

Estimating migration flows in
Northern Ireland by health
characteristics measured in the
Census:
using the NI Longitudinal Study

James Brown (University of Southampton)

John 'Mac' McDonald (Institute of Education)

BSPS Annual Conference – York Sept 2011

www.ioe.ac.uk



NI Longitudinal Study



The help provided by the staff of the Northern Ireland Longitudinal Study and the NILS Research Support Unit is acknowledged.

The NILS is funded by the Health and Social Care Research and Development Division of the Public Health Agency (HSC R&D Division) and NISRA. The NILS-RSU is funded by the ESRC and Northern Ireland Government.

The authors alone are responsible for the interpretation of the data.

Outline



- **Overview of migration and health data in the NILS**
 - The problem...
- **The classic approach: IPF**
- **Alternative approach: GLM**
- **Simulation Study**
- **Results with data from NILS**
- **Concluding Remarks**

- **Annual data published on internal migration often has limited information on the characteristics of migrants.**
 - For Northern Ireland you just get the total outflow and total inflow for each local council.

HOWEVER:

- **Understanding patterns in flows can be illuminating from a policy perspective.**
 - *Are there specific patterns in movements of individuals with and without limiting long-term illness?*
 - Potential implications for healthcare provision – *but must recognise we cannot make causal statements...*

NILS: health & migration



- **For NILS members we know their limiting long-term illness at the time of the last census (2001).**
 - For new members this is ‘missing’.
- **Through linkage to the *health card registration system* we know the internal migration movements of NILS members in each year.**
 - with their associated health status (from the linkage to the 2001 Census)...
- **At the population level, the total in-flow and out-flow (from the health card registration system) is published.**
 - Available by local council area.

Published data



- **We get the margins of a five by five table using the health and social care trusts in Northern Ireland for mid-2007 to mid-2008.**
 - Combines migration between local councils.
 - Movement within a HSC trust is movement across a local council boundary but staying within the HSC.

Start HSC Trust * Finish HSC Trust Crosstabulation

Count

		Finish HSC Trust					Total
		Belfast	South Eastern	Southern	Northern	Western	
Start HSC Trust	Belfast						10,777
	South Eastern						7,204
	Southern						6,210
	Northern						9,909
	Western						4,301
Total		10,114	7,873	6,314	10,287	3,813	38,401

NILS data



- From NILS we get the full cross-tabulation of migrant flows for the same period by health status (limiting long-term illness).
 - Further collapsing of the two smallest HSC Trusts (cell counts).
 - Three tables for ‘yes’, ‘no’, and ‘not in the census’.

Start HSC Trust * Finish HSC Trust Crosstabulation^a

Count

		Finish HSC Trust				Total
		Belfast	South Eastern	Northern	Southern & Western	
Start HSC Trust	Belfast	95	94	59	29	277
	South Eastern	57	82	26	27	192
	Northern	36	16	139	26	217
	Southern & Western	25	21	38	113	197
Total		213	213	262	195	883

a. Reported existence of a limiting long-term illness in the 2001 Census

The problem



- **We want to estimate flows between (and within) HSC by health status.**
 - Need to be consistent with the overall in-flow and out-flow data.
- **The data from the NLS estimates the underlying pattern in any flows *by health status*.**

THEREFORE:

- **Need to use the NLS structure to estimate flows consistent with the published totals.**

Classic approach: IPF



- **Iterative Proportional Fitting (IPF) sequentially scales an observed table to fit a set of margins.**
 - Referred to in survey estimation as raking.
 - A widely used technique.
 - *Difficult* to get SEs for the resulting cell estimates.
 - Not a standard application as we have a three-way table but only two of the margins...
- **IPF preserves to internal structure of the data (in this case the interaction structures observed in the NLS).**
 - While ensuring the resulting cell counts sum to the required marginal totals.

Alternative approach: GLM



- **Goes back to Willikens (1999) and others.**
 - More recent applications by Raymer *et al* (2007, 2009, 2010).
- **Essentially we fit a model to the observed data for the margins.**
 - In our case that is a simple independence model as we just have the flow margins.
 - In work by Raymer *et al* they explore more complex models.
 - The NILS cell counts enter as an offset term in the model forcing the required interaction structure.
 - Get SEs by fitting two models (*the model for the NILS structure and the model for the margins*) and then combine the two sources of error.

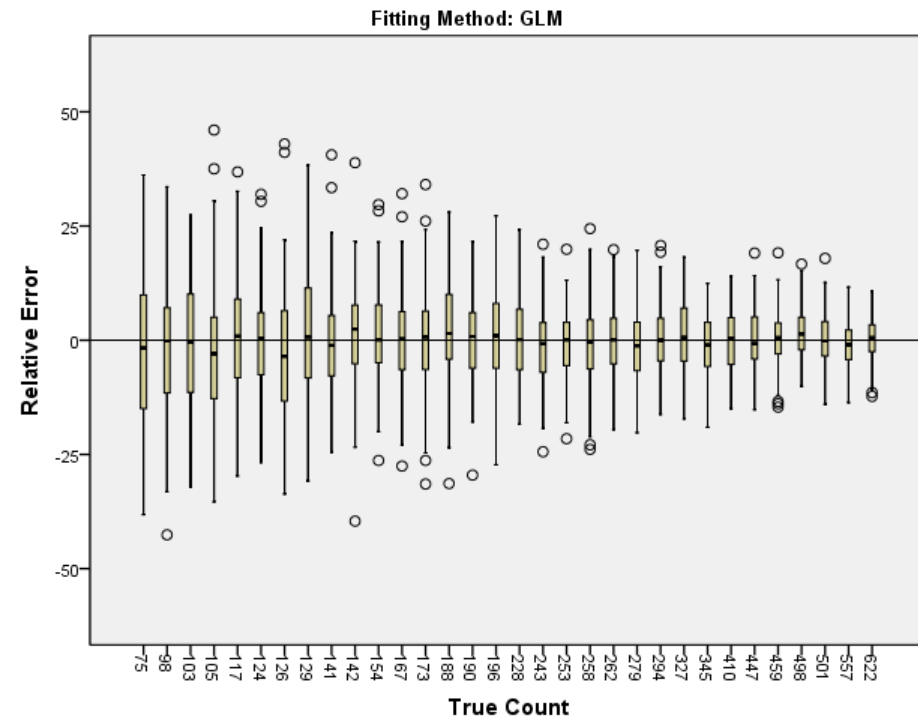
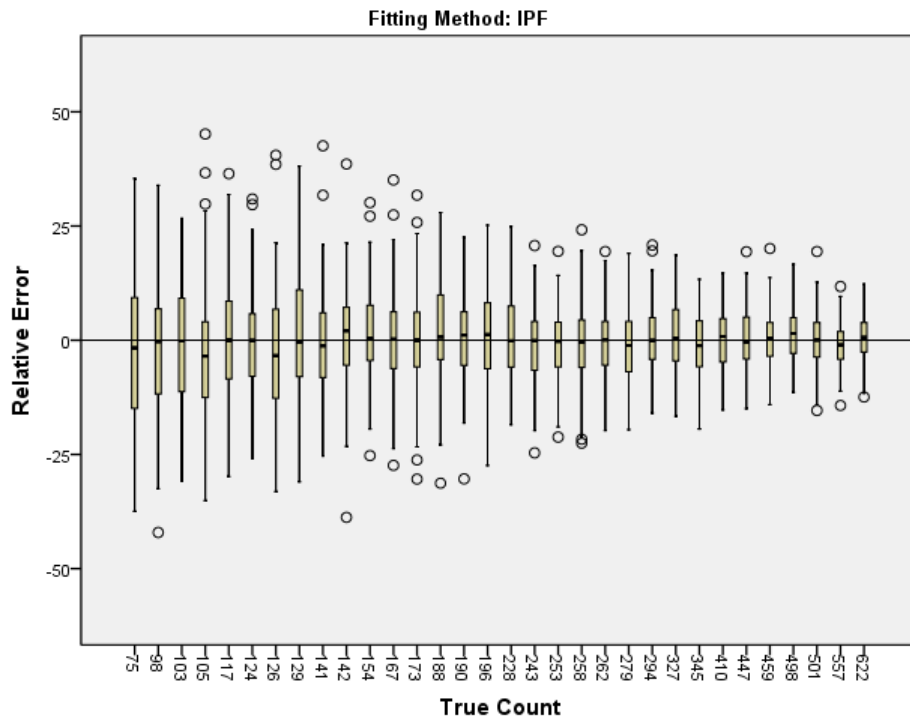
Simulation study



- **To explore the approaches, we treat NLS as a population and select samples of around 30%.**
 - Use gender (two categories) rather than health status of migrants to get reasonable cell count sizes.
 - Selected 100 samples (*one rejected due to cell counts less than 10*).
- **Can use the simulation to explore bias and variance of the two approaches.**
 - Can look at the estimation of SEs for the GLM approach.

Simulation study: results

- Distribution of the relative error on estimating each cell count.



Simulation study: results

- **What about preserving relationships within the table?**

<i>Data for Males</i>		Finish HSC	
		Belfast	South Eastern
Start HSC	Belfast	279	459
	South Eastern	262	243

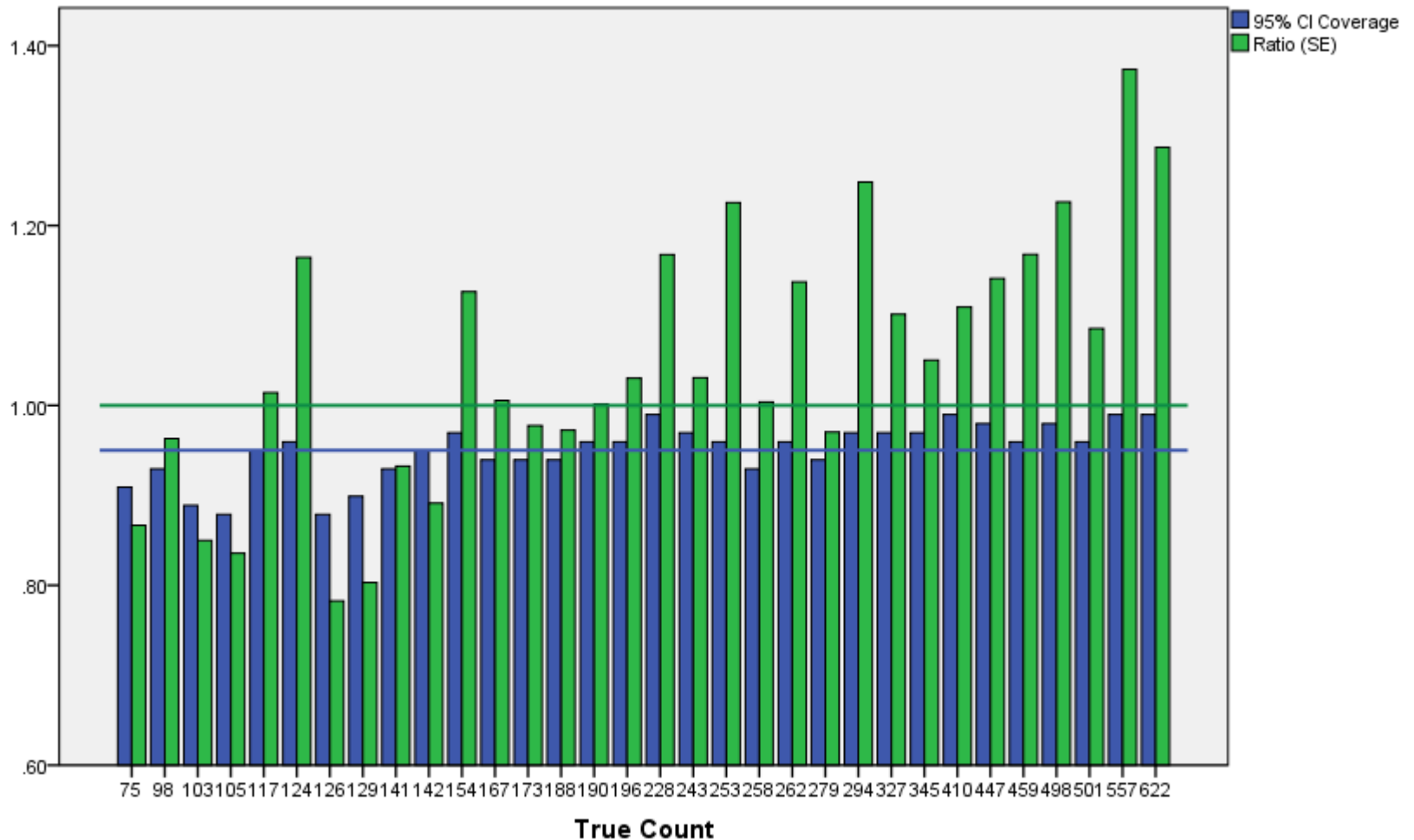
- **The odds ratio for this corner is 0.564.**
 - The average across the 99 samples drawn from the population is 0.559.
 - Both methods perfectly capture the observed pattern in each sample.

Simulation study: results



- **What about standard errors and confidence intervals?**
 - These do not come straight from the modelling as we must allow for the uncertainty in the NILS structure we are forcing to the margins.
- **Overall we get good coverage properties.**
 - Some under estimation for smaller cell counts and over estimation for larger counts.
 - Sensitive to the model used for getting the model SEs for the NILS structure.
 - » too simple = under-estimates, too complex = over-estimates

Simulation study: results



Modelling the 2007/08 data

- **Looking for a parsimonious model based on three factors:**
 - starting HSC (4), finishing HSC (4), limiting long-term illness (3).
- **Model with main effects and the start.finish interaction fits reasonably well.**
 - some evidence of the lti.finish term improving the fit.

		Finish			
		Belfast	SE	N	S & W
Start	Belfast	0.603	0.918	0.556	0.379
	SE	0.579	0.563	0.188	0.280
	N	0.415	0.200	1.210	0.277
	S & W	0.396	0.218	0.289	1.000

Flows by health status

- **Estimated flows for migrants with limiting long-term illness**

		In												
		Belfast			SE			N			S & W			
Out	Belfast	494	(442, 552)	408	(370, 451)	276	(246, 309)	147	(129, 167)				1325	
	SE	297	(266, 333)	357	(319, 401)	122	(103, 144)	137	(119, 158)				913	
	N	198	(175, 225)	74	(63, 87)	689	(628, 756)	140	(121, 161)				1101	
	S & W	156	(138, 178)	110	(94, 129)	214	(186, 247)	689	(626, 759)				1169	
		1145			949			1301			1113			4508

- **Not much evidence of particular areas attracting those with LLTI.**
 - Reflected in the lack of interaction terms in the NLS data...

Concluding remarks



- **Using the NILS data we can get updated patterns of flows by characteristics.**
 - Work by Raymer *et al* imposed the previous census structure.
- **In terms of estimating flows there is no gain from the alternative approach over using the IPF.**
 - Alternative easy to fit within a standard package.
 - » No programming needed.
 - » SEs possible (*still needs more thought*).
- **Can now build-up annual flows going forward from the Census using the NILS and published marginal flows.**