ARTICLE IN PRESS
CEUS-01074; No of Pages 11
Computers, Environment and Urban Systems xxx (2016) xxx–xxx
Contents lists available at ScienceDirect
Computers, Environment and Urban Systems
journal homepage: www.elsevier.com/locate/ceus

SimAthens: A spatial microsimulation approach to the estimation and analysis of small area income distributions and poverty rates in the city of Athens, Greece

Anastasia Panori [a,*], Dimitris Ballas [b,c], Yannis Psycharis [a]

[a] Department of Economic and Regional Development, Panteion University, 136 Syngrou Avenue, 176 71 Athens, Greece
[b] Department of Geography, University of the Aegean, University Hill, Mytilene 81100, Greece
[c] Department of Geography, University of Sheffield, Winter Street, Sheffield S10 2TN, UK

ARTICLE INFO

Article history:
Received 16 October 2015
Received in revised form 13 July 2016
Accepted 2 August 2016
Available online xxxx

Keywords:
Spatial microsimulation
Small-area microdata
Small-area income data
Inequalities

ABSTRACT

Published during a severe economic crisis, this study presents the first spatial microsimulation model for the analysis of income inequalities and poverty in Greece. First, we present a brief overview of the method and discuss its potential for the analysis of multidimensional poverty and income inequality in Greece. We then present the *SimAthens* model, based on a combination of small-area demographic and socioeconomic information available from the Greek census of population with data from the European Union Statistics on Income and Living Conditions (EU-SILC). The model is based on an iterative proportional fitting (IPF) algorithm, and is used to reweight EU-SILC records to fit in small-area descriptions for Athens based on 2001 and 2011 censuses. This is achieved by using demographic and socioeconomic characteristics as constraint variables. Finally, synthesis of the labor market and occupations are chosen as the main variables for externally validating our results, in order to verify the integrity of the model. Results of this external validation process are found to be extremely satisfactory, indicating a high goodness of fit between simulated and real values. Finally, the study presents a number of model outputs, illustrating changes in social and economic geography, during a severe economic crisis, offering a great opportunity for discussing further potential of this model in policy analysis.

1. Introduction

Over the past four decades, there have been a rapidly growing number of multidisciplinary efforts to investigate the main aspects of poverty through a wider perspective. In particular, since the 1980s, a new framework has been developed, based on the theoretical work of Sen (1983, 1992, 1999), who investigated poverty under the perspective of capability approach. According to this approach, poverty is considered to be caused by not only economic factors but also various components covering wider notions of development such as health, education, and living conditions.

Moving away from conventional poverty measures that include purely income-based indicators and examining poverty under a multidimensional perspective can be challenging due to the paucity of suitable secondary data sets at the small-area level. In order to develop new indicators that attempt to combine and interpret different dimensions of poverty, there is a need for detailed socioeconomic micro-data sets to be collected via social surveys. An example of a comprehensive survey of this type is the census of population, which generates very useful socioeconomic and demographic information for small areas and which has been typically the basis for the development of widely used indexes of deprivation such as the Townsend indicator (Townsend, 1979). Nevertheless, the census questionnaires cover a relatively limited set of themes and in most cases, they do not include any information on income, wealth, and other variables reflecting socioeconomic circumstances in order to preserve confidentiality and minimize nonresponse (Marsh, 1993).

Spatial microsimulation has been gaining prominence as an appropriate method of estimation of small-area microdata that can be used for the analysis of interdependencies between different household and individual characteristics. This method is particularly suitable for bridging the gap between the innovative, human-based theoretical frameworks to examine poverty and the difficulty of implementing it at a small-area level, because of lack of data. Adding geographical information to microlevel data with the use of spatial microsimulation analysis allows for a small-area approach in policy analysis. By using this approach, the distributional impact of implementing different socioeconomic policies could be estimated at a microlevel (Ballas, Rossiter, Thomas, Clarke, & Dorling, 2005; Ballas, Clarke, & Dewhurst, 2006; Callan, 1991).

Microsimulation models have a long history and tradition in economics, originating in the work of Orcutt (1957) who indicated the

* Corresponding author.
E-mail addresses: an.panori@panteion.gr (A. Panori), d.ballas@aegean.gr, d.ballas@sheffield.ac.uk (D. Ballas), psycharis@panteion.gr (Y. Psycharis).

http://dx.doi.org/10.1016/j.compenvurbsys.2016.08.001
0198-9715/

Please cite this article as: Panori, A., et al., SimAthens: A spatial microsimulation approach to the estimation and analysis of small area income distributions and poverty rates ..., *Computers, Environment and Urban Systems* (2016), http://dx.doi.org/10.1016/j.compenvurbsys.2016.08.001

importance of determining the relationship between parameters used in a socioeconomic model and the aggregate results. During the 1970s, the first microanalytic models were built to simulate socioeconomic systems and investigate their behavior under various policy implication scenarios (Kain & Apgar, 1985; Orcutt, Caldwell, & Wertheimer, 1976).

Although these initial efforts offered a whole new perspective on the way in which aggregate data should be approached, they were aspatial, as they did not include any geographical dimension or perspective. The necessity of incorporating the spatial context was first highlighted by Hägerstrand (1957), who treated time and space as inseparable notions that affect individual's decision making at daily, yearly, or lifetime scales of observation (PRED, 1977). The first implementation of such type of model is the work of Wilson and Pownall (1976), which inspired a series of surveys focusing on the field of regional development (Birkin & Clarke, 1988, 2011). Other domains in economics where spatial microsimulation has been implemented are labor market (Campbell & Ballas, 2013; Ballas et al., 2005; Ballas et al., 2006) and education (Kavroudakis, Ballas, & Birkin, 2012; Kavroudakis & Dimitris Ballas, 2013).

Of particular relevance to the work presented is this study is a comparative study of the social geography of two major cities in Japan and Britain, which involved an estimation of small-area microdata using spatial microsimulation (Ballas et al., 2012). Also of relevance is the work of researchers who developed and implemented statistical small-area estimation approaches involving complementing social survey microdata, such as the European Union Statistics on Income and Living Conditions (EU-SILC) with administrative sources, in order to calculate income and poverty measures based on the idea of utilizing regression models (Fay & Herriot, 1979; Elbers, Lanjouw, & Lanjouw, 2003; Nagle, Sweeney, & Kyriakidis, 2011; Pereira & Coelho, 2013; Fabrizi et al., 2014).

The EU-SILC database has proven to be an effective tool, which works as a basis upon which various microsimulation models have been developed. In most cases, the spatial level of analysis remains at a country level, and microdata are used to assess the effect of policy changes, especially referring to national tax benefit systems (Sutherland & Figari, 2013; Betti, Donatiello, & Verma, 2010; O'Donoghue, Loughrey, & Morrissey, 2013).

This study describes the development of the first spatial microsimulation model in Greece that combines census and social survey data, followed by an extensive external validation process, using labor market and occupations structure data. This model aims at estimating small-area poverty measures, including a multidimensional poverty index (MPI) within Athens, before and after the economic crisis. An indicative analysis of the main poverty components, as long as the way in which they are affected by recent ongoing economic crisis in Greece, is also performed. This analysis highlights the necessity of using innovative techniques and methodologies for simulating and assessing policy decisions at a microlevel.

2. Methodology and data

As mentioned in the previous section, there have only been a handful of studies that attempted to investigate poverty under a multidimensional context at a geographical scale lower than that of countries or NUTS1 regions (Alkire, Apablaza, & Jung, 2014; Miranti et al., 2010; Harding et al., 2006; Tanton et al., 2009). In most cases, the results illustrate that geography matters even at a high level of spatial analysis. There is also a growing literature that refers especially to urban poverty and the importance of structural clusters that are geographically defined within large spatial agglomerations (Amis, 1995; Glaeser, 1998; Wratten, 1995; Satterthwaite, 1997; Moser, 1998). Thus, the need to focus our research on smaller area levels becomes evident, highlighting the necessity of implementing innovative techniques and methodologies.

2.1. Data, methods and scales of analysis

In order to build a static spatial microsimulation model, two main sources of data are essential: aggregate data at the spatial level to be used for the analysis and nonspatial microdata. The main idea is to use the existing high-quality aggregate data that have a high degree of accuracy and reliability, such as those derived by national censuses, based on which small-area microdata fitting is acquired, resulting in resynthesized small-area populations.

In this study, the metropolitan area of Athens is used as the main case study. Aggregate data for its 59 municipalities are derived from the last two national censuses (2001 and 2011). The choice of municipalities as the main areal unit is based on the fact that municipality is the lower administrative level at which aggregate data can be found in national censuses. Their size is appropriate to perform spatial analysis in most cases, as they are not too small, leading to a large number of areal units and thus high complexity in calculations. On the contrary, using larger areal units for the analysis could potentially gloss over spatial differentiations within the metropolitan area, reducing its accuracy.

The EU-SILC database was the most appropriate main source of microdata, because of the nature of this research, high quality of the database, and the relative paucity of other relevant survey microdata in Greece. This database contains a large number of parameters referring to economic and social conditions of EU countries. However, although it offers an extremely rich variety of variables measured every year that are suitable for poverty analysis, it does not provide geographical information when descending to lower spatial levels of analysis, limiting its use to a country or NUTS1 level.

In order to make a comparative analysis before and after the economic crisis affecting Greece over the past six years, 2006 and 2011 were chosen as the two reference years, and thus the corresponding EU-SILC waves were used. Furthermore, an initial assumption that has to be mentioned is that the 2001 census data were used as aggregate basis for 2006 microdata. It has long been argued (e.g., see Rees, Martin, & Williamson, 2002) that when it comes to accuracy and geographical coverage, census data are considered to be the "gold standard." Thus, in order to take advantage of this important strength of census data, it is assumed that between 2001 and 2006, there would have been only small changes in demographic characteristics of the areal units being used here, leaving aggregate data almost unaffected.

The choice of variables being used to constrain a spatial microsimulation model is one of the key factors that play an important role in the process of model building. Different constraint variables may lead to considerable variation in the synthetic populations being produced and thus different results (Edwards et al., 2010; Ballas et al., 2007; Burden & Steel, 2015). The first step when making a decision regarding which small-area variables (known as "small-area constraints") should be used as constraints in spatial microsimulation is to examine the extent to which there is a correlation between these variables with the so-called "target" variables of the simulation (outputs – e.g., income).

Within the context of the research presented in this study, the main variables used for constraining the spatial microsimulation model are age/sex, marital status, education, and main economic activity status. All of them were selected in terms of creating a comprehensive picture of spatial units' demographic characteristics and because they are widely accepted as good indicators of an individual's socioeconomic condition.

2.2. Spatial microsimulation

Static spatial microsimulation methods generate simulated microdata for small-area populations at a given time point. Depending on the reweighting method upon which they are based, they can be categorized as either probabilistic or deterministic approaches (Ballas et al., 2005). In the first case, a combinatorial optimization approach