6$)	6.051	6.465	11.835	9.528
Observations	48,384		2609	

Reported summary statistics are unweighted.

Our control variables include geographic determinants of migration, destination and origin populations, and immigrant stocks in each state. The geographic variables include the distance between world capitals and U.S. state geographic centers using the Haversine distance formula and latitude/longitude data from the CEPII Research Center[19] and the U.S. Census. Population estimates are from the World Development Indicators (for countries) and the U.S. Census (for U.S. states). Immigrant stock is calculated by measuring the number of immigrants in each state from each country (of all education levels, including men and women, and those in and out of the labor force). The mean immigrant stock is 4766. All of our regressions control for the natural log of 1 plus this stock value so that we do not lose observations with zero values.[20]

4. Results

Less-Educated Male Immigrant Flows

We model the flow of less-educated immigrant labor from origin country o to destination state d as specified in Equation 11, using independent variables that are lagged one year. Table 5 presents the baseline results for men; standard errors are clustered by state*country dyad. In column 1, we include all immigrant flows by adding 1 to the flow variable before taking the natural log and then employing OLS (i.e., SOLS); in column 2, we use the Eaton and Tamura technique; in columns 3 and 4, we use the two-part model that first estimates a probit

[19] http://www.cepii.fr/anglaisgraph/bdd/distances.htm

[20] We do this to maximize the number of observations in each regression. All of our results are comparable when we do not add 1 to the immigrant stock variable.

Table 6. Baseline Results: Less-Educated Immigrant Men

Independent Variable	$\ln(1 + M_{t+1,d,o})$	Eaton and Tamura	Two-Part Model	
			Probit	$\ln(M_{t+1,d,o})$
$\hat{Y}_{t,d} - \hat{Y}_{t,o}$: GDP trend	0.012	0.010	0.0006	0.022
	(0.003)***	(0.010)	(0.0003)**	(0.019)
$\Delta Y_{t,d} - \Delta Y_{t,o}$: GDP fluctuations	0.022	0.034	0.0011	0.002
	(0.005)***	(0.014)**	(0.0004)***	(0.023)
$\ln(\text{Dist}_{d,o})$	−0.710	−0.548	−0.0110	−0.564
	(0.122)***	(0.112)***	(0.0027)***	(0.176)***
$\ln(\text{Pop}_{t,o})$	0.539	1.317	0.0314	1.059
	(0.170)***	(0.511)**	(0.0131)**	(0.910)
$\ln(\text{Pop}_{t,d})$	−0.201	−0.455	−0.0033	−1.752
	(0.271)	(0.538)	(0.0142)	(0.913)*
$\ln(1 + \text{Stock}_{t,d,o})$	0.037	0.140	0.0039	0.086
	(0.003)***	(0.016)***	(0.0003)***	(0.021)***
Observations	48,384	48,384	47,520	2609
R^2	0.326	—	—	0.584

Cluster-robust standard errors are in parentheses. Results incorporate time, destination, and origin fixed effects.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

(in column 3) and then estimates the nonzero immigrant flows using truncated OLS (in column 4). The sample size is much smaller in column 4 compared with columns 1–3 because observations with zero immigrant flows are dropped.[21]

SOLS results (column 1) suggest that both trend GDP and GDP fluctuations are significant determinants of the flows of less-educated immigrant men into the United States. Coefficients indicate that a \$1000 differential in GDP fluctuations between the destination state and origin country leads to a 2.2% immigrant flow (significant at the 1% level). Similarly, a \$1000 increase in the trend GDP differential between the destination state and origin country induces a significant 1.2% increase in immigration. This is particularly striking given that the model is estimated with a full set of country and state fixed effects. The coefficient is identified only by differences in trend growth rates across states and countries (b_cT in the Eqn. 10 construction of our trend variable), not by differences in permanently high levels of per capita GDP (a_c in Eqn. 10). We should also note that our array of fixed effects would absorb all of the variation in trend GDP if we had restricted growth rates (b_c) to be equal across states and countries. Thus, the GDP trend coefficient in Table 6 is only identifiable because we allow for state- and country-specific trends.

These baseline results are consistent with those of previous studies. Ortega and Peri (2009), for example, find a significantly positive relationship between GDP differentials and bilateral immigrant flows using OECD data. Their OLS specification is similar to our SOLS specification in that they add 1 to both immigrant flows and immigrant stocks while also including observations with zero flows. Their magnitudes are not directly comparable to ours because they use a different database, cover a cross-section of source countries, and do not

[21] In the probit specification, dummy variables for two countries (Antigua–Barbuda and Finland) perfectly predict the zero outcome. Hence, these observations are dropped in the estimation, leading to 864 fewer observations (2 countries × 48 states × 9 years).

distinguish between trend and cyclical effects. Nonetheless, they find that a \$1000 GDP differential (in levels) leads to a 10 percentage point increase in bilateral immigration flows across OECD countries.

Our control variables have the expected signs when significant. Distance is negatively associated with higher flows of less-educated immigrant men, and larger origin countries send more immigrants. Both results are consistent with the literature (i.e., Karemera, Oguledo, and Davis 2000; Lewer and Van den Berg 2008; Ortega and Peri 2009). We also find that immigrant stocks are highly positively correlated with immigrant flows. This network effect has been frequently documented in the literature (Bartel 1989; Zavodny 1997; Clark, Hatton, and Williamson 2007; Mayda 2010; Grogger and Hanson 2011).

Column 2 uses Eaton and Tamura threshold tobit estimation, which we prefer to SOLS because it allows the scalar added to flow values to be estimated by the data itself (as opposed to simply adding 1 before taking logs). Our results from employing these two strategies are similar but with important differences. First, effects from short-run GDP differentials increase somewhat. A \$1000 differential in GDP fluctuations between the destination state and origin country leads to a 3.4% increase in less-educated male immigration flows.[22] More interestingly, however, the coefficient on trend GDP loses significance. As with our SOLS specification, interpretation of this result must come with the caveat that much of the immigration effect from differences in long-term GDP are absorbed in the fixed effects.

The two-part model in columns 3 and 4 separates the likelihood of a country sending any immigrants from the magnitude of the immigrant flow response among existent bilateral immigration routes. The probit model in column 3 suggests that long-run and short-run GDP differentials both matter in determining which countries send positive less-educated male immigrants to the United States. Neither income measure, however, is important in determining the size of the flows in the truncated OLS model of column 4.

Our preferred interpretation of the collective results is that the significant coefficients in the SOLS and Eaton and Tamura specifications are driven by the discrete jump of going from zero to positive flows, not in changing the magnitude of flows within existent bilateral migration channels. Expressed another way, GDP fluctuations are associated with the flow of less-educated immigrant men as long as zero-flow values are included in the estimation. This is robust across our three different estimation techniques, as each suggests that a rise in short-run GDP will lead to an increase in the flow of less-educated immigrant men into that state. The evidence for long-run GDP differentials is slightly less robust. Simply adding 1 to the dependent variable leads to a significant coefficient on long-run GDP differentials. When using the Eaton and Tamura method, which estimates a scalar to add to the dependent variable, long-run GDP differentials are no longer significant.

Robustness Checks

In this section, we consider various robustness checks to determine whether our results depend on the sample being analyzed. First, we explore the possibility that the insignificant coefficients in column 4 of Table 6 are simply due to the smaller sample size of observations with positive bilateral

---

[22] We follow Head and Ries (1998) in interpreting coefficients as percent changes, but caution that the coefficients may not be true elasticities because of the parameter λ and because this is a tobit regression. The relative magnitudes of the coefficients across the specifications are less important than their sign and significance.

migration flows. We adopt two methods to explore this possibility, both of which support the conclusion that short-run GDP determines the existence of positive flows, but not the magnitude of those flows. Evidence on what is driving coefficients on long-run GDP is not conclusive.

First, we consider the effects of sample size on standard error calculations. The full data set has the potential for 5376 clusters (48 states × 112 countries) and 48,384 observations. The column 4 sample with positive flow values results in 1118 clusters and 2609 observations. If that sample had been equal in size to the full data set but had exhibited the same variation as present in the actual column 4 data, then the standard errors in that column would have been approximately half as large as the standard errors displayed in the table.[23] Note that this implies that the coefficient on GDP trend would become positive and significant, but the coefficient on GDP fluctuations would not.

Our second method of exploring the role of sample size involves bootstrapping. We begin by sampling (with replacement) 2609 observations with positive flow values and then estimating SOLS, Eaton and Tamura, and truncated $\ln(M)$ models. We perform this procedure 1000 times to assess how often the models are able to uncover positive and significant coefficients. This effectively provides a p-value for the null hypothesis that the GDP coefficients are positive and significant. We find that each method identifies positive and significant coefficients on trend GDP more than 18% of the time, so we fail to reject that null. It is possible that with a different sample of 2609 bilateral immigration channels exhibiting positive flows, each model would produce positive and significant coefficients on trend GDP. On the other hand, the coefficient on GDP fluctuations is positive and significant in only 3% of the trials. Because this occurs so rarely, we reject the null hypothesis. We believe that the insignificant coefficient on short-run GDP in column 4 arises from the nature of the data itself—GDP fluctuations are not associated with immigration flows among bilateral routes experiencing positive flows.[24] To augment this claim, we repeat the bootstrapping routine by sampling (with replacement) 2609 observations from all of the available data (including zero-flow values). SOLS and Eaton and Tamura routines uncover positive and significant coefficients on GDP fluctuations more than 20% of the time, thereby failing to reject the null hypothesis that the coefficient is positive and significant. Sample size alone is not able to rule out the potential for the coefficient on GDP fluctuations to be positive and significant, but the exclusion of zero-flow values does eliminate this possibility. Thus, we believe the inclusion of observations with zero flows is crucial for the ability of regression models to identify significant determinants of migration.

Next we analyze the role of Mexican immigrants. Because immigrants from Mexico represent a large share of flows, we drop Mexico from our sample of origin countries to test for robustness. This reduces the total number of available observations to 47,952 and the number of positive flow observations to 2246. Nonetheless, the results displayed in the top panel of Table 7 are almost identical to our previous baseline results, with strong evidence that GDP fluctuations are positively correlated with immigrant flows and weaker evidence for an effect from GDP trends. Although we have suppressed the remaining control variables from the table, they also have the same signs and significance as in the baseline case. Thus, the inclusion of Mexican immigrants in our full sample is not driving our results.

[23] That is, the smallest standard errors possible, assuming the same variation in the data but having the same number of clusters that are available in the entire data set, would be roughly $\sqrt{1118}/\sqrt{5376} = 0.46$ times the size of those reported in column 4.

[24] Note that SOLS, Eaton and Tamura, and general OLS specifications produce very similar estimates when each regression is restricted to use only observations with positive flow values. SOLS and Eaton and Tamura results for regressions using the 2609 actual observations with positive flows are available in the online appendix.

Table 7. Robustness Checks: Less-Educated Immigrant Men

Independent Variable	$\ln(1 + M_{t+1,d,o})$	Eaton and Tamura	Two-Part Model: Probit	Two-Part Model: $\ln(M_{t+1,d,o})$
	Flows from Mexico Dropped			
$\hat{Y}_{t,d} - \hat{Y}_{t,o}$: GDP trend	0.012	0.011	0.0006	0.026
	(0.003)***	(0.010)	(0.000)**	(0.020)
$\Delta Y_{t,d} - \Delta Y_{t,o}$: GDP fluctuations	0.022	0.024	0.0011	−0.000
	(0.005)***	(0.011)**	(0.000)***	(0.026)
	GDP Interacted with Mexico			
$\hat{Y}_{t,d} - \hat{Y}_{t,o}$: GDP trend	0.012	0.010	0.0006	0.022
	(0.003)***	(0.010)	(0.000)**	(0.019)
$\Delta Y_{t,d} - \Delta Y_{t,o}$: GDP fluctuations	0.021	0.035	0.0011	−0.014
	(0.005)***	(0.014)**	(0.000)***	(0.025)
GDP trend * Mexico	0.040	0.001	0.0000	0.011
	(0.044)	(0.010)	(0.000)	(0.016)
GDP fluctuations * Mexico	0.204	−0.021	0.0005	0.131
	(0.105)*	(0.032)	(0.001)	(0.060)**
	GDP Interacted with Distance			
$\hat{Y}_{t,d} - \hat{Y}_{t,o}$: GDP trend	0.154	0.115	0.0028	0.166
	(0.024)***	(0.031)***	(0.001)***	(0.057)***
$\Delta Y_{t,d} - \Delta Y_{t,o}$: GDP fluctuations	0.644	0.153	0.0094	0.567
	(0.123)***	(0.132)	(0.004)**	(0.215)***
GDP trend * ln(Distance)	−0.017	−0.013	−0.0003	−0.017
	(0.003)***	(0.003)***	(0.000)***	(0.007)***
GDP fluctuations * ln(Distance)	−0.073	−0.014	−0.0010	−0.069
	(0.014)***	(0.016)	(0.000)**	(0.027)***

Cluster-robust standard errors are in parentheses. Results incorporate time, destination, and origin fixed effects.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

The second panel of Table 7 assesses whether Mexican immigrants react to GDP differentials differently than immigrants from other countries by interacting a dummy variable for Mexico with the two GDP variables for the full sample. We find that the estimated coefficients on GDP trend and GDP fluctuations are entirely comparable to those in the baseline case, both in magnitude and significance. One new insight is that the interaction term of Mexico with GDP fluctuations is weakly significant and positive. This provides some support to the idea that male immigrants from Mexico tend to react to GDP fluctuations in a way that is structurally different from immigrants from other countries.

The final panel of Table 7 explores whether distance affects the relationship between GDP and immigrant flows. This could arise if migrants from countries with lower migration costs are more sensitive to income differentials. Interestingly, we find that whereas GDP trend and GDP fluctuations themselves are both positively associated with immigrant flows, income has a muted migratory effect for more distant (and costly) migration routes (notice the negative signs on the interaction terms). Trend GDP coefficients are jointly significant (evaluated at average distance) in SOLS and probit models, whereas short-term GDP coefficients are jointly significant in SOLS, probit, and Eaton and Tamura models, just as in the baseline regression.[25]

[25] The p-values of joint significance are available in the online appendix.

Altogether, the coefficients on short- and long-term GDP in Table 7 are comparable to those in baseline regressions. This gives us greater confidence that our baseline results are not driven by the inclusion of Mexico or immigrants from countries in close proximity to the United States.

We have performed several additional robustness checks that we have made available in an online appendix for the sake of brevity. Perhaps most notably, we consider two substitute procedures for measuring GDP: one that estimates the trend component by projecting 1990–1999 GDP growth onto the subsequent decade[26] and another that computes trends using an HP Filter on the 1990–2009 period. Both of these methods unfortunately reduce the number of observations since GDP is not available for all countries in our sample during the early 1990s, so we ultimately prefer the results from our reported trend construction that calculates trend GDP using data from 2000–2009. Nonetheless, the results from these alternative procedures continue to find that short-run GDP fluctuations are significantly correlated with immigrant flows. Moreover, they robustly find that the GDP trend differential is a significant predictor of immigrant flows. Other tables available in the online appendix include results from regressions that drop the final year of the data (to eliminate effects driven by the Great Recession), those employing country × state dyad fixed effects, Poisson specifications, regressions controlling for lagged immigrant flows in addition to stocks, and specifications for well-educated male labor. Overall, our results and conclusions are robust to these alternative methodologies.

Less-Educated Female Immigrant Flows

Thus far we have considered only less-educated male immigrants in the labor force. This is standard in the immigration literature when trying to isolate immigrants who move for economic purposes. Women are often disregarded in the immigration literature because of "a dearth of data on women and migration [that] makes it difficult to assess the full implications of international migration for women" (United Nations 2004, p. 4). Nonetheless, female labor migration is increasingly important in the United States and around the world, with women representing 47% of new immigrant flows in the United States between 2000 and 2009 (Table 1) and 49% of worldwide migrant flows in 2000 (United Nations 2004). Fortunately, the ACS/ Census data allow us to distinguish between men and women.

Estimates of Equation 11 for less-educated female immigrant labor flows are in Table 8. SOLS, Eaton and Tamura, and probit results for women echo those for men, but with muted and sometimes insignificant effects from the GDP variables. For example, the SOLS regression (in column 1) finds that a \$1000 increase in GDP trends will lead to a 0.8% increase in immigrant women, compared with 1.2% for immigrant men. Similarly, a \$1000 increase in GDP fluctuations leads to a 1.6% increase in immigrant women and a 2.2% increase in immigrant men. In the Eaton and Tamura model for women, the insignificant coefficient on GDP trend (0.005) is half the size of the effect from the male regression. The also insignificant coefficient on GDP fluctuations is less than one-third the size. Probit coefficients are similarly between one-half and one-third the size in female regressions. Altogether, these results suggest

[26] The method for projecting GDP based upon 1990–1999 GDP growth is to first calculate average GDP growth within states and countries during this period. We then take year 1999 GDP as fixed and allow states and countries to grow at these specific exponential growth rates for each subsequent year. The method assumes that countries were at trend in 1999 and that people expected growth to continue at the same rate as it did in the previous decade. Observations are lost because data are not available for all of our origin countries in early years of the 1990s.

Table 8. Less-Educated Immigrant Women

			Two-Part Model	
Independent Variable	$\ln(1 + M_{t+1,d,o})$	Eaton and Tamura	Probit	$\ln(M_{t+1,d,o})$
$\hat{Y}_{t,d} - \hat{Y}_{t,o}$: GDP trend	0.008 (0.003)***	0.005 (0.009)	0.0002 (0.0002)	0.076 (0.019)***
$\Delta Y_{t,d} - \Delta Y_{t,o}$: GDP fluctuations	0.016 (0.004)***	0.010 (0.012)	0.0004 (0.0003)	−0.017 (0.026)
$\ln(\text{Dist}_{d,o})$	−0.609 (0.114)***	−0.515 (0.107)***	−0.0097 (0.0024)***	−0.542 (0.180)***
$\ln(\text{Pop}_{t,o})$	0.240 (0.157)	0.625 (0.442)	0.0145 (0.0106)	−0.281 (0.908)
$\ln(\text{Pop}_{t,d})$	−0.030 (0.240)	−0.043 (0.450)	−0.0011 (0.0106)	1.526 (0.920)*
$\ln(1 + \text{Stock}_{t,d,o})$	0.027 (0.003)***	0.109 (0.017)***	0.0027 (0.0002)***	0.070 (0.018)***
Observations	48,384	48,384	47,520	2159
R^2	0.240	—	—	0.482

Cluster-robust standard errors are in parentheses. Results incorporate time, destination, and origin fixed effects.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

that female migration decisions are less sensitive to economic conditions than male decisions are. That is, the response of female immigrant labor flows to long-run GDP differentials is smaller than male flows, as long as the observations with zero values are included. The truncated model departs from this regularity by finding a large and significant coefficient on trend GDP for women, and a negative but insignificant coefficient on GDP fluctuations.

It is well documented that women might migrate for different reasons than men. For example, Taylor (2006, p. 20) notes that "men are more likely to make the move for purely economic reasons, while women are more likely than men to be 'tied movers'." Our result that less-educated female immigrants are less responsive to macroeconomic factors than male immigrants is consistent with this hypothesis. However, our findings also indicate that once female immigrant flows are nonzero, differences in long-run GDP affect the magnitude of the immigrant flow. The significantly positive relationship between long-run GDP differentials and recent female immigrant flows (for the truncated sample) suggests that economic forces are relevant. Still, more work remains to be done on identifying differences in the determinants of migration between men and women.[27]

Push and Pull Factors: Less-Educated Male Immigrants

The specification in Equation 11 assumes that the coefficients on the destination and origin country GDP are the same. Empirically, it is not necessary to impose this restriction. Equation 16 can help clarify the source of correlations and tell a more precise story about

[27] For an extended discussion of these issues, we refer the reader to the United Nations (2004) report.

immigrant flows by disaggregating the trend and fluctuation components of GDP by destination (pull factors) and origin (push factors).

$$\begin{aligned}\ln(M_{t+1,d,o}) = {} & \alpha + \beta_1 \hat{Y}_{t,d} - \beta_2 \hat{Y}_{t,o} + \beta_3 \Delta Y_{t,d} - \beta_4 \Delta Y_{t,o} \\ & + \delta \ln(\mathrm{Dist}_{d,o}) + \mathrm{FE}_d + \mathrm{FE}_o \\ & + \gamma \ln(\mathrm{Stock}_{t,d,o}) + \eta \ln(\mathrm{Pop}_{t,d}) + \mu \ln(\mathrm{Pop}_{t,o}) + \mathrm{FE}_t + \varepsilon_{t+1,d,o}\end{aligned} \tag{16}$$

Estimated coefficients for β_2 and β_4 will indicate whether origin income pushes emigrants out of their home countries, whereas estimates for β_1 and β_3 will identify whether destination income pulls immigrants into host states. When combined, these components represent dollar measures of GDP, which are commonly employed in gravity regressions. Although this is consistent with methods found in the literature (and are useful for exploring whether a potential migrant needs the promise of a dollar-amount gain to pay for the fixed costs of migration), they have limitations for assessing variation across countries. Dollar deviations from trends are higher for high-income observations relative to low-income ones, and a \$1000 deviation from trend could mean substantially more in terms of relative living conditions for poor countries than for rich ones. To address this, we perform push and pull regressions where we replace GDP fluctuations ($\Delta Y_{t,d}$ and $\Delta Y_{t,o}$) with short-run GDP measured relative to potential ($\tilde{Y}_{t,d} = \Delta Y_{t,d} / \hat{Y}_{t,d}$ and $\tilde{Y}_{t,o} = \Delta Y_{t,o} / \hat{Y}_{t,o}$). These percentage deviations from trend may be a better measure of the departure from typical (or expected) living standards in origin countries and destination states.

The results for less-educated male immigrants are reported in Table 9. The top panel displays coefficients using short-run GDP measured in dollars, whereas the bottom panel uses percentage deviations from potential GDP. The two analyses provide qualitatively equivalent results.

Although Tables 6–8 presented mixed evidence on the influence of long-term GDP differentials on migration decisions, both panels in Table 9 demonstrate that if such an effect exists, it is clearly driven by long-term GDP in the origin country and not by income in the destination state. The coefficient on state GDP trend is insignificant in all four specifications, while the coefficient on origin country GDP trend is negative and significant in all but the truncated OLS specification (in column 4). According to the Eaton and Tamura method in column 2 of the top panel, a \$1000 increase in origin country trend GDP leads to a 4.4% reduction in less-educated male immigrants. Recall that fixed effects absorb much of the trend GDP variation so that coefficients are being identified by differences in the growth portion of trend GDP. Thus, we see that countries experiencing more long-term GDP growth send fewer immigrants to the United States than slow-growth countries do, but state GDP trends are not a determinant in attracting immigrants.[28]

Different mechanisms appear to govern immigration's relationship with short-run GDP fluctuations. Unlike trend GDP, state-level GDP fluctuations attract immigrants, but fluctuations in origin-country income are always an insignificant determinant of immigrant flows. Similar to the baseline regressions in Table 6, the results in Table 9 argue for effects from short-run GDP only in specifications accounting for zero-flow values (columns 1–3). However, we now see that such effects arise primarily because economic booms in U.S. states attract

[28] We do caution, however, that the coefficient of variation for trend GDP data is nearly 5.5 times greater for countries than for states (see Table 5) and that insignificant coefficients on long-run state GDP might be attributable to small variations in the variable.

Table 9. Push and Pull Factors for Less-Educated Immigrant Men

			Two-Part Model	
Independent Variable	$\ln(1 + M_{t+1,d,o})$	Eaton and Tamura	Probit	$\ln(M_{t+1,d,o})$
		GDP Fluctuations ($)		
$\hat{Y}_{t,d}$: State GDP trend	0.004	−0.006	0.0000	0.012
	(0.005)	(0.013)	(0.0003)	(0.023)
$\hat{Y}_{t,o}$: Origin GDP trend	−0.027	−0.044	−0.0018	−0.053
	(0.005)***	(0.016)***	(0.0004)***	(0.033)
$\Delta Y_{t,d}$: State GDP fluctuations	0.025	0.048	0.0015	0.015
	(0.006)***	(0.015)***	(0.0004)***	(0.026)
$\Delta Y_{t,o}$: Origin GDP fluctuations	−0.001	0.030	0.0008	0.062
	(0.009)	(0.030)	(0.0008)	(0.054)
		GDP Fluctuations (%)		
$\hat{Y}_{t,d}$: State GDP trend	0.004	−0.005	0.0001	0.013
	(0.005)	(0.013)	(0.000)	(0.023)
$\hat{Y}_{t,o}$: Origin GDP trend	−0.027	−0.040	−0.0017	−0.046
	(0.005)***	(0.015)***	(0.000)***	(0.031)
$\Delta Y_{t,d}$: State GDP fluctuations	0.974	1.971	0.0624	0.530
	(0.273)***	(0.740)***	(0.020)***	(1.253)
$\Delta Y_{t,o}$: Origin GDP fluctuations	−0.109	−0.277	−0.0067	0.391
	(0.099)	(0.330)	(0.008)	(0.591)

Cluster-robust standard errors are in parentheses. Results incorporate time, destination, and origin fixed effects.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

immigrants from abroad. For example, the Eaton and Tamura results in the top panel of column 2 argue that a $1000 increase in short-run GDP in a particular U.S. state will lead to a 4.8% increase in male immigrants to that state. The bottom panel argues that a one–percentage point short-run deviation of state GDP from its trend will lead to a 1.97% increase in male immigrants to that state.

Altogether, we argue that long-run (or trend) GDP determines which countries send immigrants to the United States, whereas short-run fluctuations determine which U.S. state they move to. That is, immigrants are pushed out of poor (or slow-growth) countries and pulled into states that have experienced recent booms. These results are broadly consistent with what Mayda (2010, p. 1252) calls "a familiar puzzle"—theoretical models of migration generally predict push and pull factors to have equal but opposite effects, but empirical work often uncovers asymmetries. Mayda goes on to provide three possible explanations for her finding that pull factors are positively associated with higher emigration rates for a panel of OECD countries while push factors are rarely significant. These possibilities aid in understanding our results.

First, Hunt (2006) argues that young workers' migration decisions are not sensitive to cyclical conditions (such as unemployment) in their home countries. The behavior of young workers could overwhelm that of older workers, thus moving coefficients on push factors toward zero. At first glance, this appears to be a plausible explanation for our results because our flow variable is likely to be dominated by the activities of young workers—older immigrants arriving for family unification reasons are excluded from flows if they are not in the labor force, and circular migrants returning to the United States are similarly absent from the dependent variable. Deeper inspection of the data, however, rejects this explanation. We perform separate regressions (available in the online appendix) for flows of immigrants age 25 and younger and flows of immigrants over age 35. The push and pull effects of short- and long-term GDP are consistent with the results in Table 9 for both age groups. Most importantly, country income fluctuations fail to influence migration flows for all groups, whereas state GDP fluctuations attract immigrants in all but truncated regression models; these results hold for immigrants age 25 and under as well as those over age 35. Our estimated asymmetry in push and pull cyclical GDP effects do not appear to be explained by increased sensitivity among older immigrants to short-run GDP conditions in their home countries.

A second potential explanation is that low origin country income has two offsetting effects: it increases workers' incentives to emigrate, but it also inhibits their ability to finance migration costs through personal savings and/or by borrowing in imperfect capital markets. We find it plausible that the relative influence of these factors differ for short-term and long-term GDP. In countries that are persistently poor, the incentive to move often outweighs the significant financial and psychological costs of migration. Temporary income fluctuations are simply not important enough to change the relative influence of these factors. Trend GDP in the destination might not matter in the same way because of the nature of our data set. Each of our potential destinations is in the United States, so trend income will be relatively high regardless of the particular destination state. Note from Table 5 that the coefficient of variation of state trend GDP is just 0.176, compared with a value of 0.968 for country trend GDP. Not only might this inhibit the ability of regression models to identify significant coefficients on state trend GDP, but potential migrants might deem such differences trivial.

This second explanation does not address why short-term push and pull coefficients differ from each other. The third potential resolution to push and pull asymmetries address this issue by appealing to demand conditions. Immigrants might seek employment opportunities in states experiencing growing labor demand that cannot be met by the local native-born labor force. Moreover, economic expansions in the United States might increase the willingness of states to accept or even try to attract more immigrants. Although formal immigration policy is decided at the national level and is absorbed by time fixed effects, our model does not control for enforcement mechanisms and legislation pertaining to immigrants' interests that vary across time within states. If states become more welcoming during economic booms, then immigrant flows will be more sensitive to GDP fluctuations in states than in origin countries. Ultimately, our short-run state GDP results complement Borjas' (2001) finding that immigrants positively respond to business cycles within the United States and with relative ease since the costs of internal migration are lower.

5. Conclusion

This article adds to the literature on the determinants of immigrant flows in three ways. First, we use variation in income across U.S. states and origin countries to uncover how newly arrived less-educated immigrants respond to income differentials. Second, we decompose income differentials into short- and long-run components. Third, we employ several estimation techniques, including the threshold tobit and two-part models, to appropriately account for the large number of zero values for immigrant flows in our data set.

We studied U.S. immigration between 2000 and 2009. This period is an interesting case study because the United States experienced the largest gross inflow of new immigrants in its history, and those immigrants were more dispersed across the United States compared with recent cohorts. Additionally, the United States witnessed a severe recession in the latter half of this decade. Not only does this time period provide a great deal of macroeconomic variation, but it will also appeal to policy makers interested in the extent to which differences in trend GDP and GDP fluctuations are correlated with immigrant flows.

We find that both long-term and short-run GDP differentials significantly determine the flow of newly arrived less-educated male immigrants into U.S. states. The evidence for long-term GDP differentials is mixed, although coefficients might be difficult to identify because of the inclusion of state and country fixed effects. Additionally, the evidence for short-run differentials requires that observations of zero-flow values are included in the regression. For example, a truncated OLS specification that drops the observations with zero values (representing 95% of the sample) suggests that neither differences in GDP trends nor GDP fluctuations between the source country and destination state affect the flow of less-educated male immigrants into the United States. However, specifications that include zero values suggest otherwise, most notably in recognizing a positive relationship between GDP fluctuations and immigrant flows.

We document important differences in the response of recent immigrant flows to short- and long-run GDP components on the basis of gender and country of origin. For example, the flows of less-educated female immigrants into the United States are generally less responsive to differences in GDP fluctuations than their male counterparts. In addition, Mexican immigration, which constitutes a significant portion of all new immigrants, is not driving our results.

We also augment the immigration literature attempting to disentangle push and pull effects. We find that less-educated immigrants are pushed out of their countries by long-run GDP trends and are pulled into U.S. states by short-run upswings in economic activity. Not surprising, short-run fluctuations in the origin country do not lead to an increase in less-educated immigrant flows to the United States. It is not difficult to imagine a story consistent with these findings. People from poor countries want to immigrate to the United States, but short-term fluctuations in their country of origin are largely irrelevant for the decision to stay or leave. When deciding upon a new destination, however, an individual is likely to be enticed by a booming location and the associated promise of available jobs. From the perspective of a potential new worker, states with recent economic growth look more attractive than states with stagnant economic activity.

Our results also shed some light on the importance of empirical specification when studying immigrant flows. The truncated OLS model estimates the determinants of migration conditional on a bilateral country-by-state observation recording positive flows. It should not be taken as representative, however, for those wanting to understand all potential flows because the sample excludes 95% of all possible observations. The probit model, while useful, only estimates a dichotomous effect. That is, it identifies whether GDP differentials affect the probability that a country-by-state migration channel records positive flows. This may or may not be interesting to the policymaker. SOLS is a simple method of using the entire sample of data to estimate the effects of GDP on the quantity of immigration flows, but it accomplishes this by arbitrarily adding 1 to all flow values before taking logs. As an alternative, the Eaton and Tamura model allows the added scalar to be a value that is estimated by the data itself. It therefore permits more flexibility than simple SOLS. This flexibility should encourage researchers to prefer the Eaton and Tamura method to estimate the relationship between GDP components and immigrant flows.

References

Adsera, Alicia, and Mariola Pytlikova. 2012. The role of language in shaping international migration: Evidence from OECD countries 1985–2006. Bonn, Germany: Institute for the Study of Labor IZA Discussion Papers 6333.

Anderson, James E. 2011. The gravity model. *Annual Review of Economics* 3:133–60.

Baldwin, Richard, and James Harrigan. 2011. Zeros, quality, and space: Trade theory and trade evidence. *American Economic Journal: Microeconomics* 3:60–88.

Bartel, Ann P. 1989. Where do the new United States immigrants live? *Journal of Labor Economics* 7:371–91.

Beine, Michel, Frédéric Docquier, and Çağlar Ozden. 2011. Diasporas. *Journal of Development Economics* 95:30–41.

Bertoli, Simone, and Jesus Fernandez-Huertas Moraga. 2011. Multilateral resistance to migration. Madrid, Spain: Fundacion de Estudios de Economia Aplicada (FEDEA) Working Paper 2011-04.

Borger, Scott. 2008. *Estimates of the cyclical Inflow of undocumented migrants to the United States*. Unpublished paper, University of California San Diego.

Borjas, George. 1999. Immigration and welfare magnets. *Journal of Labor Economics* 17:607–37.

Borjas, George. 2001. Does Immigration Grease the Wheels of the Labor Market? *Brookings Papers on Economic Activity* 32(1):69–134.

Budoff Brown, Carrie. 2011. Senators look for immigration deal. *Politico*, 7 February.

Camarota, Steven. 2011. "A Record-Setting Decade of Immigration: 2000 to 2010." Washington, DC: Center for Immigration Studies *Backgrounder*, October 2011. Accessed 15 November 2011. Available http://www.cis.org/articles/2011/record-setting-decade.pdf.

Clark, Ximena, Timothy Hatton, and Jeffrey Williamson. 2007. Explaining U.S. immigration, 1971–1998. *Review of Economics and Statistics* 89:359–73.

Davis, Steve J., and John Haltiwanger. 1992. Gross job creation, gross job destruction, and employment reallocation. *The Quarterly Journal of Economics* 107:819–63.

Dodson, Marvin III. 2001. Welfare generosity and location choices among new United States immigrants. *International Review of Law and Economics* 21:47–67.

Eaton, Jonathan, and Akiko Tamura. 1994. Bilateralism and regionalism in Japanese and U.S. trade and direct foreign investment patterns. *Journal of the Japanese and International Economies* 8:478–510.

Falck, Oliver, Stephan Heblich, Alfred Lameli, and Jens Sudekum. 2012. Dialects, cultural identity, and economic exchange. *Journal of Urban Economics* 72:225–239.

Grogger, Jeffrey, and Gordon H. Hanson. 2011. Income maximization and the selection and sorting of international migrants. *Journal of Development Economics* 95:42–57.

Hainmueller, Jens, and Michael J. Hiscox. 2010. Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment. *American Political Science Review* 104:61–84.

Hanson, Gordon. 2010. International migration and the developing world. In *Handbook of Development Economics*, volume 5, edited by Dani Rodrik and Mark Rosenzweig, pp. 4363–414, Amsterdam: North-Holland.

Hanson, Gordon, and Antonio Spilimbergo. 1999. Illegal immigration, border enforcement, and relative wages: Evidence from apprehensions at the U.S.–Mexico border. *American Economic Review* 89:1337–57.

Head, Keith, and John Ries. 1998. Immigration and trade creation: Econometric evidence from Canada. *Canadian Journal of Economics* 31:47–62.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein. 2008. Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics* 123:441–87.

Hunt, Jennifer. 2006. Staunching emigration from East Germany: Age and the determinants of migration. *Journal of the European Economic Association* 4:1014–37.

Karemera, David, Victor I. Oguledo, and Bobby Davis. 2000. A gravity model analysis of international migration to North America. *Applied Economics* 32:1745–55.

Lewer, Joshua J., Gail Pacheco, and Stephani Rossouw. 2009. Do non-economic quality of life factors drive immigration? 2013. Do non-economic quality of life factors drive immigration? *Social Indicators Research* 110:1–15.

Lewer, Joshua J., and Hendrik Van den Berg. 2008. A gravity model of immigration. *Economics Letters* 99:164–7.

Mandelman, Federico, and Andrei Zlate. 2012. Immigration, remittances and business cycles. *Journal of Monetary Economics* 59:196–213.

Martin, Will, and Cong S. Pham. 2008. Estimating the gravity model when zero trade flows are frequent. Melbourne, Australia: Deakin University, Faculty of Business and Law, School of Accounting, Economics, and Finance Working Paper 2008_03.

Mayda, Anna Maria. 2006. Who is against immigration? A cross-country investigation of individual attitudes toward immigrants. *The Review of Economics and Statistics* 88:510–30.

Mayda, Anna Maria. 2010. International migration: A panel data analysis of the determinants of bilateral flows. *Journal of Population Economics* 23:1249–74.

McFadden, Daniel. 1984. Econometric analysis of qualitative response models. In *Handbook of Econometrics*, volume 2, edited by Z. Griliches and M. Intriligator. Amsterdam: Elsevier/North- Holland.

Orrenius, Pia M., and Madeline Zavodny. 2009. *Tied to the business cycle: How immigrants fare in good and bad economic times*. Washington, DC: Migration Policy Institute.

Ortega, Francesc, and Giovanni Peri. 2009. The causes and effects of international migrations: Evidence from OECD countries 1980–2005. Cambridge, MA: NBER Working Paper No. 14833.

Pedersen, Peder J., Mariola Pytlikova, and Nina Smith. 2008. Selection and network effects—Migration flows Into OECD countries 1990–2000. *European Economic Review* 52:1160–86.

Pesaran, M. Hashem. 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74:967–1012.

Rauch, James E., and Vitor Trindade. 2002. Ethnic Chinese networks in international trade. *The Review of Economics and Statistics* 84:116–30.

Ravenstein, E. G. 1889. The laws of migration. *Journal of the Royal Statistical Society* 52:241–305.

Rosenzweig, Mark. 2007. Education and migration: A global perspective. Unpublished paper, Yale University.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Available from: University of Minnesota, Minneapolis (MN).

Şahin, Ayşegül, Joseph Song, and Bart Hobijn. 2010. The unemployment gender gap during the current recession. *Federal Reserve Bank of New York Current Issues in Economics and Finance* 16.

Santos Silva, J. M. C., and Silvana Tenreyro. 2006. The log of gravity. *The Review of Economics and Statistics* 88:641–58.

Taylor, E. 2006. "International Migration and Economic Development." *International Symposium on International Migration and Development*, Accessed 16 August 2010. Available http://www.un.org/esa/population/migration/turin/Symposium_Turin_files/P09_SYMP_Taylor.pdf.

United Nations. 2004. *World Survey on the Role of Women in Development: Women and International Migration. Economic and Social Affairs*. Available http://www.un.org/womenwatch/daw/public/WorldSurvey2004-Women&Migration.pdf.

Warin, Thierry, and Pavel Svaton. 2008. European migration: Welfare migration or economic migration? *Global Economy Journal* 8:1–30.

Weisman, Jonathan. 2012. Four senators propose easing visa limits for highly skilled. *The New York Times*, 22 May 2012.

Zaiceva, Anzelika, and Klaus F. Zimmermann. 2008. Scale, diversity, and determinants of labour migration in Europe. *Oxford Review of Economic Policy* 24:427–51.

Zavodny, Madeline. 1997. Welfare and the locational choices of new immigrants. *Federal Reserve Bank of Dallas Economic and Financial Policy Review, Second Quarter* 1997:2–10.