

Çağlar Özden
Christopher Parsons
Maurice Schiff
Terrie Walmsley

The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank, its Executive Directors or the governments they represent.



Motivation

- Migration an issue of increasing international importance
- Surprisingly little data, especially when compared to financial or trade statistics
- Therefore one of the priorities of the World Bank program on International Migration and Development



Previous Work

- OECD-centric:
 - Docquier & Marfouk (2006), OECD (2005+) Brain Drain
 - Docquier et al (2007), OECD (2007+): Gendered assessment of the Brain Drain
 - Docquier & Rapoport (2007): Control for age of entry
- For the 2000 census round only:
 - Parsons et al (2005, 2007): 226*226
 - Ratha and Shaw (2007)



Who are migrants?

- Collect: Country of birth & Citizenship
- Prioritize foreign born definition, why?
 - Physical movement more appropriate
 - Dependencies
 - Changing nationality possible
 - Naturalization rates vary
 - Aggregate categories smaller in magnitude



How are migrants recorded?

- Censuses and Populations Registers
 - Commonly conducted
 - Standardized questions
 - Most comprehensive coverage



Raw Data I

- Global Migration Database UNPD custodians
 - Collaboration: UNPD, UNSD, World Bank, University of Sussex
 - 3,500 census records
 - Over 230 destinations
 - Time, gender, age



Raw Data II

- US Census Bureau
- LSE Census Library
- British Library
- Library of Congress





Raw Data III

Census Round	No. Birthplace Sources	No. Nationality Sources	No. Birthplace by Gender	No. Nationality by Gender
1960	102	68	95	63
1970	91	55	82	46
1980	126	87	112	79
1990	134	113	112	96
2000	126	120	103	96
TOTAL	579	443	504	380



Raw Data IV

Number of missing census rounds	Number of Destination Countries	Percentage of World Migration, in 2000	Percentage of World Migration in 2000, (excluding migrations within the former Soviet Union)
0	49	47	57
1	49	6	8
2	41	17	20
3	52	21	13
4	29	8	2
5	6	0	1
Total	226	100	100



Challenge I – Defining Countries

- Issue: Tracking migrants meaningfully over time
 - Break-ups: Soviet Union, Yugoslavia
 - Reunification: Germany, Yemen
 - Independence: Africa, Caribbean, Oceania
- Solution: 226*226 master list, over time
 - 203 sovereign nations +23 additions



Challenge II – Recording and recoding

- Issue: Standardizing origin regions (10,000):
 - A. Specific single entities: Vatican, Wake Island, Isle of Man
 - B. Aggregates: South America, Ex-French Africa
 - C. Miscellaneous: Born at sea, Unknown, Stateless



Challenge II – Recording and recoding

Solution:

- Relabeling $(10,000 \rightarrow 522)$
- A. Aggregate single entities (226 from Master list, 65 others)
- B. Disaggregate agglomerated categories (231)
- C. Treat consistently miscellaneous entries



- Issue: Disaggregating geographic aggregates
- Solution:
 - Propensity measures
 - Average Destination Country Shares
 - Average Regional Shares
 - Global Shares



Issue: Census dates are not harmonised

Census Round	% Censuses by birthplace	% Censuses by Nationality
1960	75	71
1970	66	73
1980	73	70
1990	76	70
2000	84	76

- Solution: Nothing
 - Different version



Challenge V – Missing Gender Splits

- Issue: Assigning gender splits to aggregate data
- Solution:
 - Propensity measures
 - Regional shares
 - Regional shares over time



Challenge VI – Combining Definitions

- Issue: Cannot easily combine Foreign Born and Nationality data
- Solution: Prioritise foreign born
 - Always choose if three or more census rounds with foreign born data (156)
 - Not so much of an issue in Middle East and East Asia



- Issue: Census rounds missing, why?
 - Lack of expertise
 - Some only very recent
 - Expensive
 - Conflict or Political upheaval
 - Politicised
 - Alternative definition



- Issue 1: Missing in-between decades
 - Interpolation (42 country-years)
- Issue 2: Missing end decades
 - Missing year same composition as closest decade for which we have data (115 country-years)



Challenge VII – Missing Census

- Issue 3: Very poor data (<3 censuses)</p>
 - Adjust to United Nation's "Trends in World Migrant Stock" data (86 cases)
- Issue 4: Missing countries
 - Afghanistan, China, Eritrea, Lebanon, the Maldives, Qatar, Somalia, the People's Republic of Korea and Vietnam.

FINAL OUTCOME

Assignment of observations by method

				Remainder	Scale &	
	Missing	Interpolation	Raw number	category	Interpolation	Total
1960	1,898	187	13,172	17,619	17,974	50,850
1970	1,898	3,470	13,428	14,349	17,705	50,850
1980	1,898	2,688	14,604	20,657	11,003	50,850
1990	1,648	2,167	17,131	22,142	7,762	50,850
2000	1,648	0	20,313	17,085	11,804	50,850

	Missing	Interpolation	Raw_number	Remainder category	Scale & Interpolation	Total
1960	4%	0%	26%	35%	35%	100%
1970	4%	7%	26%	28%	35%	100%
1980	4%	5%	29%	41%	22%	100%
1990	3%	4%	34%	44%	15%	100%
2000	3%	0%	40%	34%	23%	100%

FINAL OUTCOME

Assignment of migration numbers by method

	Missing	Interpolation	Raw_number	Remainder category	Scale & Interpolation	Total
1960	814,737	249,462	46,600,000	6,710,092	21,500,000	75,874,291
1970	1,320,981	1,803,088	52,900,000	6,413,896	18,100,000	80,537,965
1980	1,448,718	3,780,738	61,100,000	5,604,684	21,700,000	93,634,140
1990	2,227,064	2,893,973	106,000,000	7,683,809	17,500,000	136,304,846
2000	3,299,835		114,000,000	13,000,000	29,400,000	159,699,835

				Remainder	Scale &	
	Missing	Interpolation	Raw_number	category	Interpolation	Total
1960	1%	0%	61%	9%	28%	100%
1970	2%	2%	66%	8%	22%	100%
1980	2%	4%	65%	6%	23%	100%
1990	2%	2%	78%	6%	13%	100%
2000	2%	0%	71%	8%	18%	100%

FINAL OUTCOME

Reliability of raw numbers

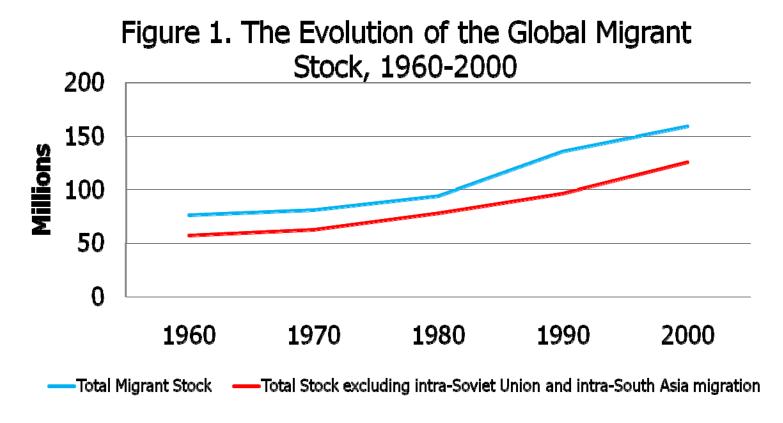
Reliability of raw numbers							
	Frequency	Percent	Cumulative				
0-50%	515	1%	1%				
50-60%	191	0%	1%				
60-70%	225	0%	1%				
70-80%	453	1%	2%				
80-90%	994	1%	3%				
90-100%	74,015	97%	100%				
Total	76,393	100%					



AND FINALLY!

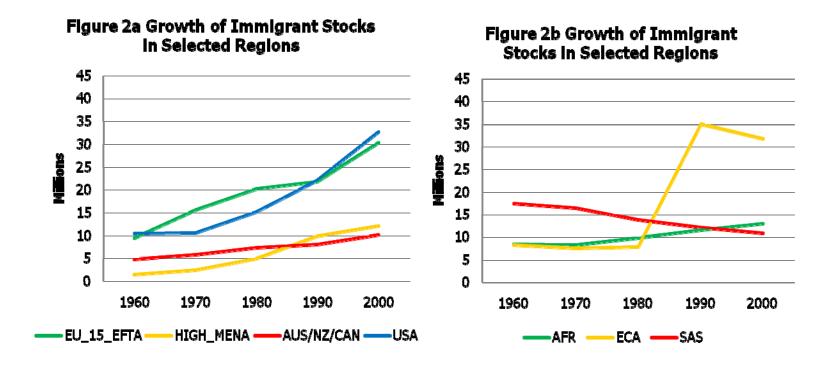


What the Data Show I



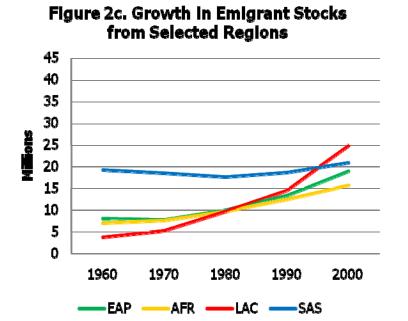


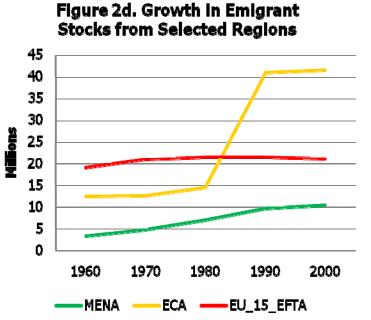
What the Data Show II





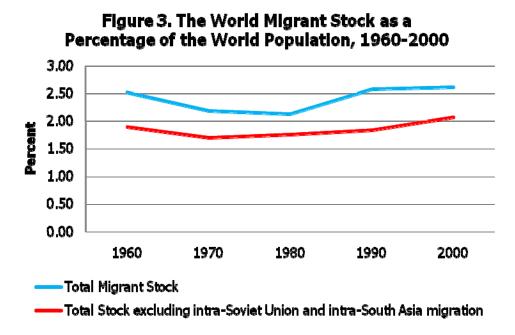
What the Data Show III







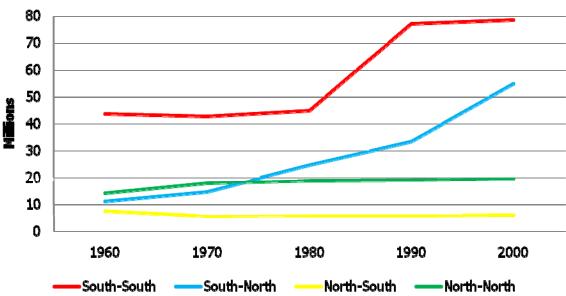
What the Data Show IV





What the Data Show V

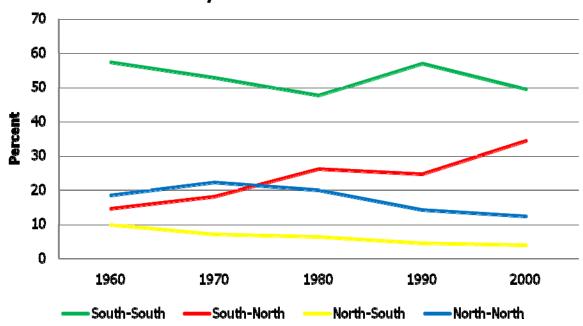




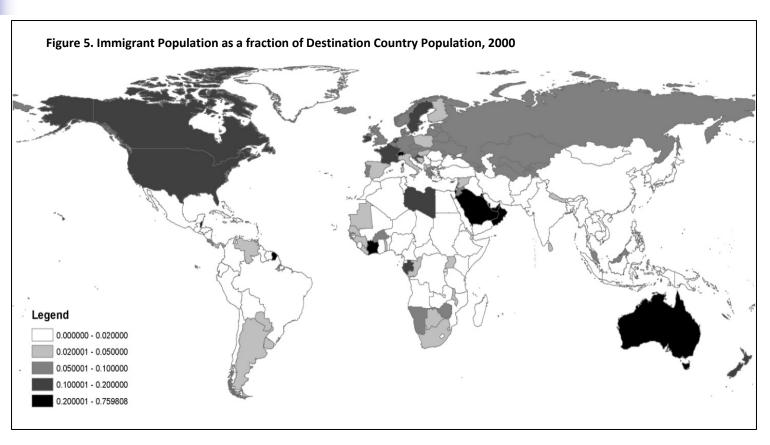


What the Data Show VI

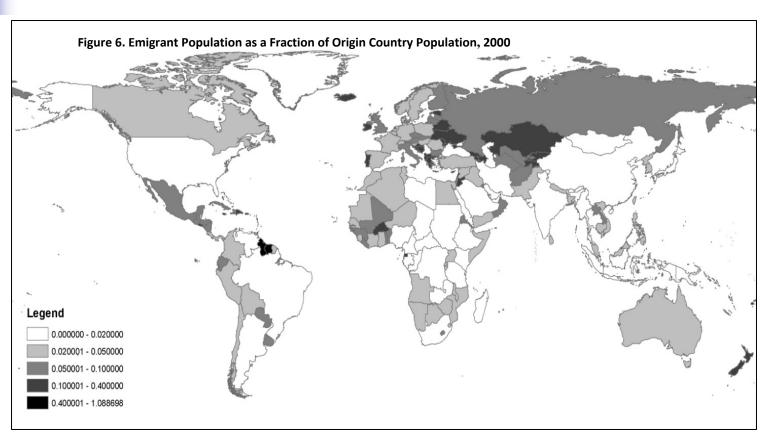
Figure 4b. The Proportion of Global Migration accounted for by the North and the South



What the Data Show VII



What the Data Show VIII





What the Data Show IX

Legend

Intra-region
10 million+

Intra-region
5 million

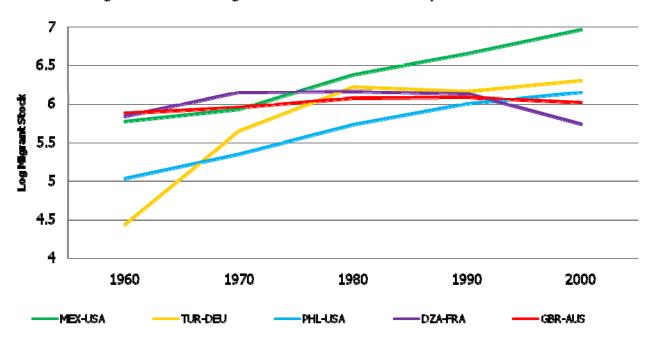
Main Inter-region
migrations

Figure 7. South-South inter- and intra-regional migration, 2000



What the Data Show X

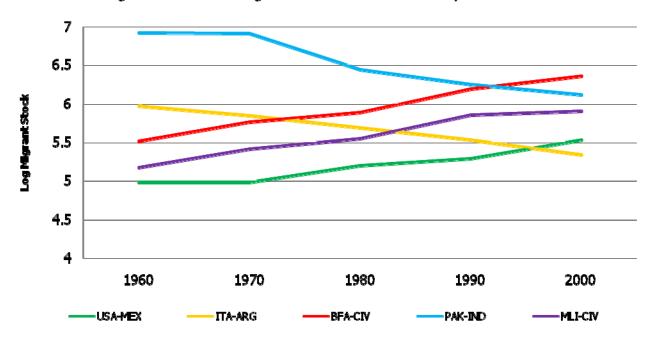






What the Data Show XI







Conclusion

- Inevitable trade-off between pragmatism and accuracy
 - Methodology clear
- Assumptions can be bettered
- Never ending story?
 - Easily updated



- Gravity model determinates
- FDI, trade and migration linkages
- Impact of diverse migration policies
- Role of Diasporas