

## **Inter-American Dialogue**

### **Policy Research on Family Remittances to Latin America and the Caribbean**

#### **Note on Data Issues**

**Caglar Ozden  
DECRC  
The World Bank**

As academics, policymakers and development specialists, we are all aware of the increasing importance of migration flows. There are numerous benefits and challenges created by the increased mobility of people – whether they are seasonal farm workers, football players or mere economists - across national borders. Given the social, economic and political issues that arise from increased migration flows, no policy discussion should take place in the absence of reliable and comprehensive data. Unfortunately, even though we are block as away from the Capitol Hill - almost all of the debates are taking behind a veil of ignorance. Let me give an example. I can easily find the number or the price of short-sleeved cotton t-shirts that came from El Salvador through the Miami port during December 2006. But there is no reliable statistic on the number of people from El Salvador in Miami, let alone detailed demographic or social information about them. Thus, the importance of data!!

Let me start with several well-known (almost) facts. The recent years have witnessed an increasing number of international migrants across the world. It is estimated that some 175 million people – 3% of the world’s population – are living in countries in which they were not born. The stock of migrants to high-income countries, for example, increased by an annual growth rate of about 3% from 1980 to 2000, up from the 2.4 percent pace in the 1970s (World Bank, 2005). Similarly, there has been a dramatic increase in migrant remittances to developing countries. Officially recorded remittances, measured as the sum of workers’ remittances, compensation of employees, and migrant transfers, are estimated to have increased from US\$58 billion in 1995 to US\$167 billion in 2005

(World Bank, 2005). These numbers are expected to increase in the following decades. Diverging demographic trends between the North and the South, and rapidly declining transportation and telecommunication costs are making it increasingly difficult to restrain migration flows through government policies. Migration and remittance flows will therefore be among the most important factors shaping our economic, social and cultural profile in the near future.

Let me briefly describe what types of data are required at the aggregate and microeconomic levels and provide the examples of World Bank – Research Group's (DECRG) recent efforts for overcoming some of the data problems.

### ***Aggregate-level Data***

Aggregate-level data on bilateral migration flows are required to analyze the migration-related issues, and especially to perform cross-country comparisons. The recently available dataset of Docquier & Marfouk (2006) presents the most comprehensive and rigorous database to date. Not only it contains the data on bilateral flows of migrants, but it also covers the educational dimension of migrants. This is particularly useful for examining the brain drain issue, one of the most hotly debated issues related to migration. However, the most serious caveat of the Doquier & Marfouk data is that it only covers migration flows to OECD countries. In other words, it cannot provide us any information on migration to, for example, the Gulf countries, one of the main destination regions for migrants from developing countries. Another recently available dataset, the Global Migrant Origin Database constructed by the Development Research Centre on Migration, Globalisation and Poverty at the University of Sussex, in contrast, overcomes this problem.<sup>1</sup> It contains the international bilateral migration stock data that were collected for 226 by 226 countries and dependent territories. Nevertheless, this Sussex database does not contain any data on the education level of migrants.

---

<sup>1</sup> The detailed description of the migration database can be found in Parsons *et al.* (2007).

Among other important dimensions of migrants for migration and remittance-related analyses are the gender of migrants, their age at the time of migrating, where they complete their education, and their occupations in destination countries. The gender dimension is especially important for analyzing the determinants of remittances. The availability of such information allows us to examine, for instance, whether female migrants are likely to remit more to their families, or whether migrants remit less if they are accompanied by their spouse. The age profile and the information regarding where migrants complete their education are useful for assessing the degree of brain drain. If the migrant goes abroad as a child and obtains a college degree in the destination country, then it is not clear whether he/she will constitute the brain drain of his/her origin country. Furthermore, the data on the performance of migrants in the labor market of the destination country are required for analyzing whether we observe any brain waste, another important issue that has also received an increasing attention of policymakers and researchers. Unfortunately, none of these issues is covered by the existing datasets mentioned above. The availability of such dimensions at the aggregate level would certainly advance our understanding of migration and remittances at the global level.

### *Microeconomic-level Data*

In addition to the aggregate-level data, it is important to have information on migration and remittances at the microeconomic, or household/individual, level. In order to examine the socio-economic and welfare impacts of migration and remittances, the availability of such data is essential. The three key issues in migration research are selection, selection, selection. Who are the migrants? Why does one person migrate, but not his friend, neighbour or even brother? What economic, social and personal differences explain how the migrants are selected from the underlying potential migrant population in a sending country? In order to answer these questions and formulate appropriate policy responses to the challenges created by migration, we need household data. Then, we can undertake a behavioral analysis of migration and remittances decisions – e.g. what factors influence them? While the household survey data including

the Living Standards Household Surveys (LSMS) are often available for developing countries, the information on migration and remittances is usually absent. Even when it is included, the available data are not detailed enough to undertake a comprehensive and systematic analysis. In particular, the LSMS datasets almost always do not have information on migrants themselves – e.g. where they are, when they left, how long they have been away, what they do, their age and education level, whether they sent any remittances, how they send money, etc.

In order to fill this gap, DECRG has been making efforts in recent years. For example, a set of detailed migration and remittance-related questions were incorporated in the usual LSMS in Ghana. Moreover, a tailor-made household survey is currently being conducted among the Brazilians of Japanese descent (*Nikkeijin*) in Brazil for a Japan-Brazil corridor study. Let me spend a minute on this unique survey. One of the challenges of household surveys are the costs. (A decent LSMS costs hundreds of thousands of dollars. Thus, ideally, migration modules should be attached to ongoing household surveys.) There are shortcuts one can implement in surveying, but they introduce sample selection biases. One of the non-trivial contributions of this survey will be the assessment of different survey sampling methodologies to collect representative data through household survey questionnaires for applied microeconomic analysis of migration and remittance corridors.<sup>2</sup>

The survey will firstly implement a “random stratified sampling frame” based on the available census data. It will then implement two alternative methodologies, namely “snowball sampling” and “intercept sampling” techniques. The former is a technique where a sampling frame will be built up from a certain number of “seed” *Nikkei* households, i.e., these seed households are asked to provide referrals to other households with the *Nikkeijin* community. In contrast, the latter involves interviews taking place during set time periods at a number of pre-specified well-defined locations that are frequented by *Nikkeijin*. These two alternatives are a non-probability sampling technique.

---

<sup>2</sup> See McKenzie and Mistiaen (2007) for more detailed discussion on the experiment of different sampling methodologies.

March 12, 2007

It is obviously more ideal to obtain data based on a probability sampling. However, it is generally a difficult task to construct a sampling frame for sub-groups of population who migrate, send or receive remittances. The fact that these sub-groups are typically “rare elements” in the population makes this task hard. The findings of the survey will therefore help us assess how representative data the alternative sampling approaches can generate in case the probability-sampling technique is not feasible or too costly.

**References:**

- Docquier, F. and A. Marfouk (2006) “International Migration by Education Attainment, 1990-2000”, in Ozden, C. and M. Schiff (eds) *International Migration, Remittances and the Brain Drain*, Washington, D.C.: World Bank and Palgrave Macmillan.
- McKenzie, D. and J. Mistiaen (2007), “Surveying Migrant Households: A Comparison of Census-Based, Snowball, and Intercept Surveys”, mimeo, Washington, D.C.: World Bank.
- Parsons, C. R., R. Skeldon, T. L. Walmsley, and L. A. Winters (2007) “Quantifying International Migration: A Database of Bilateral Migrant Stocks”, Washington, D.C.: World Bank.