Evaluating the fiscal and distributional impact of public policies at local level using EUROMOD and spatial microsimulation techniques

Francesco Figari (U Insubria) and Manos Matsaganis (Polytechinc U Milan)
with Dimitris Ballas (U Groningen) and Anastasia Panori (Aristotle U Thessaloniki)

EUROMOD Research Workshop 2021

6 October 2021

Background

- Relevance of social outcomes and policy impacts at local level
  - “The revenge of the places that don’t matter” (Rodrìguez-Pose 2018)
  - Socio-spatial effects and patterns of macro-economic shocks and tax-benefit policies with impact on inequality and social cohesion (Cassiers and Kesteloot, 2012)
  - Geographical outcomes of government actions and forecasts at the local level (Openshaw, 1995: 60)
  - Existing gap between geography and social policy (Whitworth, 2019) but increasing devolution of fiscal and social responsibilities at local levels

Motivations

- There is no dataset of individuals and households which can be used both

  (i) to explore spatial variations in living conditions and behaviour

  (ii) to monitor the effects of changes in taxation and social policies at local level

  - EU-SILC mainly at NUTS-1 (or NUTS-2 but not more disaggregation)

Potential solutions

- Spatially disaggregated population microdata through
  - A massive new survey

    - very costly

    - confidentiality problems
  - Spatial microsimulation modelling

    - synthetically reproduce households which look as similar as possible to the real ones

    - construction of small area microdata from combining samples, surveys and small area data

    - spatial microsimulation makes possible the analysis of public policies and their reforms through microsimulation at local level

Challenges

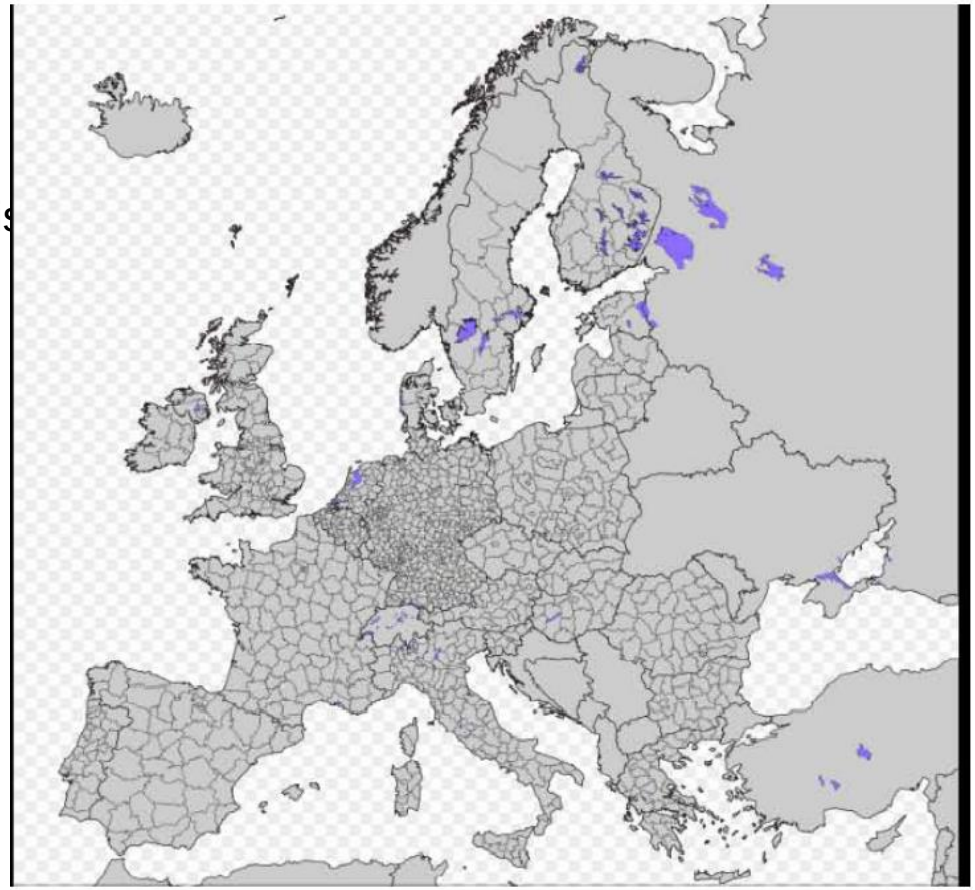
- Computational intensity: the incorporation of geography into standard microsimulation models increases significantly the computational demand
- Concerns with simulation accuracy
- Belief that geography is not important
- Unfamiliarity with geographical data and methods

Opportunities

- Linking the existing spatial microsimulation approach (Ballas et al. 2007, 2017) to tax-benefit microsimulation using EUROMOD
  - Apply in a cross-country perspective already existing open-source spatial microsimulation methodology (Lovelace and Ballas 2013)
  - Increase the scope of EUROMOD in terms of policy simulation (i.e. tax-benefit policies at local, sub-regional, level)
  - Enhance knowledge of distributional and policy effects at local level (e.g. municipality)

What does “local level” refer to?

- Local level or Small area as domain of interest, for which the sample size is not adequate to produce reliable direct estimates – in EU lower than NUTS2 level
- NUTS-3 level
  - Italy: 110 Provinces
  - Germany: 429 Districts (Kreis



Small Area Estimations

- Imputing into population census data an outcome variable (e.g. equivalised disposable income; poverty status) from household survey data—which has a sample to small for small area disaggregation
  - World Bank method based on regressions (Elbers et al. 2003)
  - M-quantile approach (Chambers and Tzavidis 2006; Giusti et al.)
  - Empirical Best Prediction approach (Molina and Rao, 2010)

- However, in order to adapt the multiple outcomes of a tax-benefit microsimulation model to small areas we need to retrieve the whole information set from surveys

(Static) Spatial Microsimulation

- Two sources of data
  - High-quality aggregate data with a high degree of accuracy and reliability (i.e. census)
  - Survey, nonspatial microdata with relevant information on income and tax-benefit system (i.e EU-SILC in EUROMOD)
- Estimation of small-area microdata through reweight of survey data to fit in small-area descriptions based on census using demographic and socioeconomic characteristics as “small-area constraint variables”
- Main “small-area constraint variables” (correlated with the target variables of the micro-simulation, e.g. income)
  - Age
  - Sex
  - Marital status
  - Education
  - Main economic activity

(Static) Spatial Microsimulation

- Reweighting approaches
  - Every record in the survey has a new weight for every small area to make it 'representative' of that small area.
    - E.g. the whole EU-SILC sample for Greece (or a sub-sample for Attica) is 'reweighted' to create a small area microdata set for every small area (matching the EU-SILC variables to the census variables)

(Static) Spatial Microsimulation

- Reweighting approaches
  - probabilistic, which typically reweight an existing national microdata set to fit a geographical area description on the basis of random sampling and optimisation techniques
  - deterministic, which reweight a non geographical population microdata set to fit small area descriptions, but without the use of random sampling procedures. Such approach uses the iterative proportional fitting (IPF) technique to give a weight to each individual, by adjusting for each constraint variable the initial weight through a reweighting algorithm

Examples of Spatial Microsimulation applications

- SimLeeds (Ballas, 2001)
- SimBritain (Ballas et al., 2005)
- SimAlba (Campbell and Ballas 2011
- SimKyoto (Ballas et al., 2011)
- Transport geography (Lovelace and Ballas, 2013)
- SimSheffield (Broomhead, Ballas and Baker, 2017)
- SimAthens (Panori, 2017)

SimBritain main data sources

- Census of UK population

- 100% coverage
- Fine geographical detail
- Small area data available only in tabular format with limited variables to preserve confidentiality
- Cross-sectional

- British Household Panel Survey:

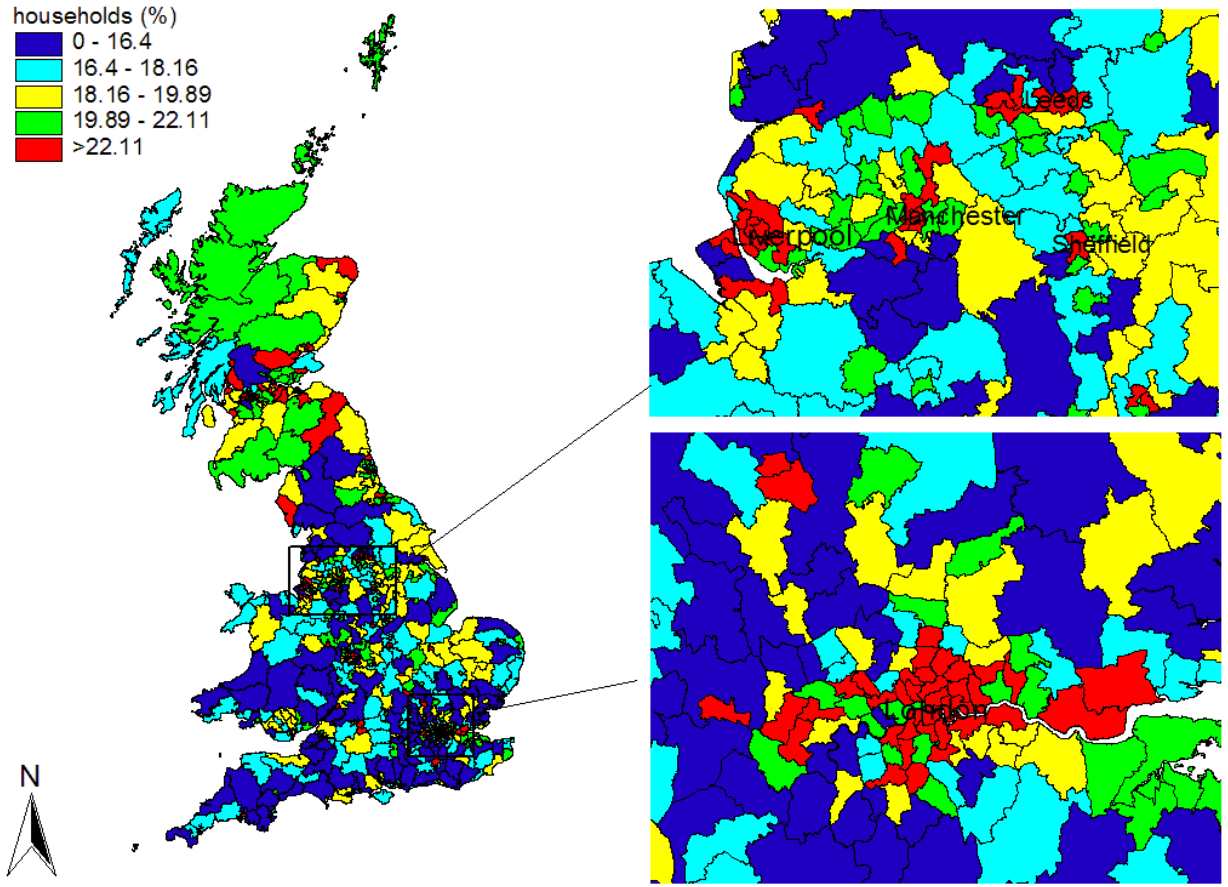
- Sample size: more than 5,000 households
- Annual surveys (waves) since 1991
- Coarse geography
- Household attrition

Ballas, D. , Clarke, G.P., Dorling, D., Eyre, H. and Rossiter, D., Thomas, B (2005) SimBritain: a spatial microsimulation approach to population dynamics, *Population, Space and Place* 11, 13–34 (http://dx.doi.org/10.1002/psp.351)

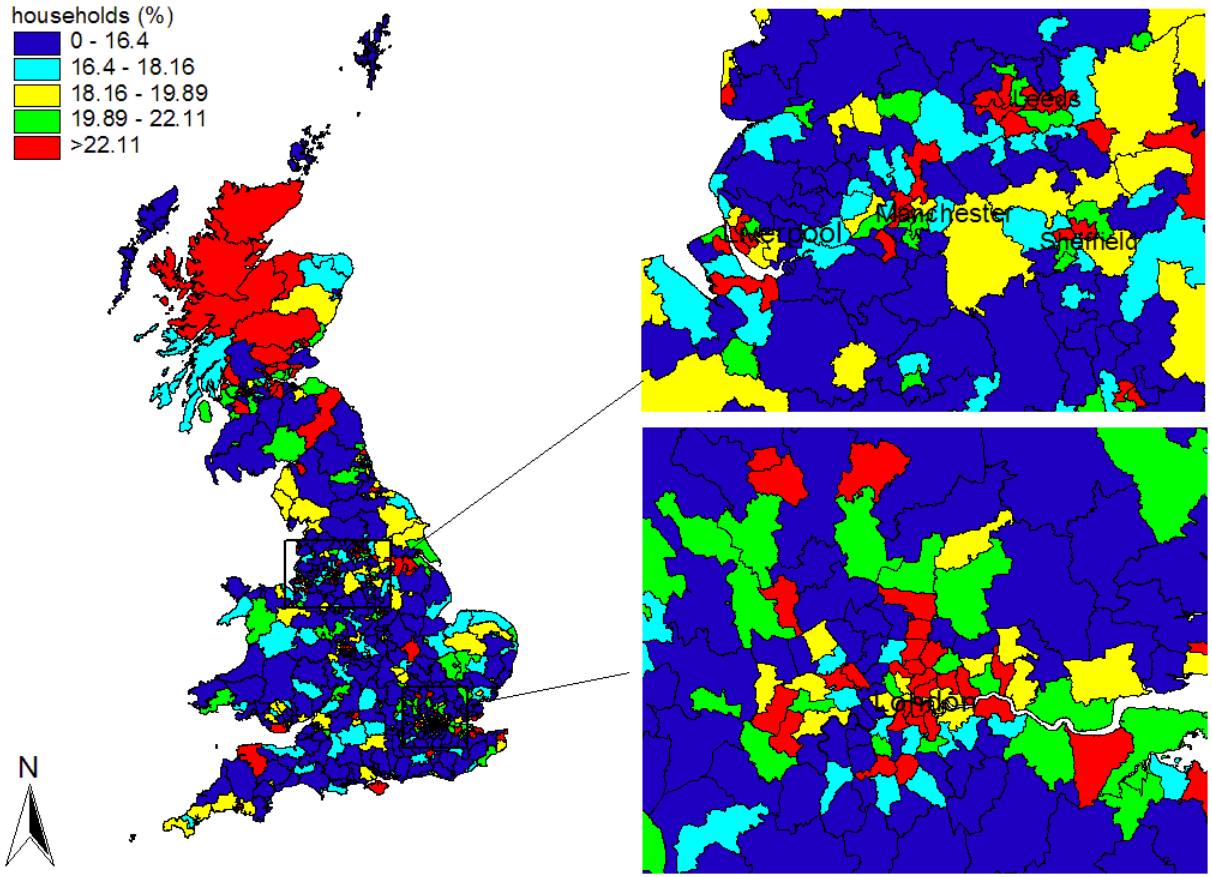
SimBritain

- Aim: reweight the first wave of the BHPS data to fit small areas
- Dynamically simulate this population for the years 2001, 2011, 2021
- Multiple outputs (e.g. spatial distribution of “poor” households)

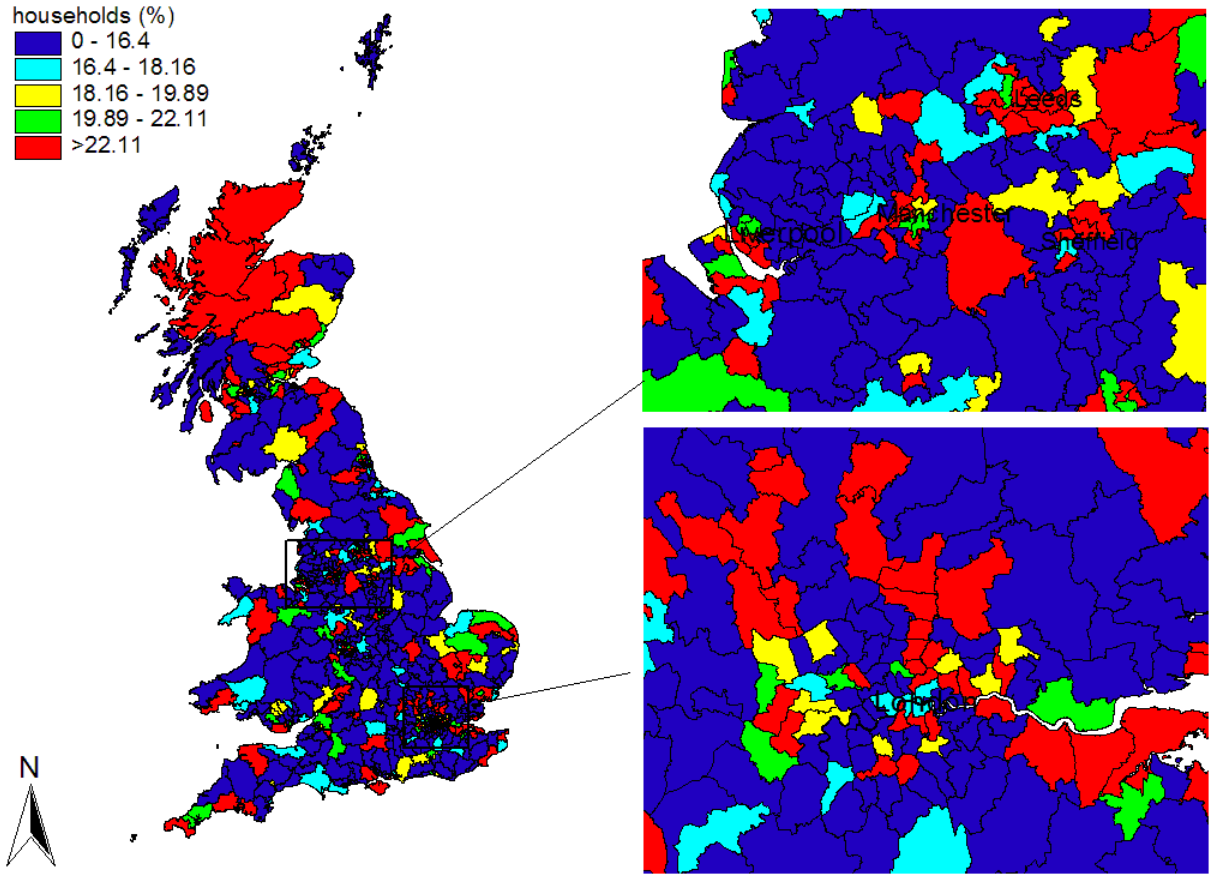
SimBritain: spatial distribution of "poor" households, 1991



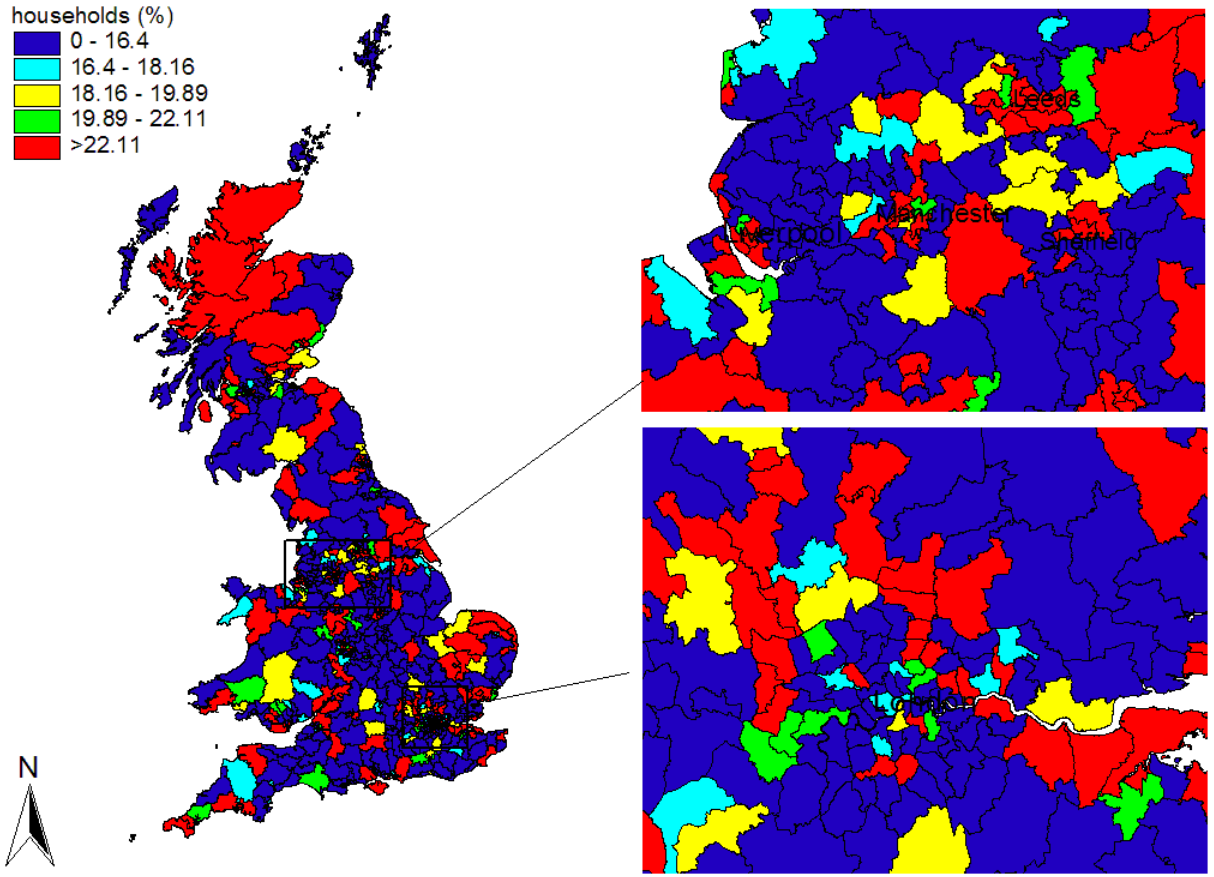
SimBritain: spatial distribution of “poor” households, 2001



Spatial distribution of “poor” households, 2011



Spatial distribution of “poor” households, 2021



SimAthens

- Main features:

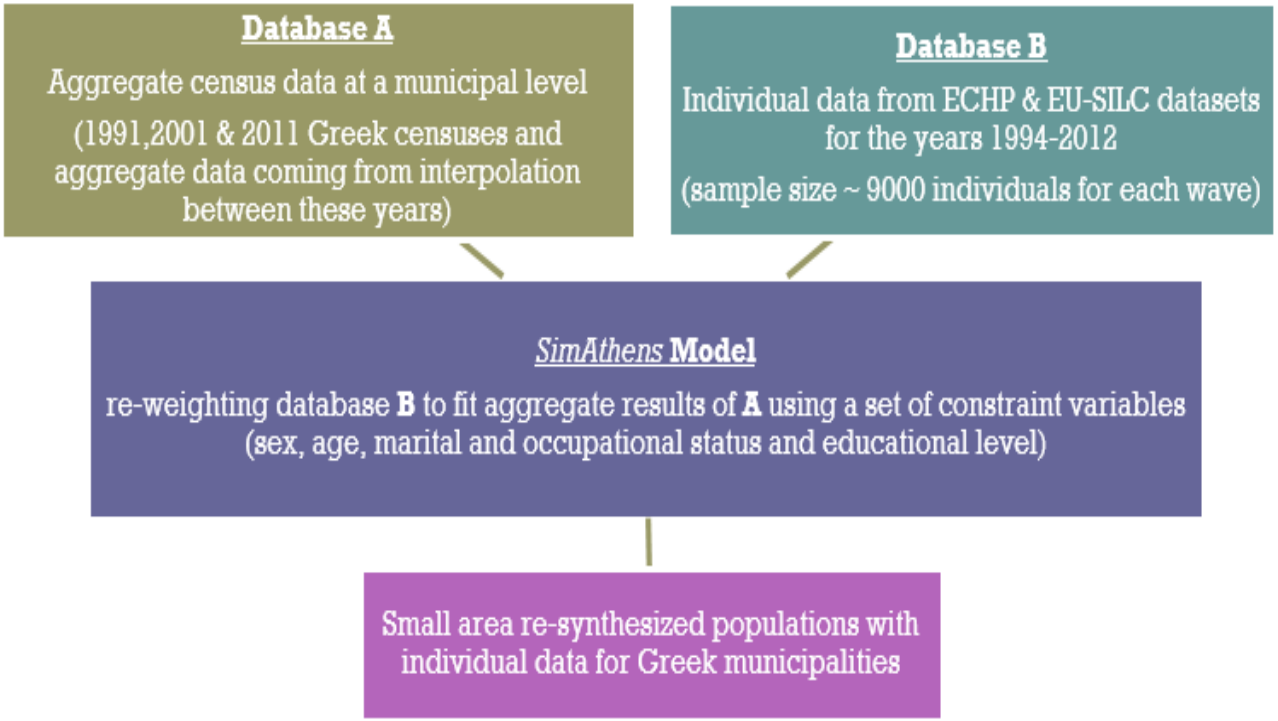
- Methodology: Static spatial microsimulation model

- Code: written in R – modification of already existing code produced by Lovelace and Ballas (2013).

- Data: combination of individual EU-SILC data (2006, 2011) with aggregate census data (2001, 2011).

Panori, A., Ballas, D., and Psycharis Y. *SimAthens*: *A spatial microsimulation approach to the estimation and analysis of small area income distributions and poverty rates in the city of Athens*, Computers, Environment and Urban systems (2016), http://dx.Doi.Org/10.1016/j.Compenvurbsys.2016.08.001

SimAthens

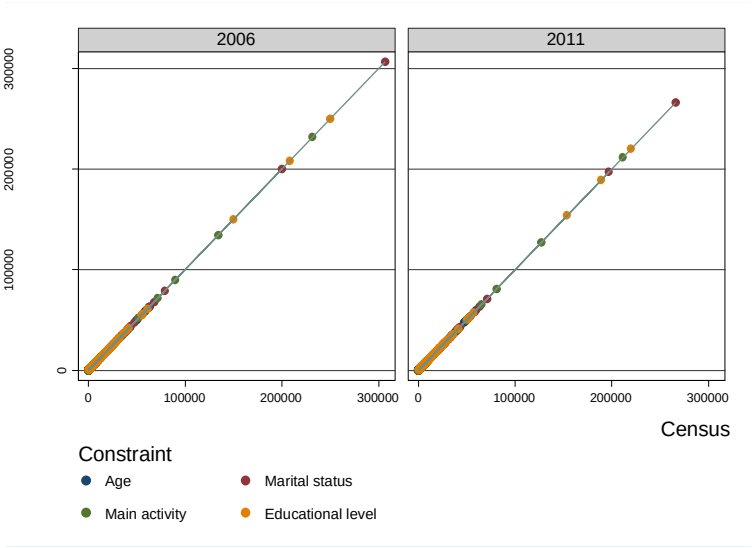


Panori, A., Ballas, D., and Psycharis Y. *SimAthens*: *A spatial microsimulation approach to the estimation and analysis of small area income distributions and poverty rates in the city of Athens*, Computers, Environment and Urban systems (2016), http://dx.Doi.Org/10.1016/j.Compenvurbsys.2016.08.001

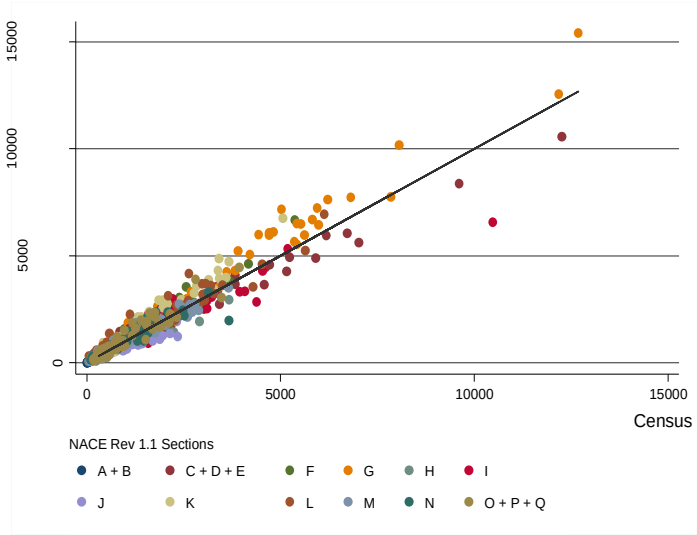
Model validation – Internal

Two types of validation have been used:

- Internal: Uses the constraint variables in order to check for the model fitting. It is usually affected by the selected re-weighting method.
- External: Uses a set of external variables, not used as constraints when building the model, in order to check for the fitting.



Model validation – External I
Labor market structure



Simulated versus actual census shares for labor market structure in Athens metropolitan area.

2006				2011			
NACE Rev. 1.1 sections	Census (%)	SimAthens (%)	Diff.	NACE Rev. 2 sections	Census (%)	SimAthens (%)	Diff.
$A+B$	0.54	0.53	−0.01	A	0.66	0.53	−0.13
$C+D+E$	15.29	14.56	−0.73	$B-E$	11.27	14.88	3.61
F	7.99	7.61	−0.38	F	6.51	7.28	0.77
G	18.40	21.41	3.01	G	19.04	21.96	2.92
H	4.98	4.09	−0.89	H	7.01	4.71	−2.3
I	9.35	9.22	−0.13	I	5.84	4.44	−1.4
J	4.63	3.21	−1.42	K	4.30	5.52	1.22
K	9.15	10.64	1.49	$L-N$	10.74	9.92	−0.82
L	9.80	10.42	0.62	O	10.32	8.69	−1.63
M	6.60	6.32	−0.28	P	7.17	7.42	0.25
N	5.94	5.53	−0.41	Q	7.02	5.50	−1.52
$O+P+Q$	7.32	6.46	−0.86	$R-U+J$	10.12	9.15	−0.97

Source: Authors' calculations.

Model validation – External II
Occupational structure

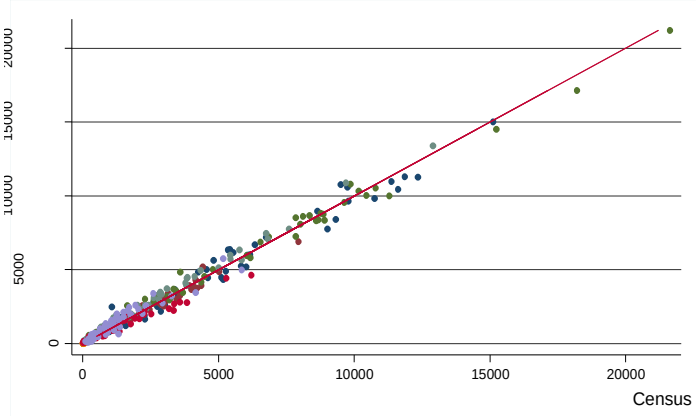
Simulated versus actual census shares for occupational structure in Athens metropolitan area.

ISCO - 88	2006		
	Census (%)	SimAthens (%)	Diff.
Legislators, senior officials and managers & Professionals	26.76	26.97	0.21
Technicians and associate professionals	11.21	11.13	−0.08
Clerks & Service workers and shop and market sales workers	30.25	30.42	0.17
Skilled agricultural and fishery workers	0.71	0.69	−0.02
Craft and related trade workers	15.61	16.87	1.26
Plant and machine operators and assemblers	6.96	6.25	−0.71
Elementary occupations	8.50	7.67	−0.83

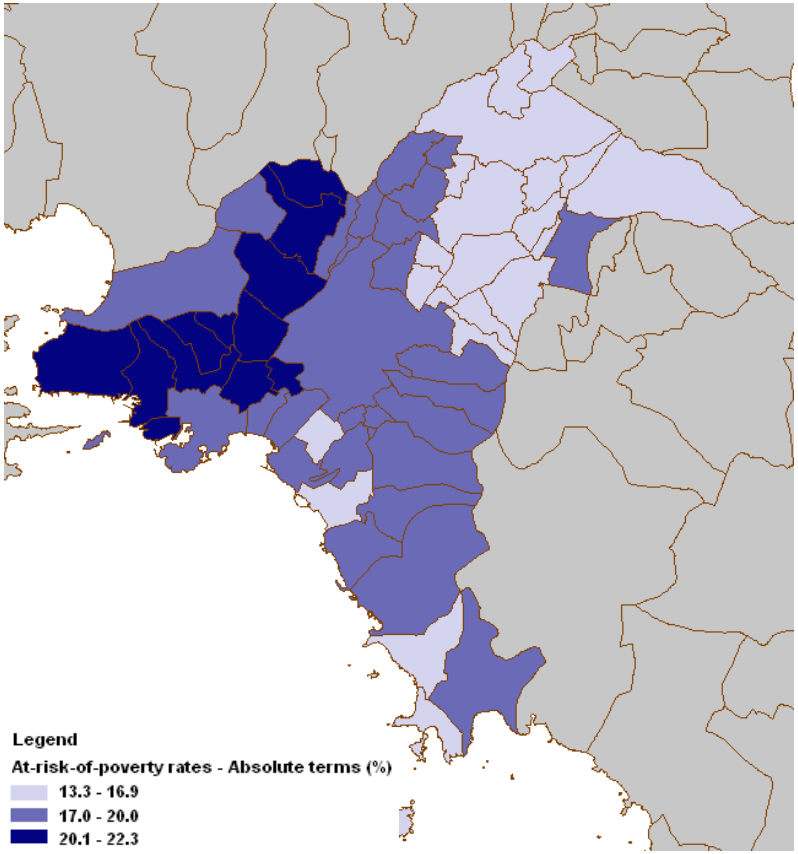
Source: Authors' calculations.

ISCO - 88	2011		
	Census (%)	SimAthens (%)	Diff.
Legislators, senior officials and managers & Professionals	28.00	23.45	−4.55
Technicians and associate professionals	11.63	9.29	−2.34
Clerks & Service workers and shop and market sales workers	33.23	33.45	0.22
Skilled agricultural and fishery workers	0.93	1.74	0.81
Craft and related trade workers	11.46	14.46	3.00
Plant and machine operators and assemblers	6.07	8.08	2.01
Elementary occupations	8.68	9.53	0.85

Source: Authors' calculations.



At-Risk-of-Poverty (%) – 2006



Absolute terms: poverty line set as 60% of the median equivalized income of the total metropolitan are of Athens

A forward looking approach

- What if we combine EUROMOD with Spatial Microsimulation approach?
  - Is it relevant in some\all EU countries?
    - Spatial disaggregated socio-economic inequalities
    - Devolution of tax-benefit policies
  - Is there an interest from national teams\JRC?
  - Are census data easily accessible ?