Prospects for the comparative study of international migration using quasi-longitudinal micro-data

Mao-Mei Liu
Mathew J. Creighton
Fernando Riosmena
Pau Baizán

©2016 Liu, Creighton, Riosmena & Baizán.

This open-access work is published under the terms of the Creative Commons Attribution NonCommercial License 2.0 Germany, which permits use, reproduction & distribution in any medium for non-commercial purposes, provided the original author(s) and source are given credit. See http://creativecommons.org/licenses/by-nc/2.0/de/
## Contents

1. Introduction
2. Review of longitudinal and quasi-longitudinal micro-surveys of international migration
3. Introducing the Mexican Migration Project (MMP) and the Migration between Africa and Europe Project – Senegal (MAFE)
4. Background
5. Survey instruments
6. Basic definitions
7. Comparison of key measures of MAFE household and MMP
8. Characteristics of family migration experience in Mexico and Senegal
9. Promises and limitations of comparative research
10. Discussion

References
Prospects for the comparative study of international migration using quasi-longitudinal micro-data

Mao-Mei Liu¹
Mathew J. Creighton²
Fernando Riosmena³
Pau Baizán⁴

Abstract

BACKGROUND
Longitudinal micro-level data on international migration behavior is notoriously difficult to collect, but data collection efforts have become more frequent in recent years. However, comparative research on the patterns and processes of international migration remains quite rare, especially that which compares across regions.

OBJECTIVE
We highlight the promises and difficulties of comparative international migration research, by offering a detailed comparison of two prominent data collection efforts.

METHODS
We systematically review existing sources of longitudinal and quasi-longitudinal individual-level and household-level data on international migration. We then compare two widely used data sources: the Mexican Migration Project (MMP) and the Migration between Africa and Europe project (MAFE).

RESULTS
Data collection efforts are increasingly diverse, yet public accessibility to data remains limited. Also, comparability of data collected across settings can be complicated. In our MMP-MAFE analysis we show some ways in which comparability can be achieved.

¹ Brown University, Providence, Rhode Island, USA. E-Mail: mao-mei_liu@brown.edu.
² University College Dublin, Dublin, Republic of Ireland.
³ University of Colorado, Boulder, Colorado, USA.
⁴ Universitat Pompeu Fabra, Barcelona, Spain.
CONCLUSIONS
A primary roadblock to international comparative research is that, with some exceptions, the public accessibility of data remains low. Even when data is public and surveys are modeled after one another, comparability is not easy due to necessary trade-offs in adapting surveys to local settings and to developments in the field.

CONTRIBUTION
We demonstrate that, despite great strides in collecting quasi-longitudinal data on international migration, limited data accessibility still hinders the study of migration. With regards to comparability, our article provides important lessons for future data collection and analysis efforts that could improve comparability and thus advance understanding of the complex dynamics of international migration.

1. Introduction

Compared to many other life events and transitions of interest in the social sciences, data on migration – especially international migration – is notoriously difficult and costly to collect (Black and Skeldon 2009; Willekens et al. 2016) and use (Beauchemin and Schoumaker 2016; Riosmena 2016). Migrants, by definition, are on the move and thus elusive. Most efforts to collect data either focus on origin or destination. At destination, representative samples of migrants are difficult to collect because migration is a relatively rare occurrence, migrants are difficult to locate (especially those with irregular status), and appropriate and efficient sampling frames are usually unavailable (González-Ferrer and Beauchemin 2011). At origin, surveys rely on return migrants’ experiences or proxy reports on current migrants elicited from kin left behind. Perhaps more problematic is that efforts to sample migrants can suffer ‘left censoring’ when entire households migrate abroad or outside the survey area. Despite shortcomings, origin-based surveys can be useful in situations of circularity when migrants and/or members of their social networks frequently enter the geography of the context of origin.

Despite the difficulties of migration data collection, micro-level longitudinal data remains essential for understanding the drivers of migration, and two main strategies dominate data collection: 1) prospective longitudinal surveys and 2) quasi-longitudinal...
retrospective surveys. The intent of many prospective longitudinal data collection efforts, often referred to as panel surveys, is not migration. Instead, the focus is on the representativeness of a dynamic population, and information is ascertained prospectively by locating and re-interviewing individuals across successive panels over time. However, most fail to follow migrants who, by definition, are more likely to leave the survey area or country. One notable exception is the nationally representative Mexican Family Life Survey (MxFLS), which has invested heavily and been successful in locating and re-interviewing a large majority of respondents who migrated to the United States between waves (Goldman et al. 2014; Rubalcalva et al. 2008). That said, longitudinal approaches are resource-intensive and, in the absence of a large migration flow (e.g., Mexico–United States), their usually short panels might yield a relatively small number of migration transitions to study.

A second, somewhat more common approach is a quasi-longitudinal design, which is also known as a retrospective or life history survey. Rather than engaging in prospective tracking the goal is to collect comprehensive life histories, of which migration is a potential component, during a single interview. This approach is more in line with a standard cross-sectional survey in terms of sampling and thus can be less costly than a prospective longitudinal approach. Although retrospective surveys capture information that is essential for understanding the determinants of migration – information about the time before migration (Bilsborrow et al. 1997) – recall bias can be a problem. That said, evidence suggests that salient life changes like a long-distance move or an international migration are less problematic (Smith and Thomas 2003), and various survey instruments have been developed to increase the accuracy of retrospective data collection (see Beauchemin and Schoumaker 2016). Another concern, which can also be the case for prospective panels, is the absence of households who have entirely migrated out of the study area prior to interview, which can potentially bias estimates of emigration downward.

6 For a comprehensive account of the issues related to the production and use of such data, see Beauchemin and Schoumaker (2016).
7 In order to include a viable number of immigrants, some destination-based panel surveys employ a migrant ‘boost sample’ or oversample – e.g., the German Socio-Economic Panel (Dustmann 2003). Data collection efforts aim to reduce these biases by including a supplemental sample in destinations, or by gathering information on family members (most notably, the children of the household head who do not live in the sampled dwelling at the time of survey, but also siblings and parents of the household head) for the purpose of aiding the indirect estimation of international migration (Zaba 1987).
8 Data collection efforts aim to reduce these biases by including a supplemental sample in destinations, or by gathering information on family members (most notably, the children of the household head who do not live in the sampled dwelling at the time of survey, but also siblings and parents of the household head) for the purpose of aiding the indirect estimation of international migration (Zaba 1987).
The lack of long- and short-panel longitudinal data and the relative paucity of information from retrospective, quasi-longitudinal data gathering efforts hamper our ability to understand many emerging (and older) migration circuits. This results in general theories and patterns of migration being based on locations where data is available, regardless of the distinctive qualities of many heretofore unsampled sending and receiving contexts. As a result, theoretical and empirical understanding of international migration has been largely built on the case of Mexican migration to the United States, which is known to be fairly exceptional in terms of migrants’ sociodemographic profiles, as well as the broader context (Massey and Riosmena 2010; Massey and Sana 2003; Passel 2006; Riosmena 2010). This scarcity of information, along with the aforementioned shortcomings of existing surveys, make comparative research particularly difficult, challenging our ability to test and refine migration theories and understand the broad applicability of immigration policies (also see Riosmena 2016).

Strictly comparative studies are few and far between. Latin American Migration Project (LAMP) and Migration between Africa and Europe (MAFE) are prominent examples, but both focus on one continent of origin and destination. The International Association for the Promotion of Cooperation with Scientists from the Independent States of the Former Soviet Union (INTAS), which surveyed Armenia, Belarus, Georgia, Moldova, and Ukraine, is another example. The Push and Pull (Factors of International Migration) Project, which was completed in the late 1990s, is one of the earlier multi-origin and multi-destination data collection efforts and focused on North-South migration flows to Europe. Technically this effort, with the inclusion of contexts of origin/reception in Europe, Africa, and Asia (i.e., Anatolia), included three continents, but it did not nor was it intended to offer insight into the Americas. None permits comparative cross-continental study of international migration. As a result, it is difficult to understand and test how universal findings for Mexican‒US migration are in other, very different contexts.

In this paper we respond to a recent call for a comprehensive and disciplinary-boundary-crossing approach to international migration research, published in Science (Willekens et al. 2016), and argue that the comparative study of international migration is necessary for understanding migration’s causes and consequences. We outline the prospects of such study by primarily focusing on international migration between less developed and more developed countries9 (UNPD 2013), which represent nearly one-

---

9 We employ the United Nations Population Division’s designations for less-developed and more-developed countries (UNDP 2013), while recognizing the heterogeneity of countries and their diverse trajectories. More-developed countries include all regions of Europe plus Northern America, Australia, New Zealand, and Japan, while other countries are ‘less-developed’.
third (Abel and Sanders 2014) to 53% (Ratha and Shaw 2007) of the international migrant stocks in the world. First, we present an exhaustive (to the best of our knowledge) review of micro-level longitudinal and quasi-longitudinal surveys of international migration and identify key elements. Then, focusing on two – the Mexican Migration Project (MMP) and the Migration between Africa and Europe (MAFE) Household data – we discuss their important characteristics (sampling coverage, measurement of key instruments) and comparability, and finally present some pertinent descriptive statistics of Mexican and Senegalese households and international migration.

2. Review of longitudinal and quasi-longitudinal micro-surveys of international migration

Here we review what we believe to be the most prominent sources of micro-level longitudinal and quasi-longitudinal quantitative data on ‘South-North’ international migration (Table 1). As mentioned above, we restrict the review to international migration between low/medium-income countries and high-income countries. The unit of interest is the individual or a member of the household or family. To understand the determinants of migration behavior, surveys must include information about both migrants and non-migrants at origin. Surveys failing to meet one or more of these characteristics are excluded (e.g., German Socio-Economic Panel, Longitudinal Survey of Immigrants to Australia, Longitudinal Survey of Immigrants to Canada, New Immigrant Survey). Information about migration may be elicited directly from return migrants or through proxy reports by household members. Most surveys are origin-based, but a few are multi-sited at origin and destination (Beauchemin 2014). In Table 1 we identify principal characteristics: project name, years, origin/destination, survey type, sampling strategy, the principal investigator, funding sources, data accessibility,
and sample publications. Then in the following section we present a cross-continental comparison of two data sources of international migration.

Table 1 shows that data collection efforts of quasi-longitudinal and longitudinal micro-data of international migration have grown more frequent and diverse over time. In the 1980s the Philippine Migration Study and the Mexican Migration Project were the only projects of the kind. In the 1990s, six new projects started and/or were completed. In the 2000s, 14 new projects were developed. So far in the 2010s there have been three new projects. At the same time, the geographical range has expanded considerably. To date, at least six different surveys have been utilized to examine Mexican migration to the United States. Other migrations of interest are out-migrations from or within Latin America, Asia (East Asia, Central Asia, South Asia), Europe (Central Europe, South-East Europe), Africa (North Africa, sub-Saharan Africa), and the South Pacific.

While most surveys rely on retrospective information about migration, the Family Life Surveys (FLS) are prospective. The Indonesian FLS is probably the best known (http://www.rand.org/labor/FLS/IFLS.html), while the Mexican Family Life Survey and the Chitwan Valley Family Life Survey appear to be making substantial progress in collecting prospective information on international migration.

Unfortunately, while nearly all projects received significant public financing, only data from 11 of the 25 projects appear to be publicly accessible now. Of these, six data sets cover Mexico, and another covers other Latin American nations (Puerto Rico, Dominican Republic, Nicaragua, Costa Rica, Haiti, Guatemala, El Salvador, Peru, Ecuador, and Colombia, all surveyed at different points), leaving only four accessible data sets of migration from the rest of the world (Push-Pull Project, Albania 2005 Living Standards Measurement Survey, Migration between Africa and Europe, and the Chitwan Valley Family Life Survey). This severely limits the ability of scholars to replicate findings or pursue new comparative research.

13 In correspondence, principal investigators have cited consent (INTAS) or legal (Polish Migration Project) issues as barriers to public data availability.

14 Besides the 11 publicly available data sets, data from at least four other projects are likely to be available through the project PI. Two other projects (BEMS and TEMPER) are still collecting and/or preparing data.
<table>
<thead>
<tr>
<th>Year/s</th>
<th>Project</th>
<th>Origin</th>
<th>Dest.</th>
<th>Survey type</th>
<th>PI</th>
<th>Primary funding source/s</th>
<th>Data public?</th>
<th>Sample publications + website</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980−</td>
<td>Philippine Migration Study</td>
<td>Philippines (Ilocos Norte)</td>
<td>USA (Honolulu)</td>
<td>Household survey (including list of former HH members and immediate family who live outside Barangay) + Individual survey</td>
<td>Fred Arnold, Gordon F. De Jong, James T. Fawcett (East-West Population Institute) + others</td>
<td>National Institute of Child Health and Development (NICHD), Population Center Foundation (Manila)</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>1982−</td>
<td>Mexican Migration Project (MMP)</td>
<td>Mexico (selected communities)</td>
<td>USA</td>
<td>Household survey (of head with retrospective migration info of household members, children of HH head)</td>
<td>Douglas S. Massey (Princeton) and Jorge Durand (Universidad de Guadalajara)</td>
<td>NICHD, Hewlett</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>1996−</td>
<td>Push-Pull Project</td>
<td>Egypt, Ghana, Morocco, Turkey, Italy, Spain (Selected Regions)</td>
<td>Italy, Spain (Selected Regions)</td>
<td>Household survey (answered by HH head + included info of migrants at destination)</td>
<td>Jeanette Schoorl (Netherlands Interdisciplinary Population Studies)</td>
<td>Demographic European Commission</td>
<td>YES</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Review of quasi-longitudinal and longitudinal micro-data sources of international migration
<table>
<thead>
<tr>
<th>Year/s</th>
<th>Project</th>
<th>Origin</th>
<th>Dest.</th>
<th>Survey type</th>
<th>Sampling strategy</th>
<th>PI</th>
<th>Primary funding source/s</th>
<th>Data public?</th>
<th>Sample publications + website</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>United States–Nigeria Migration Study</td>
<td>Nigeria</td>
<td>USA (Chicago)</td>
<td></td>
<td>Random sampling of destination households + linked origin households in Nigeria</td>
<td>Uka Okonkwo (Chicago)</td>
<td>Social Science Research</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Osili (Indiana University)</td>
<td>Council, Ford Foundation</td>
<td></td>
<td>Osili 2007</td>
</tr>
<tr>
<td>1997–2000, 2003</td>
<td>PROGRESA/Opportunidades</td>
<td>Mexico</td>
<td>USA</td>
<td>Census + Household panel, with retrospective migration data on current HH members</td>
<td>Experimental design + Random sampling (ENCEL, follow-up surveys)</td>
<td>PROGRESA (Programa de Educación, Salud y Alimentación)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998–present</td>
<td>Latin American Migration project (LAMP)</td>
<td>Perú</td>
<td>USA</td>
<td>Household survey (of head) with retrospective information</td>
<td>Random sampling of households at origin</td>
<td>Nicholls, Durand (Universidad de Guadalajara)</td>
<td>Nicholls, Massey, Fischer, and Capoferro 2006, Nicholls, Massey and Riosmena 2010, Nicholls and Massey 2005; lamp.opr.princeton.edu</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: (Continued)
<table>
<thead>
<tr>
<th>Year/s</th>
<th>Project</th>
<th>Origin</th>
<th>Dest.</th>
<th>Survey type</th>
<th>Sampling strategy</th>
<th>PI</th>
<th>Primary funding source/s</th>
<th>Data public?</th>
<th>Sample publications + website</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>Paraguay–Argentina Migration Project</td>
<td>Paraguay (Buenos Aires)</td>
<td>Argentina</td>
<td>Household survey (of head) with retrospective info</td>
<td>Random sampling of households at origin + nonrandom sampling of migrants at destination</td>
<td>Emilio Parrado (UPenn) and Marcela Cerruti (CENEP)</td>
<td>MacArthur, Pan American Health Organization contact PI</td>
<td>Parrado and Cerruti 2003</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>IPUMS 10.5% sample of the 2000 Mexican census + International Migration Supplement</td>
<td>USA (97.3% of cases)</td>
<td>Mexico</td>
<td>Census survey + Proxy reports of last trip household member</td>
<td>Stratified cluster design</td>
<td>Mexico Instituto Nacional de Estadística y Geografía</td>
<td>México Instituto Nacional de Estadística y Geografía</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>2002, 2005−2006, 2009−2012</td>
<td>Mexican Family Life Survey</td>
<td>Mexico</td>
<td>USA</td>
<td>Prospective panels of ALL adult members of households + follow-ups with migrants, including those in the United States</td>
<td>Random sampling of households at origin</td>
<td>Graciela M. Teruel (Universidad Iberoamericana) and Luis N. Rubalcava (Centro de Análisis y Medición del Bienestar Social)</td>
<td>NICHD, US National Science Foundation (NSF), Fondo Análisis y Medición del Bienestar Social (Conacyt)</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>2009−</td>
<td>Family Life Project</td>
<td>China (Fujian province)</td>
<td>USA (New York City)</td>
<td>* Household survey (of head) with retrospective info at origin and destination</td>
<td>Random sampling of households at origin + nonrandom sampling of migrants at destination</td>
<td>Zai Liang,</td>
<td>NICHD, Ford Foundation, NSF PI</td>
<td>Liang et al. 2008, Song and Liang 2013, 2014; Liang and Li 2014, albany.edu/cimp/</td>
<td></td>
</tr>
</tbody>
</table>

* Household survey (of head) with retrospective info at origin and destination
*Community-level (village) questionnaire
Table 1: (Continued)

<table>
<thead>
<tr>
<th>Year/s</th>
<th>Project</th>
<th>Origin</th>
<th>Dest.</th>
<th>Survey type</th>
<th>Sampling strategy</th>
<th>PI</th>
<th>Primary funding source/s</th>
<th>Data public?</th>
<th>Sample publications + website</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002−</td>
<td>Gender, Migration and Health among Hispanics study</td>
<td>Mexico (Durham, N.C.)</td>
<td>USA (Durham, N.C.)</td>
<td>Individual survey of randomly selected household member</td>
<td>Targeted random sampling of households at destination + sampling of top sending communities at origin</td>
<td>Emilio Parrado (UPenn)</td>
<td>NIH</td>
<td>contact PI</td>
<td>Flippen and Parrado 2015; Parrado, McQuiston, and Flippen 2005</td>
</tr>
<tr>
<td>2005</td>
<td>Polish Migration Project</td>
<td>Poland (4 communities)</td>
<td>Germany</td>
<td>Household survey (of randomly selected HH member) with retrospective migration histories of household members, close kin and extended kin</td>
<td>Random sampling of households at origin</td>
<td>Deutsche Forschungsgemeinschaft</td>
<td>German Research Foundation</td>
<td>contact PI</td>
<td>Kalter 2010; Massey, Kalter, and Pren 2008</td>
</tr>
<tr>
<td>Years Project</td>
<td>Origin</td>
<td>Dest.</td>
<td>Survey type</td>
<td>Sampling strategy</td>
<td>PI</td>
<td>Primary funding source/s</td>
<td>Data public?</td>
<td>Sample publications + website</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>-------</td>
<td>-------------</td>
<td>-------------------</td>
<td>----</td>
<td>--------------------------</td>
<td>-------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>2005‒2010</td>
<td>Mexico</td>
<td></td>
<td>Rotating panel survey of ALL individuals in households for five consecutive quarters, reason for absence noted</td>
<td>Random sampling of households at origin</td>
<td>Mexico Instituto</td>
<td>Villarreal and Blanchard 2013; inegi.org.mx/est/contenidos/Proyectos/encuestas/hogares/regulares/enoe/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>Georgia</td>
<td></td>
<td>Household survey with information about most recent migration of current and former household members at origin</td>
<td>Random sampling of households at origin</td>
<td>GeoStat (Georgia National statistics office)</td>
<td>Hofmann and Buckley 2013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>Armenia, Belarus, Moldova, Georgia, Ukraine</td>
<td></td>
<td>Household survey, including information of return migrants</td>
<td>Random sampling of households at origin</td>
<td>Nikolai Genov (Free University of Berlin)</td>
<td>European Commission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTAS Project (Ethnosurvey)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years</td>
<td>Project</td>
<td>Origin</td>
<td>Dest.</td>
<td>Survey type</td>
<td>Sampling strategy</td>
<td>PI</td>
<td>Primary funding source/s</td>
<td>Data public?</td>
<td>Sample publications</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------</td>
<td>-------------------------</td>
<td>----------------</td>
<td>-------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>--------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>2007‒2009</td>
<td>Development on the Move</td>
<td>Colombia, Fiji, Georgia, Ghana, Jamaica, Macedonia, Vietnam</td>
<td></td>
<td>Household survey with information about prior migration for current members and current migration for former members</td>
<td>Random sampling of households at origin, nationally representative in all countries and mostly random sampling of migrants at destination</td>
<td>Cris Beauchemin (INED, Paris), Robert Tchaidze and Karine Torosyan (Finland), Min. of Foreign Affairs (Luxembourg), Robert Tchaidze and Karine Torosyan (France)</td>
<td>Agency for Development, Cooper (Norway), Min. of Foreign Affairs (Spain), Dept. for International Dev. (UK).</td>
<td>Yes</td>
<td>Baizan et al. 2012, 2014; Beauchemin et al. 2014; González-Ferrer and Beauchemin 2012; Liu 2013, 2015; Mezger, Kveder and Beauchemin 2015; Toma and Vause 2013, 2014; Vickstrom and González-Ferrer 2016</td>
</tr>
<tr>
<td>Year/s</td>
<td>Project</td>
<td>Origin</td>
<td>Dest.</td>
<td>Survey type</td>
<td>Sampling strategy</td>
<td>PI</td>
<td>Primary funding source/s</td>
<td>Data public?</td>
<td>Sample publications + website</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>-------</td>
<td>-------------</td>
<td>-------------------</td>
<td>----</td>
<td>-------------------------</td>
<td>-------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>2009</td>
<td>Nepali Migrants to the Gulf Study</td>
<td>Nepal (Chitwan Valley)</td>
<td>GCC (Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, UAE) + else-where</td>
<td>* Longitudinal survey of migrants, same questionnaire as Chitwan Valley FS, plus migrant module. *3 interviews (one year apart) for Gulf migrants, 2 interviews for others</td>
<td>Contact family members reported to be living in GCC (and elsewhere abroad) by Chitwan Valley Family Life Study respondents (2009)</td>
<td>Arland Thornton (Univ. of Michigan)</td>
<td>NICHD, Georgetown University Qatar Center for International and Regional Studies</td>
<td>?</td>
<td>Williams, Thornton, and Young-DeMarco 2014</td>
</tr>
<tr>
<td>2013</td>
<td>Moroccan Migration Project</td>
<td>Morocco</td>
<td>Spain (Navarra)</td>
<td>Destination-based?</td>
<td>Katharine Donato (George-town)</td>
<td>US Office of Naval Research</td>
<td>Donato et al. 2016a, 2016b; vanderbilt.edu/ISEEBangladesh/about.php</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-2014</td>
<td>Bangladesh Environment and Migration Survey (BEMS)</td>
<td>Bangladesh (Southwest)</td>
<td></td>
<td>* Household surveys with retrospective migration information</td>
<td>Random sampling of households at origin</td>
<td>Amparo González-Ferrer (CSIC, Madrid)</td>
<td>European Commission in collection temperproject.eu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>Temporary vs. Permanent Migration (TEM-PER)</td>
<td>North Africa (Morocco)</td>
<td>France, Italy, Spain, UK</td>
<td>Origin-based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. Introducing the Mexican Migration Project (MMP) and the Migration between Africa and Europe Project – Senegal (MAFE)\(^\text{15}\)

In order to illustrate some challenges to and the potential of comparative international migration research employing longitudinal/quasi-longitudinal data, we use two publicly available data sources that are relatively comparable, the 2000–2012 surveys of the Mexican Migration Project and the 2008 Senegalese samples of the Migration between Africa and Europe Project. The MAFE project collected two surveys at origin: a household survey and an individual biographical survey (for a summary of how topics overlap between MAFE HH, MAFE BIO, and MMP, see Table 2). Although the MAFE individual survey contains rich amounts of retrospective life history information, the MAFE household survey is more suitable for use in comparison with the MMP, for two primary reasons. Both the MMP and MAFE are household surveys. Both use the household head as the primary reference for other individuals on the household roster (household members and children of head, independent of residence). Future work, however, could also explore using the MMP and MAFE individual surveys. Data and supporting documentation (including questionnaires and sampling criteria) of both the Mexican Migration Project and the Migration between Africa and Europe project are freely available on their respective websites (http://mmp.opr.princeton.edu/ and http://mafeproject.site.ined.fr/en/).

4. Background

Since 1982 the Mexican Migration Project has collected socio-economic information on Mexico–United States migration (MMP 2016). The project initially started in Western Mexico but has since expanded (MMP 2016). In each wave, several communities are chosen using anthropological methods, so that each community includes some migration to the United States (MMP 2016). In every community, several locations (by level of urbanization) are surveyed, and the household ethno-survey is administered to a representative sample of households. To maintain as comparable a period of observation as possible with MAFE, this paper analyzes recent MMP data from 2000–2012. We focus on 12,530 Mexican households in 83 communities.

In 2008 the Migration between Africa and Europe project collected socio-economic information about Senegalese migration to Europe and other countries by

\(^{15}\) MAFE-Senegal investigates Senegalese migration to France, Italy, and Spain, while the larger MAFE project also examines Ghanaian migration to the United Kingdom and The Netherlands and migration from the Democratic Republic of Congo to Belgium and the United Kingdom.
administering household surveys in Senegal and retrospective biographical questionnaires in Senegal and to Senegalese living in France, Italy, and Spain (MAFE 2016; Beauchemin 2012; Beauchemin et al. 2014a). In the greater Dakar region, first-stage sampling was based on the 2002 census and systematically selected census districts with a probability proportional to their estimated population. In each selected district, households were then stratified according to the presence of return migrants and household members abroad, versus households without either current or return migrants. In a final stage, one or several respondents were selected from each household. The stratification in Dakar aimed to obtain sufficiently large samples of households with migrants—a rare group—and thus resulted in their overrepresentation in the samples (Schoumaker and Mezger 2013). The Dakar region is home to about a quarter of the national population, and is the origin of 31% of international migrants reported by Senegalese households in the 2001–2002 ESAM-II survey (Agence Nationale de la Statistique et de la Démographie 2004). In this paper we focus on 1,141 household questionnaires administered in Senegal in 2008 to 458 non-migrant households, 205 households with at least one return migrant, 617 households with at least one current migrant, and 139 households with both return and current migrants. The household response rate was 86.4%.

16 The project was extended in 2009 to include the Democratic Republic of Congo and Ghana, as well as the primary European destinations for migrants: Belgium and the United Kingdom for the Congolese, the Netherlands and the United Kingdom for the Ghanaians. The methodology used in these countries was similar to that employed by the Senegalese study: Indeed, the household and biographical questionnaires were nearly identical in content and format (Beauchemin 2012; Beauchemin et al. 2014a).
5. Survey instruments

Table 2 illustrates the general topics covered by the MAFE and MMP questionnaires.

Table 2: Comparison of main survey topics of the MAFE (household, biographical) and MMP surveys

<table>
<thead>
<tr>
<th>MAFE Household Survey</th>
<th>MMP (ethnosurvey V, 2007–2011)</th>
<th>MAFE Biographical Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household roster information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current Household Roster, including all children of head of household</td>
<td>Retroactive histories of household rosters for each residence where interviewee ever lived; non-resident children of HH head not included</td>
<td></td>
</tr>
<tr>
<td>Basic demographics for household members (e.g., sex, age, relationship to head, etc.)</td>
<td>Relationship to interviewee</td>
<td></td>
</tr>
<tr>
<td>Household/Migrant network migration experience</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household migration out of Senegal</td>
<td>Household migration to United States, Canada, within Mexico</td>
<td>Migration out of Senegal</td>
</tr>
<tr>
<td>Information about 1st and last migration trip of all household members</td>
<td>Information about household head’s parents and siblings’ 1st United States trips of and whether currently in United States</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Count of ever and current migrants in each category: uncles, cousins, nieces/nephews, friends, siblings-in-law, children-in-law, parents-in-law, friends</td>
<td></td>
</tr>
<tr>
<td>Current document status for household migrants</td>
<td>Specific documents of household migrants’ 1st and last trips to United States</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Document history for household head and spouse at each job in United States /Canada</td>
<td></td>
</tr>
<tr>
<td>Migrant contact and transfers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nature of migrant-HH contact</td>
<td>For household head or other migrant, details of remittances to household (frequency, channel, last amount, how spent)</td>
<td>Histories of interviewee’s regular transfers (start/end years, destination country)</td>
</tr>
<tr>
<td>Details of remittances to HH (frequency, channel, last amount, how spent)</td>
<td>For household head or other migrant, details of remittances to household (purpose, average amount)</td>
<td></td>
</tr>
<tr>
<td>Transfer of goods and use</td>
<td>Savings brought to Mexico and use</td>
<td></td>
</tr>
</tbody>
</table>

17 The MAFE household questionnaire also elicited information about three other categories of individual: migrant partners of household members; migrant parents of household members under 18 years old; and other migrant kin with whom the household head or spouse had regular contact over the past 12 months.

18 This information was only collected for selected communities.
6. Basic definitions

Migrant. The MMP collects information about migrations to the United States (some later waves have separate modules for migration to Canada), while MAFE collects information about international migrations in general (a migration is defined as at least one year lived abroad).

Country of origin. In general, the country of origin is Senegal for MAFE-Senegal respondents and Mexico for MMP respondents and usually represents individuals’ country of birth. The MAFE household questionnaire includes information about individuals’ birth country, year of immigration to Senegal (if born abroad), as well as first out-migration from Senegal for all individuals.

Children of Household Head. All children reported by the household head, regardless of place of residence at time of survey.

Household Membership. MAFE and MMP identify household membership slightly differently (see Figure 1a). MAFE identifies HH members as those living in the household for the last six months or who have the intention of living there for at least six months, while MMP identified HH members as those “eating from the same pot”
Both surveys include all children of the household head, whether or not they are co-resident. Also, as seen in Figure 1a, the MAFE household questionnaire includes all migrant spouses of current household members, all migrant parents of minor children living in the household, and other kin with whom the head or spouse have been in regular contact over the previous 12 months. As a result, MAFE data can be used to fit a strict (actual or intended coresidence) or a broader (including migrants) definition of household.

**Timing of migration.** Year and destination of first migration are reported for all household members and children of the household head for first migration to the United States (MMP) or outside Senegal (MAFE). Year of first return to origin is reported for all return migrants.

### 7. Comparison of key measures of MAFE household and MMP

**What individual information is included and for whom?** In Figure 1a, we see that MAFE HH includes a broader group of individuals than the MMP. Both surveys included detailed information about household members, as well as all children of the household head and spouse. In addition, MAFE collected detailed information about different migrants who are linked to the household at the time of survey: partners of any current household member, parents of any minor child who is living in the household, and migrant kin with whom the head or spouse have been in regular contact. While the basic demographic information collected by both surveys is identical (Figure 1b), MAFE HH also captures information about individuals’ ethnicity, nationality, and labor market activity.

**What do we know about household migrations?** Both surveys also collect information about household member’s migrations (Figure 2), including similar basic information about the first migration (year, destination). In addition, the MMP collects legal status, marital status, and occupation for both first and last (or current) migrations. Given the possible heterogeneity of household members and diversity of migration flows, MAFE HH collects basic information about immigration (to Senegal), first out-migration (from Senegal), first return migration (to Senegal), and more detailed information about current migration (legal status, purpose of migration, whether household support received).
Figure 1a: Individuals for whom detailed information is collected

- Migrant partners of HH members
- Migrant parents of minor children in HH
- Migrant relatives of HH head or partner, with whom have had regular contact in past 12 months

MAFE HH

MMP

Figure 1b: Demographic information collected for current household members and children of household head

- Whether partner is migrant
- For minor children, whether parent/s is migrant
- Nationalities
- Ethnicity

MAFE HH

- Sex
- Birth year
- Birth place
- Tie to HH head
- If HH member
- Marital status
- Education
- Occupation
- Year of death (children)

MMP

Notes: MAFE and MMP elicit education information differently. For example, MMP identifies years of education, depending on a scheme that includes adult education. MAFE codes specific types and levels of formal education, excluding Koranic school, basic literacy, and national language school (MAFE household survey). Furthermore, differences between the educational systems of Mexico, DR Congo, Ghana, and Senegal should be considered.

Current activity or job information is collected slightly differently. MMP identifies ‘occupation’ and asks interviewers for ‘specification’. MAFE also identifies ‘occupation’, but asks interviewers to identify ‘socio-professional category’ (intellectual/higher-level wage-earner, skilled employee, unskilled employee, employer, self-employed, apprentice, family help). MAFE also identifies whether individuals are unemployed, students, homemakers, retired, or otherwise inactive.
Figure 2: Household members’ migration histories

- Current trip (start year)
- Reason for current trip
- Current legal status
- HH economic or other support for migration
- Nature of HH contact in last 12 months, (freq., visits)
- Remittances (freq, channel, quantity, how spent)
- Transfer of goods

- 1st trip (yr)
- 1st destination (country)
- 1st return (yr)
- Current trip (country)

MAFE HH

MMP

Notes: MAFE identifies migration as at least one year abroad in any country outside country of origin (DR Congo, Ghana, and Senegal), while MMP identifies migration trips to the United States as those involving “work, an active job search, or a reasonably stable residency” (MMP 2012 Interviewer’s Manual: 13). Year of 1st return can be gleaned from MAFE’s question about the 1st return – at least one year at country of origin – or calculated from MMP’s question about duration of the 1st trip. Legal status is also captured differently. MAFE offers “yes/no/does not need/don’t know” answers to whether the individual currently has “the residence permits/official documents that would allow him/her to stay in the county where he/she is”, while the MMP offers a whole range of documents and “undocumented”. Whether the individual holds destination citizenship can be calculated from MAFE and is available from the MMP household survey.

What do we know about migrant networks? Both the MMP and MAFE collect information about migrant networks beyond migrant spouses or migrant children of the household head. The MMP collects more limited information for a larger group of individuals and MAFE collects more comprehensive information for a more select group of network members. Figure 3a shows the network members reported by each survey. Specifically, the MMP has a broad migrant network roster (details of any parent or sibling of HH head migration, and summary information of other categories), while the MAFE household survey’s network information is limited to relatives of the household head and spouse who are currently abroad and with whom the household has been in regular contact over the 12 months previous to the survey. All other individuals are selected by the quality of their current relationship to the household. As a result, the MMP includes return migrants in select kinship categories, while MAFE HH includes no return migrants who live outside the household. Figure 3b displays the information collected about migrant networks.
Figure 3a: Network members

- Any relative of HH head or spouse who is currently living abroad and who has been in regular contact in past 12 months
- Parent of HH head
- Sibling of HH head
- Counts of current and previous migrants: uncles/aunts, cousins, nieces/nephews, siblings-in-law, children-in-law, parent-in-law, friends

MAFE HH  MMP

Figure 3b: Information about (non-spouse, non-children) network members

- Year started current trip
- Reason for current trip
- Current legal status
- Whether received HH support for migration
- Whether in regular contact with HH in past 12 mos. (freq, visits)
- Remittances received, (freq, channels, quantity, how spent)
- Transfer of goods
- Sex
- Birth year
- Year at 1st Trip
- 1st destination country
- Year of death (if applicable)
- Current residence (country)
- Legal status (last trip)
- Year received green card or citizenship (if applicable)
- Current residence

MAFE HH  MMP
8. Characteristics of family migration experience in Mexico and Senegal

Table 3 displays the descriptive statistics for individuals and households in Mexico and Senegal. Mexican results are also stratified by type of region: ranchos, small towns, large towns, and metropolitan areas. For Senegal, the region of Dakar, although mostly metropolitan, also includes several small towns and rural areas. Overall, Senegalese households are larger than Mexican households: Average household sizes are 7.9 in Senegal and 5.9 in Mexico. This reflects Senegal’s higher fertility, with a total fertility rate (TFR) of 5.3 in 2005, while Mexico’s TFR is estimated at 2.4 births per woman (Agence Nationale de la Statistique et de la Démographie 2013; United Nations 2011). The complexity of household structures in Senegal is reflected in its larger households, where several family nuclei are commonly found living together. According to the 2002 Senegal Census, a quarter of all marriages are polygamous unions. Most of the population belongs to ethnic groups where, upon marriage, the wife usually moves to the home of her husband’s family and shares the household chores and caring tasks with other women of the family, including co-spouses if her husband is polygamous and her new sisters-in-law (Poiret 1996). As was expected in the discussion of sampling design (the assignment of household headship in MMP to absent males), more than a third of households are female-headed in MAFE-Senegal, while only between a tenth and a fifth of MMP households are female-headed. These figures are also related to the male-predominant nature of migration out of Senegal, particularly migration to Europe (Liu 2013; Schoumaker et al. 2013; Toma and Vause 2014), while by 2010 nearly half of Mexico-born individuals in the United States were female (Donato and Gabaccia 2015). Moreover, among both internal and international Senegalese migrants, living apart from their partners is a frequent and long-lasting situation (Baizan, Beauchemin, and Gonzáles-Ferrer 2014a). Findley estimated that between 43% and 68% of couples in Senegal experience this situation at some point during their lives (Findley 1997: 125).

---

19 Urban areas had a 35% lower fertility than rural areas (Agence Nationale de la Statistique et de la Démographie 2013).
Table 3: Family migration and household characteristics in Mexico and Senegal

<table>
<thead>
<tr>
<th></th>
<th>Mexican Migration Project</th>
<th>MAFE-Senegal</th>
<th>Living in HH + children of HH head</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ranchos</td>
<td>Small towns</td>
<td>Large towns</td>
</tr>
<tr>
<td>Households</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. people in household</td>
<td>6.5</td>
<td>5.9</td>
<td>5.7</td>
</tr>
<tr>
<td>No. children of household head</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
</tr>
<tr>
<td>% Female-headed</td>
<td>12.6</td>
<td>13.3</td>
<td>13.3</td>
</tr>
<tr>
<td>% of HH with ever migrant on HH roster</td>
<td>49.6</td>
<td>40.1</td>
<td>44.4</td>
</tr>
<tr>
<td>% of HH with current migrant on HH roster</td>
<td>27.1</td>
<td>18.6</td>
<td>24.3</td>
</tr>
<tr>
<td>Individuals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Living in the household</td>
<td>66.9</td>
<td>68.9</td>
<td>69.4</td>
</tr>
<tr>
<td>% Female</td>
<td>50.2</td>
<td>50.5</td>
<td>51.0</td>
</tr>
<tr>
<td>% Ever migrated</td>
<td>14.9</td>
<td>12.7</td>
<td>14.9</td>
</tr>
<tr>
<td>% Currently abroad</td>
<td>8.5</td>
<td>6.7</td>
<td>8.6</td>
</tr>
<tr>
<td>Age at first migration</td>
<td>24.1</td>
<td>23.1</td>
<td>22.8</td>
</tr>
<tr>
<td>N (communities/census districts)</td>
<td>19</td>
<td>34</td>
<td>14</td>
</tr>
<tr>
<td>N (households)</td>
<td>2,094</td>
<td>3,895</td>
<td>1,881</td>
</tr>
<tr>
<td>N (individuals)</td>
<td>13,571</td>
<td>31,358</td>
<td>13,039</td>
</tr>
</tbody>
</table>

Notes: Weights were applied to the Senegalese data. Weights for the Dakar region rely on computing sampling probabilities at each stage of sampling (census districts and households, as explained above). By applying weights for the different stratified groups they become proportional in the analyses to their real number in the population (Schoumaker and Mezger 2013).

Individual-level indicators also illustrate differences between the Mexican and Senegalese cases. Nearly a quarter of the individuals included in the MAFE household survey do not live in the household. This reflects both survey inclusion rules (discussed above) and differences in migration prevalence and nature. Nearly a quarter of Mexican households have US migration experience, compared with only one-eighth of Senegalese households. At the same time, it appears that more Senegalese are currently abroad than Mexicans. On the one hand, Senegalese-Europe migration differs from that of Mexican–United States migration in that it is less often circular, leading to lower return migration rates (González-Ferrer et al. 2014). On the other hand, relatively little is known about Senegalese migration to other African countries, although it is thought to be much more short-term and to involve greater return migration (Adepoju 2004; Lucas 2006). Finally, Senegalese appear to migrate abroad at younger ages than...
Mexicans. Previous work provides some insight, but more systematic research is needed to fully explain this. For instance, studies of Mexican migration to the United States describe an inverted ‘U’ shaped relationship between a husband’s migration and the family life cycle (Lindstrom and Giorguli-Saucedo 2007; Massey et al. 1987). A husband’s migration is least likely at the start of marriage and prior to the arrival of children, and then rises with parenthood and as the income needs of the household grow. No such effect has been reported for Senegalese migration to Europe. By contrast, studies emphasize men’s need to accumulate resources prior to marriage, often leading to migration at an early age (Baizan and González-Ferrer 2014b).

The discussion in this section illustrates how harmonized samples for different migration flows can help to reveal substantive differentials and commonalities in the characteristics of each migration flow and in the origin population. Yet it also shows how easily these characteristics can be confused by methodological differentials in the surveys.

9. Promises and limitations of comparative research

Cross-national, cross-continental comparative research of international migration holds great promise. First and foremost, it is essential for examining how well current theories of migration hold up under scrutiny in a variety of contexts. Are the drivers of international migration similar for individuals and families from many different origins? Are individuals and households driven by similar motivations to migrate? Do they pursue similar strategies? How well does the current scholarly literature help us understand and examine current flows of migration? What are the limits of current migration theory? In which ways do theories of migration need to adapt or expand or get specific? Under what conditions do specific theories/factors become more/less relevant? Second, while helping us identify possibly universal aspects of international migration, comparative research also enables us to begin identifying and analyzing the importance of context-specific characteristics like gendered norms, household expectations, labor market institutions, or specific policies. Such research could be important for contextualizing influential in-depth ethnographies, case studies, and non-comparative quantitative work. In a world where globalizing influences are on the rise, understanding whether and how local contexts influence international migration is particularly important.

There are multiple limitations to comparative research on international migration (see Riosmena 2016). First, different migration flows have different levels of circularity, and circularity is influenced by the maturity of migration streams and public
policy. As a result, the predominantly origin-based or retrospective natures of surveys are particularly troublesome. Who do we capture? Which migrants are missing?

Second, since most migration surveys are not nationally representative of origin contexts, differing sampling frames may inhibit comparative research. The two surveys analyzed here, the MMP and MAFE, have notable differences in scope. While most MMP surveys sample predominantly rural areas in Mexico, the MAFE-origin samples focus on the major urban areas of the Democratic Republic of Congo, Ghana, and Senegal. Because processes of urbanization, internal migration, and international migration are intimately intertwined, the absence of rural areas in MAFE and over-emphasis on rural context in MMP can inhibit efforts to carry out comparative analysis. The study of the (macro) context emphasized in the previous paragraph and its possible interaction with micro variables involves the specification of relevant contexts (national, regional, local) for data collection. As a result, robust comparative research is complex for both theoretical and empirical reasons.

A third difficulty refers to the definition and measurement of meso-level contexts in a longitudinal perspective, in particular social networks and households. There are well-known difficulties in defining and analyzing households across time (e.g., Adato, Lund, and Mhlongo 2007; Bauman 1999; Duncan and Hill 1985). Existing data usually provides only a fragmentary view of the household context: its composition, economic exchanges, etc. For instance, MMP focuses on the migration and labor market trajectories of the household head (current) spouse, but provides little other longitudinal information about other household members. In addition, there is even less agreement on how social networks are defined and data collected in existing surveys. Which categories of individuals are included? How to characterize the (strength of the) relationships? What individual characteristics should be collected? This, again, can be illustrated by the comparison of MAFE and MMP, which followed different strategies to collect information about social networks.

10. Discussion

We have highlighted some of the challenges facing scholars of international migration (Willekens et al. 2016). Despite recent calls to collect multi-sited quantitative data (Beauchemin 2014), we expose other concerns regarding longitudinal micro-data on international migration.

First, in a context of increasingly diverse data collection efforts, the public accessibility of collected data remains low. Our inquiry suggests that even many publicly funded data collection efforts have failed to make or keep data public and available to researchers, even after a reasonable time frame, and that in some cases
there have been legal impediments to doing so. Anticipating and resolving legal restrictions on data publication before data collection and publishing migration data on sites like the Interuniversity Consortium for Political and Social Research (ICPSR) are necessary. The preparation of data for publication is very resource-intensive, and is affected by fieldwork delays and strict grant timelines. As a result, a grant (and access to funds) can end before data preparation and publication are complete. If these are not remedied, past and future investments are lost.

Second, even when data is public and surveys are modeled on one another, comparability is not easy. Our research note examining the compatibility of MMP and MAFE household surveys shows that even simple descriptive statistics are precious, and may reflect diverse decisions in data collection (see also Riosmena 2016). Thus researchers are well advised to anticipate theoretically important research questions that would benefit from comparative analysis and incorporate these into the survey instruments.

Finally, despite these difficulties, there is promise for cross-comparative analysis of households and migration across different contexts. Using the MAFE household and MMP surveys, Liu, Riosmena, and Creighton (2015) examine the gendered role of family position and network-derived social capital in how international migration experience is distributed within Mexican and Senegalese families and find evidence of family obligations and differential investments in children. A comparative lens and accessible and longitudinal micro-data are essential for understanding the true and complex dynamics of international migration.
References


Liu et al.: Prospects for the comparative study of international migration using quasi-longitudinal micro-data