Computers, Environment and Urban Systems 67 (2018) 68–79

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Origin-destination geodemographics for analysis of travel to work flows

David Martin [a,*], Christopher Gale [a], Samantha Cockings [b], Andrew Harfoot [b]

[a] Administrative Data Research Centre – England, University of Southampton, Southampton SO17 1BJ, United Kingdom
[b] 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.

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 (D. Martin).

https://doi.org/10.1016/j.compenvurbsys.2017.09.002
0198-9715/

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

TTWAs. He develops a six-way urban-rural classification of wards which is then applied to describe both origins and destinations. Analysis proceeds by looking at the pattern of the dominant, then second and third order flows into each TTWA, the latter tending to capture the longer commuting patterns. This is an interesting approach, allowing the strength of association between different area types to be measured, but is limited by the relatively large areal units and the same residential-based area classification being applied to places of both residence and employment. Nevertheless, the analysis reveals insights into the structure of the regional labour market. Area classification as a means of interpreting travel to work flows is most fully implemented by Manaugh et al. (2010) whose Montreal study uses a combination of travel survey and census data. They apply a factor-cluster analysis using a range of variables covering both socioeconomic characteristics and measures of urban form. They achieve a fine spatial scale, applying the classification to 150 x150m grid cells and undertake separate factor and cluster analyses on home and work locations, albeit using the same pool of input variables. However, the locations classified are limited to those cells containing home or workplace of individuals in the travel survey. The specific objective of their analysis is to model trip length using regression methods. We here propose a more generally applicable approach to the analysis of travel to work flows by combination of a residential classification of travel to work origin areas and a new classification of workplace destination areas, implemented at a national scale using the smallest available areal units for each.

3. Methodology and data

Our empirical study is based on the 2011 Census of England and Wales. It exploits the 2011 OAC and a new family of data products for a second set of small areas known as workplace zones (WZs), developed specifically for the reporting of census workplace data (Martin, Cockings, & Harfoot, 2013). Using these workplace data, we have developed, in collaboration with ONS, a new geodemographic classification of workplace zones, known as the Classification of Workplace Zones – England and Wales (COWZ-EW). COWZ-EW comprises one half of the data input for our origin-destination geodemographic analysis. In this section we set out the data used and methods employed, including some necessary explanation of areal units.

OAs are the smallest units for which residential census data are published in England and Wales. They were originally generated in 2001 using an automated zone design procedure, and updated where necessary for 2011 (Cockings, Harfoot, Martin, & Hornby, 2011) to reflect changes in local government boundaries or underlying population distribution, although 97.4% were unchanged. They were explicitly designed to exceed residential thresholds of 100 persons and 40 households and the resulting zones have a mean residential population of 325. For the 2011 census, a second, entirely new, set of 53,578 WZs (compared with 181,408 OAs) was created, based on respondents' reported places of work, as described in Martin et al. (2013) and Mitchell (2014). Some WZs have the same boundaries as OAs but many OAs have been aggregated or subdivided to produce WZs with minimum thresholds of 200 workers and three workplace postal codes.

Although place of work questions have been included in previous censuses (and used to generate travel to work flow matrices), the very small workplace populations in many 2001 OAs restricted the publication of workplace population data to just four univariate tables, insufficient for meaningful small area classification. The introduction of WZs in 2011 ensured sufficiently sized workplace populations to permit publication of 21 data tables at the WZ level, covering a broad range of demographic and socioeconomic characteristics, as well as Standard Industrial Classification of workplaces and journey to work characteristics such as mode and distance travelled. These data provided the opportunity to undertake a new classification of WZs, termed COWZ-EW. Cockings, Martin, and Harfoot (2015) provide a more detailed account than possible here, including full listing of source variables.

As already reviewed, geodemographic classification may be achieved using a wide range of methods. In order to provide consistency between the two classifications, COWZ-EW adopted essentially the same approach as the official 2011 OAC, employing a sequence of domain identification, selection and refinement of variables, normalisation, standardization and k-means cluster analysis. At each stage, different options were evaluated, such as the inclusion/exclusion of variables, alternative methods of normalisation and standardization, and various values for k in the cluster analysis. Given that the WZs and workplace data were entirely new in 2011, with no existing body of analysis, a great deal of detailed exploration was required to identify and understand the key patterns in the data prior to and during creation of the classification.

Four broad domains (Table 1) were identified with regard to the main purposes of the classification, developed with ONS and key users of workplace statistics. While broadly similar to those of the 2001 and 2011 OAC in terms of demographic, socio-economic and employment characteristics, COWZ-EW focuses explicitly on the characteristics of the workplace population and workplaces, rather than the residential population and households. There are thus no direct equivalents of the OAC household composition and housing domains. The additional built environment domain is intended to capture the mix of workplace and residential activities in WZs. Travel to work variables are included in the socio-economic domain rather than creating a separate, narrowly focused, transportation domain.

An initial pool of 504 candidate variables were selected and summary statistics, histograms, Q-Q plots (to assess normality), correlation matrices (Pearson's Product Moment Correlation coefficient) and mapping at various geographical scales were employed to objectively and subjectively explore patterns and relationships within and between them. Following Vickers and Rees (2007), variables which provided good spatial and/or statistical differentiation were retained, redundancy and unnecessary correlation minimised, and variables representing very small population percentages and adding little to the classification removed. Log, Box-Cox and inverse hyperbolic sine (IHS) normalisations were evaluated. The Box-Cox method provided the most consistently satisfactory performance across the wide diversity of variable distributions. The 504 variables were finally reduced to 48 (summarised by domain in Table 1), and range standardization was applied before clustering.

K-means clustering was implemented in the R language to group WZs into a two-level classification. Standard statistical and graphical diagnostics were employed to evaluate the solutions, including compactness of cluster solutions and homogeneity of cluster size (shown for Supergroups in [Supplementary Figs. 1 and 2]), visualization of the degree of robustness throughout the hierarchy, mapping of cluster distributions and evaluation of profiles for both known and unknown locations, supported by aerial and ground-based online photography. For the top level of the hierarchy (termed Supergroups as per the 2011 OAC), solutions with 2 to 12 clusters were evaluated. For each solution from 5 to 8 clusters, separate solutions with 2 to 6 clusters within each Supergroup were generated as candidates for the second (Group) level, and similarly evaluated. The final version of COWZ-EW has two levels, with 7 Supergroups and 29 Groups. Unlike the 2011 OAC, a third level of the hierarchy was not produced as this led to excessively small cluster sizes and fragmented clusters. Cockings et al. (2015) describes the full process, including a comprehensive listing of Supergroups and Groups.

COWZ-EW successfully discriminates workplace area types such as industrial estates, business and retail parks and town centres. Berry, Newing, Davies, and Branch (2016) is the first study to utilise the COWZ-EW classification, focusing on retail store performance in relation to the characteristics of work-time catchment population characteristics, but their analysis employs the workplace data independently of interaction flows.

Our intention in the following analysis is to provide exemplar analyses of travel to work in England and Wales in terms of travel between

Table 1
Summary of COWZ-EW input census variables (full listing in Cockings et al., 2015).

Domain	Composition of workplace population	Composition of built environment	Socio-economic characteristics of workplace population	Employment characteristics of workplace population
Variable groups (number of variables in each group)	Population density (1)	Ratio of output areas to workplace zones (1)	Qualifications (3)	Employment type and hours worked (7)
	Age categories (4)		National statistics socioeconomic classification (6)	Industry of employment (15)
	Ethnicity and country of birth (5)		Distance travelled to work (4)	
			Mode of travel to work (2)	

the clusters of the residential 2011 OAC and the workplace COWZ-EW. These effectively represent the types of residential and workplace neighbourhoods between which commuters travel. We have obtained counts of persons travelling to work from each OA of residence (origin) to WZ of primary employment (destination) from 2011 SWS Table WF02EW "Location of usual residence and place of work (with outside UK collapsed) (OA/WPZ level)" (ONS, 2014a). Some initial processing is required to allocate all journeys to OA-WZ pairs. For the vast majority of journeys the origin OA and destination WZ reported in the SWS may be used directly. However, individuals who work from home or have no fixed place of work are deemed by ONS to both live and work at the same location (their residential OA) and home-working is itself an important characteristic of some neighbourhood types. Where these OAs and WZs have identical boundaries, allocation is straightforward but where they partially overlap, it is necessary to impute a WZ to be associated with the OA for analysis. There are several possible ways to do this, for example the OA may be assigned to the nearest WZ based on the distance between their population-weighted centroids, or on the basis of the greatest overlap of geographical area or cluster assignment. We have assessed all three options and found little difference in terms of the overall relationship between 2011 OAC and COWZ-EW clusters, hence the nearest centroid method has been used here. COWZ-EW does not include individuals who live in England and Wales but work offshore or outside the two countries, as these destinations necessarily fall outside the geographical scope of the classification.

All the census statistics, geographical references, 2011 OAC and COWZ-EW are published (either as census outputs or by the authors) and may thus be freely downloaded under open data licences. Relevant URLs are listed in Table 2.

Journeys to work can be characterised in terms of the 2011 OAC residential cluster from which they originate and the COWZ-EW workplace cluster where they end. Based on the 8 2011 OAC and 7 COWZ-EW Supergroups we have a 56-way classification of journeys to work. At the Group level, flows between 26 2011 OAC Groups and 29 COWZ-EW Groups result in a 754-way classification, although there are small numbers in some of these combinations. Naming of geodemographic clusters is contentious but can provide intuitive shorthand labels for the area types, albeit not without risks of misrepresentation (Goss, 1995; Parker, Uprichard, & Burrows, 2007; Vickers & Rees, 2007). For consistency we here use the previously published 2011 OAC and COWZ-EW numbers and names. Full multivariate profiles, including textual and statistical description of each Supergroup and Group are available at: http://www.ons.gov.uk/methodology/geography/geographicalproducts/areaclassifications/2011areaclassifications/penportraitsandradialplots and http://cowz.geodata.soton.ac.uk/download/.

Using this combination of a new workplace classification with existing official statistics, many substantive analyses could be undertaken and we are here only able to demonstrate a range of analytical possibilities rather than extensively examine travel to work patterns. In the following section we firstly compare the overall national distributions at the Supergroup level by independent mapping of the two classifications. We then explore the pattern of flows between Supergroups by cross-tabulation, considering distances travelled and focusing particularly on flows which are much larger or smaller than expected. If there were no pattern to the relationship between residential and workplace classifications, the expected flow between any two clusters would simply be the product of the origin and destination cluster totals divided by the overall flow total. To better understand the anatomy of the flow matrix, we calculate (observed-expected)/expected flows and consider the pattern of interactions which are both larger and smaller than expected. We also calculate median distances for all the interactions in each cell. It is not feasible to represent all 754 interactions at the Group level in a conventional table, so we instead use a graphical table based on the difference between observed and expected flows, as calculated above, for both the national dataset and a large regional example. We then move to mapping the pattern of workplaces to which residents travel, revealing the 'home geography' of commuters using the workplace classification to explore first, second and third order flows in the manner of Hincks (2012). We use a small example area to demonstrate the insights to be gained by examining both the 2011 OAC and COWZ-EW classifications at both origin and destination locations. Finally, we explore the spatial intersection of the two classifications and again use a graphical table to show the extent to which types of residential and workplace neighbourhoods are co-located.

4. Results

Figs. 1 and 2 map the 2011 OAC and COWZ-EW at the Supergroup level, with insets showing Greater London. The detailed pattern of individual OAs and WZs is not visible at these scales and mapping at the Group level is impractical to reproduce in a static national map. However, interactive online mapping of both classifications is available at: http://oac.datashine.org.uk and http://cowz.datashine.org.uk respectively. Figs. 1 and 2 use colour palettes consistent with these websites in order to permit direct visual comparison. The Supergroups with the

Table 2
Download locations of input data sources.

Dataset	Source URL (all accessed 24.02.17)
2011 OAC	http://www.ons.gov.uk/methodology/geography/geographicalproducts/areaclassifications/2011areaclassifications/datasets
2011 COWZ-EW	http://cowz.geodata.soton.ac.uk/download/
2011 OA boundaries	http://geoportal.statistics.gov.uk/datasets/09b8a48426e3482ebbc0b0c49985c0fb_0
2011 WZ boundaries	http://geoportal.statistics.gov.uk/datasets/a399c2a5922a4beaa080de63c0a218a3_0
2011 SWS	https://www.nomisweb.co.uk/census/2011/wf02ew

highest population densities cover very small land areas. Nevertheless, some important general observations help to introduce key features of the two classifications. A greater proportion of WZs are classified as Rural in COWZ-EW than OAs are classified as Rural Residents in the 2011 OAC but the two are highly spatially overlapping and dominate the land surface. All other Supergroups in both classifications are essentially urban or suburban in nature. The 2011 OAC displays patterns of residential classification which are broadly similar to many other geodemographic classifiers. The COWZ-EW, by contrast, displays a much closer relationship to recognized industrial structures, with the Top Jobs Supergroup appearing mostly in the innermost areas of London and major metropolitan areas, whose outer areas are primarily classified as Metro Suburbs. Within both classifications, there are frequently-repeated concentric patterns across many towns and cities but London and the major conurbations display extreme concentrations, particularly reflecting the presence of ethnic diversity and the highest status employment. Northern and western regions of the country broadly display greater proportions of the lower status residential and workplace Supergroups.

Table 3 and Fig. 3 present for the first time the geodemographic structure of 26,568,764 individual journeys to work from 2011 OAC Supergroups to COWZ-EW Supergroups (Table 3) and 2011 OAC Groups to COWZ-EW Groups (Fig. 3), reflecting the complex commuting patterns between and within the clusters mapped in Figs. 1 and 2. Considering Table 3 first, the row and column headings show the Supergroup numbers and names. The first value in each cell contains the number of people in employment travelling from all OAs in one 2011 OAC Supergroup to all WZs in one COWZ-EW Supergroup, the second shows the (observed-expected)/expected value and the third is the median travel distance in km, with a mean of 474,442 journeys in each cell. At the Group level (mean 35,387 per cell), it is not feasible to reproduce all the values in conventional table format, but the graphical table in Fig. 3 provides a rich visualization of the flow structure, referenced by the Group level numbers and names. The symbol shading represents quintiles of (observed-expected)/observed, with the black icons representing the top quintile (values above +0.45) and the white ones the lowest quintile (values below −0.73). Any cells with expected flows of less than 3524 (10% of mean) have been suppressed to reduce the visual impact of

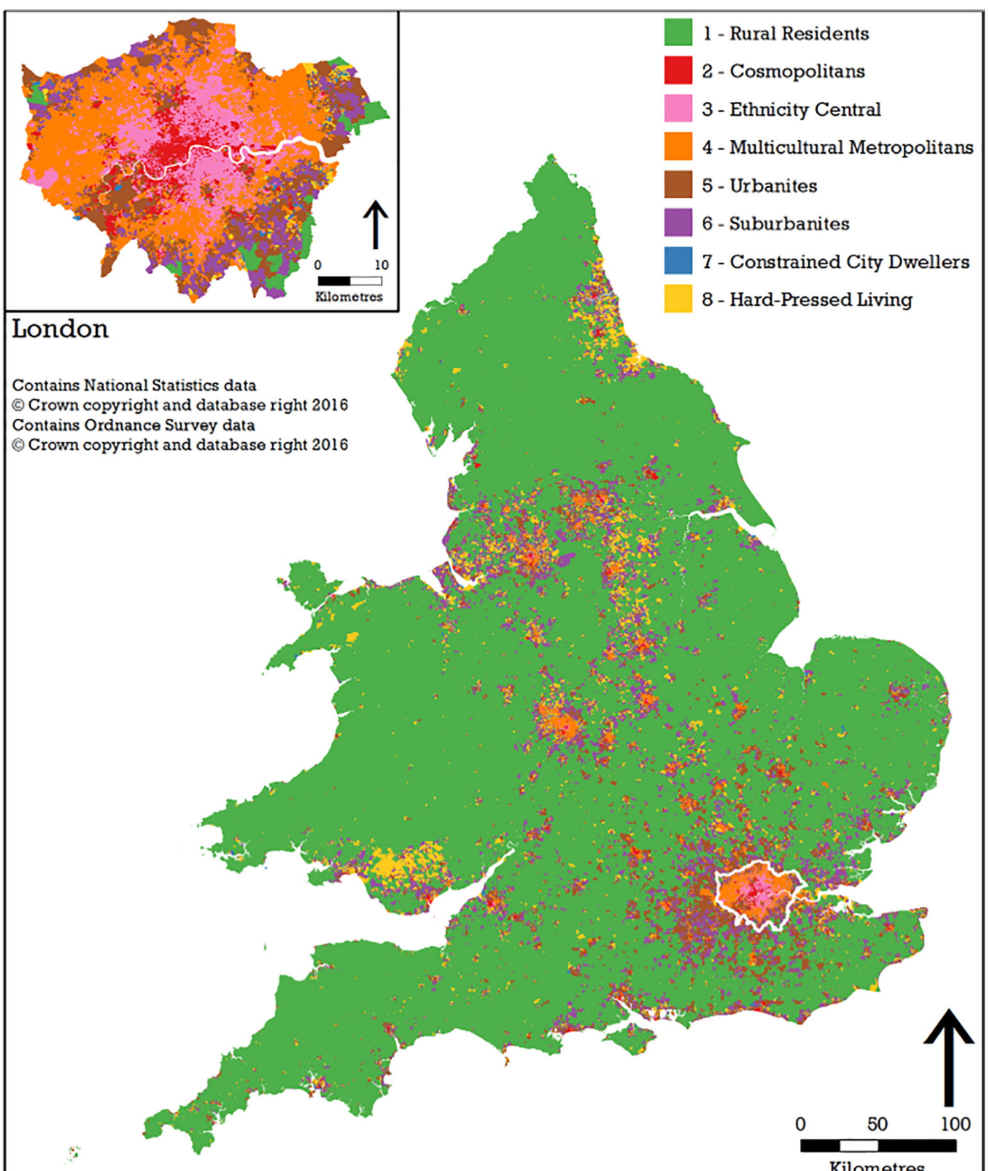

Fig. 1. 2011 OAC Supergroups, England and Wales with London inset.

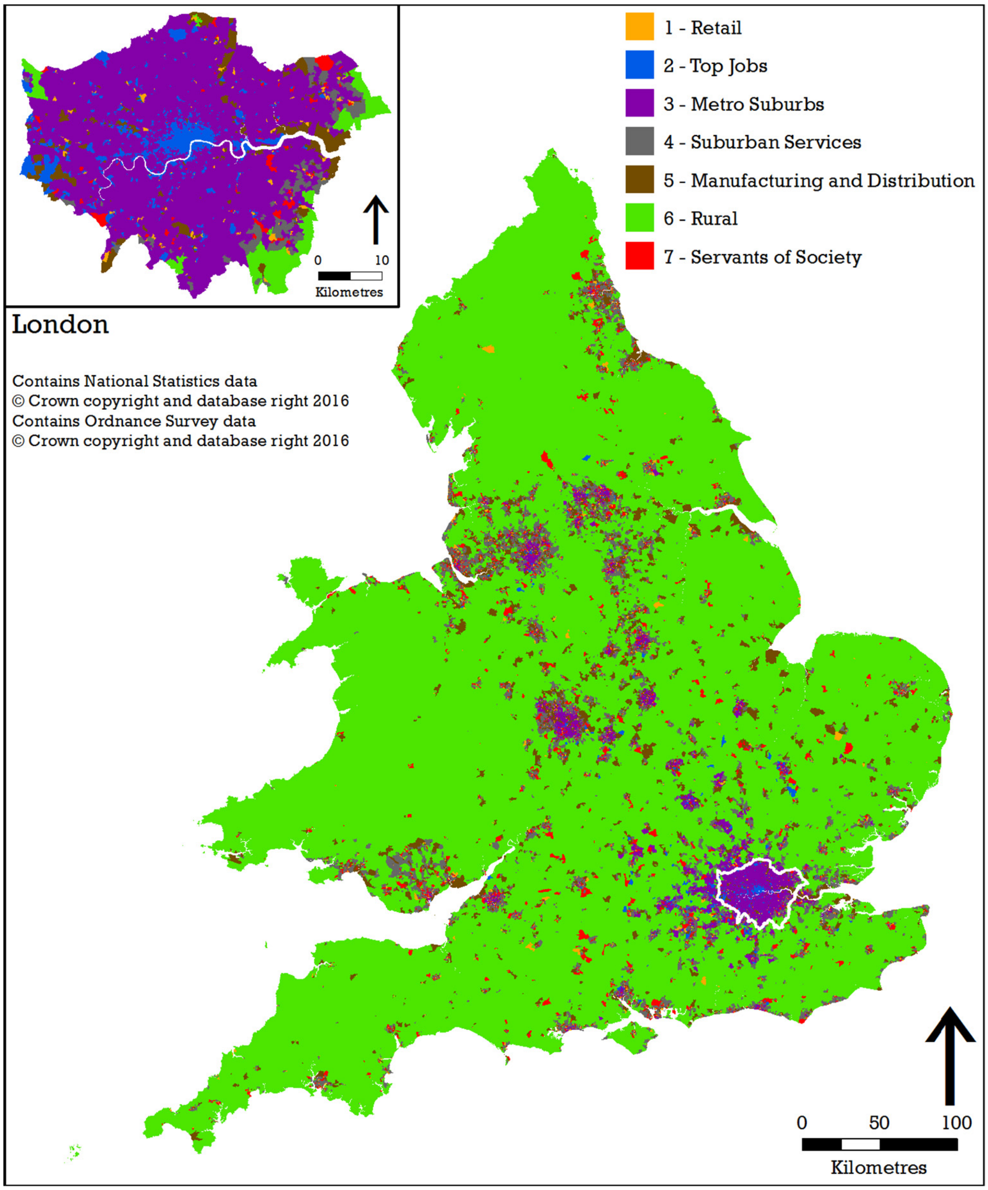

Fig. 2. COWZ-EW Supergroups, England and Wales with London inset.

small numbers, which particularly affect 2011 OAC Groups 3c: Ethnic Dynamics and 7b: Constrained Flat Dwellers. These tabular and graphical representations are effectively new geodemographic classifications of journeys to work, with 56 and 754 clusters respectively.

As with the maps, a striking feature of Table 2 is the relative self-containment of rural areas (2011 OAC Rural Residents to COWZ-EW Rural flows 2.25 times greater than expected), with Rural Residents underrepresented in every other workplace type, especially Top Jobs and Metro Suburbs. These patterns are intuitively meaningful, with very few residents of most urban areas travelling to work in rural areas. Fig. 3 reveals this even more starkly, with each individual Group level flow within the rural residential and workplace Supergroups being in the top quintile.

Three of the residential 2011 OAC Supergroups: Cosmopolitans, Ethnicity Central and Multicultural Metropolitans have very few rural workplace destinations but differ in other respects. Numerically, the smallest flow in Table 3 is from the 2011 OAC Cosmopolitans to the COWZ-EW Rural, with 42,827 trips, but at the same time Cosmopolitans Supergroup residents are underrepresented in Suburban Services, Manufacturing and Distribution and very strongly overrepresented in Top Jobs (1.76). Again, the group level reveals additional structure, for example the striking overrepresentation of the 2011 OAC Group 2a: Students Around Campus in 1e: Shop until you drop and 1f: Eat, drink, shop and be merry (5.34 times larger than expected). By contrast, 2011 OAC Group 2d: Aspiring and Affluent is underrepresented in all types of Retail workplace, but reflected in commuting from some of the highest status residential areas, with the group-level flow to the COWZ-EW Global Business Group (7.01 times larger than expected), being the largest value underlying Fig. 3.

The Ethnicity Central 2011 OAC Supergroup, which is itself concentrated in London and major metropolitan areas, is the most starkly divided of the 2011 OAC Supergroups in terms of workplace destinations, with strong overrepresentation in Top Jobs (2.09) and Metro Suburbs (2.13) and underrepresentation across all other COWZ-EW Supergroups and this is also seen across most of its Groups. The Urbanites and Suburbanites residential Supergroups are the least patterned in terms of workplace destinations, with flows the closest to expected values, demonstrating the diverse and complex trips undertaken by 11 million urban and suburban residents into all types of working neighbourhood. The last two rows of Table 2 represent the Constrained

Table 3
Observed flows 2011 OAC to COWZ-EW Supergroups: count of flows, (observed-expected)/expected, median distance (km).

2011 OAC (rows) and COWZ-EW (columns)	1: Retail	2: Top jobs	3: Metro suburbs	4: Suburban services	5: Manufacturing and distribution	6: Rural	7: Servants of society	Total
1: Rural residents	236,321 (−0.24) 9.63	**147,438** **(−0.67)** **36.76**	**49,501** **(−0.87)** **18.79**	254,100 (−0.43) 7.37	432,406 (−0.16) 11.48	**1,467,315** **(2.25)** **0.00**	429,776 (−0.09) 13.06	3,016,857 (−0.03)
2: Cosmopolitans	162,125 (0.09) 1.38	**590,897** **(1.76)** **5.19**	274,534 (0.49) 0.18	**83,889** **(−0.61)** **2.45**	**97,674** **(−0.61)** **6.80**	**42,827** **(−0.8)** **14.78**	200,909 (−0.11) 2.07	1,452,855 (0.03)
3: Ethnicity central	**63,040** **(−0.62)** **3.18**	**726,986** **(2.09)** **6.30**	**633,787** **(2.13)** **0.91**	**20,789** **(−0.91)** **6.55**	**79,809** **(−0.71)** **6.09**	**12,154** **(−0.95)** **22.06**	**59,022** **(−0.76)** **3.27**	1,595,587 (0.04)
4: Multicultural metropolitans	320,418 (−0.14) 3.06	**823,151** **(0.55)** **10.00**	**1,226,815** **(1.68)** **0.96**	194,758 (−0.63) 3.24	566,335 (−0.08) 4.66	**94,989** **(−0.82)** **12.01**	381,947 (−0.32) 3.10	3,608,413 (0.03)
5: Urbanites	588,729 (0.07) 2.98	748,044 (−0.05) 15.18	535,750 (−0.21) 2.33	943,624 (0.21) 0.60	867,347 (−0.05) 6.40	661,155 (−0.17) 3.57	987,326 (0.19) 4.46	5,331,975 (0)
6: Suburbanites	594,766 (0) 4.70	578,403 (−0.32) 18.06	380,406 (−0.48) 4.39	1,177,434 (0.39) 0.79	1,099,125 (0.11) 7.51	876,906 (0.01) 1.69	1,085,371 (0.2) 6.50	5,792,411 (−0.01)
7: Constrained city dwellers	210,815 (0.5) 2.19	**87,213** **(−0.57)** **11.51**	88,406 (−0.49) 3.03	279,078 (0.4) 0.76	312,423 (0.34) 4.43	140,888 (−0.31) 5.45	245,127 (0.15) 3.42	1,363,950 (0)
8: Hard-pressed living	555,561 (0.23) 3.52	**211,293** **(−0.67)** **14.97**	**180,380** **(−0.68)** **4.85**	939,961 (0.46) 0.84	1,085,488 (0.44) 5.42	676,008 (0.03) 3.07	758,025 (0.10) 5.22	4,406,716 (−0.01)
Total	2,731,775	3,913,425	3,369,579	3,893,633	4,540,607	3,972,242	4,147,503	26,568,764

Cells in bold indicate flows for which (observed-expected)/expected is below −0.5 or above 0.5.

City Dwellers and Hard-Pressed Living residential Supergroups, which are notable for their relative absence from the highest status workplace Supergroups Top Jobs and Metro Suburbs. These origins serve primarily lower status workplaces including Retail, Suburban Services and Manufacturing and Distribution, with some overrepresented Groups such as 1a. Low Density Retail and Wholesale, 4a. Non-metropolitan Suburban Areas, 5a. Mining and Quarrying Facilities and 5d. Manufacturing, Energy and Utilities. The Constrained City Dwellers Supergroup is the smallest overall and some of the intersections with its constituent Groups are suppressed due to small numbers of flows. It is notable that COWZ-EW Supergroup Servants of Society draws widely from across 2011 OAC Supergroups, reflecting the widespread range of occupations within public service employment such as education and health care.

Overall, we see a pattern of high self-containment of rural areas, complex self-contained mixing of lower status suburban areas and highly selective interactions between the more cosmopolitan and ethnically diverse suburban areas and central business districts. The

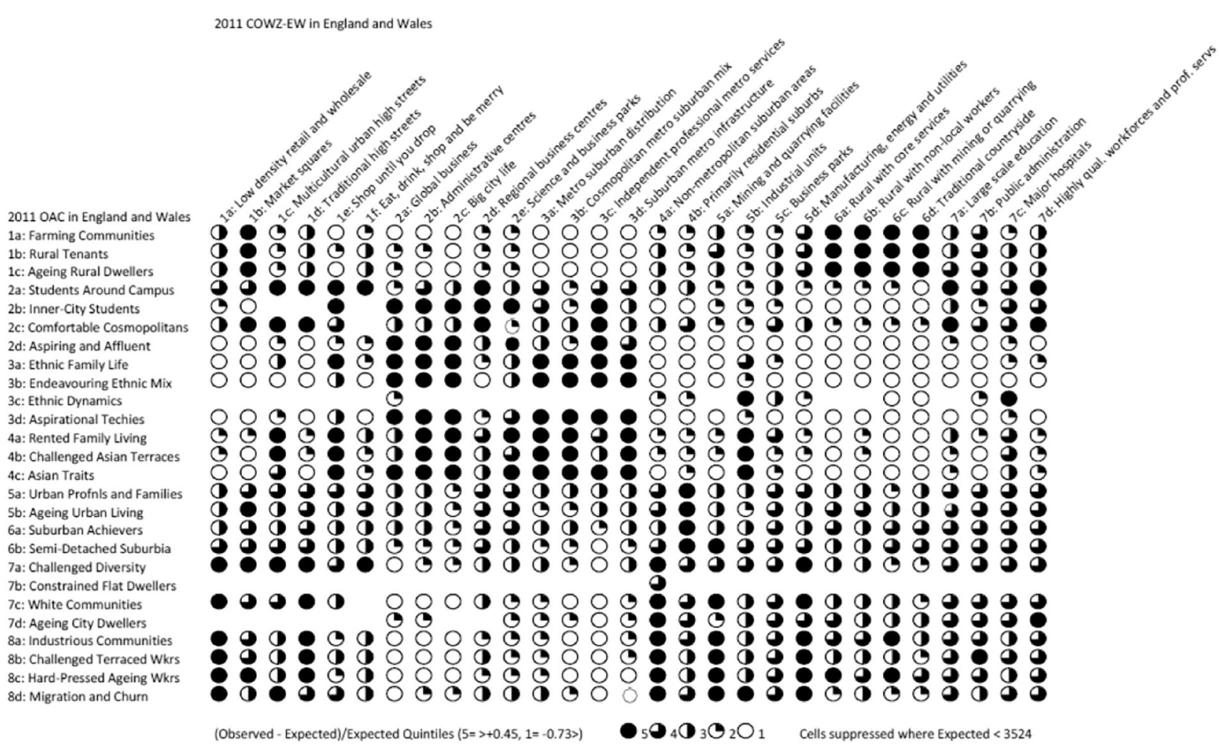

Fig. 3. 2011 OAC Group to COWZ-EW Group flows: quintiles of (observed-expected)/expected flows for England and Wales.

underlying data permit all these patterns to be explored in far greater detail, for example in terms of the trip lengths shown in Table 3. At the Supergroup level the longest median flow distance is from the 2011 OAC Rural Residents to COWZ-EW Top Jobs which, while less common than expected, still accounts for 147,438 journeys to work with a median journey length of 36.76 km. In general, the high status Top Jobs COWZ-EW Supergroup commands consistently long commuting distances, while some of the lowest distances are found among the suburban clusters. Also notable is the median distance of zero between the 2011 OAC and COWZ-EW Rural Supergroups reflecting the prevalence of working from home or at no fixed address within these areas. Those travelling into Rural COWZ-EW Supergroups from Cosmopolitans and Ethnicity Central are among the smallest numbers, and the longest distances. At the Group level we begin to see a pattern whereby some types of residential areas are almost entirely without interaction with some types of workplace, even when setting aside small absolute Group sizes. The wholly white symbols represent a dimension of population segregation not usually considered in conventional analyses of residential location.

Another way of considering the flows between places of residence and work is to map residential areas according to the workplace destinations of their principal flows. In Fig. 4 we map OAs in terms of the COWZ-EW Supergroup destinations of their first, second and third largest commuting flows. The legend includes flows to destinations outside England and Wales and thus outside the scope of COWZ-EW although there are very few of these. Multiple destinations refer to OAs where more than one destination COWZ-EW Supergroup has the same value. We here consider a much wider area than the Greater London inset in Figs. 1 and 2. The first order flows reveal the spatial consequences of rural activities being the primary workplace destination in most rural areas, and demonstrate that workers in Top Jobs are drawn both from the very central residential areas and also from the outer urban fringes. For the remaining majority of London's built-up area, Metro Suburbs are the primary work destination. The importance of Servants of Society in urban areas outside London is apparent, as is the corridor to the east of London, where Suburban Services and Manufacturing and Distribution Supergroups dominate. The second order map is completely different, revealing the huge spatial range of high status commuters, with almost the entire mapped region contributing to Top Jobs workplaces, from up to 100 km distant (made possible by an extensive suburban commuter rail network). In the very centre of London this is replaced with commuting out to Metro Suburbs. Beyond London commuting, the Servants of Society Supergroup is widely evident, but Manufacturing and Distribution and Retail begin to gain prominence in local areas. The third order map is very different again, reflecting complex localised flows of workers into generally lower status employment, with Retail and Manufacturing and Distribution evident across the map, now that the major effects of rural employment and long-distance commuting have been filtered out. Fig. 4 shows the capital city surrounded by an almost unbroken outer belt of workers who commute into Metro Suburbs employment.

Fig. 5 is a graphical table of the differences between observed and expected Group-level flows, constructed just for the 8,613,027 trips ending in the London and South East regions, broadly equivalent to the area mapped in Fig. 4. This type of analysis could be undertaken for any aggregation of census output areas and workplace zones. Compared to Fig. 3, more cells are suppressed and there is greater variation within Supergroups. Some of the national patterns in Fig. 3 are strongly influenced by London and its surroundings, but when considering these areas in isolation, the detailed structure of interactions is more complex. Groups within the Top Jobs and Metro Suburbs COWZ-EW Supergroups, for example, appear quite differently here compared to the national table in Fig. 3.

Fig. 6 demonstrates the insights possible when mapping both classifications at both trip ends, although here we can present only a very local example. We have chosen a 2 km × 2 km area of central Basingstoke, a town approximately 70 km south west of London. It is both an employment centre and a residential origin for commuters to London. To aid visual interpretation, the maps in Fig. 6 are masked to show shading only within built structures (both residential and non-residential), with principal roads and railways overlaid, from Ordnance Survey Open Data. Our four-map matrix shows the two standard classifications in Fig. 6(a) and (b) and the classification of their primary flows in Fig. 6(c) and (d). The two maps on the left are based on OAs, while those on the right are based on WZs. The Group level maps of the same area are provided in [Supplementary material]. This combination

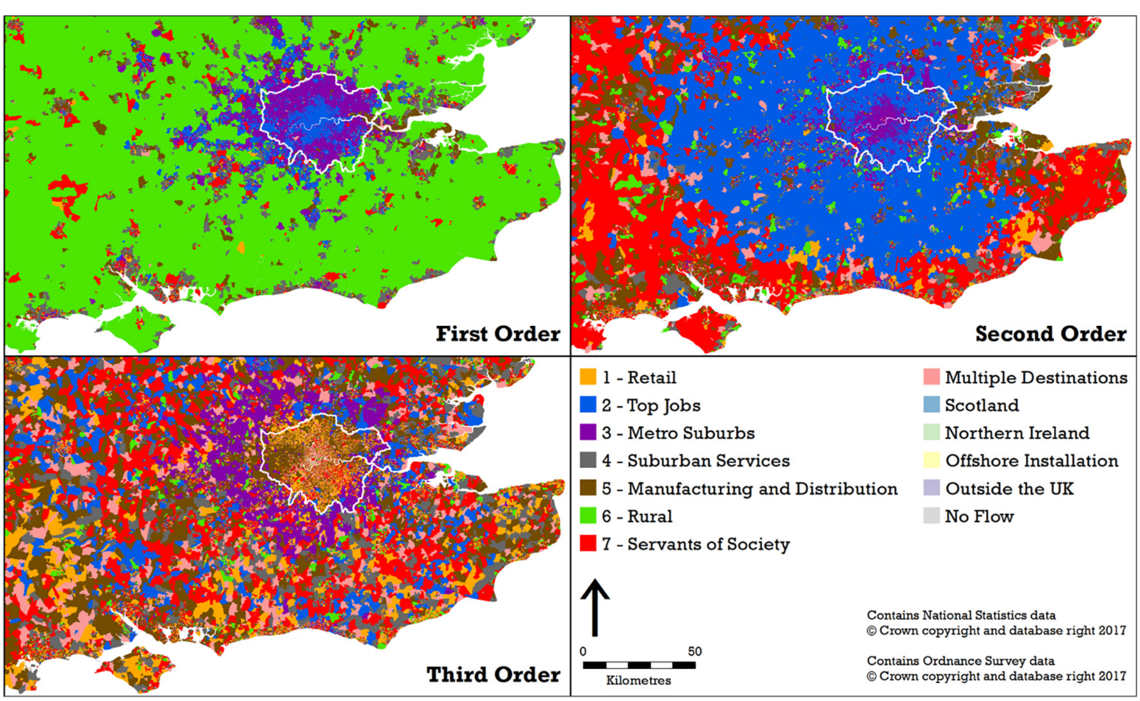

Fig. 4. First, Second and Third order dominant COWZ-EW destination Supergroups by output areas.

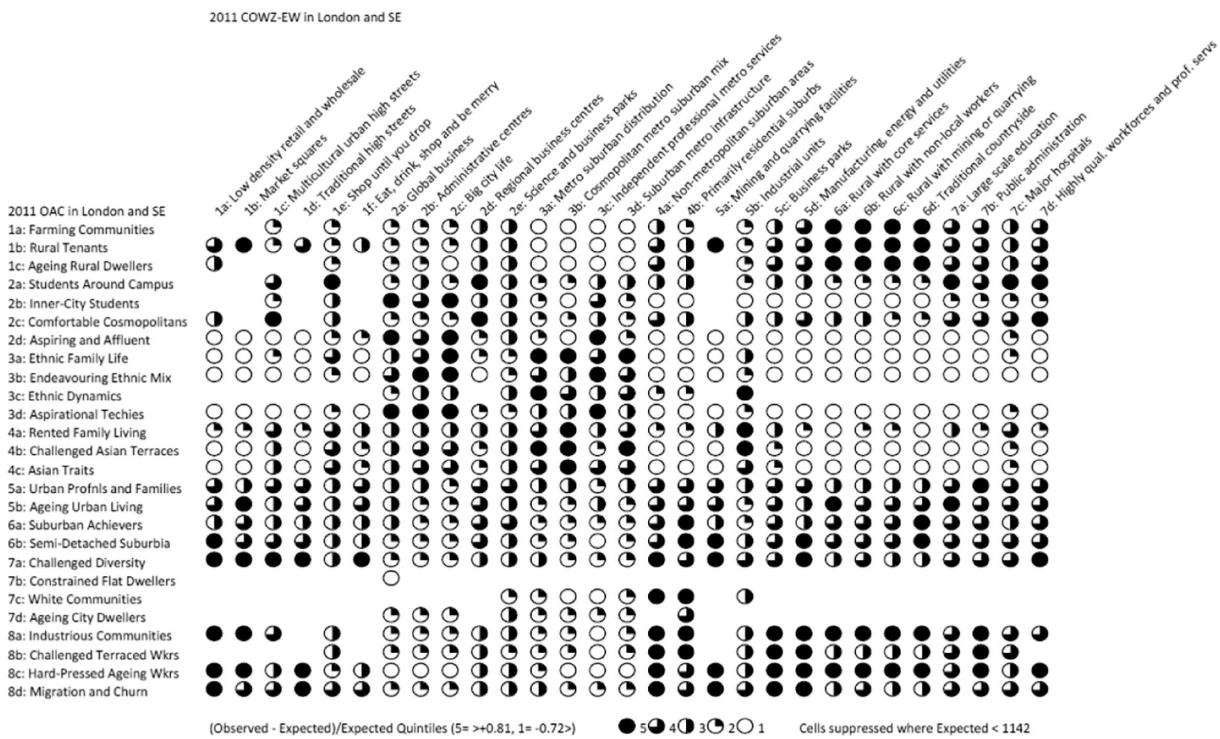

Fig. 5. 2011 OAC Group to COWZ-EW Group flows: quintiles of (observed-expected)/expected flows for trips ending in London and South East regions.

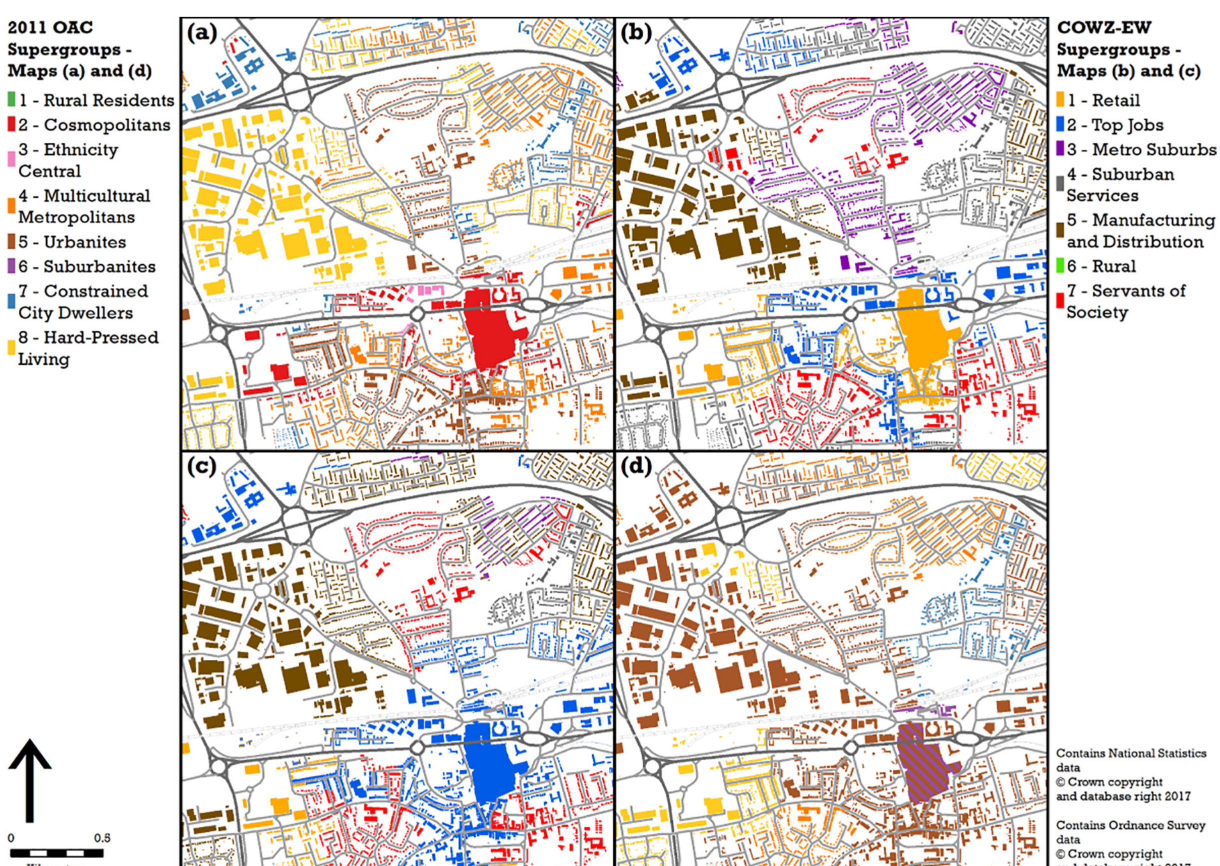

Fig. 6. Classifications by origin and destination in central Basingstoke: (a) 2011 OAC Supergroups by output area; (b) COWZ-EW Supergroups by workplace zone; (c) primary COWZ-EW destination Supergroup by output area; (d) primary 2011 OAC origin Supergroup by workplace zone. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

of perspectives reveals the complex differences between those living and working in the same neighbourhoods.

Fig. 6(a) shows 2011 OAC residential Supergroups. In the central southern area we see 2011 OAC Cosmopolitans (red), contrasting with Hard-Pressed Living in the west (yellow), surrounded by a mixture of urban and metropolitan Supergroups. Fig. 6(c) shows the COWZ-EW Supergroups of the primary workplace destination from each OA (equivalent to Fig. 4), revealing a wide area in the central southern part of the town from which residents are most likely to work in areas classified as Top Jobs (blue). Fig. 6(b) shows COWZ-EW Supergroups with Retail (orange) being the central workplace classification, encircled by Top Jobs (blue) and Manufacturing and Distribution (brown) in the northwest. There are also extensive areas classified as Servants of Society (red). Fig. 6(d) shows the 2011 OAC Supergroup of the primary origin of those working in each WZ, revealing that across many of the central and northwest areas they are most likely to have come from the Urbanites (brown) 2011 OAC Supergroup, who typically travel much shorter distances than the residents of these areas.

A further possibility for exploiting the combined classifications is to consider how COWZ-EW adds information to the 2011 OAC, already established as the official classification of residential areas. The graphical table in Fig. 7 shows the proportions of OAs within each 2011 OAC Group (rows) that are spatially coincident with each COWZ-EW Group. As expected, where commuting distances are low, for example in rural areas, the pattern is very similar to travel to work in Fig. 3 as people are essentially living and working in the same places. However, different 2011 OAC Groups are related in quite different ways to workplace activity. 2c. Comfortable Cosmopolitans is broadly spread across many COWZ-EW Groups, but absent from a few such as 5a. Mining and Quarrying Facilities and 5d. Manufacturing, Energy and Utilities, while others such as the 2011 OAC Group 3d. Aspirational Techies are heavily concentrated in very few COWZ-EW Groups, in this case dominated by 3c. Independent Professional Metro Services. The structures of Figs. 3 and 7 are related, being similar where commuting distances are small, but quite different where longer-distance commuting leads to large transfers between distant residential and workplace Groups. It is visually striking how Groups belonging to the four 2011 OAC Supergroups at the bottom of the table are strongly concentrated in the two Groups belonging to the COWZ-EW Supergroup Suburban Services. These types of pattern offer important insights to those who would seek to use conventional residential classifications for example to target the delivery of services, as further differentiation is clearly possible based on those likely to be working as well as living within these areas. Further analysis would readily begin to identify those areas where daytime (working) and night time (residential) populations are broadly similar in social characteristics and others where they are largely different populations.

5. Discussion

Geodemographic classification is widely used as a means of multivariate data reduction for the differentiation of residential areas. Travel between home and work continues to be an important area of research and policy interest having many implications for transportation planning, social mixing and energy use. Although microdata are sometimes accessed by researchers these are generally sampled census or survey records, coded to large geographical areas due to their potentially disclosive nature. Even where researchers have made forays into classification of origin and destination areas (Debenham et al., 2003; Manaugh et al., 2010), the same geographical areas and the same, mostly residential, characteristics have been used to classify both trip ends. These approaches fail to capture the very different dimensions that characterise places of residence and work: geodemographic classifications which are good at differentiating housing types are not well-suited to differentiating between retail parks and universities, for instance. In response, a new classification of small area workplace data has been combined with an existing official classification, to develop the concept of origin-destination geodemographic classification. It is worthy of note that these residential and workplace classifications are not entirely independent – in our application, both are sourced from the same census and have partially overlapping sets of input variables. No definitive validation of a specific classification is possible as it represents a data reduction rather than a directly measurable empirical reality. However, the stark clarity of some of the resulting patterns, for example in the structure of Table 3 and Fig. 3, suggests that much important structural

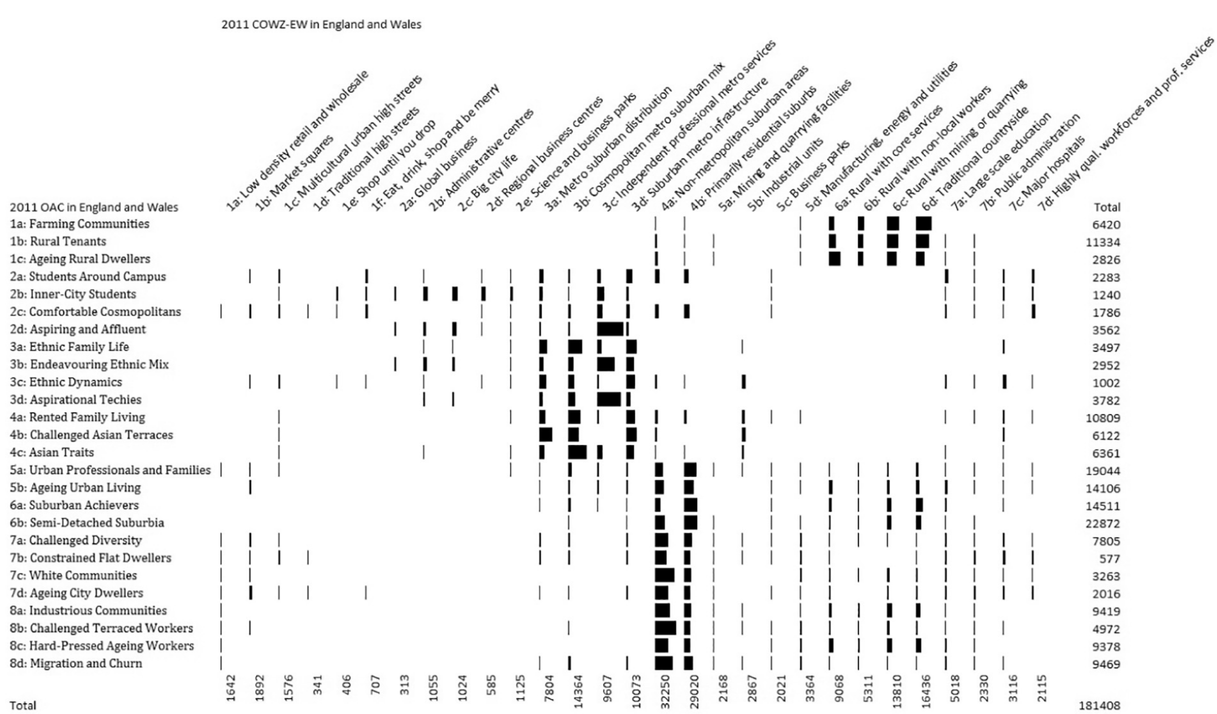

Fig. 7. Spatial coincidence of 2011 OAC Groups with COWZ-UK Groups: proportions of 2011 OAC Groups.

information has been captured through the intersection of these two classifications.

We have used these paired area classifications for a series of illustrative explorations of travel to work, but it is the underlying approach rather than the specific empirical observations which are the most important feature of this paper. This method for classifying travel to work flows offers multiple avenues for investigation, exploring the characteristics of flows between residential clusters and workplace types, identifying the varying levels of connection between the different groups and the distances travelled between them. There are many potential extensions, including investigation of the exclusivity of some of these relationships in terms of narratives on equality of access to employment, the structure of urban systems, modes and costs of travel and relationships between wages and housing costs in different areas. The 2011 England and Wales census did not include a question about travel times, but an application meriting further work would be to use the origin-destination classifications as a means of stratifying analyses of travel survey data, which might also be extended to consider provision and barriers to public and private transport between the 2011 OAC and COWZ-EW clusters.

The quality of the classifications is here dependent on the original census data, and although questions relating to work and workplaces are found by quality surveys to be among the least accurately answered among census questions (ONS, 2014b), they still offer a unique combination of high population coverage and rich socioeconomic detail compared to alternative sources. Further, there are many alternative classification methods, none of which can be considered definitive and our approach could be implemented using entirely different classifications of both origin and destination neighbourhoods. For example, non-census workplace data of the type used by Debenham et al. (2003) could be combined with a workplace geography to undertake analysis similar to that introduced here.

The approach presented here is compatible with the increasing potential to derive flow magnitudes from the tracking of mobile devices or public transport journeys. Chow, Santacreu, Tsapakis, Tanasaranond, and Cheng (2014) undertake a multisource data analysis of urban traffic congestion in London, including Oyster public transport travel card data. An exciting range of studies are beginning to exploit mobile telephone data, including Novak et al. (2013) focusing on Tallinn, Estonia, Iqbal, Choudhury, Wang, and González (2014) on Dhaka, Bangladesh and Alexander, Jiang, Murga, and González (2015) on the Boston, USA metropolitan area. These studies estimate daily trip origins and destinations from mobile phone call records. Both Novak et al. (2013) and Alexander et al. (2015) identify travel to work trips by inference from analysis of apparent 'home' and 'work' locations. In isolation, these new forms of data provide a rich picture of movement patterns, but little detailed characterisation of either the commuters or the neighbourhoods of residence and employment. Our approach which characterises trips based on the geodemographic characteristics of their origin and destination areas, for which non-census-based products are already available, is equally applicable to flows derived from these new forms of data as to conventional census-based small area and travel to work data.

6. Conclusions

We have demonstrated an approach to the investigation of travel to work by using purpose-specific geographical areas and geodemographic classifications for places of residence and employment. The availability of a new workplace geography and associated classification overcomes recognized data limitations and analytical challenges, and potentially offers new dimensions to our understanding of the complexity of workplaces and travel to work. The COWZ-EW classification recasts the census datasets in terms of the working population at their places of work and thus provides a rich description of the small area geography of employment as a counterbalance to traditional geodemographic classifications of residential areas, including insights into the differences between the residential and working populations of neighbourhoods. Although demonstrated using England and Wales census data, there is nothing nationally specific about these methods. The key innovations are the construction of a national small area workplace classification and particularly the combination of the origin and destination classifications as a basis for exploring the structure of commuting flows. We anticipate the publication of a workplace classification for the whole of the UK and comparable datasets could be constructed in many countries. More generally, origin-destination classifications have power for the reduction and visualization of travel to work data.

This paper comes at a time when there are major changes taking place in censuses globally, and increasing moves towards the replacement of census data with administrative information (Ralphs and Tutton, 2011). Rather than being limited to the analysis of census aggregate data, the approach proposed here has the potential to offer a common framework for the interpretation of trip magnitudes derived from many new forms of data. This approach will continue to have considerable merit in characterising trip origins and destinations: the key point is that classifications need to be based on the most appropriate geographical areas and data, relevant to the differing characteristics of residential and workplace neighbourhoods.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.compenvurbsys.2017.09.002.

Acknowledgements

The authors gratefully acknowledge the assistance of staff at ONS Geography, and the support of Economic and Social Research Council award ES/L007517/1.

References

Adams, J. S., VanDrasek, B. J., & Phillips, E. G. (1999). Metropolitan area definition in the United States. *Urban Geography*, *20*(8), 695–726.
Alexander, L., Jiang, S., Murga, M., & González, M. C. (2015). Origin–destination trips by purpose and time of day inferred from mobile phone data. *Transportation Research Part C: Emerging Technologies*, *58*, 240–250.
Australian Bureau of Statistics (2011). Census dictionary 2011. http://www.abs.gov.au/ausstats/abs@.nsf/mf/2901.0, Accessed date: 22 September 2016.
Berry, T., Newing, A., Davies, D., & Branch, K. (2016). Using workplace population statistics to understand retail store performance. *The International Review of Retail, Distribution and Consumer Research*, *26*(4), 375–395.
Chow, A. H. F., Santacreu, A., Tsapakis, I., Tanasaranond, G., & Cheng, T. (2014). Empirical assessment of urban traffic congestion. *Journal of Advanced Transportation*, *48*(8), 1000–1016.
Cockings, S., Harfoot, A., Martin, D., & Hornby, D. (2011). Maintaining existing zoning systems using automated zone design techniques: Methods for creating the 2011 Census output geographies for England and Wales. *Environment and Planning A*, *43*(10), 2399–2418.
Cockings, S., Martin, D., & Harfoot, A. (2015). A classification of workplace zones for England and Wales (COWZ-EW): User guide. http://cowz.geodata.soton.ac.uk/download/, Accessed date: 24 February 2017.
Coombes, M. G., & Bond, S. (2008). *Travel-to-work areas: The 2007 review*. London: Office for National Statistics.
Debenham, J., Clarke, G., & Stillwell, J. (2003). Extending geodemographic classification: A new regional prototype. *Environment and Planning A*, *35*(6), 1025–1050.
Dennett, A., & Stillwell, J. (2011). A new area classification for understanding internal migration in Britain. *Population Trends*, *145*, 146–171.
Duke-Williams, O. (2010). Mapping the geodemographic classifications of migrants' origins and destinations. *Journal of Maps*, *6*(1), 360–369.
Federal Highway Administration (2010). TAZ delineation business rules (March 2010). http://www.fhwa.dot.gov/planning/census_issues/ctpp/data_products/tazddbrules.cfm, Accessed date: 24 February 2017.
Gale, C. G., & Longley, P. A. (2013). Temporal uncertainty in a small area open geodemographic classification. *Transactions in GIS*, *17*(4), 563–588.
Gale, C. G., Singleton, A. D., Bates, A. G., & Longley, P. A. (2016). Creating the 2011 area classification for output areas (2011 OAC). *Journal of Spatial Information Science*, *12*, 1–27.
Goss, J. (1995). We know who you are and we know where you live. The instrumental rationality of geodemographic systems. *Economic Geography*, *71*(2), 171–198.
Hincks, S. (2012). Daily interaction of housing and labour markets in North West England. *Regional Studies*, *46*(1), 83–104.
Iqbal, M. S., Choudhury, C. F., Wang, P., & González, M. C. (2014). Development of origin–destination matrices using mobile phone call data. *Transportation Research Part C: Emerging Technologies*, *40*, 63–74.

Kim, C., Sang, S., Chun, Y., & Lee, W. (2012). Exploring urban commuting imbalance by jobs and gender. *Applied Geography*, *32*(2), 532–545.

Leventhal, B. (2016). *Geodemographics for marketers: Using location analysis for research and marketing*. London: Kogan Page.

Manaugh, K., Miranda-Moreno, L. F., & El-Geneidy, A. M. (2010). The effect of neighbourhood characteristics accessibility home–work location and demographics on commuting distances. *Transportation*, *37*(4), 627–646.

Martin, D., Cockings, S., & Harfoot, A. (2013). Development of a geographical framework for census workplace data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(2), 585–602.

Mitchell, B. (2014). Workplace zones: A new geography for workplace statistics. http://webarchive.nationalarchives.gov.uk/20160105160709/https://geoportal.statistics.gov.uk/Docs/An_overview_of_workplace_zones_for_workplace_statistics_V1.01.zip, Accessed date: 24 February 2017.

Niedzielski, M. A., O'Kelly, M. A., & Boschmann, E. E. (2015). Synthesizing spatial interaction data for social science research: Validation and an investigation of spatial mismatch in Wichita, Kansas. *Computers, Environment and Urban Systems*, *54*, 204–218.

Novak, J., Ahas, R., Aasa, A., & Silm, S. (2013). Application of mobile phone location data in mapping of commuting patterns and functional regionalization: A pilot study of Estonia. *Journal of Maps*, *9*(1), 10–15.

Office of Management and Budget (2010). 2010 standards for delineating metropolitan and micropolitan statistical areas notice. https://www.gpo.gov/fdsys/pkg/FR-2010-06-28/pdf/2010-15605.pdf, Accessed date: 24 February 2017.

O'Kelly, M. E., Niedzielski, M. A., & Gleeson, J. (2012). Spatial interaction models from Irish commuting data: Variations in trip length by occupation and gender. *Journal of Geographical Systems*, *14*(4), 357–387.

ONS (2014a). 2011 census origin-destination data user guide. http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-prospectus/release-plans-for-2011-census-statistics/subsequent-releases-of-specialist-products/flow-data/origin-destination-data–user-guide.pdf, Accessed date: 24 February 2017.

ONS (2014b). 2011 census quality survey. http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-user-guide/quality-and-methods/assessing-accuracy-of-responses–census-quality-survey-/index.html, Accessed date: 24 February 2017.

ONS (2015). Methodology note for the 2011 area classification for output areas ONS. http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/ns-area-classifications/ns-2011-area-classifications/methodology-and-variables/methodology-oa.pdf, Accessed date: 24 February 2017.

ONS (2016). Travel to work area analysis in Great Britain: 2016. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/traveltoworkareaanalysisingreatbritain/2016, Accessed date: 24 February 2017.

Parker, S., Uprichard, E., & Burrows, R. (2007). Class place and place classes geodemographics and the spatialization of class. *Information Communication and Society*, *10*(6), 902–921.

Poppie, K., & Miller, D. (2016). Case study: US market for geodemographics. In B. Leventhal (Ed.), *Geodemographics for marketers: Using location analysis for research and marketing* (pp. 163–166). London: Kogan Page.

Rae, A. (2016). *The geography of travel to work in England and Wales: Extracts from the 2011 census applied spatial analysis online early access*. . https://doi.org/10.1007/s12061-016-9196-0 (Accessed 24.02.17).

Ralphs, M., & Tutton, P. (2011). *International modelsfor census taking: Current processes and future developments.* http://webarchive.nationalarchives.gov.uk/20120819102859/http://www.ons.gov.uk/ons/about-ons/what-we-do/programmes-projects/beyond-2011/news-and-events/reports-and-publications/beyond-2011-international-models-for-census-taking.pdf, Accessed 18 September 2017.

Singleton, A., Pavlis, M., & Longley, P. A. (2016). The stability of geodemographic cluster assignments over an intercensal period. *Journal of Geographical Systems*, *18*, 97–123.

Sohn, J. (2005). Are commuting patterns a good indicator of urban spatial structure? *Journal of Transport Geography*, *13*(4), 306–317.

Stillwell, J., Duke-Williams, O., & Dennett, A. (Eds.). (2010). *Technologies for migration and commuting analysis: Spatial interaction data applications*. Hershey PA: IGI Global.

Vickers, D., & Rees, P. (2007). Creating the UK National Statistics 2001 output area classification. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *170*(2), 379–403.

Webber, R. J. (1977). *An introduction to the national classification of wards and parishes*. London: Planning Research Applications Group Technical Paper, Centre for Environmental Studies.