It's the Journey, not the Destination: Estimating Refugee Flows from Stock Data

Molly Roberts and Kara Ross Camarena

Harvard University

Motivation

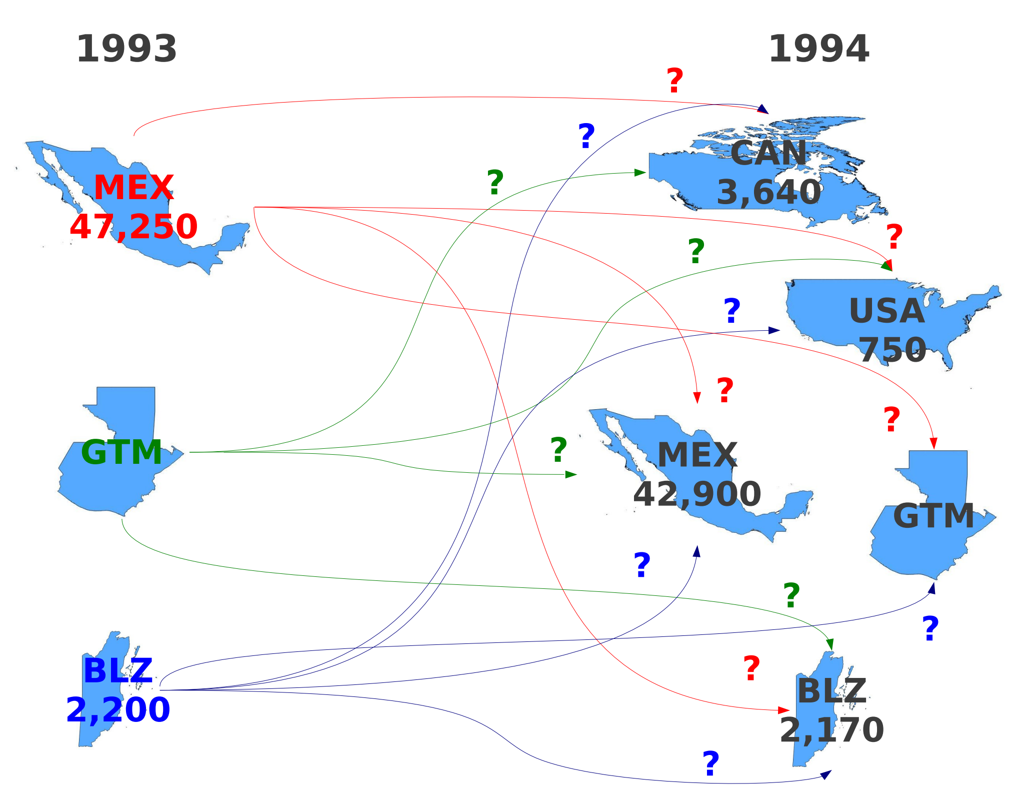
- Current Approach: use stock data (the number of refugees from a given country in another country for a given year), and model flows as net changes.
- Country Directed Dyad Year Assumption: every refugee returns to his country of origin before journeying to a third country.
- The Problem:
  - for resettlement, refugees are moved from their country of asylum to their country of resettlement, never returning to their country of origin.
  - refugees often migrate through countries of asylum to countries further away or between countries of asylum as conditions change.

The Solution: Drop Country Directed Dyad Year assumption and model the flows of refugees, as the real quantity of interest.

Working Example: Guatemala

- The roots of the Civil War in Guatemala began in the 1950s, targeted violence increased in the late 1970s.
- By early 1980s thousands of Guatemalans had fled to Mexico, Honduras, and Belize.
- By mid 1980s Guatemalan refugees were arriving in the US and Canada.
- Refugees stayed outside of Guatemala well into the 1990s.
- Through the 1980s and into 1990s US and Canadian news sources report refugees from Guatemala arriving in the US and Canada by way of Mexico and Honduras.

Unknowns



Bounds: the "Known Unknowns"

- Simple example: if Mexico were the only country to ever have Guatemalan refugees, all refugees in Mexico came directly from Guatemala.
- More complicated: # of refugees travel to a country in 1994 + # that stay 1993 to 1994 = total # of refugees in 1994
- Most general: use "Origin-Destination" Routing Matrix, example for U.S., Mex, Guatemala:

$$\begin{bmatrix} US_{t-1} \\ US_t \\ Mex_{t-1} \\ Mex_t \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} US \rightarrow G \\ US \rightarrow Mex \\ US \rightarrow US \\ Mex \rightarrow G \\ Mex \rightarrow Mex \\ Mex \rightarrow US \\ G \rightarrow Mex \\ G \rightarrow US \end{bmatrix}$$

Y: Observed = A:Identity × X: Unobserved

The Model

- "Network tomography" problem, follow Tebaldi and West (1998)
- Problem: find $p(X|Y)$ given $Y = AX$.
- $X_a \sim \text{Poisson}(\lambda_a)$
- X can be partitioned into X_1 and X_2
- The posterior has the form: $p(X_1|X_2, \Lambda, Y)p(X_2|\Lambda, Y)$, where
  - $p(X_1|X_2, \Lambda, Y)$ is degenerate: $X_1 = A_1^{-1}(Y - A_2X_2)$
  - $p(X_2|\Lambda, Y) \propto \prod_{a=1}^{c} \frac{\lambda_a^{X_a}}{X_a!}$

Metropolis Within Gibbs Algorithm

1. Draw Λ from $p(\lambda_a|X_a)$
2. For each $i = r+1 \ldots c$, sample new X_i from $p(X_i|X_{2,-i}, \Lambda, Y) \propto \frac{\lambda_i^{X_i}}{X_i!} \prod_{a=1}^{r} \frac{\lambda_a^{X_a}}{X_a!}$
3. Reevaluate X_1 at each step of 2.
4. Return to step 1 and iterate.

Results