



Small Area Estimation of At Risk of Poverty Rates

TiPSE — Territorial Dimensions of Poverty and Social Exclusion

Andrew Copus – The James Hutton Institute, Aberdeen

Mike Coombes – University of Newcastle

ESPON Scientific Conference

University of Luxembourg – 13th September 2013



What is the AROP, and why is it important?

- The **A**t **R**isk **o**f **P**overty rate = Percentage of the population living in households which have <60% of the (national) median disposable income, after social transfers.
- The **AROP** rate is one of three components of the composite EU2020 indicator **AROPE** (E = “Exclusion”). 2020 goal is to reduce number of people at risk of poverty and social exclusion by 20m.
- More immediately, the new programming period for Cohesion policy presents some possibilities for (regionally specific) poverty reduction programmes.
- However, at present Eurostat publishes AROP rates only at NUTS 2 (at best – some MS only NUTS 1 or NUTS 0)

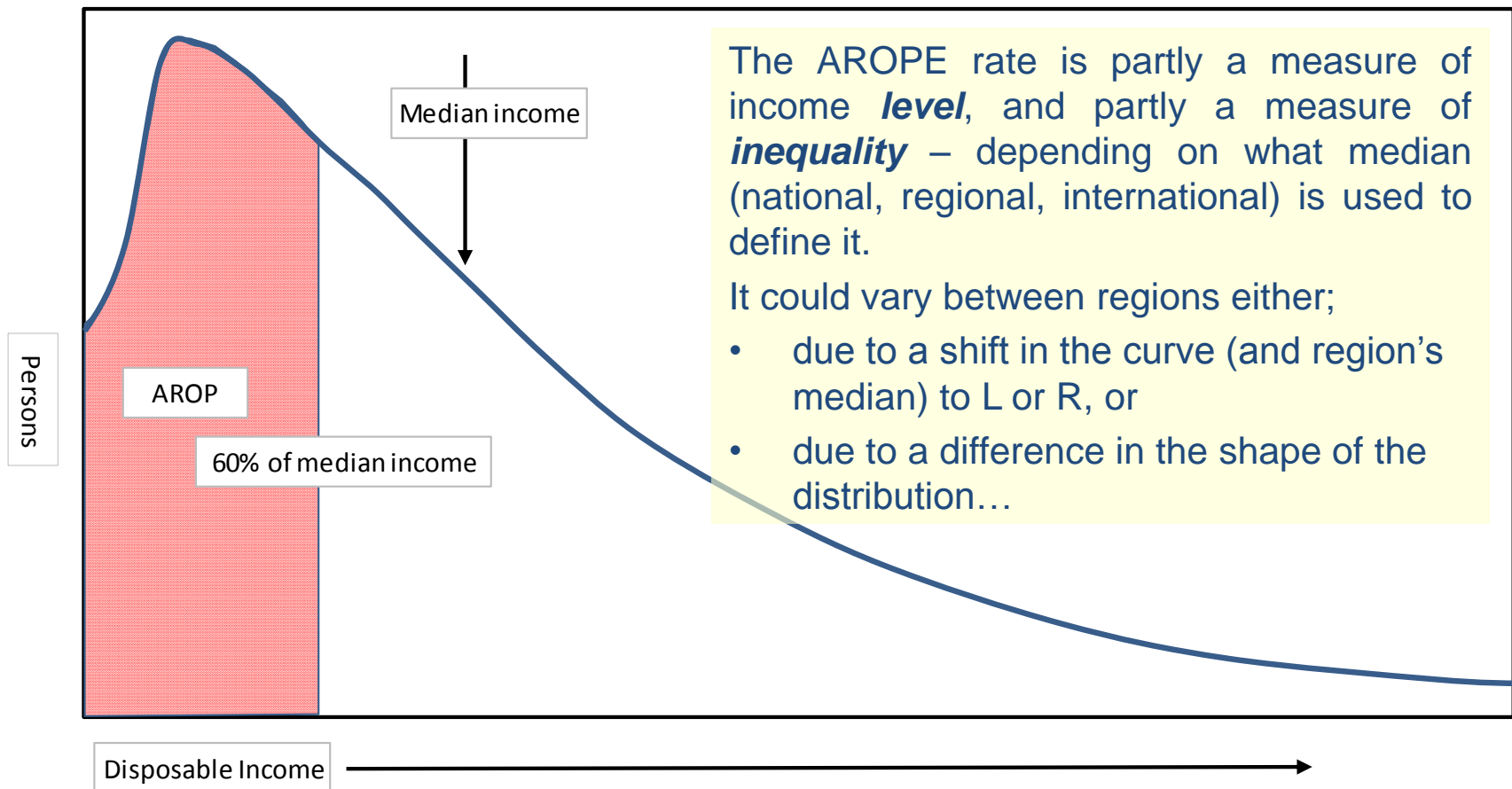


How are AROP rates calculated or estimated?

- The Eurostat AROP rates are derived from the EU-SILC (Survey of Income and Living Conditions). Sample sizes in many NUTS3/2 regions mean that sampling errors are unacceptable.
- However in some MS (Nordics, NL etc) regional AROP rates can be generated from register data.
- There are a variety of methodologies for estimation of regional AROP rates, of varying degrees of sophistication. They are generally quite demanding in terms of data requirements, and some of them require higher levels of econometric expertise.
- The most well known of these is the PovMap model of the World Bank (WB) which combines survey and census microdata.



A tricky indicator to estimate...



Why a “Plan B”?

- TiPSE project spec. required estimation of AROP rates at NUTS 3 (or smaller) for as many as possible of ESPON countries, excluding those already being worked on by WB (i.e. NMS12 CEECs).
- TiPSE proposed using the PovMap model to maximise commonality with WB maps for the CEECs.
- Unanticipated problems call for exploration of a Plan B:
 - (a) Because European population census have a separate purpose/history from EU-SILC survey, the matching of variables required by PovMap is problematic.
 - (b) 2011 Census microdata publication delays.
- **Therefore** we have been experimenting with simpler estimation approaches, based upon regional data.



Objectives and Overview of the Approach

The objective is an approach which:

- Can use area Census data (available earlier than microdata).
- Is relatively simple and transparent, not requiring high levels of expertise, and “parsimonious” in terms of data requirements.
- Results in NUTS 3 AROP estimates which are consistent with Eurostat rates published at NUTS 2 or 1.

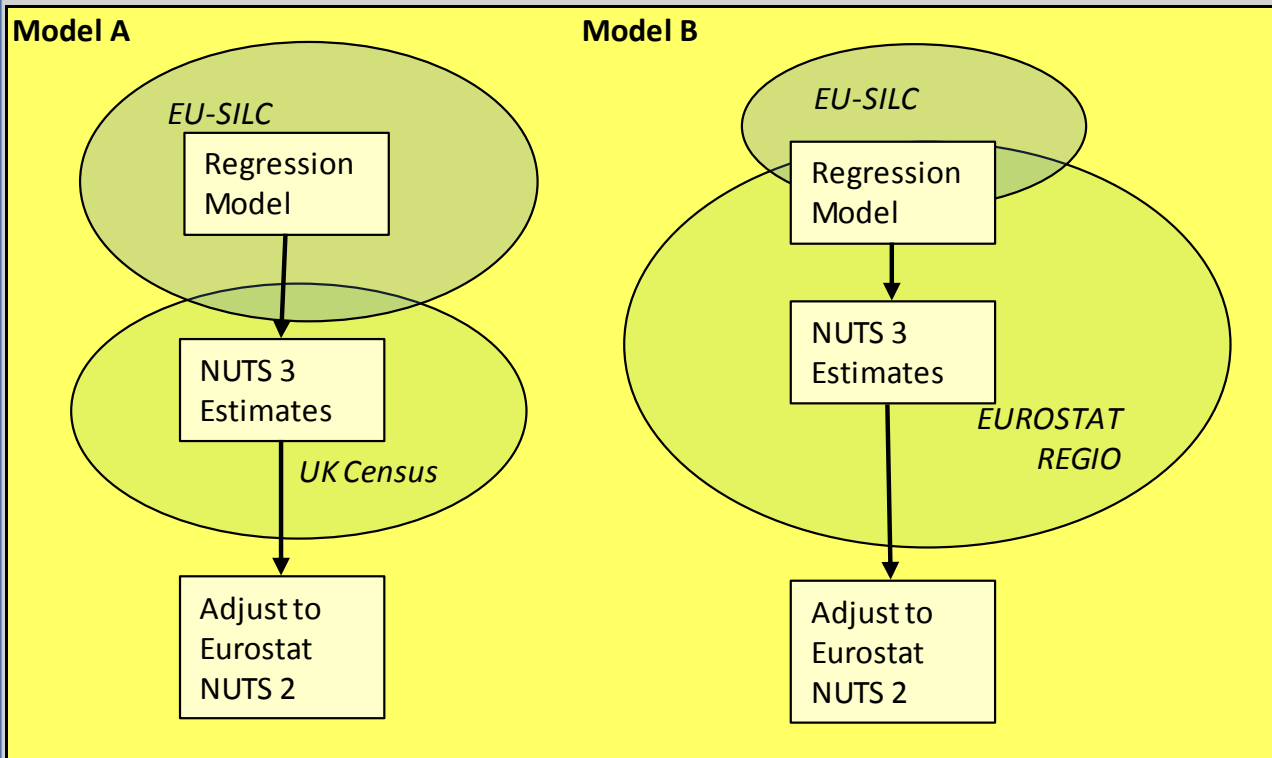
NB Not intended to replace PovMap, except in MS where microdata is not available, elsewhere seen as a short/medium-term solution...

3 stage approach:

- a) Estimate relationship between AROP rates and a selection of socio-economic indicators at NUTS 2 using basic OLS regression.
- b) Estimate (first round) NUTS 3 AROP rates by applying coefficients to NUTS 3 data for the same independent variables.
- c) Adjust these (first round) NUTS 3 AROP rates so they conform with rates published by Eurostat for larger (i.e. NUTS2) regions.



Outline of the approach and different data sources



Model A: Easier to get a closely fitting regression model, but more difficult to match SILC and Census for NUTS 3 estimates

Model B: More difficult to estimate the regression model, but NUTS 3 estimates are based upon exactly matching variables (only the scale is different).

Summary of the models

Model A

Percentage of Household Representative Persons:-	Coeff	p-value
economically inactive, excluding retired	0.565639	0.0007245
living in a detached dwelling	0.201623	0.0001403
in households comprising a couple and 2 or more children	-0.43404	0.0134319
who are married	-0.38073	0.0011657
who are occupied as Legislators, senior officials and managers	-0.69571	0.0001825
who are occupied as Service Workers, Shop and Market Sales	-0.33027	0.0243429
who work within Public administration and defence.	-0.43935	0.0482613
who have good or very good health	-0.14941	0.1171187
Constant	54.72232	0.0000009
Adjusted R-square	0.86	

Model A – tested >100 potential independent variables (IV) from EU-SILC

Model B – experimented with almost 50 potential IV – mostly from Eurostat Regio.

Model B

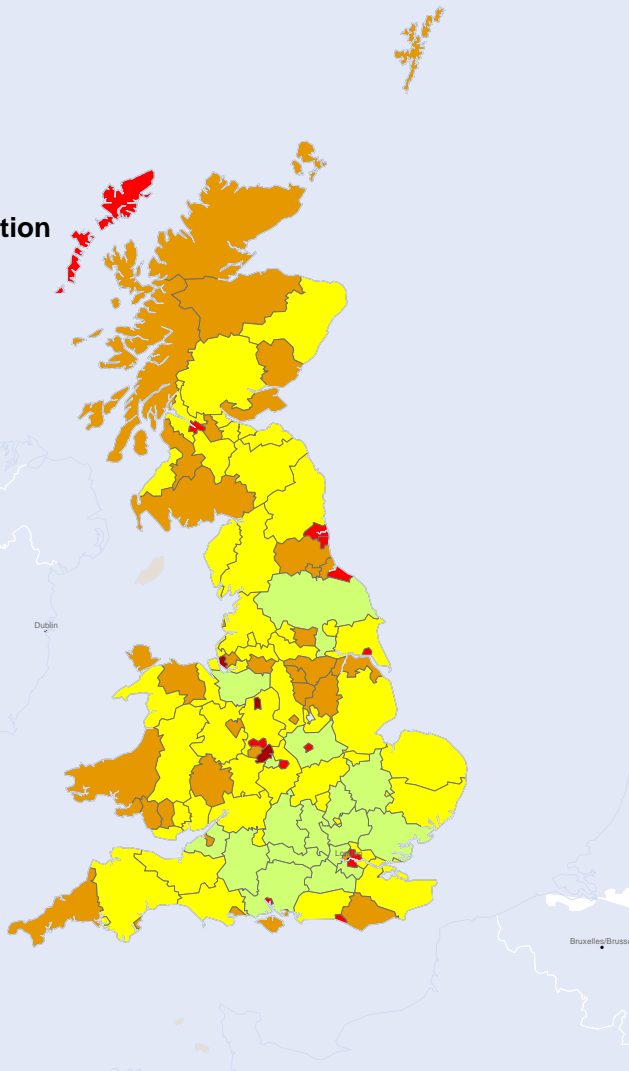
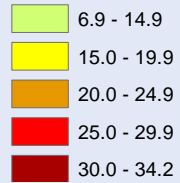
Variable	Source	Coeff	p-value
Population Density	Eurostat Regio	0.003	0.000015
Percentage of population with only Primary Education	Eurostat Regio	0.498	0.000174
Percentage of Lone Parent Households	Eurostat Regio	-1.124	0.065994
Gross Domestic Household Income (UK = 100)	UK Office for National Statistics	-0.233	0.003620
Multi-modal Accessibility Index	ESPON Territorial Observations No 2	-0.052	0.105950
Constant		44.923	0.000068
Adjusted R-Square	0.65		



The resulting maps

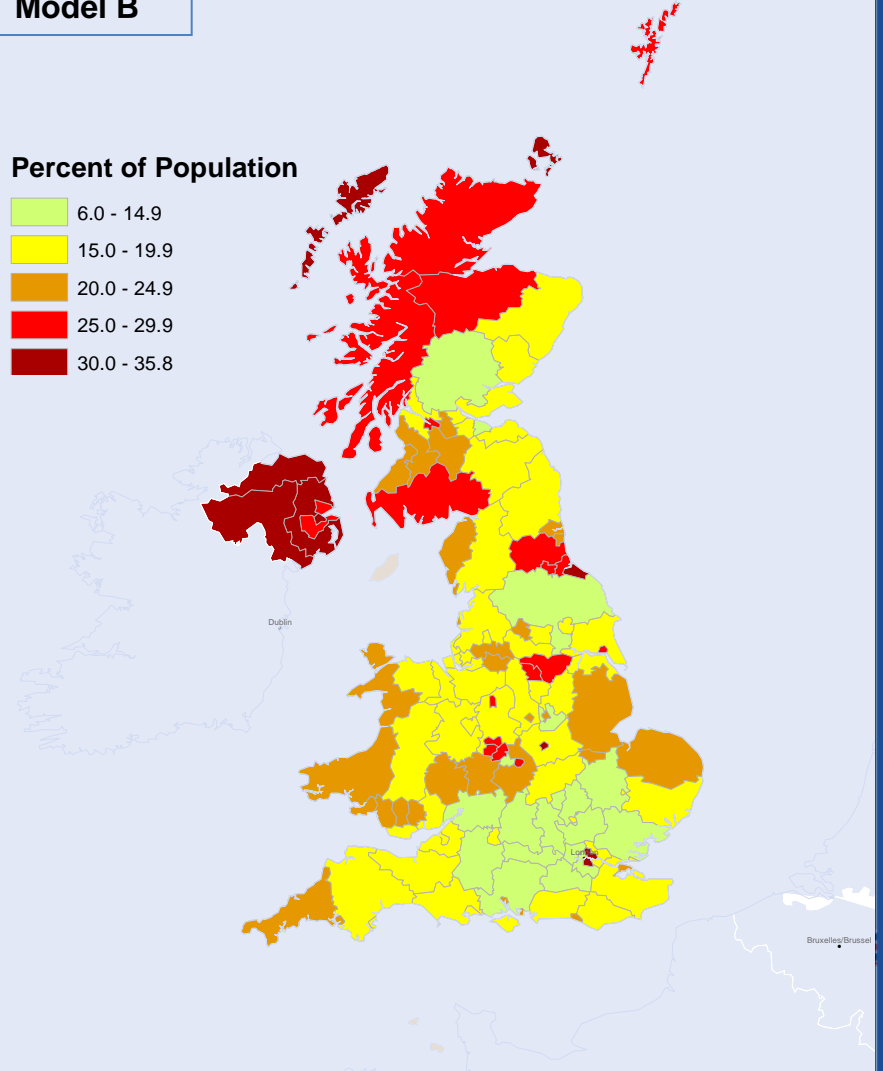
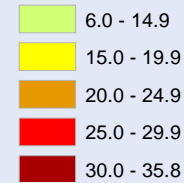
Model A

Percent of Population

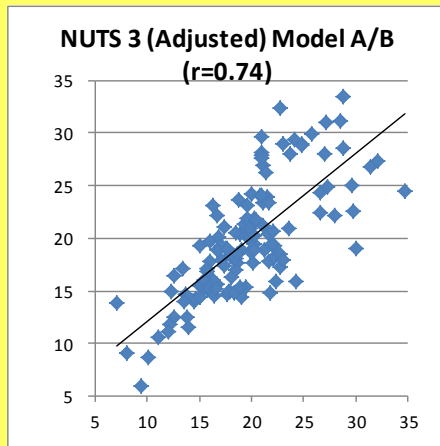
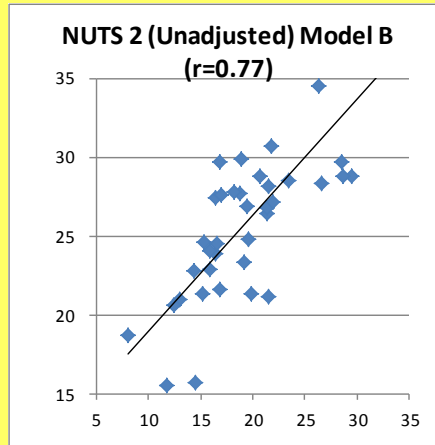
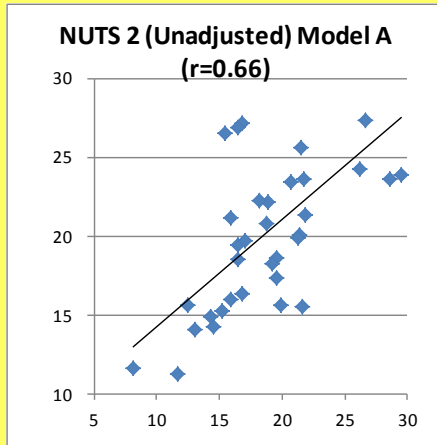


Model B

Percent of Population



Preliminary (Technical) Assessment



- Visually, the patterns shown by the two maps show many similarities.
- At NUTS 2 level Model B (unadjusted) is a closer fit to the original SILC data.
- AT NUTS 3 relationship between (adjusted) Model A and Model B estimates is $r=0.74$.
- Key test will be to compare with the results from WB PovMap.

The Way Ahead...

- Compare with WB PovMap results
- 3 year average SILC data
- Move to 2011 analysis
- Other countries – Macro region mapping?
- Comparison with Register-based maps in Sweden...

Policy Utility:

- Could provide timely intelligence on regional patterns
- Microdata publication dates may be too late for the policy requirement
- A trade-off between sophistication and transparency of methodology



And finally...

Key message: “Sledgehammers to crack nuts”

Wherever possible, without compromising the quality and integrity of outputs, researchers have a responsibility to provide the policy community with methodological nutcrackers (rather than sledgehammers), which are as transparent and easy to replicate as possible.

A trade-off between sophistication of analysis and timeliness, transparency and transferability?

Is 1 Rolls Royce < 28 VW Golfs?

Thank you for your attention.

