

Origin-specific models for analysing commuting flows in Northern Ireland: scale effects and other problems

C. D. Lloyd, I. G. Shuttleworth, G. Catney

School of Geography, Archaeology and Palaeoecology, Queen's University, Belfast, BT7
1NN
Tel. +44 (0)28 9097 3478; Fax +44 (0)28 9097 3212
Email: c.lloyd@qub.ac.uk

KEYWORDS: commuting, spatial interaction, origin-specific models, Poisson regression

1. Introduction

Understanding commuting flows and their characteristics has obvious economic significance and, as such, has been the subject of much academic research. A large proportion of existing research is based on official data sources such as censuses of population. This analysis entails the application of a spatial interaction modelling framework to explore the size of commuting flows between different wards (administrative zones) in Northern Ireland. Initially, global unconstrained and constrained models and the residuals from these models were explored (Lloyd et al., 2007). In the present paper, an origin-specific model is used to model flows from each individual location. The analysis explores the relationship between commuting flows and a host of explanatory variables including the distance from the origin zone, number of workers by ward of workplace (destination), the absolute difference between the percentage of Catholics at the origin and destination wards, a measure of employment opportunities at the origin and destination wards and finally a measure of destination competition. Selected results are summarised and interpreted after which some problems with this approach are then considered and some ideas for future work outlined.

2. Data and methods

This paper makes use of data provided as outputs from the 2001 Northern Ireland Census of Population. The main focus is on counts of workers moving between wards. There are 582 wards and thus the flow matrix comprises 582 by 582 entries (338724 possible interactions). The small cell adjustment procedure, adopted prior to release of data from the 2001 Census of Population, converted small counts to values of zero or three (see Duke-Williams and Stillwell, 2007 for an assessment of the effects of small cell adjustment on interaction data). The sparsity of the matrix of flows is apparent from the fact that some 87% of the flows have a value of zero. Some 6.4% of flows have a value of three, and this is a function of small cell adjustment. The maximum flow value is 1727. The explanatory variables are listed and described in Table 1. The road distances between Output Areas (OAs; nested within wards) were provided by the Northern Ireland Statistics and Research Agency (NISRA) and the averages of these distances between wards were computed for the purposes of spatial interaction modelling.

Variable name	Description
DistKM	Road distance between wards in KM
WrkTotal	Number of workers by ward of workplace
CathDiff	Absolute diff. between origin Catholic % and destination Catholic %
EmpScoMult	Origin and destination employment access scores multiplied together
Centrality	Destination competition measure

Table 1. Variable names and definitions. Diff. is difference.

The division of the population of Northern Ireland into two groups, Protestants and Catholics, and the study of the degree to which members of the two groups live and work together has received much attention in academic research (see, for example, Smith and Chambers, 1991). The absolute difference between the origin ward and destination ward Catholic percentages is included to enable exploration of the degree to which commuting flows are segregated. The employment accessibility score is defined as the level of job opportunities at location j estimated as the number of jobs in a given area down-weighted in proportion to the distance of the area from location j (Coombes and Raybould, 2000). Centrality is a measure of destination competition as defined by Fotheringham and O’Kelly (1989).

Aggregate interactions between places are often explored in a spatial interaction modelling framework (Haynes and Fotheringham, 1984; Fotheringham and O’Kelly, 1989). This paper makes use of Poisson regression (see Flowerdew and Aitkin, 1982, for an introduction) to model flows from each origin ward to all destinations; this is termed an origin-specific model (Fotheringham and O’Kelly, 1989). For a given ward, the flow to all wards from that ward is the dependent variable while distances from that ward to all other wards and the other variables listed in Table 1 are the independent variables.

3. Results

The minimum and maximum origin-specific coefficients (for the 582 sets of values) are summarised in Table 2. The signs of the coefficients are as expected. That is, all of the distance coefficients are negative, indicating that an increase in distance corresponds to a reduction in the size of flows. The coefficient for the number of workers at the destination ward is positive – the number of workers (the number of filled jobs) is, not surprisingly, an attraction. The coefficients for the absolute difference between the origin ward Catholic percentage and the equivalent figure for the destination ward range from large negative to large positive values. In words, for some origins if the religious composition of a destination ward is different to that of the origin ward then that destination is likely to attract less workers than destination wards with more similar compositions to the origin ward. The employment score multiplier coefficients also vary – there are large positive values in Belfast and neighbouring areas. This is probably because there are more extensive employment opportunities in Belfast than in other (particularly rural) areas (if there are many large flows between an origin and destinations with similarly large employment opportunities then this would result in large positive values of this coefficient). The coefficients for centrality have negative values in some locations and positive values in others.

Variable name	Minimum	Maximum
DistKM	-4.466	-0.406
WrkTotal	0.559	2.061
CathDiff	-8.000	4.584
EmpScoMult	-4.829	7.253
Centrality	-0.333	0.605

Table 2. Origin-specific model coefficients.

The distance coefficient for the origin-specific model is mapped in Figure 1. Distance decay values tend to be larger in the rural west. This suggests that individuals in (for example) the rural west are deterred from travelling large distances, whereas those in urban areas such as Belfast are deterred to a lesser degree; this is clearly counter-intuitive. The competing destinations model of Fotheringham (Fotheringham and O’Kelly, 1989) has been offered as a potential solution to this problem. However, the model employed here is a competing destinations model – a centrality variable has been included. A variety of characteristics of the

flows were explored in an attempt to better understand this problem of counter-intuitive distance coefficients.

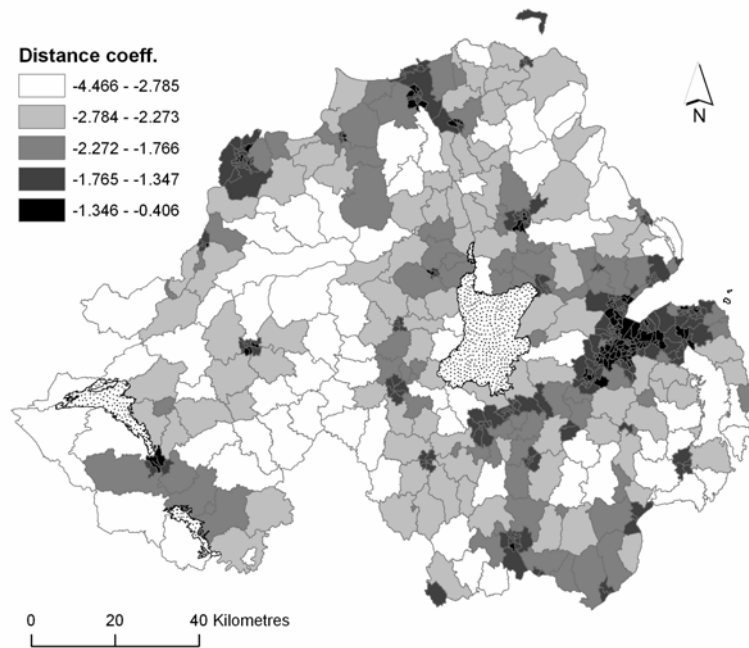


Figure 1. Origin-specific model distance coefficient. Northern Ireland Census of Population data — © Crown Copyright. Reproduced under the terms of the Click-Use Licence.

Figure 2 shows mean distance commuted from each ward against internal flows as a percentage of all flows from a ward. The relationship between the two variables is positive and this is counter-intuitive. The plot points to a key problem with the kinds of data employed in this analysis. In large rural wards often many people live and work in the same ward. The internal distance in this ward (generated using OA centroids, as described above) may be larger than many distances in the neighbourhood of a small urban ward. In effect, using these data, the commuting distance assigned to people living in large rural wards is always larger than the minimum commuting distance in most (if not all) urban areas. The destination choices available for residents of rural areas are likely to be, simply because of large zone sizes in these areas, at larger distances than those choices available to residents of urban areas.

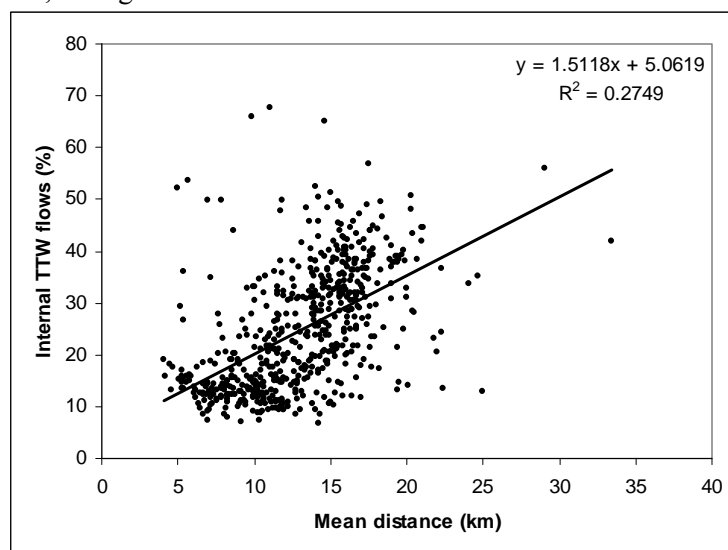


Figure 2. Mean distance against internal flows as a percentage of all flows from a ward.

Figure 3 shows competing destinations origin-specific model distance coefficients against internal flows as a percentage of all flows from a ward. This plot suggests that, on average, residents of areas with a large percentage of internal flows are less likely to travel large distances to work. This seems intuitive; with the plot in Figure 2, this indicates that the size of zones (not surprisingly) has a major impact on results.

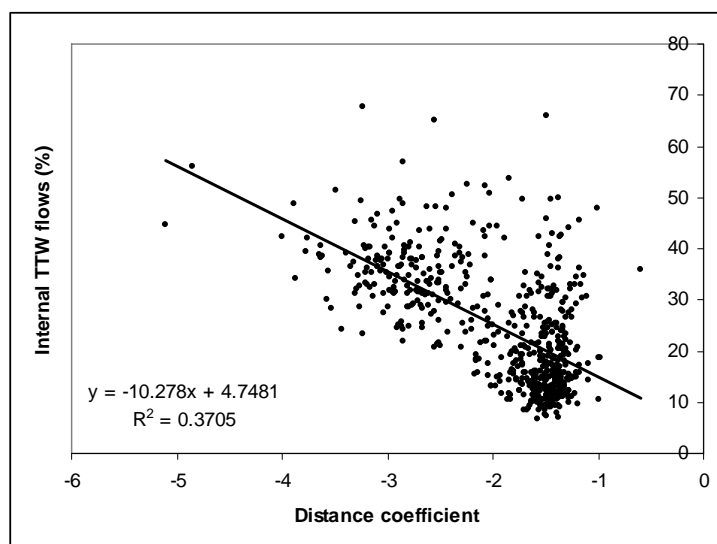


Figure 3. Distance coefficient against internal flows as a percentage of all flows from a ward.

4. Conclusions and future work

This analysis raises a number of issues which must be addressed to make best use of the available data for understanding commuting flows in Northern Ireland. Working with large areas is unavoidable in sparsely-populated areas – counts over small areas may not be released for confidentiality reasons. Various attempts have been made to make use of information on intra-zone flows to help overcome this problem. The use of geographically-weighted approaches (see Nakaya, 2001) may also offer benefits. Such approaches will be explored. In addition, this paper makes use of only origin-specific models. A logical extension of the work would be to make use of destination-specific models to allow exploration of propulsiveness of origin wards, rather than attractiveness of destination wards as examined here.

5. Acknowledgements

The Northern Ireland Statistics and Research Agency (NISRA) are thanked for the provision of 2001 Northern Ireland Census of Population data. The ward-to-ward road distances were calculated using data provided by NISRA.

References

- Coombes M and Raybould S (2000) Policy-relevant surfaced data on population distribution and characteristics *Transactions in GIS* 4 pp319–342.
- Duke-Williams O and Stillwell J (2007) Investigating the potential effects of small cell adjustment on interaction data from the 2001 Census *Environment and Planning A* 39 pp1079–1100.
- Flowerdew R and Aitkin M (1982) A method of fitting the gravity model based on the Poisson distribution *Journal of Regional Science* 22 pp191–202.

Fotheringham AS and O'Kelly ME (1989) *Spatial Interaction Models: Formulations and Applications*. Kluwer Academic Publishers, Dordrecht.

Haynes KE and Fotheringham AS (1984) *Gravity and Spatial Interaction Models*. Sage Publications, Beverly Hills.

Lloyd CD, Shuttleworth IG and Catney G (2007) Commuting in Northern Ireland: Exploring spatial variations through spatial interaction modelling. In Winstanley AC (Ed.) *GISRUUK 2007. Proceedings of the Geographical Information Science UK Conference*. National Centre for GeoComputation, National University of Ireland Maynooth, 11th–13th April 2007. Maynooth: National University of Ireland Maynooth, pp258–263.

Nakaya T (2001) Local spatial interaction modelling based on the geographically weighted regression approach *GeoJournal* **53** pp347–358.

Smith D and Chambers G (1991) *Inequality in Northern Ireland*. Clarendon Press, Oxford.

Biography

Chris Lloyd is Lecturer in GIS at Queen's University, Belfast. His research interests include spatial data analysis, the analysis of Census data, and historical geography.

Ian Shuttleworth is Lecturer in Human Geography at Queen's University, Belfast. His research interests include Census analysis, labour markets and labour mobility, and the political economy of transnational migration.

Gemma Catney is a PhD research student in Geography at Queen's University, Belfast. Her research interests are in migration, residential segregation and population change, and in methodologies for contemporary empirical research in population and social geography.