4-1-2008

A Gravity Model of Immigration

Joshua Lewer
Bradley University, Peoria, IL, jlewer@bradley.edu

Hendrik Van den Berg
University of Nebraska - Lincoln, hvan-den-berg1@unl.edu

A Gravity Model of Immigration

Joshua J. Lewer [1] and Hendrik Van den Berg [2]

[1] (*Corresponding author*) Department of Economics, Bradley University,
1501 Bradley Avenue, Peoria, IL 61625, USA; tel 309 677-2298, email jlewer@bradley.edu
[2] Department of Economics, University of Nebraska–Lincoln,
Lincoln NE 68588-0489, USA; tel 402 202-6997, email hvan-den-berg1@unl.edu

Abstract
This paper develops a gravity model of immigration. Tests of the model using panel data for 16 OECD countries for 1991–2000 confirm the model's high explanatory power, and examples illustrate its usefulness for testing other hypothesized determinants of immigration.

Keywords: immigration, gravity model, hypothesis testing

1. Introduction

Immigration is a controversial issue in many countries, and economists are increasingly called on to explain its causes and consequences. Most economic studies of immigration, such as Friedberg and Hunt (1995), Card (2001), and Borjas (2003), use a standard labor market model in which immigrant workers respond to differences in wages between countries. Many other factors influence immigration, however. This paper offers an adaptable regression model, based on the popular *gravity model* of international trade, with which to test hypothesized influences on immigration.

The gravity model of trade specifies trade as a positive function of the attractive "mass" of two economies and a negative function of distance between them. Defining TRADE_{ij} as total trade between countries i and j, DIST_{ij} as the distance between the two countries, and the gravitational "mass" as the product of gross domestic products of countries i and j, the gravity model of trade is

$$\text{TRADE}_{ij} = f\left[(\text{GDP}_i \cdot \text{GDP}_j)/\text{DIST}_{ij}\right]. \qquad (1)$$

Showing natural logs in lower case, the regression equation is commonly specified as

$$\text{trade}_{ij} = a_0 + a_1(\text{gdp}_i \cdot \text{gdp}_j) + a_2(\text{dist}_{ij}) + u_{ij}. \qquad (2)$$

Researchers using the gravity model to explain trade often include variables to control for demographic, geographic, ethnic/linguistic, and economic conditions, as for example

$$\begin{aligned}
\text{trade}_{ij} = {} & a_0 + a_1(\text{gdp}_i \cdot \text{gdp}_j) + a_2(\text{pop}_i \cdot \text{pop}_j) \\
& + a_3(\text{dist}_{ij}) + a_4\text{BLOC}_{ij} + a_5\text{LANG}_{ij} \\
& + a_6\text{CONT}_{ij} + a_7\text{LINK}_{ij} + u_{ij}.
\end{aligned} \qquad (3)$$

In (3), BLOC, LANG, CONT, and LINK are dummy variables for pairs of countries that share membership in a free trade area, a common language, a contiguous border, and colonial links, respectively, and $\text{pop}_i \cdot \text{pop}_j$ is the log of the product of the populations.

Tinbergen (1962) first used the gravity model to explain international trade patterns, and economists have consistently found it to explain a large proportion of the variation in international trade flows, making the model attractive for testing the marginal influence of other hypothesized variables on international trade. Theoretical justifications for the model have been provided by Linnemann (1966), Anderson (1979), and Deardorff (1998).

2. A Gravity Model of Immigration

Immigration, like international trade, is driven by the attractive force between immigrant source and destination countries and impeded by the costs of moving from one country to another. The labor market model of immigration suggests that the attractive force between immigrant source and destination countries depends on the difference between labor incomes in the two countries. Population size also matters; *ceteris paribus*, the more people there are in a source country, the more people are likely to migrate, and the larger the population in the destination country, the larger is the labor market for immigrants. Like trade, migration costs are likely to be correlated with the physical distance between countries. These considerations suggest the gravity equation

$$\text{imm}_{ij} = a_0 + a_1(\text{pop}_i \cdot \text{pop}_j) + a_2(\text{rely}_{ij}) + a_3(\text{dist}_{ij}) + u_{ij}, \quad (4)$$

in which imm_{ij} represents the log of immigration to destination country i from source country j, and rely_{ij} is the ratio of destination to source country per capita incomes. The expected signs of the coefficients are $a_1 > 0$, $a_2 > 0$, and $a_3 < 0$.

Researchers will want to control for other influences on immigration. Evidence shows that current immigration is correlated with earlier immigrant flows because the cost of adapting to a new society is mitigated by the presence of compatriots familiar with both the source and destination country cultures. For example, Kahan (1978), Murayama (1991), Rephann and Vencataawmy (2000) find distinctive ethnic concentrations of immigrants in the United States, and Zawodny (1997) finds that family ties overwhelm other factors in determining immigration. Evidence also shows that immigration is larger, *ceteris paribus*, when the language and culture in the destination country is familiar. These considerations suggest the augmented immigration gravity equation

$$\begin{aligned}
\text{imm}_{ij} = {} & a_0 + a_1(\text{pop}_i \cdot \text{pop}_j) + a_2(\text{rely}_{ij}) + a_3(\text{dist}_{ij}) \\
& + a_4(\text{stock}_{ij}) + a_5\text{LANG}_{ij} + a_6\text{CONT}_{ij} \\
& + a_7\text{LINK}_{ij} + u_{ij} \quad (5)
\end{aligned}$$

in which stock_{ij} is the number of source country natives already living in the destination country.

3. Econometric Methodology for the Gravity Model

In the regression model (5), each variable is *bilateral* in that it applies to both countries i and j. However, researchers often want to test the influence on immigration of *unilateral* variables that reflect characteristics in only the source or destination country. Redding and Venables (2004) and Rose and van Wincoop (2001) show that gravity model estimates are likely to be biased by standard error clustering when some variables in the model apply to only one of the two countries in each observation. Feenstra (2004) shows that adding fixed effects to the model eliminates this bias.

A second source of bias is related to the fact that many variables in the gravity equation model (5) are natural logs, which means standard regression methods require omitting observations with zero values. Immigration between pairs of countries may be zero in a substantial percentage of observations, and omitting those zero observations biases the regression results. Fortunately, all observations can be included by applying the scaled ordinary least squares (SOLS) method first applied by Wang and Winters (1992) and Eichengreen and Irwin (1995).

Finally, heterogeneity may plague a gravity model. Cheng and Wall (2005) advise including an *error ranking*

Table 1. Applications of the gravity model

Independent variables	(1) Eq. (3): Gravity model of trade	(2) Eq. (5): Gravity model of immigration	(3) Eq. (6): Testing for institutions' effect on immigration	(4) Eq. (7): Testing effect of source country education on immigration
Constant	− 5.493 (− 13.14)**	4.218 (13.90)**	4.054 (13.19)**	3.914 (11.78)**
$\text{gdp}_i \cdot \text{gdp}_j$	0.691 (51.00)**			
$\text{pop}_i \cdot \text{pop}_j$	0.081 (3.99)**	0.221 (14.48)**	0.222 (14.45)**	0.221 (14.51)**
dist_{ij}	− 0.589 (− 16.53)**	− 0.261 (− 8.79)**	− 0.269 (− 9.01)**	− 0.245 (− 8.01)**
CONT_{ij}	0.346 (3.59)**	− 0.091 (− 1.09)	− 0.106 (− 1.26)	− 0.095 (− 1.13)
LANG_{ij}	0.499 (5.24)**	0.275 (3.34)**	0.308 (3.71)**	0.249 (3.00)**
BLOC_{ij}	0.768 (13.03)**			
LINK_{ij}	0.474 (4.63)**	0.288 (3.21)**	0.272 (3.03)**	0.304 (3.38)**
rely_{ij}		0.00004 (2.31)**	0.00003 (1.82)*	0.00005 (2.48)**
stock_{ij}		0.401 (33.13)**	0.402 (33.22**	0.403 (33.21)**
rlaw_{ij}			0.001 (1.86)*	
property_{ij}			0.131 (2.88)**	
human_i				0.002 (2.24)**
Buse R^2	0.731	0.662	0.663	0.663
Observations	2710	2710	2710	2710

Figures in parentheses are heteroskedasticity-consistent *t*-statistics. **indicates significant at the 95% level and *at the 90% level. The joint hypothesis of the cross-section units having a common intercept is rejected (H_o: $\gamma_2 = \gamma_3 = \ldots = \gamma_{16} = 0$, Fcalc = 8.93 > Fcrit = 1.30).

variable calculated by first running regressions with the data ordered alphabetically by country and then rank ordering the average residual for each country pair.

4. Estimating the Gravity Model

Table 1 reports the estimates of four gravity regressions using panel data on total legal immigration to each of 16 OECD destination countries from all source countries throughout the world for the ten years 1991–2000. In order to compare the gravity models of trade and immigration, column (1) of Table 1 reports the estimated coefficients for the gravity model of trade from Equation (3), and column (2) reports the results from estimating the basic gravity model of immigration specified in Equation (5). The similarities between the two models are obvious. Most variables are highly significant, and the *R*-squares indicate that the model explains much of the variation in the respective dependent variables, trade and immigration. However, geographic contiguity is not significant in the latter regression, suggesting that people move more easily across multiple borders than do goods. The significant immigrant stock coefficient confirms that immigration is indeed path dependent.

Table 1 also reports regression results for two examples of how researchers can use the gravity Equation (5) to test for specific hypothesized influences on immigration. Column (3) shows the results from estimating the augmented immigration gravity equation

$$\text{imm}_{ij} = a_0 + a_1(\text{pop}_i \cdot \text{pop}_j) + a_2(\text{rely}_{ij}) + a_3(\text{dist}_{ij})$$
$$+ a_4(\text{stock}_{ij}) + a_5\text{LANG}_{ij} + a_6\text{CONT}_{ij}$$
$$+ a_7\text{LINK}_{ij} + a_8(\text{rlaw}_{ij}) + a_9(\text{property}_{ij}) + u_{ij} \quad (6)$$

in which rlaw_{ij} and property_{ij} are the logs of the ratios of indexes quantifying how well destination and source countries, respectively, adhere to the rule of law and protect property rights. The positive coefficients suggest that people are more likely to immigrate the greater is the expected improvement in their institutional environment.

Column (4) reports the test of a unilateral source country characteristic, namely the level of human capital (secondary school enrollment) in source countries *j*, in the equation

$$\text{imm}_{ij} = a_0 + a_1(\text{pop}_i \cdot \text{pop}_j) + a_2(\text{rely}_{ij}) + a_3(\text{dist}_{ij})$$
$$+ a_4(\text{stock}_{ij}) + a_5\text{LANG}_{ij} + a_6\text{CONT}_{ij}$$
$$+ a_7\text{LINK}_{ij} + a_8(\text{human}_j) + u_{ij}. \quad (7)$$

As discussed above, a unilateral variable in the regression equation requires adding fixed effects to deal with error clustering bias. The positive coefficient for source country education levels confirms Carrington and Detragiache (1998) and Docquier and Marfouk (2002), who find that more educated people are more likely immigrate. The coefficient, no doubt, also reflects

OECD countries' preference for educated immigrants over uneducated immigrants. This result suggests that improved education in source countries serves to increase the "brain drain."

5. Conclusions

The gravity model of international trade is a useful and popular regression model for testing hypothesized influences on trade flows between pairs of countries. Immigration is likely to respond to gravitational forces and distance in a similar fashion. This paper shows that a gravity model of immigration can be used to test the marginal influence of additional variables on immigration.

Appendix A: Data Sources

Annual data on legal immigration (immigration_{ij}) and the 1985 stock of immigrants (stock_{ij}) are from the OECD's International Migration Database. While the annual flows of immigrants include legal immigration only, the stock of immigrants, based in part on census data, may for some countries include some illegal immigrants as well as legal immigrants. Bilateral trade data (trade_{ij}) are from Direction of Trade Statistics Yearbook, 2002. GDP in U.S. dollars ($\text{gdp}_i \cdot \text{gdp}_j$), population ($\text{pop}_i \cdot \text{pop}_j$), and common trading block (BLOC_{ij}) are from 2002 World Development Indicators. The distance between capital cities (dist_{ij}) is from the U.S. Geological Survey. The dummies for common borders (CONT_{ij}), languages (LANG_{ij}), and colonial histories (LINK_{ij}) are from the CIA World Factbook 2002. Human capital (human_j) is the gross secondary education enrollment ratio from the UNESCO Statistical Yearbook in the source country. Rule of law index (rlaw_{ij}) is from Kaufmann et al. (1999) and property rights (property_{ij}) are from Gwartney and Lawson (2002).

References

ANDERSON, 1979 — J.E. Anderson, A theoretical foundation for the gravity equation, *American Economic Review* 69 (1979), pp. 106–116.

BORJAS, 2003 — G. Borjas, The labor demand curve is downward sloping: reexamining the impact of immigration on the labor market, *Quarterly Journal of Economics* 118 (2003), pp. 1135–1174.

CARD, 2001 — D. Card, Immigrant inflows, native outflows, and the local labor market impacts of higher immigration, *Journal of Labor Economics* 19 (2001), pp. 22–64.

CARRINGTON AND DETRAGIACHE, 1998 — W.J. Carrington and E. Detragiache, How big is the brain drain?, *IMF Working Paper WP/98/102, July* (1998).

CHENG AND WALL, 2005 — I.-H. Cheng and H.J. Wall, *Controlling for Heterogeneity in Gravity Models of Trade and Integration* vol. 87, Federal Reserve Bank of St. Louis Review (2005), pp. 49–63.

DEARDORFF, 1998 — A. Deardorff, Determinants of bilateral trade: does gravity work in a classical world?. In: J.A. Frankel, Editor, *Regionalization of the World Economy*, University of Chicago, Chicago (1998), pp. 7–22.

DOCQUIER AND MARFOUK, 2002 — F. Docquier and A. Marfouk, Measuring the International Mobility of Skilled Workers (1990–2000), *World Bank Working Paper* vol. 3381 (2002).

EICHENGREEN AND IRWIN, 1995 — B. Eichengreen and D.A. Irwin, Trade blocs, currency blocs and reorientation of world trade in the 1930s, *Journal of International Economics* 38 (1995), pp. 1–24.

FEENSTRA, 2004 — R.C. Feenstra, Advanced International Trade: Theory and Evidence, Princeton University Press, Princeton, NJ (2004).

FRIEDBERG AND HUNT, 1995 — R.M. Friedberg and J. Hunt, The impact of immigrants on host country wages, employment and growth, *Journal of Economic Perspectives* 9 (1995), pp. 23–44.

GWARTNEY AND LAWSON, 2002 — J. Gwartney and R. Lawson, Economic Freedom of the World 2002 Annual Report, Fraser Institute, Vancouver (2002).

KAHAN, 1978 — A. Kahan, Economic opportunities and some pilgrims' progress: Jewish immigrants from eastern europe in the U.S., 1890–1914, *Journal of Economic History* 38 (1978), pp. 235–251.

KAUFMANN ET AL., 1999 D. Kaufmann, A. Kraay and P. Zoido-Lobato, Aggregating Governance Indicators, *World Bank Working Paper* vol. 2195 (1999).

LINNEMANN, 1966 — H. Linnemann, An Econometric Study of International Trade flows, North-Holland, Amsterdam (1966).

MURAYAMA, 1991 — Y. Murayama, Information and immigrants: interprefectual differences of Japanese emigration to the pacific northwest, 1880–1915, *Journal of Economic History* 51 (1991), pp. 125–147.

REDDING AND VENABLES, 2004 — S. Redding and A.J. Venables, Economic geography and international inequality, *Journal of International Economics* 62 (2004), pp. 53–82.

REPHANN AND VENCATAAWMY, 2000 — T.J. Rephann and C.P. Vencataawmy, Determinants of the spatial mobility of immigrants in Sweden, *Review of Regional Studies* 10 (2000), pp. 189–213.

ROSE AND VAN WINCOOP, 2001 — A.K. Rose and E. van Wincoop, National money as a barrier to international trade: the real case for currency union, *American Economic Review* 91 (2001), pp. 386–390.

TINBERGEN, 1962 — J. Tinbergen, Shaping the World Economy, Twentieth Century Fund, New York, NY (1962).

WANG AND WINTERS, 1992 — Z.K. Wang and A. Winters, The trading potential of eastern Europe, *Journal of Economic Integration* 7 (1992), pp. 113–131.

ZAWODNY, 1997 — M. Zawodny, *Welfare and the Locational Choice of New Immigrants*, Federal Reserve Bank of Dallas Economic Review (1997), pp. 2–10 Second Quarter.