



## Origin-destination geodemographics for analysis of travel to work flows



David Martin <sup>a,\*</sup>, Christopher Gale <sup>a</sup>, Samantha Cockings <sup>b</sup>, Andrew Harfoot <sup>b</sup>

<sup>a</sup> Administrative Data Research Centre – England, University of Southampton, Southampton SO17 1BJ, United Kingdom

<sup>b</sup> Geography and Environment, University of Southampton, Southampton SO17 1BJ, United Kingdom

### ARTICLE INFO

#### Article history:

Received 5 November 2016

Received in revised form 24 February 2017

Accepted 7 September 2017

Available online xxxx

#### Keywords:

Travel to work

Workplace zones

Area classification

Census

Geodemographics

England and Wales

### ABSTRACT

This paper introduces a novel approach to the analysis of travel to work flows by combining separate geodemographic classifications of origins and destinations. A new classification of workplace areas is used in combination with an established official classification of residential areas. The approach is demonstrated using an empirical analysis of 26 million commuting flows in England and Wales, measured between the smallest residential and workplace areas in the 2011 census. The analysis demonstrates potential insights to be gained by this approach, revealing clear patterns in the structure of travel to work flows between geodemographic clusters. Our broad approach is not limited to use in specific countries and has potential application for use with data from non-census sources.

© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### 1. Introduction

In this paper we demonstrate a novel method for the analysis of spatially detailed, aggregate travel to work data by combining separate geodemographic classifications of trip origins and destinations and explore the flows between the clusters in these classifications. The generation and use of a separate, workplace-based, classification is, in itself, new. We apply this origin-destination approach to the 26 million travel to work flows recorded in 2011 census data for England and Wales. The representation and analysis of travel to work is important for understanding local labour markets, economic development, transport planning, daytime service delivery and more general insight into the factors underlying population mobility. Traditional sources such as censuses provide enormously powerful and complex travel to work data (Stillwell, Duke-Williams, & Dennett, 2010) which present challenges for analysis. These types of interaction data are available in many countries where censuses include a question about place of work, thereby providing a second georeferencing frame, additional to the place of usual residence which forms the basis for most small area census statistics. Sources of travel to work data range from long-established travel surveys and census microdata to new forms of real-time data with exciting potential, but none of these presently offer the combination of open access with full population coverage and detailed socioeconomic and spatial characteristics provided by census interaction data.

Complexity in travel to work data from all sources arises primarily from the large number of interactions, set within a very sparse origin-destination matrix. Rae (2016) presents an overview of the geography of travel to work using 2011 census data, focusing on the geovisualization of around 2.4 million small area interactions in England and Wales. The key challenge is one of data reduction, and spatial visualizations are constructed which allow important features of the data to be more readily seen and understood. The results clearly present the geometry and magnitude of interactions but do not capture the social characteristics of the origin and destination areas. Geodemographic classification is a form of area classification and a powerful data reduction tool (Leventhal, 2016) which has found many applications, including market analysis and service planning. In this case, complexity arises primarily from the large pool of variables which are available to describe each small area. The use of geodemographic classification has become widespread but it has almost always been applied only to data for residential neighbourhoods.

Some previous studies (Debenham, Clarke, & Stillwell, 2003; Manaugh, Miranda-Moreno, & El-Geneidy, 2010) have attempted to characterise travel to work flows by geodemographic classification of origin and destination areas. These have however been limited to classifications that are based on the same source data and spatial units for the locations of both residence and employment. The novelty of the approach proposed here is to combine separate classifications of places of work and residence, each based on the most appropriate geographical units and variables, exploiting separate small area geographies optimised for residential and workplace-based data products. We have used the official geodemographic classification of residential

\* Corresponding author.

E-mail address: [D.J.Martin@soton.ac.uk](mailto:D.J.Martin@soton.ac.uk) (D. Martin).

areas and undertaken an additional classification of workplace zones and here, for the first time, use the two classifications in combination.

The following section provides a brief review of geodemographic classification and analysis of travel to work, with particular reference to combination of relevant methods. The third section describes a new Classification of Workplace Zones for England and Wales (COWZ-EW), and the data and methods to be used here. We then present a series of example analyses made possible by combining residential and workplace-based classifications. The aim is not a comprehensive analysis of travel to work patterns, but demonstration of a new way to present and understand travel to work using origin-destination geodemographics.

## 2. Geodemographics and travel to work

Geodemographics refers to the classification of small areas based on reduction of multivariate aggregate data, typically involving a series of data preparation steps such as transformation and standardization, prior to application of some form of cluster analysis (Leventhal, 2016). Use of the term can be traced from the 1970s (Gale & Longley, 2013). We here follow the convention adopted in ONS (2015), Leventhal (2016) and Gale, Singleton, Bates, and Longley (2016) and appearing in the documentation of our data sources, which refers generically to geodemographic data products as “classifications” and individual groupings as “clusters”, regardless of the specific algorithm employed in their construction. These classifications seek to aggregate areas based on their similarities and emerged as a methodological solution for handling highly dimensional census data (Webber, 1977). They are conventionally applied to data representing the characteristics of residential populations, reflecting the fact that most geographical referencing is to residential addresses, whether from censuses, commercial or administrative sources. Geodemographics systems are internationally widespread and Poppie and Miller (2016) review the contemporary US scene, identifying a similar development trajectory to that in the UK. Output Area Classifications (OAC) for 2001 and 2011 were developed entirely from UK small area census data and published as open data by the Office for National Statistics (ONS), the national statistical organisation for England and Wales (Gale et al., 2016; Vickers & Rees, 2007). In addition to these official statistics, many classifications have been produced as commercial data products incorporating non-census sources, such as electoral registers, vehicle registrations, county court judgements, credit reference agency and lifestyle data (Leventhal, 2016), often in conjunction with census data. There are many data-informed but subjective design decisions required for any such geodemographic system, and the detailed methods and source datasets are not usually published for the commercial data products.

The 2011 OAC is a geodemographic classification of residential areas created using 2011 census data for the smallest available output areas (OAs). Detailed accounts are provided in ONS (2015) and Gale et al. (2016), describing a similar methodology to that used for the 2001 OAC (Vickers & Rees, 2007). For the 2011 OAC multiple methods were tested and evaluated, with particular emphasis on exploring how interactions between different methods and techniques influenced the final cluster solutions. Predicating this was a decision to use only open source software and the open release of all outputs and code through the website <http://www.opengeodemographics.com>. In total, 60 variables covering demographic, household composition, housing, socio-economic and employment domains were used. These were standardized, normalised and clustered using a *k*-means algorithm to create a three-tiered nested structure comprising 8 Supergroups, 26 Groups and 76 Subgroups. The 2011 OAC is a key input to the analysis presented here.

Interest in analysis of commuting patterns to better understand aspects of, for example, urban spatial structure, employment and gender imbalances, home and work locations, commuting distances or excess commuting is internationally widespread. Recent studies employing a

range of approaches to these issues include Sohn (2005), Kim, Sang, Chun, and Lee (2012) and Niedzielski, O’Kelly, and Boschmann (2015) in the USA, Manaugh et al. (2010) in Canada, Novak, Ahas, Aasa, and Silm (2013) in Estonia, O’Kelly, Niedzielski, and Gleeson (2012) in Ireland and Hincks (2012) in England. All pursue, in different ways, the description and analysis of travel to work flows between residential and business districts. Further, while many residential areas host very little work activity, in other areas work activity can reach enormously high concentrations, with many thousands of workers and sometimes quite different work activities occupying the same spatial units. A common challenge when using aggregate data is the inherently different spatial characteristics of places of residence and work. Data which differentiate residential areas well, including for example housing and household characteristics, are not the most appropriate for characterisation of workplaces. Most studies have used census residential geographical units to analyse variables for both origins and destinations of work trips, limiting the ways in which workplace destinations are demarcated.

Several families of purpose-specific geographical units such as Travel to Work Areas (TTWAs) in the UK (Coombes & Bond, 2008; ONS, 2016), POWCAR commuting catchments in Ireland (O’Kelly et al., 2012) and Metropolitan and Micropolitan Statistical Areas in the USA (Adams, VanDrasek, & Phillips, 1999; Office of Management and Budget, 2010) have been developed to capture the structure of travel to work, using statistical measures to demarcate regions in which there is a degree of self-containment between resident and workforce populations. Australian Bureau of Statistics (2011) and Federal Highway Administration (2010) describe additional geographical units (destination zones and traffic analysis zones, in Australia and US, respectively) created as more appropriate destination zones for transportation modelling than the standard residence-based census output units. However, these areas are all intended to encompass local commuting systems and therefore are designed at a scale too coarse to characterise differences between very small origin and destination areas.

It is attractive to use data which directly describe commuters, rather than the areas to and from which they are travelling. Kim et al. (2012) and Niedzielski et al. (2015) make use of US 2000 Census Transportation Planning Package (CTPP) data allowing commuting patterns to be broken down by occupation and gender or race/ethnicity and income respectively. Debenham et al. (2003) employ UK Special Workplace Statistics (SWS) and limited non-census workplace data but focus on a single set of medium-sized spatial units, in this case postcode sectors (mean population approximately 5100). CTPP and SWS-type products are effectively the richest of the aggregate datasets available from conventional census-based sources but do not contain the full range of census characteristics found in small area aggregate data. Where data about individual commuters are available in census microdata files such as the Irish POWCAR data (O’Kelly et al., 2012) or ONS controlled access microdata samples these are subject to tightly controlled access conditions and can rarely be used at the most detailed spatial scales. We propose that there is much still to be learned from further analysis of the rich aggregate data about flows between small areas.

Duke-Williams (2010) and Dennett and Stillwell (2011) describe a bespoke area classification for the study of migration flows, while Singleton, Pavlis, and Longley (2016) address the challenge of comparing geodemographic classifications between two censuses. In both of these cases, the focus is on understanding interactions between classifications of residential areas. Travel to work is subtly different, in that it relates to flows from residential areas, typically characterised by housing and family characteristics, to workplace areas, characterised by employee and business characteristics and it is with these flows that our interest lies here. Very few studies have attempted to employ geodemographic classification as a means of describing the origins and destinations of travel to work flows.

Hincks (2012) aggregates census travel to work data from the 2001 census in North West England to sub-regional housing market areas and

Download English Version:

<https://daneshyari.com/en/article/4965107>

Download Persian Version:

<https://daneshyari.com/article/4965107>

[Daneshyari.com](https://daneshyari.com)