

Informing Migration Policies

A Data Primer

Calogero Carletto

Jennica Larrison

Çaglar Özden



WORLD BANK GROUP

Development Research Group

Poverty and Inequality Team and Trade and Integration Team

November 2014

Abstract

Researchers in many fields, such as demography, economics, and sociology, have established various data collection methodologies and principles to answer a range of academic and policy questions on migration. Although the progress has been impressive, some basic challenges remain. This paper addresses some basic, yet fundamental, questions on identification of international migrants and how their various demographic, personal, and human capital characteristics are captured via different data sources. The

critical issues are the construction of proper sampling frames in censuses, registers, and surveys and the design of questionnaires in household, labor market, and other relevant surveys. The paper discusses how these data sources can be used to answer policy questions in areas such as labor markets, education, or poverty. The focus is on how some of the existing shortcomings in availability, quality, and relevance of migration data can be overcome via improvements in data collection methods.

This paper is a product of the Poverty and Inequality Team and Trade and Integration Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at gcarletto@worldbank.org and cozden@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Informing Migration Policies: A Data Primer¹

Calogero Carletto, Jennica Larrison and Çağlar Özden²

JEL classification: F22, J61, O15,

Keywords: migration, development, survey design, data collection, data dissemination

¹ A revised version of this paper is forthcoming as chapter 2 in *International Handbook of Migration and Economic Development*, edited by Robert E. B. Lucas and published by Edward Elgar Publishing, Inc. in 2015.

² Part of the analysis in the paper has been generously supported by the Knowledge of Change Program (KCP)

I. INTRODUCTION

While establishing his theory of human migration, which forms the basis of much of modern research on migration, Ernst Georg Ravenstein (1885, 1889) relied on various data sources from over 20 countries in Europe and North America to back up his assertions with facts. One of his main conclusions was the importance of high-quality primary data, mainly from national sources, for demographic and geographic research. Since then, researchers in numerous fields such as demography, economics, sociology and political science have established various data collection methodologies and principles via censuses, administrative registers, nationally representative or special-purpose surveys to answer a wide range of academic and policy questions. On one hand, the progress has been quite impressive in terms of the variety and quality of data sources available. On the other hand, some of the basic challenges from the nineteenth century still haunt us with all their vigor.

This paper has two purposes. First, it answers some fundamental questions on who international migrants are and how their numbers as well as various demographic, personal and human capital characteristics are captured via different data sources. The critical issues are the construction of proper sampling frames in censuses, registers, surveys and the design of questionnaires in the relevant surveys. These are among the challenges with which collectors and users of migration data continue to struggle from the days of Ravenstein to date. Second, the paper discusses how these data sources on migration can be used to answer different policy questions in various areas, such as labor markets, education policies or economic welfare. The focus of the discussion will be on

how some of the existing shortcomings in terms of availability, quality and policy relevance of migration data can be overcome via improvements in data collection methods.

II. DATA SOURCES AND CHALLENGES³

Sources of Data

The majority of migration data come from destination countries, as it is often easier to capture people where they currently are rather than where they left. Destination countries use a wide range of tools to count and analyze characteristics of migrants within their boundaries. Among these data sources are (i) censuses that are aimed at capturing all people within borders at a given point, (ii) various surveys, such as labor market or specialized and multi-topic surveys, that sample a smaller portion of the population but ask more detailed questions, (iii) population registers, common in certain countries, and (iv) various administrative data sources such as border statistics, employment and residency permits, as well as naturalization records. Such data sources are used in quantifying migration patterns, especially between country pairs, as well as in identifying demographic, economic, social and cultural characteristics of migrants within a country. In addition, the impact of migration on destination countries' labor markets or various other social and economic outcomes can be assessed using these data in conjunction with other relevant data sources.

1. Some of the sections in this paper draw from Carletto and de Brauw (2008) and de Brauw and Carletto (2012).

‘Easier’ data collection in destination countries in no way means ‘easy’, as collecting information from migrants presents many challenges and shortcomings that prevent researchers from answering many important policy questions. Most of these are concerned with the impact of emigration on the families and communities that migrants have left behind in their home countries. These effects range from the poverty alleviation impact of remittances to the decline in health and education services when doctors and teachers emigrate. While censuses and administrative records in origin countries may provide clues on these issues, most relevant data come from surveys with special migration questions or modules. The main challenge is that questions about migrants need to be answered by a proxy, generally a family member, which introduces many imperfections, as we discuss below.

Censuses

Censuses survey an entire population at a single point in time and are generally conducted decennially, even though there are plenty of exceptions. Everyone needs to be counted and the staff of the relevant National Statistics Office uses their expertise to reach everyone with the same, short questionnaire on mainly demographic variables. The main goal of a census is not to collect data on the migrant population, but on the whole population; the data on migration are generally a ‘by-product’.

An important distinction is between a ‘*de jure*’ census, which aims to count all ‘usually resident’, and a ‘*de facto*’ census, which targets all physically present at the time of census. *De facto* censuses in an origin countries would generally, by definition, fail to

count the emigrants abroad, while *de jure* censuses might capture some recent or temporary emigrants. Thus it is the censuses from destination countries that can be reliably used in migration research.

Universal coverage is the main advantage of a census, and ‘census rounds’ are generally conducted within ten-year periods from the middle of each decade. For example, the 2010 round censuses are conducted between 2005 and 2014. Censuses are expensive; resources and expertise of the national statistical agencies determine the quality of the data and the results. Many statistical agencies just publish cross-tabulations of the main variables of interest as they process the data according to a priority list. Migration data, unfortunately, are among the least crucial for many countries and are published with a significant lag. While census data provide important snapshots of migration over time, they have two key shortcomings. First, they include only basic variables – such as gender, age, place of birth/nationality and maybe education – and thus cannot be used to answer many policy questions. Second, they cannot be used to analyze recent and nuanced trends. However, almost every country conducts a census and the questions are relatively homogeneous.⁴ This degree of standardization and geographical breadth means that census data have become the backbone of most of the available global databases on migration stocks, especially in bilateral corridors.⁵ A key challenge arises when the census tries to capture undocumented migrants who might be reluctant to respond due to concerns that their data will be used for identification and subsequent

4. The United Nations Department of Economic and Social Affairs Statistical Division (UNDESASD) reports that 202 countries have completed a census since 2005, with 26 more countries planning one.

5. Censuses may be used to calculate migration flows by asking questions pertaining to migrants’ residence in previous years and the date of their arrival in the current location.

deportation. As a result, migrant populations may be grossly undercounted in censuses, unless proper corrective procedures are implemented.

Population registers

Population registers are continuous reporting systems used to enumerate the resident population of a particular area, which typically and historically corresponds to a municipality, parish or police precinct. Popular in many parts of Europe – especially the Benelux, Scandinavian and Baltic countries – population registers record the names and addresses of the residents as well as their key demographic variables, such as births, deaths and changes in marital or residential status. A legal requirement typically exists to register and notify any alterations in status. As such, they may be used to record both internal and international migration, and provide detailed up-to-date demographic and socioeconomic information. This makes them a rich, and often fairly exhaustive, source of (migration) data, although the incentive to keep records updated varies across individuals and subgroups of the population. In- and outflows can be tracked with much greater frequency and depth. Although the extent of register data is also rather limited, the continuous updating of the information makes registers more appropriate for tracking migration and for nuanced research questions in comparison with census data.

Population registers, unfortunately, are not widely available outside Europe, and suffer from a higher degree of heterogeneity across countries in terms of registration criteria as well as data scope and quality. The Nordic countries typically implement impressively accurate registers, while the precision of those from Southern Europe is far

worse (Redfern, 1989). Undocumented migrants are, by definition, not captured at all since they would not want to register with local authorities. Another shortcoming is that departures are significantly underreported since many people avoid deregistering in order to retain residency rights, and some of the benefits that come with it (OECD, 2009).

Administrative data

There are various sources of administrative data that may provide detailed information about international migration. Again, most of these sources come from destination countries, as they collect data on the migrants themselves. The most common are residency permits, which give migrants many rights and are implemented across a wide range of countries. Many destination countries regularly publish such data, both on the stock of migrants with permits, as well as the flow of new permits issued within a given time frame. Data may be disaggregated by different qualification criteria, such as family reunification, professional qualifications, humanitarian reasons and lotteries. One has to be careful in interpreting residency permit data since issuances need not equal the number of new immigrants. Permits might be based on issuance but taken up by migrants with an extended delay. Another possibility is that permits are issued to those already in the country as they change their legal status, without a real change in the number of migrants within the country. Most importantly, the legal criteria and rights of permits vary greatly. In some countries, residency permits lead to citizenship, while in others they are granted for limited stays. In cases where free mobility agreements between countries exist, such

as in the European Union, migration might not lead to issue of new permits. Students who are issued long-term visas may or may not be included in the official statistics.

There are other potentially useable administrative data sources, such as border crossings, police records and social security data. There have been several innovative papers that use social security data from European countries to analyze return and circular migration (e.g., Borjas and Bratsberg, 1996), since all employed individuals can be tracked for extended time periods. Ultimately, the quality of the administrative data depends on the responsible national agency in charge of collection and dissemination, and comparability across sources varies by country. But the use of this kind of data is often limited by the fact that they are seldom made public by government officials.

Household surveys

The final and possibly most useful and comprehensive sources of information on migration are surveys that include, among others, Labor Force Surveys (LFS), Demographic and Health Surveys (DHS), Living Standards Measurement Study Surveys (LSMS), as well a variety of purpose-specific surveys. Specialized surveys are also fielded in many countries, often to investigate specific aspects of migration. These micro-level surveys provide a rich source of data and are essential for identifying microeconomic linkages between migration and other facets of households' livelihoods and outcomes that other data sources fail to capture. They are also useful as they better capture undercounted migrants if proper sampling frames and techniques are employed. Finally, such surveys can be conducted in both destination and origin countries, and

enable researchers to explore a wide range of issues. Most importantly, surveys are the most reliable data sources in origin countries.

There are several potential drawbacks of survey data, including the relative variance of the difficulties faced in implementing household surveys in terms of data collection and processing, and their comparatively small sample sizes. These limit their potential for assessing global trends and performing comparisons. Ethnic and other minorities, including migrants, might be underrepresented in the data and stratified sampling can only be meaningfully applied should other nationwide data with proper sampling, such as a census, become available. The variation in surveying practices across countries and their relatively small sample sizes imply that they fail to capture migrants from smaller corridors and are seldom used in the construction of databases of bilateral migration stocks. But surveys remain an indispensable tool for studying migration, its determinants and impacts. Where migration modules have been successfully integrated in household surveys using appropriate sampling techniques, they have often succeeded in capturing sufficiently large numbers of migrants for meaningful analyses. Depending on the resources devoted, use of surveys will and should increase in the future (Center for Global Development, 2009).

Data Challenges

As discussed above, countries use a range of statistical methods and tools – censuses, registers, administrative records and surveys – to document and profile their populations, including migrants. Each source has advantages and disadvantages with respect to

coverage, detail of data collected, frequency and comparability across time and countries. Changes in the questionnaires and data collection methods are two key determinants of comparability over time or within a country. Some issues are more relevant for data collected in the country of destination; others present more of a challenge when collecting data on migrants at the origin. Below we discuss some of the more systematic challenges faced in data collection and analysis.

Definition of a migrant

How a migrant is defined has been a challenge, and Ravenstein (1885) devoted a long discussion to the issue. According to the United Nations Statistics Division (1998, p. 6), a migrant is ‘any person that changes his or her country of usual residence’. The essence is movement from one geographic location to another, which is the most relevant criterion for economic analysis. In practice, however, migration in official statistics manifests itself in myriad guises and is anchored to different concepts and definitions, including the individual’s country of birth, country of citizenship, purpose of visit or visa type, place of last permanent residence, duration of stay, and even ethnicity.

Place of birth and nationality are the two most commonly used definitions. Foreign-born applies to the first generation that crosses the national boundary. Nationality, on the other hand, might apply to dependents and offspring if the destination country’s laws are relatively restrictive in granting citizenship to migrants and/or their children born there. Individuals may be classified as migrants or non-migrants, depending on the definition, even in different sources within a single country. Variation in countries’

adoption of definitions has hampered cross-country comparability of migration data. Some destination countries grant citizenship to foreign-born people who are family members of citizens or who satisfy certain legal and residency requirements. These naturalized citizens continue to be recorded as migrants under the foreign-born definition, but not under the foreign citizen definition. Many countries (for example, the USA) grant automatic citizenship to people born within their territory regardless of the citizenship status of the parents. Yet others, such as Japan, require at least one parent to be a citizen, even if children were born within the country's borders. Because of these differences in citizenship and naturalization laws, the number of migrants will be substantially higher in the USA if the foreign-born criterion is used. In Japan, on the other hand, the number of migrants is higher under the foreign citizenship criterion.⁶

Most bilateral migration databases use the criterion under which census data are collected, which tends to be the country of birth. First, this is more appropriate in analyzing physical movements and handling the cases of former colonies and dependencies. Second, while nationality can change, place of birth cannot. Third, the possibility of holding multiple nationalities complicates matters further. Fourth, naturalization laws and rates vary enormously across destination countries. Differences in laws on citizenship criteria (for both migrants and their children born in the destination country) do not affect data based on place of birth. Finally, when migrants, especially refugees and asylum seekers, cannot be assigned to a specific nationality, they are often recorded under an aggregated umbrella heading, leading to ambiguity.

6. Further confounding interpretation of the underlying definitions is the fact that countries variously apply the foreign-born definition (Docquier and Marfouk, 2006).

All of these issues plague censuses, population registers and administrative data, as well as surveys. Thus it is critical for collectors of data to ask specific questions and users to be aware of the differences. Ideally, both place of birth and citizenship status should be asked about. However, many undocumented migrants will refrain from participation in the survey for fear of identification when faced with citizenship questions, which will bias the data collected. If a choice needs to be made, place of birth is preferable.

Variation in census dates and missing censuses

Despite the attempts by multilateral institutions to achieve coordination, countries choose when to conduct a census or a survey. Even though the standard procedure is to conduct the censuses in the middle of the round, say in 2010, and every ten years, many exceptions exist to both norms. For example, the census dates in France were 1962, 1968, 1975, 1982, 1990, 1999 and 2006. Such large variation can lead to difficulties in comparison across countries, especially in global databases.

The more serious problem is that many countries fail to conduct a survey or a census at a given time or fail to include relevant questions to identify migrants. Censuses are both expensive and demanding in terms of human resources, which make them less attractive activities in poorer countries. Wars, civil conflict or simple public opposition might prevent data collection. There might be political opposition since surveys and censuses collect data on potentially controversial subjects such as ethnicity, language and religion. Even if collected, data might not be released in detail or in a timely manner. Or

worse, the data might be manipulated for political reasons. And, depending on the country's capacity and political sensitivities, adding even a minimal set of questions on migration is often deemed impractical and undesirable, as it may have adverse consequences for the quality of the census data. Researchers must therefore judge the reliability of the data source, as would be the case in other areas, before using it.

Definition of a country

Even if a survey or a census asks country of birth or citizenship questions of the participants, the definitions of countries change over time. Many countries gained independence (Eritrea, Timor-Leste, South Sudan and many other countries in Sub-Saharan Africa), dissolved into smaller states (the Soviet Union, Yugoslavia and Czechoslovakia) or unified (Yemen, Germany, Vietnam) since the Second World War. Shifting national borders creates other challenges. First, with partition, millions of migrants are artificially created without ever moving from their homes. Those born in Moscow, but residing in Kiev, would never have been classified as migrants under either of the two most commonly used definitions until August 23, 1991, but are classified as such in the following censuses. For instance, as Özden et al. (2011) show, the sudden jump in international migrants' numbers between 1990 and 2000 is mainly due to the break-up of the Soviet Union.

Changing borders pose problems when analyzing time-series data. One option is to use the countries in existence at that point in time. Migrants from Africa who came to the USA before the 1970s were recorded under different origin countries in different

censuses as their birth countries gained independence. Other changes are more subtle. The definition of an ‘Ethiopian’ included Eritreans in the 1970 census but not in 2000. Thus there is an artificial decline in Ethiopian migrants, as some have been relabeled as Eritreans in later years. Researchers need to keep these border changes in mind when performing their analysis and make the necessary adjustments.

Collecting migrant data through proxy respondents

Collecting survey data on migrants in destination countries presents a number of challenges. These difficulties center on the absence of a proper sampling frame and the high cost of tracking down individuals who tend to form a small portion of the population or might be present without proper documentation and want to avoid detection.

Nevertheless, most high-income destination countries have sophisticated national data collection mechanisms, especially for surveys for labor-force, expenditure/income or health-related issues, and can overcome these challenges in relatively satisfactory ways. Among the most prominent examples is the micro sample of the US Census and the American Community Survey, where a smaller sample of the American population fills out an extensive questionnaire that covers migration-related questions such as the year of migration, languages spoken at home and ethnicity. When these are used together with information on education, income, occupation and demographic characteristics, researchers can answer questions on labor market performance, poverty, cultural integration and social outcomes. Similarly, Demographic and Health Surveys, Labor Force Surveys or Income and Expenditure Surveys with proper sampling frames and

detailed questions to identify migrants in the data are used extensively to analyze linkages between migration and the applicable issues.

The real challenge in survey design and data collection emerges in origin countries, especially among those that lack adequate statistical capacity. The absence of migrants from the household where data are collected only complicates matters further. Such survey data collection operations require reliance on proxy respondents, in most cases a family member, to answer the questions on behalf of the absent migrants. Various concerns arise from eliciting information through proxy respondents, including whether respondents remember or even know the answers. The information collected through proxies in the household or community of origin can be complemented and cross-checked with short interviews of the migrant themselves via other means, such as phone or online interviews. Alternatively, for certain types of migration, say seasonal or circular migration, proper timing of fieldwork may enable eliciting information at the point of origin directly from the migrant. Finally, interviews at the origin can be administered directly to returnee migrants, often using long recall periods.

Deciding how to identify migrants is the first step in ensuring that desired individuals are captured in an origin-country survey. In order to assess the impact of overall migration patterns, one should try to identify (1) all current household members with past experience with international migration over a given period – return migrants, (2) all former household members who are now living abroad – current migrants, and (3) all former household members with past international migration experience who now live in the source country but in a different household – non-member returnees. As mentioned, collecting information for each group presents different challenges, especially

when combined with the necessity to use a proxy respondent (groups 2 and 3) or not (group 1).

The criterion of household membership, generally defined as all individuals who normally live and eat their meals together, affects whether migrants are captured in a survey. In the case of migration questions, additional restrictions are generally imposed to refine the concept, for example by asking about the number of months individuals have been absent over the previous 12 months. In most standard surveys, if the absence is more than 3 (or 6) months, the person will no longer be considered a household member and thus excluded from data collection past the basic household roster. However, many of these people maintain linkages – such as through remittances – and they should be considered in the survey, especially if the goal is to assess the impact of migration on those left behind. Another question is whether these groups should include all former household members (i.e. any individual who used to live in the household at any point in time) or only members of the nuclear family. Although the latter approach may result in an underestimation of the total number of international migrants, it may be preferable. This method has been applied to internal migration in a nearly nationally representative survey collected in China (de Brauw et al., 2002), and again in the Mexican National Rural Household Survey (Richter and Taylor, 2005).

An alternative method takes advantage of the fact that many household surveys already contain a fertility module, in which information on all children ever born from all female members of reproductive age is collected. A drawback is that it will miss all children of women no longer in the household or who have passed away. A similar approach is to list in a separate module all adult children of the head of the household

and/or his/her spouse regardless of when they left, especially if the mother is absent or no longer alive. These methods were used in various Albania Living Standards Measurement Surveys (INSTAT, 2002, 2005; Carletto and Azzarri, 2007).

In all of these techniques, the critical issue is double-counting of migrants, especially those who can be claimed as members in other households' rosters. The problem of double-counting is even more acute if household rosters are further extended to include any former household member irrespective of their relationship to the household head. Constructing the list based on clearly defined familial relationships, such as for children or siblings, renders the identification and recall of potential migrants simpler and more accurate, and the sample more demographically representative.

Collecting information to assess the impact of migration

Data collection needs to be designed according to objectives. Since migration, by its very definition, is a selective process, any analysis needs to control for the determinants of migration and collect the necessary information for identification of both migrant (treatment) and non-migrant (control) individuals and their households. In addition, data on pre-migration conditions are needed. Assuming a migrant is identified based on departure within a year, the pre-migration timing corresponds to the prior year. For longer reference periods, ideally one would want to collect information for each single year as the factors affecting migration are likely to have changed over time (Bilsborrow et al., 1997).

In terms of impact of migration, one must first decide where the impact occurs – on the migrant while abroad or on return, on the household left behind or on the community. Second, the outcome of interest needs to be collected properly. For example, if the topic of interest is poverty, the survey must collect consumption or income data from the household. Another important issue is the identification strategy, as unobservable factors affecting migration decisions are also likely to be correlated with the outcome of interest. Ideally, one would rely on an experimental design in which the treatment is randomly assigned, and before and after information are collected. However, given the nature of migration, this is hardly ever the case.⁷ Sasin and McKenzie (2007) discuss these issues extensively.

The information collected will depend on a number of factors, including the length of the questionnaire, the capacity and training of fieldworkers, but, most importantly, whether information is gathered directly from the migrant or through a proxy respondent. Use of a proxy might severely constrain the ability to ask questions in depth. However, a minimum set of questions can easily be asked about the emigrants, including their basic demographic characteristics, education level, occupation abroad, country (and location) of current residence, the year of first (and last) migration and remittance behavior. Other questions may be asked about legal status, marital status, the basic demographic composition of the household abroad, frequency and nature of contact with the household, and occupation prior to migrating. Further questions that relate to the specific objectives of the survey can also be added. An attempt could be made to collect

7, An exception is work by McKenzie et al. (2006), who take advantage of the random allocation of New Zealand visas to Tongan residents.

more extensive information on past migration episodes, including timing and country of destination.

Finally, to ensure that household conditions prior to migration are recorded, it is useful to collect other selected information within the recall period. For example, one could collect information on occupations or assets in the household prior to, during and after the migration spell of any household member who had migration experience. Other information, often subjective, might also be worth collecting for individuals who may be return migrants, such as whether or not they plan to leave the household again; and the reasons for return. Some potential reasons for return, such as health of household members, can be corroborated in other parts of the survey. While the design of a questionnaire is important to ensuring that the ‘right’ information is gathered for a study, a survey is only as good as its sample and sampling frame.

Tracking migrants

Another approach to collecting more exhaustive information about emigrants would be to track them. This requires detailed contact information to reach the migrant. Such methods have been used to examine the impacts of internal migration on welfare in Tanzania via a health and development panel data set (Beegle et al., 2011), and in Ethiopia combining a rural household survey with a migrant tracking survey (de Brauw et al., 2013). Given the high mobility of migrants, the tracking survey would ideally occur within weeks of the survey at the sending destination. Internationally, tracking surveys of this type have been carried out in a few countries, including between Mexico and the USA, and between

Albania and Greece. Alternatively, one can first carry out a survey in destination countries and, using a similar approach, track down the original household in the sending countries. An example is the aforementioned study between New Zealand and Tonga (McKenzie et al., 2006).

Tracking surveys can also be used as validation of information that was gathered in the original household through proxy respondents, as well as to measure differences in perceptions between migrants and household members left behind. While allowing for direct interviews with the migrants, tracking presents a number of problems, which might outweigh the benefits. It can be too costly and is characterized by a high level of attrition, particularly when the share of undocumented migration is high.

Sampling design

Within migration surveys, a traditional probability sample based on a multi-stage cluster design will not succeed in finding many migrants, whether this is done in origin or sending countries. A probability sample, by assigning a known non-zero probability of selection to each sampling unit, allows for making inferences to the whole population. The foundation of a proper sample is an updated sampling frame. However, this is the main stumbling block in the design of migration-focused surveys, in origin and destination countries. As mentioned earlier, migrants with certain characteristics or from certain origin countries might not be present in large enough numbers to be properly captured in the sample. Furthermore, certain types of migrants might be reluctant to participate for cultural and legal reasons. The Census Department in the USA, for

example, uses various techniques to adjust for these biases. These include conducting preliminary surveys to determine the extent of undocumented migration in certain areas, and then adjusting their weight in the sample after taking their response likelihood into account. In other cases, using listing from NGOs working with migrants is often the only, even if incomplete, frame from which to draw a sample. Snowballing and other quasi-random techniques are often used, but the resulting samples are seldom representative of the migrant population as a whole.

In origin countries, most available frames do not contain any information on the exposure to past or current migration of the listed households, preventing *ex ante* stratification of the sample based on migration status. As discussed at the beginning of this paper, neither the population census nor available administrative records provide adequate sampling frames for selecting emigrants in a given sending country. Nor, in most cases, do they provide information on previous migration experience to help to identify temporary migrants and returnees.

As such, migration is considered a rare event, defined as an infrequent statistical occurrence. In a normal clustered sample design typical of multi-topic surveys, the expected number of households associated with emigration may be very low. However, a several techniques have been proposed in the literature to better identify rare events, such as migration (Kish, 1965; Huang, 2005). Two such approaches deemed more appropriate in capturing data on migrants, particularly if used in combination, are disproportionate sampling and two-phase sampling. However, both sample designs require some prior knowledge of migration in the population.

Disproportionate sampling implies that primary sampling units (PSUs) with higher migration rates are identified and oversampled. Thus the PSUs known to have a low rate of emigration would be allocated less probability of selection than PSUs with high migration. Representativeness would be regained through weighting. One drawback to this method is that the migration rate might still be too low within each PSU to use simple random sampling or systematic sampling to select households within each PSU.

Alternatively, one can initially select PSUs using the standard method, and then, within each PSU, oversample households known to be migrant households relative to other households. This method is known as two-phase sampling. A random draw of households within a PSU is unlikely to be an efficient way to select a sufficiently large number of migrant households, even in high-migration areas. In this case, a listing operation to clearly identify households with migrants may be a more cost-effective way to select a more balanced sample of migrant and non-migrant households. Listing operations are generally not very expensive and, except in special circumstances, they add up to only about 10–15 percent of the total survey budget (Muñoz, 2007), and the benefits may greatly outweigh the costs.

Finally, one could combine the two methods by initially giving more weight to PSUs with higher migration levels and then oversampling migrant households within each selected PSU. Whichever approach is used, the primary goal is to ensure that a large enough number of migrant households is drawn. It is important to note that using any of these methods is predicated on having prior information about the prevalence of migration in the population at either the area or household level. While this may be the case if one is interested in sampling immigrants in a destination country, it is rarely the

case for the study of emigration in a source country. For methods 1 and 3, one needs information on the relative prevalence of emigration by PSU, and for methods 2 and 3 one needs information about emigration within PSUs.

One further decision to make, assuming that migrants can be properly identified in the frame, is whether to select based on the proportion of households with migrants out of the total number of households or to select based on the proportion of migrants over the population in the reference area. Given that a significant amount of analysis on migration is performed at the household level, the first option may be preferable (Bilsborrow et al., 1997).⁸

For the specific purpose of using surveys to learn about emigrants in sending countries, the lack of a suitable sampling frame would still be an obstacle to implementing a disproportionate sampling design. One possible modification, but a departure from a full probability sample, would be to use alternative data sources to identify high-emigration areas in a country. These sources may include, for example, expert opinions, qualitative surveys or surveys in destination countries where, in addition to the immigrant's country of origin, the specific location of departure is asked. However, the last method is not recommended unless most emigration from the source country has a specific destination and all of these main destinations are covered.

Lacking a proper sampling frame, a less than perfect alternative would be to select all area sampling units (or clusters) at different stages with probability proportional to the estimated size (PPES) of the overall population (or the number of occupied dwelling units) and carry out a full listing operation only in the area sampling unit last selected.

8. For a worked example of a three-stage disproportionate sample of immigrants using a suitable sampling frame, see Bilsborrow et al. (1997, pp. 280–83).

The method would be appropriate only in the unlikely event that the shares of migrants or migrant households were similar across area units, but finding a sufficient number of migrants in the select units might still be a challenge.

Other non-probability sampling techniques may also be used to capture rare events; for example, multiplicity methods such as snowballing have been used in the migration literature. One use of snowballing is to gather information on undocumented migrants, using as a starting point, or ‘seed’, a list of members of a diaspora organization or a list of migrants assisted by an NGO in destination countries. The ‘seed’ household is used to identify additional migrant households of the same country of origin, and so on until the necessary number of observations is reached. However, using data from Senegal, Beauchemin and González-Ferrer (2011) found snowballing to present a number of selection biases, including overrepresentation of migrants with close ties to the sending country.

Finally, techniques such as random walks and aggregation point intercept methods can be used to identify rare events, such as migration. Random walks using selected households in a community may act as a starting point to identify migration occurrences. A Brazilian survey of Nikkei population used the aggregation point intercept method, which, together with snowballing, was compared with more traditional census-based random sampling (McKenzie and Mistiaen, 2009). In all cases, when using these non-probabilistic methods, it is crucial to collect ancillary information on the implementation of the sample to be able to identify the reference population in an attempt to make ‘educated inference’ to a larger population group. This is particularly important given McKenzie and Mistiaen (2009) findings that non-probability methods, such as the

aggregation point intercept, are unlikely to provide representative samples and tend to overestimate the migrant population. However, they also show that reweighing the intercept point estimates to account for visits by the same individual to multiple aggregation points may generate estimates rather close to the census-based method.

III. MIGRATION AND POLICY: DATA REQUIREMENTS AND LIMITATIONS

With an understanding of the issues involved in measuring the migration process, we now turn to the specific policy questions that are often pursued with regard to migration. Our focus here is on examining the data requirements and identifying the data limitations analysts are likely to face when analyzing these issues.

Welfare Impact: Poverty and Income Distribution

The interactions of migration, poverty and income distribution are of primary interest to researchers and policy makers, and have been extensively studied in the past (e.g. Lipton, 1980; Stark et al., 1986; de Haan, 1999). Income differences are among the key determinants of migration between countries and, conversely, migration flows tend to change income levels and distributions in both origin and destination economies. Yet because migration is difficult to identify statistically, few studies have been able to convincingly demonstrate a causal relationship between migration and either poverty or inequality. There are several fundamental challenges.

First, income levels and migration influence each other. Low-income levels are among the main causes of migration. On the other hand, extreme poverty may hinder migration as very poor people lack financial and other resources. For example, in Nicaragua, Murrugarra and Herrera (2011) find that the socioeconomic level of a migrant affects his choice of destination, with poorer migrants choosing destinations that have lower transaction costs, and fewer potential gains. While income levels and poverty may influence migration decisions, migration also leads to poverty reduction and changes in the income distribution, especially in origin countries. In Guatemala, Adams (2005) finds that remittances reduce the degree, depth and severity of poverty in the country. In addition, numerous studies examine expenditure or consumption patterns of households with migrants abroad currently or previously (Taylor and Mora, 2006; Yang, 2008; Adams and Cuecuecha, 2010). As a result of data requirements and the difficulties involved in identifying migration, research on the dynamics of the relationship between migration and inequality or income growth remains relatively limited (McKenzie and Rapoport, 2007; de Brauw and Giles, 2008; Gibson et al., 2009; Lokshin et al., 2010).

The primary requirement to study migration, poverty and income distribution is an accurate measure of well-being: the preferred measure is consumption, which is an integral part of any multi-topic survey. Deaton and Grosh (2000) provide a detailed description of the issues one faces in measuring consumption, and Deaton and Zaidi (2002) discuss issues related to the computation of a consumption aggregate in household survey data in detail. In order to measure the causal relationship between migration and welfare, one would need to estimate what the migrant household's per capita consumption level would have been if the migrant had remained within the household.

Panel data with information on pre- and post-migration can be most useful. However, panel data are seldom available and cross-sectional data sets must be relied on. Although the counterfactual is difficult (if not impossible) to ascertain in a cross-sectional study, one can attempt to learn about it by collecting information on pre-migration conditions, such as measures of asset holdings, that can be reconstructed using recall methods.

The endogenous relationship between migration decisions and income outcomes requires a proper identification strategy. Except where the migrants are chosen randomly, for example in the case of Pacific Islanders in New Zealand (McKenzie et al., 2006), the general approach is instrumental variables (IV) estimation. The challenge of finding valid instruments that influence migration decisions but not the outcomes of interest, especially economic outcomes such as income, is one of the most severe challenges faced by researchers. Household surveys or labor force surveys with migration indicators are the main data sources used for poverty and income distribution analysis. The instruments are also constructed with the data in these surveys. Among the most common variables used in the literature are current and historical social networks and diaspora linkages (Hanson and Woodruff, 2003; Hildebrandt and McKenzie, 2005; De, 2008). In cases where purpose-specific surveys are used or the researcher has the option to design the questionnaire, certain questions can be inserted that will lead to construction of appropriate variables to be used as instruments. For example, McKenzie and Mistiaen (2009) use the generation of the migrant when analyzing Japanese–Brazilian migration to Japan.

Different types of migration may have differing effects on poverty and inequality (see Germenji and Swinnen, 2005). Furthermore, measurements at the household level

must take into account the migrant's absence (Barham and Boucher, 1998). If emigrants are not accounted for in poverty estimates, poverty for some original group of individuals can also be overestimated (Clemens and Pritchett, 2008).

Migration may also have general equilibrium effects on the within-community income distribution. When migration occurs, one expects local wages, either explicitly or implicitly, to rise for the types of workers most likely to migrate, leading to more complex effects on the income distribution than the direct effects on inequality or income. In a cross-section, one could use questions about wages found in the labor modules or labor force surveys to investigate whether communities with more migration have higher increases in wages than other communities, if information on wages in the previous period are also collected. Otherwise, panel data are necessary to investigate general equilibrium effects.

Welfare Impact: Health and Education

The relationship between migration and human capital has also been widely examined, particularly the effects of migration on education and health outcomes of the families left behind (Kanaiaupuni and Donato, 1999; Cox Edwards and Ureta, 2003; Mansuri, 2006; Acosta et al., 2007; Nobles, 2007; Amuedo-Dorantes and Pozo, 2009). These are among the main development benefits of migration in labor-sending lower-income countries. Migration may have positive effects on educational attainment of children in migrant families through the provision of remittances to alleviate credit constraints. Calero et al. (2008) find that remittances increase school enrollment and act to ensure education

through financial shocks in Ecuador. However, parental migration may also have negative consequences if the child must now fill the role of the absent parent by working more at home, or if there is loss of parental involvement. In Mexico, McKenzie and Rapoport (2007) report a decline in school attendance and attainment for boys aged 12–18 and girls aged 16–18 with a parent abroad. Similarly, in China, de Brauw and Mu (2011) find that the status of underweight children aged 7–12 is positively correlated with parental migration.

Type of migration and migrants affect welfare outcomes. Macours and Vakis (2010) find that maternal seasonal migration positively affected early cognitive development in Nicaragua. Gibson et al. (2011) show that children who migrate with parents experience improvements in diet reflected in improved weight-for-age and height-for-age, while children who remain have worsened diets and declining weight-for-age and height-for-age.

In examining health outcomes, a household survey could be combined with anthropometric indicators to examine issues of nutritional status and food security as related to migration (see Antón, 2010; and Carletto et al., 2011). Similarly, for education outcomes, purpose-specific surveys as well as general household surveys with relevant questions can be used. Information specifically on remittance income is also needed.

Last, but not least, migrant selectivity is again a key concern. Health and education outcomes are closely related to income levels, which have an endogenous relationship with migration decisions, as discussed above. Furthermore, in many surveys, migrants state that they choose to migrate specifically to provide the resources for their children's education and family's health expenditures. Thus a natural endogenous

relationship can emerge between migration decisions and health/education outcomes. Since the data come from similar household survey data, parallel IV approaches can be used, as discussed in the previous section on poverty and income.

Labor Market Impact

Migration for employment can have several effects on the economy beyond the direct change in location of some workers. In countries of origin, at the household level, migration may change the household labor composition as well as labor supply decisions of family members who stayed behind. At the macro level, departure of a group of workers, especially if they are concentrated in certain skill, occupation and age categories, will have implications for absolute and relative wage levels as well as other labor market outcomes. These same issues are relevant in destination countries. Labor market implications receive extensive attention in OECD countries, where there is significant public and political opposition to migration flows on the grounds that migrants take jobs away from natives.

Different data sources can be used, depending on the questions asked. Multi-topic surveys are well suited to studying the impact of emigration on wages in specific sectors, labor mobility across sectors, and whether or not skills acquired through migration foster upward mobility among individuals. Aside from migration variables and data collected for potential identification strategies, surveys should include employment status, wage rates, job types and human capital variables in the employment module. For people with migration experience, it is particularly useful to collect information on occupations prior

to migration and at the destination, as well as the present occupation for return migrants (As examples, see Schaffner, 2000; for examples in practice see Carletto and Kilic, 2011; Mendola and Carletto, 2012). A survey that includes oversampling of migrants may be particularly useful for statistical power, especially as the number of job types increases.

There are numerous studies on the impact of emigration on local labor market outcomes in origin countries. In Albania, Miluka et al. (2010) find that international migration reduces overall labor inputs into farming. Similarly, in the Kayes area of Mali, Azam and Gubert (2006) report that individuals left behind by international migrants can exhibit shirking behavior in terms of working in agriculture. Within China, de Brauw et al. (2008) find that, as internal migration increased, overall agricultural labor inputs decreased, but the average share of agricultural labor inputs provided by women did not change. Migration can also affect child labor: on one hand, households with out-migrants may become wealthier, reducing the need for child labor; on the other hand, migration could lead to labor shortages on the farm, potentially increasing the need for child labor. Migration need not only affect the households that migrants leave; at the source, emigration can have general equilibrium effects on wages (Hanson, 2007; Mishra, 2007).

The macro impact of immigration in source and destination countries can be analyzed using aggregate data as well as labor force survey data. The big debate (Borjas, 1988, 1989, 1994 and Card, 2001, 2009) on how migration affects native workers is concerned deeply with properly identifying endogenous location choices. Migrant workers are attracted to geographic areas and sectors where there are strong labor demand and promising future prospects. On the other hand, arrival of migrants is likely to cause native workers to move to other sectors and areas, causing the effect to ripple

through the overall economy. Various authors use different methods to control for these endogenous processes, obtaining quite different results along the way. For example, Card (2001) uses geographic local markets, while Borjas (2003) chooses education-age cohort cells as the unit of analysis. Among the more recent papers, Docquier et al. (forthcoming) use the bilateral migration database of Artuc et al. (forthcoming) to analyze the impact of immigration and emigration on labor market outcomes – both wages and employment levels – of native workers of different skill levels in OECD countries. The analytical framework is typically based on a general equilibrium model of labor markets and production where the substitution elasticities between different types of labor, especially between migrants and natives, play a critical role. Most of these studies rely on data aggregated from labor force surveys and construct their instruments from macro data. Özden and Wagner (2014) use the age-distribution data from origin countries interacted with historical migration patterns as instruments of current migration patterns in Malaysia. Borjas (2014) is an excellent survey of the issues, empirical methods and data requirements in research on labor market implications of migration.

An underexplored but very important issue with development implications is return migration. Migrants obtain skills and financial resources during their time away, which can prove useful on return (Mayr and Peri, 2008; Gibson and McKenzie, 2011). The benefits may especially be concentrated in the labor market outcomes of return migrants (Bijwaard et al. (2013)). A multi-topic survey should cover several important issues in order to study the implications of return migration. In addition to identifying return migrants in the labor force, it is also important to ask whether these individuals plan to migrate again. Although plans are not the same as actual migration after the

survey, individuals who answer negatively can be labeled as truly returned migrants. It is important to ask further about the activities of a returned migrant in order to establish that the individual is truly a return migrant committed to staying.

Governments may want to encourage return migration, as migrants come back with valuable human and social capital acquired abroad. If programs exist to encourage return migration, such as job training or rural development programs, then the questionnaire should explicitly ask about whether migrants took part in any such programs. However, one might also want to know about the likelihood of return for current out-migrants, which is difficult to determine without a survey in the destination country. One can add questions to the survey about the integration of the migrant in the destination country or area. Specifically, one can ask whether or not the spouse lives with the migrant, if the migrant has children who were born or live at the destination, how often the migrant returns home, and whether or not the migrant has investments at home. If children live with migrants away from home, in cases of international migration or countries where multiple languages are spoken, one should also ask whether or not children living at the destination speak the native language.

In addition, migration itself can be affected by changing opportunities at potential destinations. As the populations of developed countries age, many developed-country governments realize that migrants can help fill gaps in their labor market that native workers would not fill, particularly prime-age workers. As migrants usually keep close ties to the households and communities they leave, policies that encourage temporary migration followed by a return can benefit both the source and destination countries.

Brain Drain, Gain and Waste

One of the most intensely debated issues in the migration literature involves the departure of highly skilled workers from poorer countries to wealthier ones. Often referred to as ‘brain drain’, loss of human capital in low-income countries that already suffer from relative skill shortages can have important negative development and growth implications. The main concern is that the positive externalities generated by high-skilled workers will disappear with their migration. The issue takes on an even more emotional dimension since the education of these migrants is funded through limited public resources and their move to high-income OECD countries implies a subsidy from the poor to the rich. Despite the extensive and old analytical literature (Bhagwati, 1979; Bhagwati and Hanson, 2009), empirical analysis has been relatively lacking. The first breakthrough was the construction of the bilateral migration databases for OECD destinations based on census data (Docquier and Marfouk, 2004, 2006; Dumont and Lemaître, 2004). These databases focused on migrant workers in the labor force, aged 15–65 generally, and disaggregated them by three education levels – primary, secondary and tertiary – since this was the most uniformly available human capital variable that could be used as a proxy for human capital and skill levels. Initially for the OECD countries, these data sets highlighted, for the first time, the extent of global skilled migration patterns.

There had been several extensions, starting with the gender dimension (Docquier et al., 2009, and Dumont et al., 2007). A more important issue is the age of migration,

which is a proxy for the location of training. The issue is whether a Jamaican doctor who migrated to the USA as a child with his family, and trained in the USA, can be considered a real ‘loss’ for Jamaica. Beine et al. (2007) construct the database for OECD countries and show that the initial ‘brain-drain’ levels decline considerably for many countries as large numbers of their migrants obtained their tertiary education at the destination and did not necessarily impose a fiscal and productivity loss on the birth countries.

The next extension was the addition of non-OECD destinations. OECD’s DIOC-e database added over 30 destination countries to the original OECD-centric data set, again using census-style sources. Artuc et al. (forthcoming) increase the number of countries to 63 in 1990 and 100 in 2000. Their main innovation is to use a gravity-type estimation model to predict the data – by gender and skill level – for the missing destination countries. The result is a truly global overview of human capital mobility. They estimate that there were slightly over 110 million working-age migrants between the ages of 15 and 64, with around half in OECD countries, half female and a quarter tertiary educated. They establish striking patterns. For example, over 70 percent of all tertiary educated migrants in the world end up in just four countries – the USA, the UK, Canada and Australia. The constructed database enables researchers to perform many analyses, including the paper by Docquier et al. (forthcoming), which evaluates the impact of both immigration and emigration in OECD labor markets.

One of the most interesting ideas to emerge in this literature is the possibility of ‘brain gain’. The main premise is that higher expected returns to human capital in destination countries will induce people to invest more in their human capital. If some of

these people end up staying at home, the final human capital level in the origin country might actually increase due to improved skilled migration opportunities. The important issue is that this likelihood of migration can neither be too high – too many people would leave and nobody would stay at home – nor too low – people would not invest in their education. Using the above-mentioned bilateral skilled migration databases, Beine et al. (2008) show that the brain-gain effect can take place if the tertiary education level in the labor force is between 2 and 8 percent and the skilled migration level is not very high. The brain-gain effect has also been shown to exist using single-country survey data. For example, Batista et al. (2009) find significant brain gain in Cape Verde, specifically through increased education attainment of non-migrants. Similarly, using survey data in Moldova, Böhme (2012) demonstrates that migration positively affects parental aspirations for children's education at the lower end of the human capital spectrum. In their survey of the 'best and brightest' in Tonga and Papua New Guinea, Gibson and McKenzie (2011) show that students who contemplated migration took additional classes. Chand and Clemens (2008) found that ethnic Indians were more likely to invest in their education relative to native Fijians after the 1987 coup, which led to violence against the Indian community and increased their desire to migrate.

The last concept to emerge in this literature is 'brain waste'. The first type of brain waste occurs in an origin country if it overinvests in a specific type of human capital (e.g., engineers) and restricts labor movement. As a result, there will be high unemployment levels for those workers and the resources spent on them will be wasted. The source economy would theoretically benefit by allowing those individuals to

emigrate and realize returns to their human capital elsewhere, as they might be expected to send remittances back to the source economy.

Second, brain waste may occur when highly skilled migrants cannot find employment commensurate with their skills in the destination country. This is sometimes exemplified by Eastern European physicists driving taxis in New York. This might arise for various reasons. Their qualifications might not be recognized, such as foreign doctors who cannot obtain licenses. They might not be trained in location-specific skills, such as lawyers trained in a legal tradition that is different from the US legal regime. Finally, migrants may have the bureaucratic qualifications (the diplomas), but lack the actual human capital since their education is below the standards demanded in the destination labor markets. Mattoo et al. (2008) show that highly skilled immigrants to the USA from Eastern Europe and Latin America tend not to be able to find employment for which they are trained, whereas immigrants from OECD countries and Asia are more likely to find such employment. Their analysis shows that the main reasons are the lower quality of education in the origin countries, as well as selection effects. Since it is easier for Latin American migrants to enter the USA illegally or through family reunification programs, their average human capital level might be lower than that of Asian migrants who are forced to enter through more demanding employment-based preferences.

In addition to census-based global bilateral databases, multi-topic surveys are well suited to study brain drain, brain gain, and brain waste from the source-country perspective. Through an education module, one can characterize the educational background of a country's (or region's) out-migrants relative to the remaining population (Glewwe, 2000). To study brain gain, one can attempt to relate past migration variables

constructed from a migration history to school enrollment. If the survey includes a detailed employment module, brain waste can be studied from the source-country perspective by examining job types held by individuals in different educational and experience cohorts. A survey that oversamples migrant households could be particularly informative in studying either brain drain or brain gain, as the distribution of educational attainment among migrants would be measured with less error. On the other hand, for learning about the brain-drain issue from the perspective of specific occupations, a specialized survey targeted at individuals with specific types of training (e.g. nursing school) would be more informative.

Trade, Foreign Direct Investment and Migration

The main premise of this area of literature is that migration can facilitate integration between countries in other areas, mainly through trade and foreign direct investment (FDI) flows. Migrants possess knowledge of specific conditions, regulations and product features in both import and export markets. By using their knowledge, they can lower transactions costs, match parties in a relationship and increase economic integration. Migrants also are consumers of specific products from their home countries. Starting with Rauch (1999), an extensive literature mentioned these information and demand channels as key to the relationships between trade, migration and FDI flows.

The main analytical framework consists of gravity models and focuses on OECD countries due to data availability. Most papers analyze a single country and its multiple trading partners (such as Head and Ries, 1998; Dunlevy and Hutchinson, 2001; Girma

and Yu, 2002; Blanes-Cristóbal, 2003; Hatzigeorgiou, 2010). The main exceptions are Felbermayr and Toubal (2008), who perform cross-section analysis, and Felbermayr and Jung (2009), who use panel data. Other papers focus on the trade and migration links between province/regions of a country and its trading partners (Wagner et al, 2002; Co et al., 2004; Bardhan and Guhathakurta, 2005; Herander and Saavedra, 2005; Dunlevy, 2006; Bandyopadhyay et al., 2008; Peri and Requena-Silvente, 2010).

The empirical analysis requires the matching of bilateral trade, migration and FDI data. Migration data are generally obtained from macro sources such as censuses, while trade and FDI data come from relevant customs and financial administrations. Single-country studies find a positive relationship, but are likely to suffer from omitted-variable biases as we cannot implement country fixed effects due to the nature of the data, an issue emphasized strongly in the literature on gravity models (Anderson, 1979).

Felbermayr and Jung (2009) control for country-pair unobserved heterogeneity (and multilateral resistances) using a panel for 1990 and 2000, but they can cover only North–South trade. These authors find a significant and positive effect as well. In a more recent paper, Parsons (2012) uses the Özden et al. (2011) database of global bilateral migration for 1960–2000. The long-term panel structure of the data, combined with the fact that they cover all countries, enables Parsons (2012) to control for pairwise heterogeneity as well as to instrument properly for migration patterns by using historical stocks.

The parallel analysis can be extended to linkages between FDI and migration flows. Similar macro sources are used for bilateral migration and FDI data. Among the important examples, Kugler and Rapoport (2005) use a standard gravity model while Javorcik et al. (2010) use the US data with education levels to analyze the role of skilled

migration. In addition, they use geographic variables and passport costs as instruments to control for potential endogeneity problems.

Forced Displacement

The study of forced displacement is often associated with people fleeing from conflict. This may include, but is not limited to, generalized violence, armed conflict or persecution. As defined by the 1951 Convention on Refugees, a refugee is ‘a person who is outside his or her country of nationality or habitual residence; has a well-founded fear of being persecuted because of his or her race, religion, nationality, membership of a particular social group or political opinion; and is unable or unwilling to avail him- or herself of the protection of that country, or to return there, for fear of persecution’ (UNHCR, 2007, p. 3). While many individuals fleeing persecution may seek assistance across international borders either in the form of asylum or through temporary assistance, others may fear being returned and choose to remain anonymous.

Accurately recording refugees in data is extraordinarily difficult. The level of accuracy often depends on the destination of the refugees, the size of the displacement, and whether the refugee chooses to integrate into the destination community. The United Nations High Commissioner for Refugees (UNHCR) maintains a statistical database of refugees across the world. The data are usually reported by local UNHCR offices and reflect host-country data. In the developing world, this information largely incorporates refugees in camps, but may not include those who have integrated into society. In 25 industrialized countries, the data are estimated based on individual-country recognition of

the refugee population over a ten-year period.⁹ Furthermore, in OECD countries, refugees are most often integrated into society and may be registered in the census.

Forced displacement is not limited to those fleeing conflict, but may also include individuals fleeing natural disasters or environmental change. Such migration may be international or domestic, and therefore, while forced, environmental migrants often do not fit the legal definition of a refugee. Individuals fleeing natural disasters, such as Haitians responding to the 2010 earthquake by fleeing to the Dominican Republic, may do so temporarily or on a longer-term basis. Documenting the temporary migration may be done through administrative data, such as information on border crossings, or through targeted surveys at the destination shortly after the natural disaster. Alternatively, longer-term migration may be captured in census data or through multi-topic surveys.

Individuals forced to migrate due to environmental change are unlikely to be able to return. Researchers are devising a number of ways to link multiple types of data in order to assess and anticipate environmental migration. In Ghana, Van der Geest et al. (2010) link census data to vegetation data to examine the connections between internal migration and vegetation dynamics. Feng et al. (2010) examine the connection between crop yields and migration by merging crop data with census data. Others are modeling climate change patterns in an attempt to understand future migration trends (Barbieri et al., 2010; Hunter et al., 2011; McLeman, 2013).

Forced migration may also include displacement due to development policies or projects. These could include the introduction of conservation areas or biosphere projects, or large infrastructure projects, such as the construction of dams or deforestation. Similar

9. See the UNHCR website (<http://www.unhcr.org/4a01417d6.html>) for a list of the 25 industrialized countries and a complete explanation of the database.

to many environmental refugees, those displaced by development projects/policies remain largely within their own country, and do not meet the legal requirements to be considered refugees. Surveys specifically targeting those who may be affected by such policies would be most beneficial. Moreover, in designing surveys to answer questions about all types of forced migration, it is important to include questions regarding where individuals lived previously and why they moved.

IV. CONCLUSIONS

The study of migration and its importance as a driver and consequence of many socioeconomic processes has long been a key component in many countries' development debates. However, the availability of data has been a major constraint hampering the formulation of proper policies in both destination and origin countries. A number of data sources exist, including censuses, various administrative data and surveys, but their thematic and geographic coverage remains inadequate and inconsistent.

Collecting migration data poses a number of technical and logistical challenges requiring special methods and techniques to properly sample and interview migrants. Furthermore, because migration is a highly selective decision, assessing the true impact of migration on the outcomes of choice remains particularly difficult and demanding in terms of data. The objective of this paper was to give readers of this volume a primer on the issues surrounding the collection of

migration data and on the data requirements to address some important questions at the center of national and global policy debates.

Each of the different data sources reviewed has its own strengths and weaknesses, depending of its specific use, but improving the linkages across these different data sources and the various available databases is probably an aspect of the migration data agenda that deserves more attention and would result in substantial benefits to policy analysts. Ultimately, no single data source will be able to meet all data demands; thus countries need to think systemically about how the various sources are integrated and complement each other.

The issue of representativeness, or the lack thereof, of many migration studies also highlights the difficulties associated with drawing suitable samples of migrants at both the origin and destination. Several techniques were highlighted, but the lack of complete and updated sampling frames is bound to continue to hamper future efforts.

Migration is innate to human nature, and both internal and international flows will undoubtedly continue to be a natural part of the policy debates of most countries. Some processes, like increasing rural–urban migration, are inevitable and come with countries' development and households' evolution. Other emerging issues, like climate change, are bound to accelerate these processes. And others, like the aging of population in developed countries, are sure to generate additional demand for unskilled and skilled labor from sending countries. However, how migration is managed and its positive impact fully exploited will ultimately depend on the availability of accurate and timely data on the different facets of the

phenomenon. From Ravenstein's day much has changed in terms of global migration patterns, and our knowledge has steadily increased. Yet some of the same data challenges remain to this day, and this requires a concerted effort among researchers and policy makers to resolve.

REFERENCES

Acosta, P., P. Fajnzylber and J.H. López (2007), 'Remittances and household behavior: evidence for Latin America', in P. Fajnzylber and J.H. López (eds.), *Remittances*

- and Development: Lessons from Latin America*, Washington, DC: The World Bank, pp. 133–70.
- Adams, R.H. (2005), ‘Remittances and poverty in Guatemala’, World Bank Policy Research Working Paper No. 3418, Washington, DC: The World Bank.
- Adams, R.H. and A. Cuecuecha (2010), ‘Remittances, household expenditure and investment in Guatemala’, *World Development*, **38**(11), 1626–41.
- Amuedo-Dorantes, C. and S. Pozo (2009), ‘New evidence on the role of remittances on health care expenditures by Mexican households’, Discussion Paper No. 4617, Bonn, Germany: Institute for the Study of Labor (IZA).
- Anderson, J. E. (1979), ‘A theoretical foundation for the gravity equation’, *The American Economic Review*, **69**(1), 106–16.
- Antón, J. (2010), ‘The impact of remittances on nutritional status of children in Ecuador’, *International Migration Review*, **44**(2), 269–99.
- Artuc, E., F. Docquier, Ç. Özden and Christopher Parsons (forthcoming), ‘A global assessment of human capital mobility’, *World Development*.
- Azam, J.P. and F. Gubert (2006), ‘Migrants’ remittances and the household in Africa: a review of the evidence’, *Journal of African Economies*, **15**(2), 426–62.
- Bandyopadhyay, S., C. Coughlin and H. J. Wall (2008), ‘Ethnic networks and US exports’, *Review of International Economics*, **16**(1), 199–213.
- Barbieri, A.F., E. Doningue, B.L. Queiroz, R.M Ruiz, J.I. Rigotti, J.A.M. Carvalho, and M. F. Resende (2010), ‘Climate change and population migration in Brazil’s Northeast: scenarios for 2025–2050’, *Population Environment*, **31**, 344–70.

- Bardhan, A. D. and S. Guhathakurta (2005), 'Global links of subnational regions: coastal exports and international networks', *Contemporary Economic Policy*, **22**, 225–36.
- Barham, B. and S. Boucher (1998), 'Migration, remittances, and inequality: estimating the net effects of migration on income distribution', *Journal of Development Economics*, **55**, 307–31.
- Batista, C., A. Lacuesta and P. Vicente (2009), 'Micro evidence of the brain gain hypothesis: the case of Cape Verde', Banco de España Working Paper No. 0902, Madrid: Banco de España.
- Beauchemin, C. and A. González-Ferrier (2011), 'Sampling international migrants with origin-based snowballing method: new evidence on biases and limitations', *Demographic Research*, **25**, 103–34.
- Beegle, K., J. de Weerd and S. Dercon (2011), 'Migration and economic mobility in Tanzania: evidence from a tracking survey', *Review of Economics and Statistics*, **93**(3), 1010–33.
- Beine, M., F. Docquier and H. Rapoport (2007), 'Measuring international skilled migration: a new database controlling for age of entry', ULB Institutional Repository 2013/10411, Université Libre de Bruxelles.
- Beine, M., F. Docquier and H. Rapoport (2008), 'Brain drain and human capital formation in developing countries: winners and losers,' *Economic Journal*, **118**(528), 631-52.
- Bhagwati, J. (1979), 'International migration of highly skilled: economics, ethics and taxes', *Third World Quarterly*, **1**(3), 17–30.

- Bhagwati, J. and G. Hanson (2009), *Skilled Immigration Today: Prospects, Problems, and Policies*, New York: Oxford University Press.
- Bijwaard, G. E., C. Schluter and J. Wahba (2013.), 'The impact of labor market dynamics on the return migration of immigrants', *The Review of Economics and Statistics*, online.
- Bilsborrow, R., G. Hugo, A.S. Oberai and H. Zlotnick (1997), *International Migration Statistics: Guidelines for Improving Data Collection Systems*, Geneva: International Labour Organization.
- Blanes-Cristóbal, J.V. (2003), 'The link between immigration and trade in Spain', Seville, Universidad Pablo de Olavide, Paper presented at XXVIII Simposio de Análisis Económico, 11–13 December, Seville.
<http://www.revecap.com/vieea/autores/B/62.pdf>.
- Böhme, M. (2012), 'Migration and education aspirations – another channel of brain gain?', Kiel Working Paper No. 1811, Kiel: Kiel Institute for the World Economy.
- Borjas, G.J. (1988), 'Self-selection and the earnings of immigrants', National Bureau of Economic Research, Inc. Retrieved from <http://ideas.repec.org/p/nbr/nberwo/2248.html>.
- Borjas, G.J. (1989), 'Economic theory and international migration', *International Migration Review*, **23**(3), 457–85.
- Borjas, G.J. (1994), 'The economics of immigration', *Journal of Economic Literature*, **32**(4), 1667–717.
- Borjas, G. J. (2003), 'The labor demand curve is downward sloping:

- reexamining the impact of immigration on the labor market', *Quarterly Journal of Economics*, **118**(4), 1335–74.
- Borjas, G.J. (2014), *Immigration Economics*, Cambridge, MA: Harvard University Press.
- Borjas, G.J. and B. Bratsberg (1996), 'Who leaves? The outmigration of the foreign-born', *The Review of Economics and Statistics*, **78**(1), 165–76.
- Calero, C., A.S. Bedi and R. Sparrow (2008), 'Remittances, liquidity constraints and human capital investments in Ecuador', *World Development*, **37**(6), 1143–54.
- Card, D. (2001), 'Immigrant inflows, native outflows, and the local market impacts of higher immigration', *Journal of Labor Economics*, **19**(1), 22–64.
- Card, D. (2009), 'Immigration and inequality', *American Economic Review*, **99**(2), 1–21.
- Carletto, C. and C. Azzarri (2007), 'Migration in Albania: trends, determinants and impact on poverty', Background Paper to the 2007 Albania Poverty Assessment, mimeo, Washington, DC: The World Bank.
- Carletto, C., K. Covarrubias and J.A. Maluccio (2011), 'Migration and child growth in Guatemala', *Food Policy*, **36**(February), 16–27.
- Carletto, C. and A. de Brauw (2008), 'Measuring migration using household surveys', Migration Operational Vehicle, Operational Note No. 2, Washington, DC: The World Bank.
- Carletto, C. and T. Kilic (2011), 'Moving up the ladder? The impact of migration experience on occupational mobility in Albania', *Journal of Development Studies*, **47**(6), 846–69.
- Center for Global Development (2009), 'Migrants count: five steps toward better

- migration data', Report of the Commission on International Migration Data, Washington, DC: Center for Global Development.
- Chand, S. and M.A. Clemens (2008), 'Skilled emigration and skill creation: a quasi-experiment', International and Development Economics Working Papers idec08-05, International and Development Economics.
- Clemens, M. and L. Pritchett (2008), 'Income per natural: measuring development for people rather than places', *Population and Development Review*, **34**(3), 395–434.
- Co, C. Y., P. Euzent and T. Martin (2004), 'The export effect of immigration into the USA', *Applied Economics*, **36**, 573–83.
- Cox Edwards, A. and M. Ureta (2003), 'International migration, remittances, and schooling: evidence from El Salvador', *Journal of Development Economics*, **72**(2), 429–61.
- De, P.K. (2008) 'Identifying the impacts of international migration on women's contraceptive use in Mexico', mimeo, New York: New York University.
- De Brauw, A. and J. Giles (2008), 'Migrant opportunity and the educational attainment of youth in rural China', Policy Research Working Paper Series 4526, The World Bank.
- De Brauw, A., J. Huang, S. Rozelle, L. Zhang and Y. Zhang (2002), 'The evolution of China's rural labor markets during the reforms', *Journal of Comparative Economics*, **30**(2), 329–53.
- De Brauw, A. and C. Carletto (2012), 'Improving the measurement and policy relevance of migration information in multi-topic household surveys', mimeo.

- http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1199367264546/Migration_Data_v14.pdf.
- De Brauw, A. and R. Mu (2011), 'Migration and the overweight and underweight status of children in rural China', *Food Policy*, **36**(February), 16–27.
- De Brauw, A., V. Mueller and T. Woldehanna (2013), 'Does internal migration improve overall well-being in Ethiopia?', Ethiopia Strategy Support Program II (No. 55). Retrieved from <http://essp.ifpri.info/2013/06/20/does-internal-migration-improve-overall-well-being-in-ethiopia/>.
- De Brauw, A., Q. Li, C. Liu, S. Rozelle, and L. Zhang (2008), 'Feminization of Agriculture in China? Debunking Myths Surrounding Women's Participation in Farming', *China Quarterly*, **194**:327–334.
- De Haan, A. (1999), 'Livelihoods and poverty: the role of migration – a critical review of the migration literature', *Journal of Development Studies*, **36**(2), 1–47.
- Deaton, A. and M. Grosh (2000), 'Consumption', in M. Grosh and P. Glewwe (eds), *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study Volume 1*. Washington, DC: The World Bank, pp. 91–134.
- Deaton, A. and S. Zaidi (2002), *Guidelines for Constructing Consumption Aggregates for Welfare Analysis*, Washington, DC: The World Bank.
- Docquier, F. and A. Marfouk (2004), 'Measuring the international mobility of skilled workers (1990–2000) : release 1.0', Policy Research Working Paper Series 3381, Washington, DC: The World Bank.
- Docquier, F. and A. Marfouk (2006), 'International migration by educational attainment

- (1990–2000)', in Ç. Özden and M. Schiff (eds), *International Migration, Remittances and the Brain Drain*, Basingstoke, UK: Palgrave-Macmillan, ch. 5.
- Docquier, F., B.L. Lowell and A. Marfouk (2009), 'A gendered assessment of highly skilled emigration', *Population and Development Review*, **35**(2), 297–321.
- Docquier, F., Ç. Özden and G. Peri (forthcoming), 'The wage effects of immigration and emigration', *The Economic Journal*.
- Dumont, J.C. and G. Lemaître (2004), 'Counting immigrants and expatriates in OECD countries: a new perspective', mimeo, Paris: OECD.
- Dumont, J.C., J.P. Martin and G. Spielvogel (2007), 'Women on the move: the neglected gender dimension of the brain drain', Discussion Paper Series No. 2920, Institute for the Study of Labor.
- Dunlevy, J. A. (2006), 'The influence of corruption and language on the protrade effect of immigrants: evidence from American states', *Review of Economics and Statistics*, **88**(1), 182–6.
- Dunlevy, J. A. and W. K. Hutchinson (2001), 'The pro-trade effects of immigration on American exports during the period 1870 to 1910', Vanderbilt University Working Paper 01-W25.
- Felbermayr, G. J. and B. Jung (2009), 'The pro-trade effect of the brain drain: sorting out confounding factors', [Economics Letters](#), **104**(2), 72–5.
- Felbermayr, G. and F. Toubal (2008), 'Revisiting the trade–migration nexus: evidence from new OECD data', mimeo.
- Feng, S.F, A.B. Krueger and M. Oppenheimer (2010), 'Linkages among climate

- change, crop yields and Mexico–US cross-border migration’, *Proceedings of the National Academy of Sciences USA*, **107**(32), 14257–62.
- Germenji, E. and J. Swinnen (2005), ‘Human capital, market imperfections, poverty and migration: evidence From Albania’, Licos Discussion Papers, No. 175, Leuven: Centre for Institutions and Economic Performance.
- Gibson, J. and D. McKenzie (2011), ‘The microeconomic determinants of emigration and return migration of the best and brightest: evidence from the Pacific’, *Journal of Development Economics*, **95**(1), 18–29.
- Gibson, J., D. McKenzie and S. Stillman (2009), ‘The impacts of international migration on remaining household members: omnibus results from a migration lottery program’, World Bank Policy Research Working Paper No. 4956, Washington, DC: The World Bank.
- Gibson, J., D. McKenzie and S. Stillman (2011), ‘What happens to diet and child health when migration splits households? Evidence from a migration lottery program’, *Food Policy*, **36**(February), 7–15.
- Girma, So. and Z. Yu (2002), ‘The Link between immigration and trade: evidence from the U.K.’, *Weltwirtschaftliches Archiv*, **138**(1), 115–30.
- Glewwe, P. (2000), ‘Human capital’, in M. Grosh and P. Glewwe (eds), *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study Volume 1*. Washington, DC: World Bank, pp. 143–76.
- Hanson, G. (2007), ‘Emigration, labor supply, and earnings in Mexico’, in G. Borjas (ed.), *Mexican Immigration*, Chicago, IL: University of Chicago Press

- and the National Bureau of Economic Research, pp. 289–328.
- Hanson, G.H. and C. Woodruff (2003), ‘Emigration and educational attainment in Mexico’, mimeo, San Diego, CA: University of California.
- Hatzigeorgiou, A. (2010), ‘Does immigration stimulate foreign trade? Evidence from Sweden’, *Journal of Economic Integration*, **25**(2), 376–402.
- Head, K. and J. Ries (1998), ‘Immigration and trade creation: econometric evidence from Canada’, *Canadian Journal of Economics*, 47–62.
- Herander, M. G. and L. Saavedra (2005), ‘Exports and the structure of immigrant-based networks: the role of geographic proximity’, *Review of Economics and Statistics*, **87**(2), 323–35.
- Hildebrandt, N. and D. McKenzie (2005), ‘The effects of migration on child health in Mexico’, *Economía*, Fall, 257–89.
- Huang, W. (2005), *Methods to Extract Rare Events* (Degree of Doctor of Philosophy). Los Angeles, CA: University of California.
- Hunter, L.M., W. Twine and A. Johnson (2011), ‘Adult mortality and natural resource use in rural South Africa: evidence from the Agincourt Health and Demographic Surveillance Site’, *Society of Natural Resources*, **24**, 256–75.
- INSTAT (2002), Albania Living Standards Measurement Survey, www.worldbank.org/lsms.
- INSTAT (2005), Albania Living Standards Measurement Survey, <http://www.worldbank.org/lsms>.
- Javorcik, B., Ç. Özden, M. Spatareanu and C. Neagu (2010), ‘Migrant networks and foreign direct investment’, *Journal of Development Economics*, **94**, 231–41.

- Kanaiaupuni, S. and K. Donato (1999), 'Migradollars and mortality: the effects of migration on infant survival in Mexico', *Demography*, **36**(3), 339–53.
- Kish, L. (1965), *Survey Sampling*, New York: John Wiley.
- Kugler, M. and H. Rapoport (2005), 'Skilled emigration, business networks and foreign direct investment', CESifo Working Paper Series 1455, ECSifo Group, Munich.
- Lipton, M. (1980), 'Migration from rural areas of poor countries: the impact on rural productivity and income distribution', *World Development*, **8**, 1–24.
- Lokshin, M., M. Bontch-Osmolovski and E. Glinskaya (2010), 'Work-related migration and poverty reduction in Nepal', *Review of Development Economics*, **14**(2), 323–32.
- Macours, K. and R. Vakis (2010), 'Seasonal migration and early childhood development', *World Development*, **38**(6), 857–69.
- Mansuri, G. (2006), 'Migration, school attainment and child labor: evidence from rural Pakistan', mimeo, Washington, DC: The World Bank.
- Mattoo, A., I.C. Neagu and Ç. Özden (2008), 'Brain waste? Educated immigrants in the US labor market', *Journal of Development Economics*, **87**(2), 255–69.
- Mayr, K., and G. Peri (2008), 'Return migration as channel of brain gain', CReAM Discussion Paper Series No. 0804, Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London.
- McKenzie, D., J. Gibson and S. Stillman (2006), 'How important is selection? Experimental versus non-experimental measures of the income gains from migration', World Bank Policy Research Paper No. 3906, Washington, DC: The World Bank.

- McKenzie, D. and J. Mistiaen (2009), ‘Surveying migrant households: a comparison of census-based, snowball and intercept point surveys’, *Journal of the Royal Statistical Society, Series A*, **172**(2), 339–60.
- McKenzie, D. and H. Rapoport (2007), ‘Network effects and the dynamics of migration and inequality: theory and evidence from Mexico’, *Journal of Development Economics*, **84**(1), 1–24.
- McLeman, R. (2013), ‘Developments in modeling of climate change-related migration’, *Climatic Change*, **117**, 599–611.
- Mendola, M. and C. Carletto (2012), ‘International migration and gender differences in the home labour market: evidence from Albania’, *Labour Economics*, **19**(6): 870–80.
- Miluka, J., G. Carletto, B. Davis and A. Zezza (2010), ‘The vanishing farms? The impact of international migration on family farming’, *Journal of Development Studies*, **46**(1), 140–61.
- Mishra, P. (2007), ‘Emigration and wages in source countries: evidence from Mexico’, *Journal of Development Economics*, **82**(1), 180–99.
- Muñoz, J. (2007), ‘Sampling: what you don’t know can hurt you’, presented at the Multi-topic Household Survey course, January, Washington, DC: World Bank.
- Murrugarra, E. and C. Herrera (2011), ‘Migration choices, inequality of opportunities, and poverty reduction in Nicaragua’, in E. Murrugarra, J. Larrison and M. Sasin (eds), *Migration and Poverty: Toward Better Opportunities for the Poor*, Washington, DC: World Bank, pp. 101–24.

- Nobles, J. (2007), 'Parental migration and child health in Mexico', Conference paper presented at the Population Association of America Meetings, New York, March 29–31.
- OECD (2009), *International Migration Outlook*. Paris OECD.
- Özden, Ç. and M. Wagner (2014), 'Immigrant versus natives? Displacement and job creation', World Bank Policy Research Working Paper 6900.
- Özden, Ç., C. Parsons, M. Schiff and T. Walmsley (2011), 'Where on earth is everybody? The evolution of global bilateral migration 1960–2000', *The World Bank Economic Review*, **25**(1), 12–56.
- Parsons, C. R. (2012), 'Do migrants really foster trade? The trade-migration nexus, a panel approach 1960-2000', Policy Research Working Paper Series 6034, The World Bank.
- Peri, G. and F. Riquena-Silvente (2010), 'Rauch The trade creation effect of immigrants: evidence from the remarkable case of Spain', *Canadian Journal of Economics*, **43**(4), 1433-59.
- Rauche J. (1999), 'Networks versus markets in international trade', *Journal of International Economics*, **48**, 7–35.
- Ravenstein, E. G. (1885), 'The laws of migration', *Journal of the Statistical Society*, **46**, 167–235.
- Ravenstein, E. G. (1889), 'The laws of migration: second paper', *Journal of the Royal Statistical Society*, **52**, 241–305.
- Redfern, P. (1989), 'Population registers: some administrative and statistical pros and cons', *Journal of the Royal Statistical Society, Series A* (153), 1–41.

- Richter, S. and J.E. Taylor (2005), 'Policy reforms and the gender dynamics of rural Mexico-to-U.S. migration', Davis, CA: University of California at Davis.
- Sasin, M. and D. McKenzie (2007), 'Migration, poverty and human capital', Migration Operational Vehicle Operation Note No. 1, Washington, DC: World Bank.
- Schaffner, J.A. (2000), 'Employment', in M. Grosh and P. Glewwe (eds), *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study Volume 1*, Washington, DC: World Bank, pp. 217–50.
- Stark, O.J., E. Taylor and S. Yitzhaki (1986), 'Remittances and inequality', *Economic Journal*, **101**, 1163–78.
- Taylor, J.E. and J. Mora (2006), 'Does migration reshape expenditures in rural households? Evidence from Mexico', World Bank Policy Research Working Paper No. 3842.
- United Nations High Commissioner for Refugees (2007), *Convention and Protocol Relating to the Status of Refugees*, Geneva: UNHCR.
- United Nations Statistics Division (1998), *Recommendations on Statistics of International Migration Revision 1*, New York: United Nations.
- Van der Geest, K., A. Vrieling and T. Dietz (2010), 'Migration and environment in Ghana: a cross-district analysis of human mobility and vegetation dynamics', *Environment and Urbanization*, **22**(1), 107–24.
- Wagner, D., K. Head and J. Ries (2002), 'Immigration and the trade of provinces', *Scottish Journal of Political Economy*, **49**(5), 507-25.

Yang, D. (2008), 'International migration, remittances and household investment: evidence from Philippine migrants' exchange rates shocks', *The Economic Journal*, **118**(April), 591–630.