

Measuring International Skilled Migration: A New Database Controlling for Age of Entry

Michel Beine, Frédéric Docquier, and Hillel Rapoport

Recent data on international migration of skilled workers define skilled migrants by education level without distinguishing whether they acquired their education in the home or the host country. This article uses immigrants' age of entry as a proxy for where they acquired their education. Data on age of entry are available from a subset of receiving countries that together represent 77 percent of total skilled immigration to countries of the Organisation for Economic Co-operation and Development. Using these data and a simple gravity model to estimate the age-of-entry structure of the remaining 23 percent, alternative brain drain measures are proposed that exclude immigrants who arrived before ages 12, 18, and 22. JEL Codes: F22, J24

Recent data sets on international skilled migration (Carrington and Detragiache 1998; Adams 2003; Docquier and Marfouk 2004, 2006; Dumont and Lemaitre 2004) define skilled immigrants as foreign-born workers with university or post-secondary training. This definition, based on the country of birth, does not account for whether education was acquired in the home or the host country and may therefore appear either too inclusive or too exclusive depending on the intended use of the data. For example, the definition would seem too narrow for measuring the extent of a country's skilled diaspora, but may be too inclusive for estimating the fiscal cost of the brain drain for the

Michel Beine is professor of economics at the University of Luxemburg and at Université Libre de Bruxelles and a research fellow at the Center for Economic Studies, Institute for Economic Research at the University of Munich; his email address is mbeine@ulb.ac.be. Frédéric Docquier is a research associate at the Belgian National Fund for Economic Research, professor of economics at the Université Catholique de Louvain (Belgium), and a research fellow at the Institute for the Study of Labor (Bonn) and the Center for Research and Analysis of Migration at University College London; his email address is docquier@ires.ucl.ac.be. Hillel Rapoport (corresponding author) is senior lecturer in economics at Bar-Ilan University, a member of EQUIPPE, Universités de Lille (EA CNRS 4018), and a research fellow at the Center for Research and Analysis of Migration at University College London; his email address is hillel@mail.biu.ac.il. This article is part of the World Bank Migration and Development Program, which provided financial support. The authors are grateful to Riccardo Faini, Mark Rosenzweig, and two anonymous referees for comments. A supplemental appendix to this article is available at <http://wber.oxfordjournals.org>.

THE WORLD BANK ECONOMIC REVIEW, VOL. 21, NO. 2, pp. 249–254
Advance Access Publication 1 June 2007

doi:10.1093/wber/lhm007

© The Author 2007. Published by Oxford University Press on behalf of the International Bank for Reconstruction and Development / THE WORLD BANK. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org

source country, which should consider as skilled emigrants only people who received post-secondary training in their home country.

This article uses immigrants' age of entry as a proxy for where education was acquired. Data on age of entry are available from a subset of receiving countries that together represent 77 percent of total immigration of skilled workers to Organisation for Economic Co-operation and Development (OECD) countries. These data are used in a simple gravity model to estimate the age-of-entry structure of skilled immigration for the remaining 23 percent. These estimates can be used to establish alternative measures of the brain drain for both 1990 and 2000 by defining skilled immigrants as those who left their home country after age 12, 18, or 22. These corrected skilled emigration rates are by construction lower than those computed without age-of-entry restrictions by Docquier and Marfouk (2006).

I. CENSUS DATA ON AGE OF ENTRY

Census and registry data were collected in a sample of OECD host countries for which information on immigrants' age of entry was available: the 1991 and 2001 Australian Censuses, the 1991 Belgian Census, the 1991 and 2001 Canadian Censuses, the 2000 Danish register, the 1999 French Census, the 2001 Greek Census, the 1991 and 2001 New Zealand Censuses, and the 1990 and 2000 U.S. Censuses. In 2000, the sampled countries hosted 77 percent of total skilled immigrants to the OECD. The sample is representative of the OECD in that it includes countries with different demographic sizes, regional locations, development levels, and immigration policy and traditions. In addition to the data on age of entry, the statistical sources provide harmonized bilateral information on migrants' age, education level, and country of birth for a total of 192 origin countries.

An age-of-entry structure was constructed for skilled immigrants aged 25 or older at the time of the census to show what proportion of this population arrived after ages 12, 18, and 22. After zeros and a few suspicious observations were eliminated, 1,580 observations remained for each age threshold (1990 and 2000 included).¹ These observations are used to forecast the age-of-entry structure among the skilled foreign-born for the rest of the OECD area for which such information was not available.

Obviously, an approach based on census data is not perfect. As Rosenzweig (2005, p. 9) explains, "information on entry year... is based on answers to an ambiguous question—in the US Census the question is 'When did you first come to stay?' Immigrants might answer this question by providing the date when they received a permanent immigrant status instead of the date when

1. Table S.1 of the supplemental appendix gives descriptive statistics on the observed proportions of skilled immigrants arrived after age J ($J = 12, 18, \text{ and } 22$). It may be seen that immigrants arriving after ages 12, 18, and 22 represent on average 85.7 percent, 78.2 percent, and 69.1 percent of the total.

they first came to the US, at which time they might not have intended to or been able to stay.”² Only surveys based on comprehensive migration histories can provide precise information about the location in which schooling was acquired. Still, the census is the only harmonized data source available. Survey data are not available for many countries, and when they are (for example, in the EU Labor Force Survey and in the European Community Household Panel), they do not provide representative cross-sectional pictures of immigrants’ characteristics. Their coverage can be very small for countries with few emigrants. And with few exceptions (such as the New Immigrant Survey in the United States) they are not explicitly designed to capture immigrants’ characteristics. Hence, extrapolating the entry age structure from national surveys can be misleading.

II. ESTIMATING THE AGE-OF-ENTRY STRUCTURE OF IMMIGRATION

To estimate the age structure of immigration for receiving countries for which information on age of entry is missing a simple gravity model is used. It aims to identify the determinants of the proportion of migrants from country i to country f with a tertiary education who arrived after age $J = 12, 18$, and 22 . These bilateral proportions are denoted by σ_{if}^J . Since the proportions of skilled migrants who arrived after a given age lie between 0 and 1, it is appropriate to use a transformation so that the dependent variable is defined on $(-\infty, +\infty)$. Therefore, $\theta_{if}^J \equiv \ln[\sigma_{if}^J/(1 - \sigma_{if}^J)]$ is used as the dependent variable. More precisely, the following equation is estimated:

$$(1) \quad \theta_{if}^J = \alpha + \sum_{k=1}^{n_{if}} \beta^{kJ} X_{if}^k + \sum_{k=1}^{n_i} \gamma^{kJ} Z_i^k + \sum_{k=1}^{n_f} \lambda^{kJ} W_f^k + \varepsilon_{if}^J$$

where X_{if}^k ($k = 1, \dots, n_{if}$) is a collection of n_{if} variables capturing proximity between origin and host countries, Z_i^k ($k = 1, \dots, n_i$) are origin country characteristics, and W_f^k ($k = 1, \dots, n_f$) are host country characteristics. These variables can affect the age-of-entry structure through self-selection mechanisms as well as through outselection mechanisms due to differences in host country immigration policies. In addition, a time fixed effect for 2000 is included to account for possible common trends in immigration policies.

Included as the proximity variables in X_{if}^k are indicators of economic disparity between the home and the host countries, indicators of geographic and linguistic distances, and dummy variables for whether the pair of countries

2. In most countries (for example, Australia, France, and New Zealand), immigrants are simply asked about year of arrival or number of years of residence. In Canada, the way the question is asked creates an upward bias: “When did you become a landed immigrant for the first time,” a landed immigrant being a person with the right to permanent residence. As one referee noted, another potential source of bias is the possibility that a person was born in country A, educated in country B, and lives in country C.

share a colonial link. Included as origin country characteristics in Z_i^k are democracy indicators and measures of public expenditures on primary, secondary, and tertiary education. And included as host country characteristics in W_f^k are indicators of social expenditures, education expenditures,³ and degree of openness to immigration. The variables used and their data sources are presented in the supplemental appendix, which also discusses several econometric issues and reports all the estimates.⁴

All coefficients are highly significant for the parsimonious models, robust across specifications, and affect the structure by age of entry in a very intuitive way. The proportion of skilled migrants that arrived after age J increases with economic disparity (as measured by the ratio of host to origin GDP per capita) and geographic distance and decreases with colonial and linguistic links. Education expenditures at destination favor family migration while social expenditures have the opposite effect. The higher the host country immigration rate, the higher the proportion of skilled migrants who arrived as children. Regarding origin country characteristics, the democracy index has no significant effect, and public education expenditures are not significant. Bringing together the census data on age of entry, which represent 77 percent of skilled immigrants to the OECD, and the estimated structure computed using the results of the parsimonious model for the remaining 23 percent⁵ provides alternative measures of the brain drain from which skilled immigrants who arrived before a given age are excluded. These are described in the next section.

III. ALTERNATIVE BRAIN DRAIN ESTIMATES

The Docquier and Marfouk (2006) data set indicates the total number of skilled emigrants from a given origin country i to a given host country f and the number of skilled residents in the home country. Denoting by M_{if} the number of skilled emigrants from country i to country f and by N_i the number of skilled residents in the home country, the skilled emigration rate is then defined as the ratio of skilled emigrants to the total number of skilled natives (residents + emigrants). The method here is to multiply M_{if} by the estimated proportions of skilled migrants who left their home country after age J ($J = 12, 18, 22$) to obtain skilled emigration rates controlled for age of entry/departure. The adjusted skilled emigration rates are then given by:

$$(2) \quad m_i^J = \frac{\sum_f \sigma_{if}^J M_{if}}{N_i + \sum_f \sigma_{if}^J M_{if}}$$

3. See the supplemental appendix for a discussion of possible reverse causality between our dependent variables and some of the explanatory variables used in the regressions.

4. See tables S.2, S.3, and S.4.

5. See column 4 in tables S.2–S.4.

where σ_{if}^J is the observed or the predicted proportion of skilled migrants who left after age J . The Docquier and Marfouk measures correspond to the special case where $J = 0$ or $\sigma_{if}^J = 1$. As σ_{if}^J decreases with J , the corrected rates for $J = 12, 18, 22$ are by construction lower than m_{if}^0 .⁶

For the 192 sending countries in the sample, the m_{if}^{12}/m_{if}^0 ratios range from 74.8 percent to 98.6 percent, the m_{if}^{18}/m_{if}^0 ratios from 59.4 percent to 97.9 percent, and the m_{if}^{22}/m_{if}^0 ratios from 48.5 percent to 95.0 percent. The correlation between corrected and uncorrected measures is extremely high. Simple regression results of m_i^J on m^0 give R^2 values of 0.9775 for $J = 22$, 0.9895 for $J = 18$, and 0.9966 for $J = 12$. Table S.5 of the supplemental appendix focuses on the countries most affected by the brain drain. As table S.5 shows, controlling for age of entry does not significantly alter the rankings by degree of brain drain intensity.

IV. CONCLUDING REMARKS

Recent data sets on international migration of skilled workers define skilled migrants by education level independently of where the education was acquired. This leads to evaluations of the magnitude of the brain drain that may appear too broad or too narrow, depending on the objective for which the data are used. This article uses immigrants' age of entry as a proxy for where education was acquired. It combines observations and estimations and proposes alternative measures of the brain drain that exclude those who left their home country before age 12, 18, or 22. The corrected rates are obviously lower than those calculated without age-of-entry restrictions. However, the correlation between corrected and uncorrected rates is very high, and the country rankings by brain drain intensities are only mildly affected by the correction. This implies that the results from recent empirical studies on the growth effects of the brain drain (for example, Beine, Docquier, and Rapoport 2001, forthcoming) are likely to be robust to the choice of corrected or uncorrected skilled emigration rates.

REFERENCES

- Adams, R. 2003. "International Migration, Remittances and the Brain Drain: A Study of 24 Labor-Exporting Countries." Policy Research Working Paper 2972. World Bank, Washington, D.C.
- Beine, M., F. Docquier, and H. Rapoport. 2001. "Brain Drain and Economic Growth: Theory and Evidence." *Journal of Development Economics* 64(1):275–89.
- . forthcoming. "Brain Drain and Human Capital Formation in LDCs: Winners and Losers." *Economic Journal*.

6. The complete data set can be found at: http://siteresources.worldbank.org/INTRES/Resources/DataSet_BDwith_age_of_entry_DocquierRapoport.xls.

- Carrington, W.J., and E. Detragiache. 1998. "How Big Is the Brain Drain?" Working Paper 98-102. Washington, D.C.: International Monetary Fund.
- Docquier, F., and A. Marfouk. 2004. "Measuring the International Mobility of Skilled Workers (1990–2000)." Policy Research Working Paper 3381. World Bank, Washington, D.C.
- . 2006. "International Migration by Education Attainment in 1990–2000." In C. Ozden, and M. Schiff eds., *International Migration, Remittances, and the Brain Drain*. New York: Palgrave Macmillan.
- Dumont, J.C., and G. Lematre. 2004. "Counting Immigrants and Expatriates in OECD Countries: A New Perspective." Social, Employment, and Migration Working Papers 25. Paris: OECD Directorate for Employment, Labour, and Social Affairs.
- Rosenzweig, M.R. 2005. "Consequences of Migration for Developing Countries." Paper prepared for the United Nations Expert Group Meeting on International Migration and Development, Population Division, July 6–8, New York.