Multi-Country Surveys on International Migration: An Assessment of Selection Biases in Destination Countries

Cris BEAUCHEMI N
INED, France

Amparo GONZALEZ FERRER
Spanish Council for Scientific Research (CSIC-IEGD)
1. Introduction

- A lack of empirical evidence to feed the policies...
  - and, therefore, a need for specific surveys
  - especially in Subsaharan Africa (Lucas, 2006)

- The «parallel sampling» requirement
  - Principle: survey both at origin & destination
  - Essential to study factors & effects of migration (Bisilborrow et al. 1997; Massey 1987)

- A challenge
  - Several countries = costly + complex organisation
  - International migration = rare phenomenon
  - At destination specifically:
    - Migrants = difficult to reach population
    - Absence of a sampling frame

\[ \Rightarrow \text{The big sampling issue} \]
2. A review of sampling options

Basically, 3 options (McKenzie et al. 2007)

A. Random samples of migrants
B. Intercept point surveys
C. Snow-balling techniques

A. Random samples of migrants

Principle: Stratified sampling using disproportionate sampling fractions (Bilsborrow et al. 1997)

Some attempts but no plain success:

- Push-Pull in Spain (Groenewold et al. 2008)
  - Sample size reached at end... but adaptations of the general design made the sample non-random
- Nikkei community in Brazil (McKenzie et al. 2007)
  - Randomness preserved but... Final sample did not reach the objective (292 long interviews + 111 short interviews against 900 expected questionnaires)
2. A review of sampling options

B. Intercept point surveys

- Principle: Random selection of individuals in selected public places
- Possibility to compute weights when recording...
  - The number of potential interviewees in each place
  - Frequencies of interviewees’ visits to the different places
  - Information on refusals (number + basic characteristics)
  - Sample of at least 400-500 individuals
- Some success: Italy (Push-Pull project, Blangiardo)
- Shortcomings:
  - Public places selection ➔ potential biases
    - McKenzie et al. (2007): individuals more connected to the Nikkei community and more related to migrants
  - No possibility for long questionnaires
2. A review of sampling options

C. Snow-balling techniques

“Hard” method: Respondent-Driven Sampling (Heckathorn 1997)
- Possibility to achieve representative samples...
- But no success when applied to migrant populations

“Soft” method at destination:
- Some success: Grasmuck et al (1991): 7 seeds (interviewers themselves), 301 individuals at the end (as expected)

“Soft” method at origin:
- Principle: contacts collected at origin, “Origin-based snow-balling”
- Does work, but not always: Mexican vs. Dominicans
- Undocumented migrants easier to reach (more than 50% of the MMP sample)
- Drawbacks
  - No real representativeness at destination
  - Possible bias: Migrants with strong ties with the origin community would be over-represented
2. A review of sampling options

Objective:

- To test the origin-based snow-balling method
  - 1 Question: how efficient is the method in quantitative terms?
  - 4 hypothesis
    - H1: success varies according to the context (period, country of destination)
    - H2: helps to recruit undocumented migrants
    - H3: Bias towards more connected and less established migrants

- With the MAFE project data
3. The MAFE project

- MAFE for « Migrations between Africa and Europe »
  - departure, return, circulation

- Objective:
  - To produce comparative & representative data…
  - … to study the factors & effects of international migration

- General design
  - Inspired by the MMP and Biographic surveys conducted in Africa and France
  - 2 questionnaires: HH + Life histories (ind.)
  - Parallel sampling
    - Accounting for multiple destination
    - But focusing on European destinations
3. The MAFE project

Case studies

 Migration system 1
  Senegal
  Spain
  Italy

 Migration system 2
  DR-Congo
  Belgium
  Great Britain

 Migration system 3
  Ghana
  Netherlands

SUB-SAHARAN AFRICA  EUROPE
3. The MAFE project

The partners

Migration system 1
- Université de Dakar & Enda-Diapol
- University Pompeu Fabra
- Fieri

Migration system 2
- Université de Kinshasa
- Univ. Catho. de Louvain la Neuve
- University of Sussex

Migration system 3
- University of Ghana
- University of Maastricht

SUB-SAHARAN AFRICA   EUROPE
3. The MAFE project

MAFE samples at destination: objectives

- **200 migrants per country, as representative as possible at the national level**
- Similar designs in the 3 countries (France, Italy, Spain)
- Strata: - 50% of males; 50% of females
  - 50% between 25-40; 50% between 41-70
- Include difficult-to-reach migrants
  - Undocumented migrants
  - Migrants living in places of lesser concentration

Constraint:

- **No available sampling frame** at the European level
  - Nor at the national level, except in Spain (Padron)
3. The MAFE project

Data – Senegalese samples at destination (fol’d)

Options

• Which 1 of the classics?
  – Stratified sampling:
    NO, too costly, un-reliable in terms of sample size objectives
  – “Respondent driven sampling” or weighted intercept survey:
    NO, samples too small
  – Origin-based snow-balling:
    NOT ALONE, Bias + number

• A multi-source recruitment as a way
  – To insure that we would reach our quantitative objectives
  – To diversify as far as possible the directions of the potential biases

• In all countries, original design = 2 sub-samples
  1. Migrants contacted through origin-based snow-balling
  2. A complementary sample, taking advantage of the opportunities in each country
### 3. The MAFE project

#### Senegalese samples at destination
According to the source of recruitment

<table>
<thead>
<tr>
<th>Method</th>
<th>France</th>
<th>Italy</th>
<th>Spain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Origin-based snow-balling</td>
<td>30</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Snow-balling at destination</td>
<td>25</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>Padron</td>
<td>0</td>
<td>0</td>
<td>188</td>
</tr>
<tr>
<td>Associations' contacts</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other methods</td>
<td>130</td>
<td>173</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>200</strong></td>
<td><strong>202</strong></td>
<td><strong>200</strong></td>
</tr>
</tbody>
</table>
4. Results on origin-based snow-balling

An efficient source of recruitment?

- to maximize the number of contacts,
  2 sources of contacts in Dakar
  - **Source 1: 248 contacts** obtained
    - through the sampled households
    - for migrants cited within the questionnaires
  - **Source 2: 116 additional contacts**
    - individuals out of the sampled HH (neighbors, friends…)
    - or from HH out of the sample (training, …)

Problem 1: a high level of attrition

<table>
<thead>
<tr>
<th>Source 1</th>
<th>Number (N)</th>
<th>Ratio (N)/(T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of migrants living in Spain, France or Italy (T)</td>
<td>783</td>
<td>-</td>
</tr>
<tr>
<td>Number of obtained contacts</td>
<td>248</td>
<td>32%</td>
</tr>
<tr>
<td>Number of correct contacts</td>
<td>128</td>
<td>16%</td>
</tr>
<tr>
<td>Number of actual interviews</td>
<td>47</td>
<td>6%</td>
</tr>
</tbody>
</table>
4. Results on origin-based snow-balling

- Hyp. 1: Results vary according to the country of destination. YES
  - Number of obtained contacts is lower where migration is recent
  - Contacts are of better quality where migrants are less stressed by the policy context (Spain)
  - Refusals are at their maximum in Italy, where the fieldwork occurred at a time of hard deportation

<table>
<thead>
<tr>
<th>Sources 1 &amp; 2</th>
<th>France</th>
<th>Italy</th>
<th>Spain</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Number of obtained contacts</td>
<td>172</td>
<td>126</td>
<td>66</td>
<td>364</td>
</tr>
<tr>
<td>(2) Number of correct contacts</td>
<td>111</td>
<td>92</td>
<td>54</td>
<td>257</td>
</tr>
<tr>
<td>Ratio (2)/(1)</td>
<td>65%</td>
<td>73%</td>
<td>82%</td>
<td>71%</td>
</tr>
<tr>
<td>(3) Number of actual interviews</td>
<td>30</td>
<td>5</td>
<td>12</td>
<td>47</td>
</tr>
<tr>
<td>Ratio (3)/(1)</td>
<td>17%</td>
<td>4%</td>
<td>18%</td>
<td>13%</td>
</tr>
</tbody>
</table>
4. Results on origin-based snow-balling

Hyp. 2: Origin-based SB helps to recruit undocumented migrants. NO.

<table>
<thead>
<tr>
<th>Legal status at the time of the survey among...</th>
<th>% of undocumented migrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>All migrants (N=783)</td>
<td>13%</td>
</tr>
<tr>
<td>Migrants with a given contact (N=248)</td>
<td>13%</td>
</tr>
<tr>
<td>Migrants with a correct contact (N=128)</td>
<td>12%</td>
</tr>
</tbody>
</table>

(source: HH survey)

Origin-based SB presents nor advantage, neither obstacle to recruit undocumented migrants at origin.
Hyp. 3: Bias towards more connected migrants. YES

<table>
<thead>
<tr>
<th>Migrants' characteristics</th>
<th>Given contact</th>
<th>Correct contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Migrants' Migratory Experience</td>
<td>no significant result</td>
<td>no significant result</td>
</tr>
<tr>
<td>Migrants' Links with the Household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever lived with the HH head?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1,00</td>
<td>1,00</td>
</tr>
<tr>
<td>No</td>
<td>0,86</td>
<td>0,83</td>
</tr>
<tr>
<td>Don't know - Missing</td>
<td>45,18</td>
<td>6,94 *</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contribution of the transfers to the household budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>no transfer</td>
</tr>
<tr>
<td>Small share</td>
</tr>
<tr>
<td>Moderate share</td>
</tr>
<tr>
<td>Large share</td>
</tr>
<tr>
<td>Don't know - Missing</td>
</tr>
</tbody>
</table>
5. Conclusion - Discussion

- Origin-based SB does not...
  - ... allow to building large samples
  - ... does not help to recruit undocumented migrants

- Quantitative results depend hugely on the context
  - Country of destination
  - Periods of data collection
    (pilot surveys vs. real survey)

- Origin-based samples are biased and may lead to an over-rating of...
  - The influence of cumulative causation
    (migration experience of the head – further analyses needed)
  - The measurement of the connections of the migrants with origin country.
Aknowledgements

Project funded under the Socio-economic Sciences and Humanities