SERC DISCUSSION PAPER **213**

Using Micro-Geography Data to Identify Town-Centre Space in Great Britain

Paul Cheshire (LSE and SERC)
Christian Hilber (LSE and SERC)
Piero Montebruno (University of Cambridge)
Rosa Sanchis-Guarner (Imperial and SERC/CEP, LSE

May 2017

This work is part of the research programme of the Urban Research Programme of the Centre for Economic Performance funded by a grant from the Economic and Social Research Council (ESRC). The views expressed are those of the authors and do not represent the views of the ESRC.

Using Micro-Geography Data to Identify Town-Centre Space in Great Britain

Paul Cheshire*
Christian Hilber*
Piero Montebruno**
Rosa Sanchis-Guarner***

May 2017

* London School of Economics and SERC
** University of Cambridge
*** Imperial College Business School and SERC/CEP London School of Economics

We thank SERC/CEP for funding (ESRC grant ES/M010341/1) and Sanchis-Guarner acknowledges support from the British Academy. We thank Mr Mark Teale from CBRE and the Department for Communities and Local Government (DCLG) for providing us with data. ABI data accessed via NOMIS under BRES notice NTC/BRES15-P0572. We benefited from comments and suggestions from participants at the SERC Work-in-progress research seminar, at the CESifo Conference on the Use of Geocoded Data in Economic Research and at the University of Birmingham City REDI seminar. Any errors remain our own.

Abstract

We often talk about 'Town Centres', but defining their location and extent is surprisingly difficult. Their boundaries are hard to pin down and intrinsically fuzzy. Nevertheless, the British government introduced very specific policies for them in 1996 – Town Centre First Policies (TCFP) – without defining them. The semi-official definitions introduced in 2004 did not cover Scotland, only England and Wales. Using a range of variables available for the whole of Great Britain that capture all the dimensions of 'town centredness', we start by replicating the definitions for England and Wales. Then, we use an alternative list of towns and cities and apply our estimated coefficients to predict their size. Our models yield high correlations between the semi-official DCLG values and our predicted values, so we then move on to identify Town Centres for all three countries of GB. Our method is a contribution in its own right but is also an essential step if there is to be a rigorous evaluation of TCFP since it makes it possible to compare changes in the 'policy treated' Town Centres of England and Wales with changes in the 'policy untreated' ones of Scotland.

Keywords: town centre, planning, retail sector, land use
JEL Classifications: L81; R12; R52

1. Introduction

Imagine you are anywhere in a city—London, Lyon, Berlin, Wolverhampton—and you know that city well. Suddenly, someone comes up to you and asks, 'Could you tell me where the town centre is?' This could appear to be a simple, even a trivial, question, but it is not. In fact, in many instances it proves to be surprisingly hard to answer. The aim of this paper is to provide a response. The question has a particular salience since, especially in Great Britain, there are influential urban policies that apply to 'Town Centres' (TC). But, if we cannot define the boundaries of these areas not only can we not identify the actual areas the policies are supposed to apply to, we cannot evaluate any effects such Town Centre policies may have on outcomes. Our aim in this paper is to design, explain, apply and test a method to answer this apparently trivial question. We are not concerned with *why* the Town Centre is sought. Instead, we explore and provide an operational answer: the extent and location of 'town-centre spaces' in Great Britain.

Our interest in identifying and predicting Town-Centre space arose as one part of an investigation into the effects of Town Centre First Policy (TCFP) on shoppers' travel patterns (Cheshire *et al*, 2017), as adopted in England in 1996 (Department of the Environment, 1996). This policy was intended to 'redirect development, not just in retailing but in all 'key Town Centre uses', including leisure, office development and other uses, such as restaurants, to Town Centres', although the policy most notably affected the location of new retail development. As was shown in Cheshire *et al* (2015), this policy had a substantial negative impact on total factor productivity in the supermarket sector. Since the intended purpose of the policy was to 'increase the sustainability of cities' and 'reduce the need to travel', a logical topic for research is what impact it in fact had on shopping trips.

To begin to assess the impacts of the TCFP it is necessary to have definitions of where and what Town Centres (TCs) are. The policy was implemented, however, with no such definitions. While for England and Wales (E&W) TCs were subsequently defined in research commissioned by the relevant government department (ODPM, 2004), these were not official nor are they enforced: "It should be noted that these areas [Areas of Town Centre Activity] have no policy status and are not town centres for policy purposes – such centres will be designated in development plans." (ODPM & CASA, 2002).

The outcome of interest in Cheshire *et al* (2017) is the change in average shopping travel costs[2] as a result of the implementation of TCFP. Our hypothesis is that, if consumers live in an area with a given exposure to TCs, an increase in the supply of grocery stores affects them differently depending on their country. For example, in England, from the date TCF policy began to take effect new shops had to be concentrated in TCs, while in Scotland, the location of new shops was far less restricted. Given the dominant mode of travel – car – and the decentralisation of population, by concentrating retail in TC areas we expect that Town Centre First policy might not have contributed to reducing the average length of shopping trips. Our research question is, therefore, whether, for a given increase in the supply of shops, concentrating them in Town Centres influenced the average length of shopping trips.

Depending on where consumers live with respect to Town Centres (where new retail openings were concentrated), their shopping choices would therefore be more or less affected by changes in the supply of stores. Because, all else equal, consumers will prefer shops which are nearer to them, for each consumer's location we calculate exposure to the restrictions within a given distance from where the consumers live. We define a 'shopping area' for each consumer, centred on their location and with a radius determined by the observed patterns of shopping trips in the pre-policy period (1998)[3]. Given that TCFP restricted the opening of new retail units to 'Town Centres', the accessibility of town-centre space for each consumer's location affects the extent to which their shopping decisions were affected by the policy. We use several definitions of exposure to TCFP restrictions based on the location of the consumer's postcode sector with respect to the surrounding TCs and the size of these[4].

We illustrate this in Figure 1. This shows the potential relationships between the residential location of consumers (in a given postcode sector), the location and size of the surrounding TCs and the radius of the shopping areas. As sub-figures A and B show, depending how close and how large the TC(s) to which consumers have access to are, their shopping choices, and thus the

---

[2] The data on the changes in residential average shopping trips travel costs (time or distance) comes from the National Survey on Local Shopping Destinations (NSLSP) – years 1998 and 2008 –, kindly provided to us by CBRE. The consumers' residential information is available at the Postcode Sector level, which are small geographical areas aggregated from the full postal sector and there are around 11,000 of them in Great Britain.

[3] We also calculate the control variables within this radius. The main results in Cheshire *et al* (2017) are obtained using a radius of 10 kilometres, which captures around 90% of the shopping trips distances.

[4] The size of a TC can be measured using its area (square kilometres) or other metrics such as initial number of retail units (CBRE), address or small businesses counts (NSPD) or workplace employment (Census Small Area Statistics).

length of their shopping trips, would be differently affected. A consumer in situation A would be less restricted by the policy, as most of the new retail openings would be in a TC which is very accessible. A consumer in situation B would be more restricted by the policy, as just a small proportion of town-centre space is accessible to them given the extent of their shopping area.

To estimate the impact of TCFP on patterns of travel for shopping[5], it is therefore key to have clear definitions of Town Centres and to be able to apply the same definitions to contexts where TCFPs were not introduced. To provide the tools for such an evaluation, the focus of the present paper is to develop a method for predicting and estimating the location and extent of town-centre space in both England and Scotland. To do this we first obtained data on Town Centres as defined for 2000 from the Department for Communities and Local Government (DCLG)[6]. Even with the caveat that they have 'no policy status', these 'official'[7] TCs definitions are the most reliable and accurate definitions of town-centre space in England and Wales. They consist of GIS shapefiles for 1,075 Town Centres, of which the majority are defined as 'Areas of Town Centre Activity' (ATCAs) and 46 as 'Retail Cores' (RCs) - which overlap and are sub-centres of the ATCAs. From this long list, we discard a small number of TCs if no stores recorded in the National Survey of Local Shopping Patterns (NSLSP) existed within a certain distance of their boundary[8]. From these shapefiles we obtain the centroids of the TCs (called DCLG TCs in what follows). This identifies the central point in each town or city. . Separately, we obtain a list of alternative TCs for all Britain, from the Towns and Cities list in the Ordnance Survey (OS) Gazetteer. Below we refer to these as OSC TCs.

To predict the extent of the Town Centres (TCs) around these points, we use abundant small scale geographical information, in a range of 1-3 km from the centroids. We calculated a long list of geographical and socio-economic factors that relate to town-centre activities and regressed them

---

[5] Or in fact, to evaluate the impact of TCFP on any local economic outcome.

[6] Data for 2004 can be accessed at https://data.gov.uk/dataset/english-town-centres-2004, but we have also data for years 2000 and 2002 provided to us by DCLG. This data was originally created by the ODPM, Office of the Deputy Prime Minister (2001-2006), predecessor of the DLCG. The methodology is described in ODPM (2004).

[7] As we have said there are only unofficial estimations of TCs by the ODPMA. This is what we call pretended 'official'.

[8] The NSLSP data we obtained from CBRE consisted of an origin (postal sector) destination (store) matrix of shopping trips. We therefore used this data to obtain a list of main (grocery) stores which existed in 1998 which had to be mentioned as a destination in the NSLSP data. We can use this to infer is a TC is relatively important. This is below illustrated in Figure 2.

on the radius of the DCLG TCs (derived from the area of the shapefiles) to replicate as closely as possible the areas of the DCLG TCs for England and Wales. We then subject the results to robustness checks and, having satisfied ourselves as to the results, use the estimated coefficients in a separate exercise to a set of locations (OSC TCs) available for all three countries of Great Britain and predict their size. By doing this, we obtain a full set of estimated Town Centres boundaries for all countries in Britain, and, in particular, Scotland on a measure consistent with that used to identify the DCLG TCs for just England and Wales.

The rest of the paper is organised as follows. In Section 2 we discuss the definition of Town Centres and some existing methods to identify their location and boundaries. In Section 3 we describe the existing data on Town Centres for England and Wales. In Section 4 we explain our methodology to predict the location and extent of Town Centres for all of Great Britain, and provide some summary statistics to check how well the method works. Finally, Section 5 concludes.

2. What is a Town Centre and how should it be identified?

As noted in the introduction, identifying the exact boundaries of town centres (TC) is a more challenging question to answer than it appears at first sight. TCs are not definite entities. They might not be located at the geometric or geographical centre of a city and they might have fuzzy or indeterminate borders. The 'ideal' TC is not a point but is represented by a space of significant dimension. As the Oxford English Dictionary (OED) defines it: 'the central part or main business and commercial area of a town'. In general conversation, people might understand a TC to be the focal point of a city where main roads converge and people congregate. Historically the town or city centre was a place where citizens met or gathered: the place of the Italians' *passegiata*. Another function of a TC, captured in the OED definition, is as a space where jobs are concentrated, a shared workplace for people who live more spatially dispersed; a centralized destination (workplace) for decentralized origins (households). Firms locate in TCs to be able to draw on a wider pool of labour. So, people commute to work in TCs. And the third main function of TCs is as a commercial hub, the space where people shop. 'High Streets' are located in the TCs.

But the space that represents a TC not only need not be at the geometric centre of a city, it does not have a unique shape. It would only be like that in a location that is constructed according to a rigidly imposed, utopian planning scheme, where all the uses and functions identified would be neatly and exclusive concentrated in only the TC, and TCs would have some uniform shape. Real TCs, in real cities, are much more messy and diverse, sometimes two or three blocks in the centre of a small town and sometimes very extensive. For example, Central London's DCLG 'designated' town centre extends over 44 square kilometres, centred around Trafalgar Square and includes many retail sub-centres, areas focused on business, and other specialised areas such as 'theatre land' or entertainment zones with a concentration of restaurants and nightlife. The diversity of real TCs certainly adds to choice and likely generates greater productivity and welfare. Left to choose for themselves, businesses and individuals will usually find superior locations to those decided on by urban planners although there are significant qualifications resulting from externalities in land use that individualistic decision-makers will tend to ignore.

If we are to reliably identify TC areas, then, we ought to give due weight to the location of all the main functions discussed above to identify the location size and shape of the TC. All three aspects of TCs tend to be problematic theoretically and empirically. Centres do not need to be at the centroid of the city or some set of central jurisdictions. The observed shapes of TCs are motley and uneven. Size is also contentious. Empirically, in this paper we try to predict radiuses using a model with over 65 explanatory variables that captures all the multiple dimensions of 'town centeredness'.

Attempts to provide operational definitions of TCs have been lead historically by what is now the DCLG (see Thurstain-Goodwin *et al* (2000), ODPM & CASA (2002) ODPM , ODPM (2004), and more recently Dolega *et al* (2016)). ODPM & CASA (2002) starts by discussing a TC definition that depends on the perspective of a particular stakeholder. For instance, a taxi-driver would have a different definition of a TC to a planner. For the taxi-driver, the areas with the highest footfall can be determinants, while for a planner, the future evolution of the area might be a priority. Moreover, ODPM & CASA (2002) make the definition of TCs relative to other features of a city, creating an open approach from which they can build their model to define TCs. The result is that their TCs are necessarily diverse. For some TCs the priority would be 'a retail core, and office centre and an area of high building density', while for others, 'a concentration of visitor attractions and associated retail outlets' would be the focus (Thurstain-Goodwin *et al*, 2000). What is meant by this is that it is essential to include multiple dimensions and functions not just focus on one. This implies that TCs are 'indeterminate objects' with fuzzy borders, extremely difficult to define and agree upon. We can add that an operational definition should be implemented with consistency over an entire set of cities, because the identification of a TC remains problematic. For example, Wolverhampton's TC has a distinct ring road, some emergency services use it as a boundary, but administrative boundaries have been set in a much more extensive area reflecting a longer-term strategic vision of how the Town Centre should evolve (ODPM & CASA, 2002).

Typically, humans can easily detect an outlier, but not as easily notice when observations are clustered (Everitt *et al, 2011)*. Estimating *kernel density functions* can help identify clusters of 'objects'. These generate surfaces similar to mountainous terrain. This is called 'smoothing' and permits discrete and clustered data to be transformed into these mountain ranges. The kernel counts the number of observations in a given two-coordinate space as a histogram would, but it uses the number of observations to amplify a pulse function (rectangular, triangular, or normal most commonly) (Everitt *et al*, 2011). Thus, waves effectively transform the discrete information of the numbers and intensities of the points into peaks and valleys. The key parameter is the bandwidth, which can be adjusted (Everitt *et al*, 2011).

A very small bandwidth creates a single point to be counted independently, resulting in a spiky, disaggregated graph. An even smaller bandwidth provokes equal-sized pulse functions independent of each other if the observations are not located in exactly the same place. A very high bandwidth includes all points in a uniform one-shaped image equal to the generating pulse kernel. Figure A1 in the Appendix (modified from Everitt *et al*, 2011) shows an example of a one-dimensional normal kernel function for extremely low, low, optimal and extremely high bandwidths. So, to be useful a researcher estimating kernel density functions needs to find a Goldilocks bandwidth neither too high nor too low. Many techniques have as a result been elaborated for finding such appropriate bandwidths. Then comes the next vital step: slicing the surfaces to get the curves or contour maps which are much easier to interpret. Thus, clustering can be detected by higher mountains, and areas where data points are scarce can be detected by lower ones.

Thurstain-Goodwin *et al* (2000) define an index of intensity of 'town centeredness' using the dimensions of property, economy, diversity and visitor attractiveness. Because the categories are different in units, they employ a z-score normalisation. The model is populated by points at the Unit Post Code (UPC) level (full postcodes), shaping town centeredness as a mass function that is sliced for visualisation. The intensity of the functions helps to delimit the border of the Town

Centres, the visualisation of which is the point of the study. The ODPM reports (ODPM & CASA (2002), ODPM (2004)) are based on this methodology.

A catchment area is the area that draws in some group – customers or workers for example. A gravity model adds some forces of attraction and repulsion. Gravity models are simple but can be empirically well-behaved and make good predictions. In the case of a retail centre, gravity models typically use square footage of retail space as a measure of size and travel time between retail centres for distance. The so called 'Huff model' (Huff, 1963) uses square footage as a directly proportional proxy of the number of products a consumer would find in each shopping centre and time as an inversely proportional proxy of the cost (including opportunity costs) of travelling to the given retail centres. Then, the more products there are and the greater quantity of a given product that is sold – represented by the square footage dedicated to a given kind of product – the greater the probability of visiting the given retail centre. And the lower the cost – measured as time – the greater the probability of visiting a given retail centre. The model has in the numerator the linear probability of the consumer choosing the retail good of a given type and in the denominator the sum of the linear probabilities of choosing all types of retail goods.

The Liverpool group, Dolega *et al* (2016), discusses a method of defining TCs based on catchment areas. In summary, their method consists of replicating a catchment area for multiple stores. They use the Huff-model (Huff, 1963, 1964, 2003) mentioned above. In this the probability, P_{ij}, that a consumer located at i chooses to shop at retail centre j is:

$$P_{ij} = \frac{A_j^{\alpha} D_{ij}^{-\beta}}{\sum_{j=1}^{n} A_j^{\alpha} D_{ij}^{-\beta}}$$

where:

A_j is a measure of attractiveness of retail centre j, as square footage.

D_{ij} is the distance from location i to shop j.

α is the attractiveness parameter to be estimated.

β is the distance decay parameter to be estimated.

Until recently the estimation of these parameters did not have known properties of large samples. Huff (2003) suggests it is necessary to explore alternative models similar to those presented in this paper. In addition, Dolega *et al* (2016) suggests calibration at a national level would be superior to a local or subnational one. We also include a national level estimation in our model.

The approach of we take is more pragmatic and, in spirit, closer to Thurstain-Goodwin *et al* (2000). We take the extent of the DCLG-defined TCs (their area-imputed radius) as 'true' on average and collate a long list of explanatory factors that we believe correlate with town-centre activities and characteristics to predict the TC radius. Then, having satisfied ourselves that the method provides sufficiently high goodness-of-fit, we use the estimated coefficients from this prediction to extrapolate out-of-sample and apply the coefficients to a different set of locations. Details of the data used for the estimates and the details of the method are explained more fully in the next two sections.

3. The existing Town Centres data for England and Wales

As explained above, the first step of methodology relies on the use of a given set of town centre locations that we believe are accurately estimated: those identified by DCLG for England and Wales and as defined for 2000. Thurstain-Goodwin *et al* (2000) and ODPM & CASA (2002) set out a methodology to identify what they call 'Areas of Town Centre Activities' (ATCAs), which was generalised to all E&W locations in ODPM (2004). In the 2000 data, there are 1,029 ATCAs, and additionally, within these ATCAs 46 'Retail Cores' (RCs), giving a total of 1,075 TCs for E&W.

The Areas of Town Centre Activities (ATCA) were constructed using information on retail floorspace (supplied by the Valuation Office Agency - VOA) and using a model which identifies concentrations of the type of activities and patterns of property development likely to be found in Town Centres. These occur where there are high levels of employment in economic activities common to TCs (including retail, offices and leisure activities), a diversity of these activities, and a high density of office and retail floorspace. The ODPM/DCLG model is derived from estimates of employment and diversity of employment from the Office for National Statistics (ONS) Annual Business Inquiry (ABI), along with retail and office floorspace data from the VOA. Estimates are mapped at the detailed unit level postcode to produce a surface of economic activity. Cutting through the peaks in the activity at a prescribed level for the whole of England and Wales gives the ATCA boundaries. Intuitively, combining employment and retail floorspace data, a 3-D data surface was constructed for different locations in England and Wales where the tallest peaks identified the largest concentrations of retail activity. Then, contours were drawn around these peaks and the resulting areas were identified as ATCAs. In a second step, the data was crossvalidated using external sources to make sure they corresponded to the main centres of activity in England and Wales.

Even if the ODPM/DCLG are not fully operational for robust policy evaluation, since they correspond to revealed TC space and not planners' TCs as used for purposes of policy, they are the best definitions available to us and their identification is based on very small geography high quality data. However, for the purposes of the evaluation in Cheshire *et al* (2017), there exists an important limitation meaning we cannot simply use this data. This critical limitation is that these TC are not defined for Scotland and to evaluate the impact of TCFP we wish to be able to compare developments in Town Centres in England and Wales, where policy was strictly applied, to those in Scotland, where it was not.. At the same time, we cannot replicate the exact methodology of ODPM/DCLG using data for Scotland because either this data is not readily available to us (for example the postcode-level information on different activities), or it does not exists for Scotland (for example the VOA data). Given these reasons, we opted to exploit the information on the size of the TCs that we can derive from the E&W set in the DCLG data, and combine it with a very rich dataset on small geography explanatory factors (including socio-economic and topological features) that can successfully explain the variation in town-centre space we observe in the data.

4. Identifying town-centre space for all locations in Great Britain: methodology

We combine data at small geographical scales from multiple sources to predict the extent of Town Centres for the whole of Great Britain. There are seven steps in the process:

1. **Select DCLG 2000 TC sample:** we start the process by exploring the DCLG list of TCs for E&W for the year 2000[9]. From their observed surfaces we find the radius representing all the TCs as circular[10]. Then we select the samples for the regressions in step 4. Out of the 1,075 TCs (1,029 ATCAs and 46 RCs), we select two main samples: (1) all ATCAs; (2) ATCAs and, for Central and West London, the RCs. From these

[9] Unfortunately, year 2000 is the earliest date at which the data is available. We do not think this is a major limitation. Even if TCFP was first implemented in 1996 it is likely that it took time for the policy to be fully in place. At the same time, the underlying information used for the construction of the DCLG TCs as in 2000 mainly comes from 1998 and 1999. Without loss of generality we will consider these TC definitions to correspond reasonably to the start of the policy period. In addition, we tried to collect the data used in the estimation described in steps 3-4 for years 1998 or before when available, but for some of the most recent datasets that was impossible. In any case, the variables with highest partial explanatory power mostly date to 1998 or before.

[10] If, instead, we use the average distance to boundaries the results do not change.

samples we drop the TCs which we consider 'too small'or that cannot be used in the estimations. To identify these we use the information in NSLSP on the location of (grocery) shops in 1998 (more details are provided below). The final samples have between 810 and 950 TCs located in E&W. The mean radius of these is slightly less than 250 meters. We then create centroids from the shapefiles of these TC.

2. **Identify alternative TC locations for all GB**: we define an alternative list of TC centroid candidates using the Towns and Cities information in the Ordinance Survey (OS) Gazetteer Towns and Cities[11]. Initially, there are 1,315 Towns and Cities in Great Britain as a whole. The list is further trimmed when we combined it with the data around the centroids. The exact location of some of these town and city centroids was 'corrected' by looking at where popular map navigation tools (such as Open Street Map or Google Maps) located the city centroid.
3. **Collection of data around the centroids of the DCLG and OSC TCs:** we collect abundant information at very small geographical scales (the largest is the Output Area and the smallest are postcode units) for the areas around the centroids of the DCLG and OSC town centres. The main results (presented here) use information around 1 kilometre (km) of the centroid, but we also calculated all the models using information around 2 and 3 km (these models had less predictive power so we favoured the ones using 1 km). We believe that these long list of socio-economic and topological features around 1 km of the centroid are sufficient to satisfactorily predict the extent of town-centre space around these centroids (remember the average DCLG TC radius is around 250m). We obtained information on multiple variables (over 100) and 66 were used for the regressions of step 4. The list of variables and their data sources appear in Table 1.
4. **Estimation of the factors determining the extent of town-centre space:** for the DCLG TCs samples selected in step 1 we estimated several models where we explained the (log) radius of the TC as a function of the large set of explanatory variables around 1 km of the centroid of the TC. The results are shown in Table 2 and discussed below. The majority of estimates are significantly different from zero and the models have high goodness-of-fit statistics (R^2 between 0.78 and 0.88), especially when the size of the sample is taken into account.
5. **Validation of the results (within DCLG sample):** We use the coefficients estimated in step 4 to predict the (log) radius of the DCLG TCs, both for the whole sample (1,001)

[11] We experimented using an alternative list based on the Towns and Cities feature of the settlements layer of the OS Strategi. The results were very similar.

and for the samples used in each of the models estimated (referred as sub-samples in the tables). We both summarise and correlate the actual and predicted radius (and derived area) and use this to check the internal validity of the methodology. The results are shown in Tables 3 (and A1) and 4, and are discussed below.

6. **Application of the model to predict town-centre space around the OSC TCs:** The results from step 5 give us sufficient confidence that the models are sufficiently accurate in their prediction of the extent of town-centres for different values of the explanatory factors. We, therefore, proceed to apply the estimated betas from step 4 to the 'out-of-sample' list of OSC TCs and calculate the predicted (log) radius and area for these locations. This generates a set of estimated surrogate TC shapefiles to cover all the TCs of Great Britain. We can compare the predicted radius for the two sets of TCs (DCLG and OSC) for the sample which is available in both datasets (e.g. E&W and England and Wales separately). This is done in the first rows of Table 5.
7. **Comparison of socio-economic variables within the DCLG and OSC TCs**: the DCLG TCs and our predicted OSC TCs differ in two dimensions: their particular size for a given set of explanatory factors (which we fit in step 4) and their specific location. The precise places where the OSC and DCLG centroids are located can differ and, in particular, there is no comparison group for Scotland. So in this step 7 we calculate some socio-economic descriptive statistics (population, number of addresses, number of shoppers, etc.) within the boundaries (or a small distance of them) of the two sets of TCs. The summary statistics for these are shown in the remaining rows of Table 5. These allows us to check whether, even when located at slightly in different places, the underlying economic factors within TC boundaries are comparable in the two samples and, additionally, to explore how different the Scottish TCs are with respect to those in England and Wales.

The DCLG 2000 TC dataset originally contained 1,075 units. When we calculate the variables included in the estimation of step 4, as these are within 1 km of the centroid, a number of TCs are dropped from the sample (for example, for the 1 km dataset we are left with 1,001 TCs). The same occurs in the OSC sample (from 1,315 we are left with 964 locations). Also, the ATCAs and RCs overlap, so we do not want to use all of them. In columns (1) to (3) of Table 2 we use all the ATCAs and none of the RCs. However, for Central and West London, the ATCAs are very large and they mask the richness of small sub-centres (towns) within them. This is depicted in Figure A2. In columns (4) to (6) of Table 2 we use the ATCAs in all E&W but for the Central and West London areas we use the RCs (purple areas)[12].

Within each of these two samples, we introduce an additional criterion to select which TCs to include in the estimations. The NSLSP 1998 data allows us to map a set of (grocery) shops (approximately 4,700) which the consumers identify as their main shopping destinations. TCs which do not contain any of these shops within a certain distance of their boundaries can be considered to be of 'less-importance'. This is illustrated in Figure 2: for areas around Manchester and Glasgow we plot (orange triangles) the NSLSP shops in 1998. We calculate, for both the DCLG and the OSC samples, the number of shops (and shoppers that choose those shops) within different distances of the TC boundary. We can choose a threshold beyond which we consider the shop too far to be part of that TC. A TC can have shops inside its boundaries, within some allowed distance of its boundary (fuzzy) or beyond an allowed distance of the boundary[13]. In the full results we used six distance tolerance levels (fuzzy boundaries): 0m (at least one shop completely within the TCs), 10m, 100m, 250m, 500m and 1 km. Without loss of generality, for the regression results provided in the paper we focus on 10, 100 and 500m. The use of this restriction is what makes the sample size in columns 1 to 6 differ from one another. In step 6, we also use the fuzzy boundary criterion to select which OSC TCs are relevant in our sample.

In step 3, we select a large number of explanatory factors to predict the extent of TCs. We choose factors that we believe relate to town-centre activities. This step involves the collection of potentially relevant variables; GIS work to geographically match the data; and then choosing what variables to include in the final empirical model mainly on the basis of intuition and goodness-of-fit. This is akin to a forecasting and descriptive process so we do not pay serious attention to multicollinearity. The variables include factors related to the concentration of retail and shopping activity (we use two dataset from CBRE, NSLSP and RETail LOCations (RETLOC),

[12] In all the maps the background geographic areas are the postal sectors.

[13] In the map for Manchester we can observe all these cases: first inside the ACTA area of Eccles there are 3 shops (orange triangles), while The Trafford Centre has a nearby shop outside its ATCA area but probably inside a 100 m and also a 500 m buffer from its boundary. Finally, Oldham Road, close to the Manchester metropolitan area has no nearby shop, so it should be dropped from our sample if one of our many restrictions applies. In the map for Glasgow there are no ATCA areas – because there are no 'official' areas defined for Scotland, only our predictions. These are shown as purple circles around Glasgow and Renfrew. Inside Glasgow's predicted TC there are 6 shops (orange triangles) and 1 nearby probably at a buffer distance of 10, 100 and 500 m.

the second being a directory of all shops in the area, not only grocery stores); size of the retail sector (units and employment); socio-economic and workplace based factors; infrastructure endowments; local amenities (cultural, consumption, institutional); postcode centrality (based on the order of the postcodes within the postal sector, district and town); location (distance to social and natural amenities); topological features (elevation and slope) and nightlights brightness intensity[14]. We calculated these features around 1, 2 and 3km of both the DCLG and OSC TCs, but in this paper we focus on the results using 1 km. The different sets of explanatory variables and their main data source are summarised in Table 1.

In step 4 we use all the variables from Table 1 to predict the (log) radius of the DCLG TCs, and after checking how good the fit is (step 5) we apply the estimated coefficient to data around the OSC TCs. Formally, our methodology is in two steps. First, we estimate the extent of TCs regressing the explanatory variables (such as shoppers, socio-economic, etc.) on the TC radius using the DCLG England and Wales TCs ($DCLG$):

$$\log(TC\ radius_{DCLG}) = \alpha + \beta_{1,DCLG} shoppers_{DCLG} + \beta_{2,DCLG}\ socioeconomic_{DCLG} + \cdots + \varepsilon \qquad (1)$$

The results of the regressions on the six DCLG 2000 samples explained above are provided in Table 2. Most of the estimates are significantly different from zero (and by groups, all the sets of explanatory variables are jointly significant) and the goodness-of-fit of the models is very high (R^2 between 0.78 and 0.88). This suggests that our models predict the extent of TCs relatively well.

Having estimated these models, we can save the resulting coefficient values and apply them to different values of the explanatory variables. We do that in steps 5 and 6. In step 5 we apply the coefficients to the DCLG sample in order to compare the predicted and actual TC radius (and area) for the estimating sample. In Step 6 we apply the coefficients out-of-sample to the set of OSC TCs to predict the extent of town-centre space for the new set of TC locations. Formally, we calculate the prediction by multiplying the estimated $\hat{\beta}$s to a different set of locations OSC:

[14] We experimented adding additional topological features related to land use (EEA Corine data) and other natural boundaries (share of land in water bodies and green spaces) but they did not add any further explanatory power to the models.

$$\log(TC\ \widehat{radius}_{OSC}) = \hat{\alpha} + \hat{\beta}_{1,DCLG}\, shoppers_{OSC} + \hat{\beta}_{2,DCLG} socioeconomic_{OSC} + \cdots \quad (2)$$

Table 3 (and Table A1 in the Appendix for England and Wales separately) summarises the actual and predicted values for the radius and area of the DCLG TCs for the whole sample (1,001, 1,075 TCs minus 74 TCs without shops within 1 km of the centroid) for each of the 6 specifications of Table 2 and for the average of the 6 predictions. In the bottom panel, for each model, we show the summary statistics of the predictions when we restrict the observations to the sample used in each of the estimated models. By comparing the numbers in each row with the actual values in the first row, we can see that on average, the actual values are very similar to the actual TC values. In Table 3 we provide correlations between the actual and predicted values for the same samples for both E&W and separately by country (Table A1). The correlations are again very high, and, in some cases (especially for predicted area) they are almost equal to 1.

Once we obtain the coefficient is step 4, in step 6 we apply them to the OSC centroids and calculate the predicted radius, and create buffers around the OSC to draw the extent of the OSC TCs in a map. Figures 4, 5 and 6 illustrate the method. The DCLG TCs are depicted in blue and the OSC TCs are depicted in purple. The background geographical boundaries correspond to the postal sectors.

Figures 4 and 5 show the steps of the prediction method in 3 boxes, one for Manchester (in England) and one for Cardiff (in Wales). Box A shows the blue DCLG's 'main' 2000 TCs around Manchester (Figure 3) and Cardiff (Figure 4). The fragmented and varied shape of the TCs identified for the DCLG's is obvious. Then in Box B, the purple points show a list of OS Gazetteer Towns and Cities around Manchester and Cardiff. Finally, in Box C, our predictions of the extent of TCs around these centroids are seen as purple circles around Manchester and Cardiff. As explained, these predictions have been made by applying the estimated coefficients from Table 2. We can see that for these two cases, the location and extent of both DCLG and OSC TCs is very similar.

Figure 5 shows the predictions around Edinburgh and Glasgow, where there is no DCLG counterpart since they are located in Scotland. As expected, the size of the circles of the two major Scottish cities is larger than those in the neighbouring smaller towns.

The differences between the DCLG and OSC samples is due to the facts that (i) the number/location of what are considered Towns differs, (ii) we do not have a comparison group for Scotland, and (iii) the shape of TCs differs. Concerning shapes, OSCs are circles, while the DCLG have various shapes. As discussed and shown in Tables 3 and 4, we compare the actual and predicted values in the DCLG sample and they are very similar.

Even if the estimated values of the R^2s and the in-sample validations make us confident that we can successfully predict the radius within a TC centroid, we could still be getting the 'location' of the TCs wrong if the OSC centroids are not located in the same place as actual TCs. For this reason, in step 7 we provide a final validation exercise, which is to compare the socio-economic characteristics of the TCs in the DCLG and OSC samples, both for the countries in which we have information for both (E&W) and also for Scotland and the whole of Britain. This is done in Table 5. The table shows the average value for a set of socio-economic and shopping variables using both the DCLG and the OSC sample (we use the criterion of one shop within 500m of the boundary to select our TCs). These values were obtained combining data from Table 1 and information of the location and extent of the TCs (the original DCLG 2000 shapefiles and the buffer OSC TCs using the average prediction for the six models of Table 2).

The average value of the variable is provided both for its level and for the by-square-kilometre values (to normalise by the size of the TCs and make them more comparable). The DCLG TCs seem to be slightly larger than the OSC ones, but in general both samples seem quite similar. The number of TCs also differs, with more TCs in England in the DCLG sample and fewer in Wales. The last columns show the values for the sample for the whole of GB and for Scotland alone. The Scottish values seem to be somewhere in between the English and the Welsh ones, but they do not look extremely different from the average British or E&W values. In a nutshell, the values in Table 5 suggest that the socio-economic and shopping density values of the DCLG and our OSC samples are quite comparable and so we can be reasonably confident that our methodology yields estimates of TCs for all three countries of GB very similar to those of DCLG for England and

Wales alone. It seems reasonable therefore, to proceed to use these estimated TCs as the foundation on which to implement our proposed strategy for comparing changes in English TCs with those in Scotland post 1996 as a methodology for evaluating the impact of Town Centre First policy.

5. Conclusions

The aim of this paper was to present a novel method for predicting the location and extent of TC space in all of Great Britain. Our method relies on four assumptions. The first assumption is that the DCLG TC definitions are good approximations of the true TCs for England and Wales. The second assumption is that the underlying socio-economic and geographic factors within a radius of around 1 km of the TC centroids are effective determinants of TCs. The validity of this assumption can be assessed by looking at the goodness-of-fit statistics of our statistical models predicting the extent of the DCLG's TCs and at the evidence provided in Tables 3 to 5. The third assumption is that the OS list of towns and cities provides a reliable set of potential TC locations. The final assumption is that the determinants of TC space in Scotland do not systematically differ from those in England and Wales, both in observed and unobserved characteristics relevant to defining TCs. If all these assumptions hold, we can satisfactorily apply the coefficients on socioeconomic and geographic variables estimated in Table 2 on Britain-wide data to yield estimates of the location and extent of TCs for England, Wales and, in particular, Scotland.

As mentioned above, this study gave an answer to the question of where the TCs are located, but there is no such thing as *the* answer. As the robustness checks and data validations suggest, the method can be considered 'successful' with a correlation of actual to predicted radius of 0.75-0.99 depending on the sample. Our predictions for England and Wales matches the actual DLCG ATCA 2000 quite accurately. In Scotland, its direct accuracy cannot be judged because there are no 'official' DCLG TCs for Scotland – to offset for which is the ultimate purpose of this study. However, the exploration of socio-economic and shopping density values in and very closely to the TCs defined with both methodologies suggests that they provide a very similar picture. Overall, our method is promising and certainly provides a useful tool to be applied for the evaluation of TCFP, and more generally, for the evaluation of any policy that applies to town centres.

References

Cheshire P., C. Hilber, and I. Kaplanis, *"Land Use Regulation and Productivity-Land matters: Evidence from a UK supermarket chain,"* Journal of Economic Geography, vol. 15, no. 1, pp. 43–73, 2015.

Cheshire P., C. Hilber, and R. Sanchis-Guarner, *"Do English Planning Policies Make Shopping More Accessible?,"* 2017.

Department of the Environment, "*Planning Policy Guidance: Town Centres and Retail Developments,"* 1996.

Dolega L., M. Pavlis, and A. Singleton, *"Estimating Attractiveness, Hierarchy and Catchment Area Extents for a National Set of Retail Centre Agglomerations,"* J. Retail. Consum. Serv., vol. 28, pp. 78–90, 2016.

Everitt B. and T. Hothorn, *"An Introduction to Applied Multivariate Analysis with R"*. 2011.

Huff D. L., *"A Probabilistic Analysis of Shopping A Probabilistic Analysis Trade Areas,"* Land Econ., vol. 39, no. 1, pp. 81–90, 1963.

Huff D. L., *"Defining and Estimating a Trading Area,"* J. Mark., vol. 28, no. 3, pp. 34–38, 1964.

Huff D. L., *"Parameter Estimation in the Huff Model,"* ArcUser, p. 3, 2003.

ODPM & Centre for Advanced Spatial Analysis (CASA), *"Producing Boundaries and Statistics for Town Centres - London Pilot Study Summary Report,"* 2002.

ODPM, *"Producing Boundaries and Statistics for Town Centres, England and Wales, 2000 Interim Report,"* 2004.

Thurstain-Goodwin M. and D. Unwin, *"Defining & Delineating the Central Areas of Towns for Statistical Monitoring Using Continuous Surface Representations,"* Cent. Adv. Spat. Anal., vol. 4, no. 18, pp. 1–11, 2000.

Tables

Table 1: List of explanatory variables included in our model

VARIABLE	*DATA SOURNCES*
Number of shoppers, shops and location of these	*CBRE: RETLOC and NSLSP data in 1998*
Number and employment in sector 52 (retail)	*Annual Business Inquiry (ABI) accessed via NOMIS*
Population, residential employment, workplace-based employment, share of occupations (in work-place employment) of different levels, residential socio-economic characteristics (age structure, unemployment rates, labour market, commuting patterns)	*Census, 2001*
Transport infrastructure (roads, rail, buses)	*Ordnance Survey (OS) Strategi, Open Street Map*
Cultural amenities (libraries, museums, art galleries, theatre, cinemas), consumption amenities (bars, restaurants), and historical amenities (landmarks, tourist info, local government)	*OS Strategi, OS Points of Interest (POI)*
Postcode centrality structure (sectors, districts and towns)	*ONS National Statistics Postcode Directory (NSPD) and Wikipedia*
Geographical location (distance to the coast, river, rail, town hall, natural park/woodland)	*OS Strategi, POI, Wikipedia*
Topological features: terrain elevation (m) and slope (degrees)	*OS Panorama 50x50*
Nightlight brightness intensity (96-97 average)	*NOAA-NGDC*[15]
Postcode centrality structure (sectors, districts and towns)	*NSPD*

[15] National Oceanic and Atmospheric Administration; National Geophysical Data Center, non-censored version.

Table 2: Town Centre Extent Prediction Model Results

DCLG 2000 TC Sample	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Type of TC	ATCA	ATCA	ATCA	ATCA	ATCA	ATCA
Fuzzy TC border	10m	100m	500m	10m	100m	500m
Central and West London	ATCA	ATCA	ATCA	RC	RC	RC
Number of observations	812	870	931	824	882	949
Adjusted R^2	0.878	0.868	0.810	0.841	0.834	0.779
R^2	0.888	0.878	0.824	0.854	0.846	0.795

	(1)	(2)	(3)	(4)	(5)	(6)
Log of grocery shoppers	0.020***	0.024***	0.031***	0.021**	0.025***	0.032***
(NSLSP98),1 km of centroid	[0.007]	[0.007]	[0.009]	[0.009]	[0.009]	[0.009]
Log number of shops (RETLOC/	0.024***	0.026***	0.021***	0.022***	0.024***	0.018***
NSLSP98), 1 km of centroid	[0.003]	[0.003]	[0.004]	[0.003]	[0.003]	[0.004]
Log average distance to shops	0.441***	0.401***	0.684***	0.435***	0.396***	0.631***
(RETLOC98), 1 km of centroid	[0.123]	[0.123]	[0.144]	[0.131]	[0.133]	[0.146]
Log average distance to shops	0.177***	0.161***	0.283***	0.175***	0.158***	0.261***
(RETLOC98), 1 km of centroid^2	[0.054]	[0.055]	[0.064]	[0.059]	[0.060]	[0.065]
Log average distance to shops	0.021***	0.019***	0.034***	0.021***	0.019**	0.031***
(RETLOC98), 1 km of centroid^3	[0.007]	[0.007]	[0.008]	[0.008]	[0.008]	[0.008]
Log average distance to shops	0.051	0.080	0.149	0.046	0.090	0.155
(NSLSP98), 1 km of centroid	[0.095]	[0.095]	[0.115]	[0.093]	[0.096]	[0.114]
Log average distance to shops	0.015	0.021	0.045	0.011	0.024	0.046
(NSLSP98), 1 km of centroid^2	[0.043]	[0.043]	[0.052]	[0.043]	[0.044]	[0.053]
Log average distance to shops	0.002	0.002	0.004	0.002	0.003	0.004
(NSLSP98), 1 km of centroid^3	[0.006]	[0.006]	[0.007]	[0.006]	[0.006]	[0.007]
Log distance to closest shop	0.054***	0.046***	0.013	0.059***	0.050***	0.018*
(NSLSP98), 1 km of centroid	[0.010]	[0.010]	[0.011]	[0.011]	[0.010]	[0.011]
Number of units retail sector	0.001***	0.001***	0.001***	0.001***	0.001***	0.002***
(ABI98), 1 km of centroid	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Number of units retail sector	0.000*	0.000**	0.000	0.000***	0.000***	0.000***
(ABI98), 1 km of centroid^2	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Employment retail sector	0.000	0.000	0.000	0.000	0.000*	0.000**
(ABI98), 1 km of centroid	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Employment retail sector	0.000	0.000*	0.000	0.000**	0.000***	0.000***
(ABI98), 1 km of centroid^2	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Log density of total population	0.362	0.211	0.457*	0.323	0.180	0.325
(Census01), 1 km of centroid	[0.254]	[0.230]	[0.251]	[0.251]	[0.228]	[0.238]
Log density of total population	0.024	0.015	0.032*	0.020	0.012	0.022

(Census01), 1 km of centroid^2	[0.016]	[0.015]	[0.016]	[0.016]	[0.015]	[0.016]
Log workplace employment	0.286***	0.284***	0.282***	0.281***	0.281***	0.279***
(Census01), 1 km of centroid	[0.024]	[0.024]	[0.028]	[0.026]	[0.026]	[0.031]
Share wempl high occupations	1.507***	1.647***	1.238**	2.391***	2.444***	2.130***
(Census01), 1 km of centroid	[0.481]	[0.465]	[0.573]	[0.505]	[0.494]	[0.553]
Share wempl medium occupation	1.838***	1.784***	1.308**	2.577***	2.434***	2.032***
(Census01), 1 km of centroid	[0.474]	[0.457]	[0.576]	[0.522]	[0.507]	[0.578]
Share wempl low occupations	2.292***	2.433***	2.104***	3.155***	3.201***	3.009***
(Census01), 1 km of centroid	[0.490]	[0.473]	[0.581]	[0.515]	[0.503]	[0.559]
Average commuting distance	0.013	0.012	0.012	0.018	0.017	0.018
(Census01), 1 km of centroid	[0.009]	[0.008]	[0.009]	[0.011]	[0.011]	[0.011]
Average commuting distance	0.000	0.000	0.000	0.000	0.000	0.000
(Census01), 1 km of centroid^2	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Percentage commuters by foot/	0.110	0.125	0.216	0.307**	0.305**	0.383**
bike (Census01), 1 km of centroid	[0.142]	[0.145]	[0.157]	[0.144]	[0.146]	[0.155]
Percentage commuters public	0.018	0.083	0.107	0.294**	0.344***	0.279**
trans (Census01), 1 km of centroid	[0.114]	[0.114]	[0.128]	[0.120]	[0.119]	[0.130]
Average household size	0.100*	0.030	0.042	0.141**	0.079	0.067
(Census01), within 1 km of centroid	[0.057]	[0.064]	[0.071]	[0.065]	[0.068]	[0.076]
Average age of resident population	0.019***	0.013*	0.009	0.028***	0.022***	0.018**
(Census01), 1 km of centroid	[0.006]	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]
Percentage of population aged 18-44 yo	0.297	0.184	0.064	0.733**	0.626*	0.571
(Census01), 1 km of centroid	[0.273]	[0.280]	[0.333]	[0.325]	[0.330]	[0.368]
Total unemployment rate	0.819***	0.815***	0.853***	1.178***	1.160***	1.298***
(Census01), 1 km of centroid	[0.285]	[0.290]	[0.326]	[0.293]	[0.295]	[0.341]
Share of students in population	0.933***	0.890***	0.663***	1.043***	0.983***	0.842***
(Census01), 1 km of centroid	[0.183]	[0.180]	[0.204]	[0.233]	[0.222]	[0.247]
Share of retired in population	0.699	0.476	0.274	0.842*	0.630	0.333
(Census01), 1 km of centroid	[0.450]	[0.453]	[0.475]	[0.478]	[0.477]	[0.501]
Log of km of all roads	0.005	0.004	0.007	0.001	0.004	0.013
(Strategy 2009), 1 km of centroid	[0.022]	[0.022]	[0.025]	[0.024]	[0.024]	[0.025]
Number of bus stations	0.001	0.000	0.001	0.005	0.004	0.004
(POI 2015), 1 km of centroid	[0.008]	[0.008]	[0.009]	[0.008]	[0.008]	[0.010]
Number of tube/tram stations	0.059**	0.057**	0.001	0.048**	0.047**	0.056***
(POI 2015), 1 km of centroid	[0.023]	[0.025]	[0.026]	[0.022]	[0.021]	[0.019]
Number of rail stations	0.003	0.000	0.012	0.000	0.002	0.010
(POI 2015), 1 km of centroid	[0.010]	[0.010]	[0.014]	[0.015]	[0.015]	[0.015]
Number of libraries	0.014	0.011	0.032**	0.028**	0.026**	0.045***
(POI 2015), 1 km of centroid	[0.010]	[0.011]	[0.015]	[0.012]	[0.012]	[0.014]
Number of museums	0.004	0.004	0.000	0.021*	0.011	0.010
(Strategy 2009), 1 km of centroid	[0.012]	[0.012]	[0.013]	[0.013]	[0.013]	[0.014]

Number of art galleries	0.010*	0.009*	0.010*	0.005	0.004	0.009
(POI 2015), 1 km of centroid	[0.005]	[0.005]	[0.005]	[0.005]	[0.005]	[0.006]
Number of cinemas and theatres	0.017***	0.016***	0.017**	0.018**	0.017**	0.019**
(POI 2015), 1 km of centroid	[0.006]	[0.006]	[0.008]	[0.008]	[0.008]	[0.008]
Number of discos and night clubs	0.004	0.003	0.002	0.022***	0.021***	0.021***
(POI 2015), 1 km of centroid	[0.004]	[0.004]	[0.005]	[0.005]	[0.005]	[0.005]
Number of landmarks	0.027*	0.020	0.005	0.035**	0.030**	0.019
(Strategy 2009), 1 km of centroid	[0.015]	[0.015]	[0.021]	[0.014]	[0.014]	[0.021]
Number of cafes, restaurants and	0.001*	0.001**	0.001**	0.001**	0.001*	0.001**
pubs (POI 2015), 1 km of centroid	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Number of B&B, hotels and motels	0.000	0.000	0.000	0.000	0.000	0.000
(POI 2015), 1 km of centroid	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Number of youth hostels	0.007	0.008	0.006	0.012	0.013*	0.008
(POI 2015), 1 km of centroid	[0.007]	[0.007]	[0.008]	[0.008]	[0.008]	[0.008]
Number of local government sites	0.015***	0.016***	0.013	0.012	0.013*	0.009
(POI 2015), 1 km of centroid	[0.005]	[0.005]	[0.010]	[0.007]	[0.007]	[0.011]
Number of tourist info offices	0.012	0.015	0.023*	0.012	0.014	0.025*
(Strategy 2009), 1 km of centroid	[0.012]	[0.012]	[0.013]	[0.013]	[0.012]	[0.014]
Number of visitor centres	0.002	0.001	0.004	0.005	0.004	0.006
(Strategy 2009), 1 km of centroid	[0.007]	[0.006]	[0.008]	[0.008]	[0.008]	[0.009]
Share of central addresses (pcsect	0.000	0.002	0.011	0.001	0.004	0.014
on pctown), 1 km of centroid	[0.031]	[0.030]	[0.031]	[0.031]	[0.029]	[0.030]
Share of central addresses (pcdistr	0.034**	0.038**	0.030	0.034*	0.038**	0.028
on pctown), 1 km of centroid	[0.017]	[0.017]	[0.019]	[0.019]	[0.019]	[0.020]
Log distance in km to first	0.002	0.004	0.009	0.002	0.003	0.008
postcode in the closest postal town	[0.005]	[0.005]	[0.006]	[0.005]	[0.005]	[0.005]
Log of distance in km	0.003	0.003	0.001	0.005	0.005	0.004
to closest town hall (Wikipedia)	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]
Log of distance in km to closest	0.002	0.000	0.010	0.006	0.004	0.011
rail or tube station (POI 2015)	[0.006]	[0.006]	[0.008]	[0.008]	[0.008]	[0.008]
Log of distance in km to closest	0.005	0.000	0.008	0.003	0.001	0.007
point in the coastline (Strategi 2009)	[0.005]	[0.005]	[0.006]	[0.006]	[0.006]	[0.007]
Log of distance in km to closest	0.008	0.000	0.008	0.013*	0.005	0.002
river or lake (Strategi 2009)	[0.006]	[0.006]	[0.006]	[0.007]	[0.007]	[0.007]
Log of distance in km to closest natural	0.001	0.003	0.003	0.001	0.004	0.003
park or woodland (Strategi 2009)	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
Standard dev of elevation	0.001	0.000	0.001	0.002	0.001	0.002
(Panorama), 1 km of centroid	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]
Mean of elevation	0.001	0.001	0.002*	0.001	0.001	0.002*
(Panorama), 1 km of centroid	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Max of elevation	0.002*	0.001	0.002**	0.001	0.001	0.002*

(Panorama), 1 km of centroid	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Range of elevation	0.002	0.001	0.002**	0.002	0.001	0.002**
(Panorama), 1 km of centroid	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Standard dev of terrain slope	0.016	0.018	0.022	0.015	0.017	0.017
(Panorama), 1 km of centroid	[0.017]	[0.016]	[0.017]	[0.018]	[0.017]	[0.017]
Mean of terrain slope	0.011	0.009	0.012	0.008	0.007	0.010
(Panorama), 1 km of centroid	[0.011]	[0.011]	[0.011]	[0.012]	[0.012]	[0.012]
Max of terrain slope	0.001	0.010	0.011	0.020	0.019	0.001
(Panorama), 1 km of centroid	[0.127]	[0.124]	[0.122]	[0.131]	[0.129]	[0.127]
Range of terrain slope	0.004	0.008	0.013	0.017	0.017	0.003
(Panorama), 1 km of centroid	[0.127]	[0.124]	[0.122]	[0.131]	[0.129]	[0.127]
Standard dev of lights brightness	0.005	0.003	0.006	0.004	0.004	0.000
(NASA 96-97), 1 km of centroid	[0.005]	[0.005]	[0.007]	[0.005]	[0.005]	[0.006]
Mean of lights brightness	0.000	0.001	0.001	0.000	0.001	0.001
(NASA 96-97), 1 km of centroid	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]
Max of lights brightness	0.001	0.000	0.000	0.000	0.000	0.000
(NASA 96-97), 1 km of centroid	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Range of lights brightness	0.002	0.001	0.003	0.001	0.001	0.001
(NASA 96-97), 1 km of centroid	[0.002]	[0.002]	[0.003]	[0.002]	[0.002]	[0.002]
Sum of lights brightness	0.000	0.000*	0.000*	0.000	0.000	0.000
(NASA 96-97), 1 km of centroid	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Constant term (excluded: occupation	4.977***	3.948***	5.621***	4.656***	3.798***	4.972***
other and commuter by motor vehicle)	[1.138]	[1.063]	[1.252]	[1.143]	[1.078]	[1.193]

Table 3: E&W TC radius/area (actual and predicted), and number of TCs, DCLG 2000 sample

	TC Derived Radius in km					TC Area in square km				
Variable	**Obs**	**Mean**	**StdD**	**Min**	**Max**	**Obs**	**Mean**	**StdD**	**Min**	**Max**
Observed value	1,001	0.247	0.153	0.113	3.187	1,001	0.265	1.047	0.040	31.905
Average all samples	1,001	0.261	0.165	0.107	2.973	1,001	0.308	0.994	0.036	28.018
Sample 1 betas	1,001	0.273	0.196	0.104	3.095	1,001	0.355	1.223	0.034	30.091
Sample 2 betas	1,001	0.270	0.195	0.106	3.076	1,001	0.350	1.203	0.035	29.718
Sample 3 betas	1,001	0.260	0.178	0.105	3.051	1,001	0.312	1.059	0.034	29.245
Sample 4 betas	1,001	0.257	0.148	0.105	2.526	1,001	0.276	0.718	0.034	20.040
Sample 5 betas	1,001	0.255	0.153	0.107	2.711	1,001	0.277	0.809	0.036	23.088
Sample 6 betas	1,001	0.249	0.163	0.106	3.382	1,001	0.278	1.184	0.035	35.929
Sub-sample 1	812	0.256	0.157	0.104	3.095	812	0.283	1.090	0.034	30.091
Sub-sample 2	870	0.250	0.153	0.106	3.076	870	0.270	1.044	0.035	29.718
Sub-sample 3	931	0.243	0.147	0.111	3.051	931	0.254	0.996	0.039	29.245
Sub-sample 4	824	0.251	0.120	0.105	1.040	824	0.243	0.306	0.034	3.397
Sub-sample 5	882	0.246	0.118	0.107	1.070	882	0.233	0.302	0.036	3.600
Sub-sample 6	949	0.239	0.112	0.108	1.030	949	0.218	0.281	0.036	3.336

Table 4: Correlation coefficients of real versus predicted values for radius and area, DCLG 2000 sample

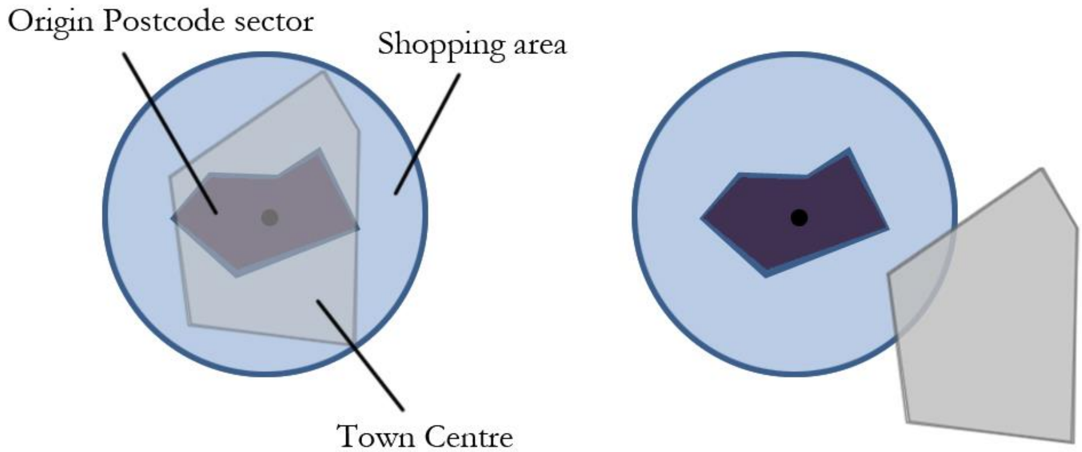
Town Centre Radius in km	All (1,001)			Sub-sample		
	All	*England*	*Wales*	*All*	*England*	*Wales*
Average all samples	0.755	0.757	0.866			
Sample 1	0.722	0.718	0.912	0.970	0.970	0.975
Sample 2	0.725	0.721	0.912	0.966	0.966	0.973
Sample 3	0.803	0.800	0.923	0.948	0.948	0.967
Sample 4	0.867	0.866	0.912	0.943	0.942	0.977
Sample 5	0.870	0.869	0.913	0.940	0.938	0.973
Sample 6	0.883	0.882	0.921	0.915	0.913	0.966
Town Centre Area in km²	All (1,001)			Sub-sample		
	All	*England*	*Wales*	*All*	*England*	*Wales*
Average all samples	0.928	0.928	0.865			
Sample 1	0.810	0.810	0.858	0.994	0.994	0.974
Sample 2	0.811	0.811	0.857	0.993	0.993	0.974
Sample 3	0.914	0.914	0.873	0.988	0.988	0.958
Sample 4	0.940	0.941	0.863	0.937	0.937	0.977
Sample 5	0.950	0.950	0.862	0.928	0.927	0.975
Sample 6	0.973	0.973	0.868	0.893	0.892	0.955
Sample size	All (1,001)			Sub-sample		
	All	*England*	*Wales*	*All*	*England*	*Wales*
Average all samples	1,001	944	57			
Sample 1	1,001	944	57	812	768	44
Sample 2	1,001	944	57	870	820	50
Sample 3	1,001	944	57	931	876	55
Sample 4	1,001	944	57	824	780	44
Sample 5	1,001	944	57	882	832	50
Sample 6	1,001	944	57	949	894	55

Table 5: Predicted size and socio-economics for OSC and DCLG's TCs (500 fuzzy boundary tolerance)

	E&W		ENGLAND		WALES		OSC NEW AREAS	
	OSC	**DCLG**	**OSC**	**DCLG**	**OSC**	**DCLG**	**ALL GB**	**SCOT**
Variable	*752*	*931*	*687*	*876*	*65*	*55*	*861*	*109*
Predicted TC radius in km	0.231	0.246	0.236	0.248	0.187	0.210	0.222	0.161
Predicted TC area in km²	0.215	0.257	0.222	0.263	0.135	0.161	0.201	0.109
NSLSP98 shops	2.20	2.37	2.23	2.37	1.92	2.33	2.20	2.17
NSLSP98 shoppers	20910.63	24085.03	21560.91	24416.07	14037.62	18914.91	20003.98	13748.96
NSPD Address counts	3183.63	4175.27	3291.85	4266.92	2039.77	2715.62	3044.47	2084.39
NSPD Small Businesses counts	419.88	479.08	433.57	487.45	275.25	345.84	398.52	251.14
NSLSP98 shops per km²	1.26	1.31	1.26	1.30	1.26	1.46	1.29	1.50
NSLSP98 shoppers per km²	11387.03	12928.03	11651.57	13032.77	8591.06	11292.24	11100.73	9125.53
NSPD Address counts per km²	1762.30	2130.30	1806.99	2162.26	1289.95	1621.23	1716.77	1402.70
NSPD Small Businesses counts per km²	208.51	217.23	212.86	218.54	162.48	196.36	200.70	146.79
Population	6004.47	8192.70	6221.02	8389.54	3715.80	5158.30	5716.36	3728.64
Residential Employment	2699.94	3708.18	2821.36	3817.85	1416.57	2017.62	2563.13	1619.28
Workplace Employment	7438.53	8828.14	7792.53	9088.64	3697.03	4812.55	6986.72	3869.62
Workplace High Occupations	3292.30	3972.25	3474.86	4113.57	1362.80	1793.74	3072.93	1559.44
Share of wemployment high occupations	0.350	0.363	0.353	0.366	0.310	0.327	0.345	0.314
Ratio wemployment over population	1.05	0.93	1.07	0.94	0.79	0.84	1.01	0.75
Population per km²	3718.12	4487.26	3785.34	4563.91	3007.71	3320.25	3634.28	3055.87
Residential Employment per km²	1658.80	2019.03	1706.66	2066.52	1153.01	1295.91	1616.10	1321.48
Work Employment per km²	3501.33	3366.00	3611.98	3417.46	2331.83	2582.36	3340.43	2230.32

Figures

Figure 1: Illustration of different exposure to town-centre space within shopping-area



A Town Centre is substantially co-incident with shopping area **B** Town Centre only slightly co-incident with shopping area

Figure 2: Shops inside and outside an ATCA, Manchester and Glasgow areas

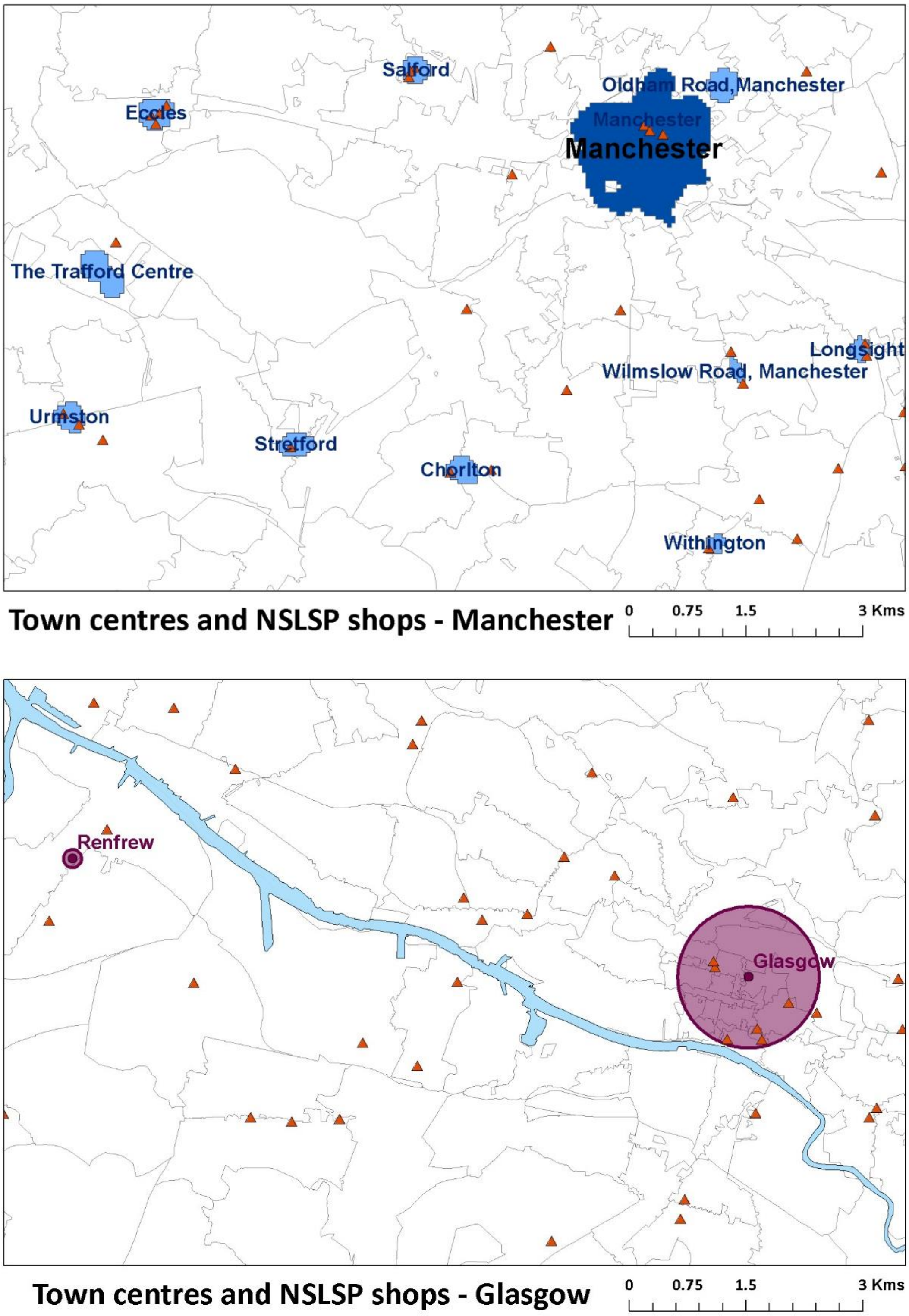


Figure 3: TCs prediction, Manchester, England (Step-by-step)

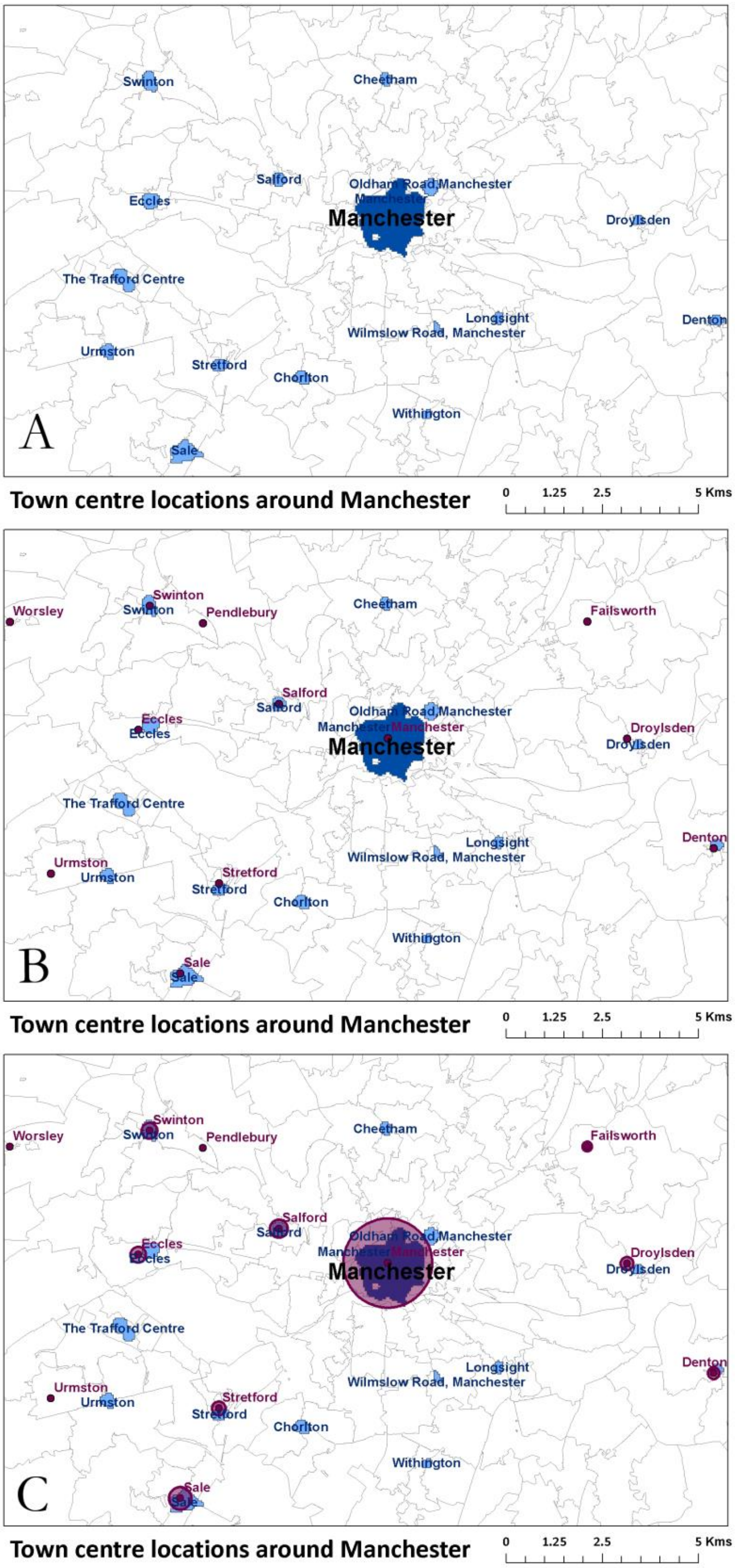


Figure 4: TCs prediction, Cardiff, Wales (Step-by-step)

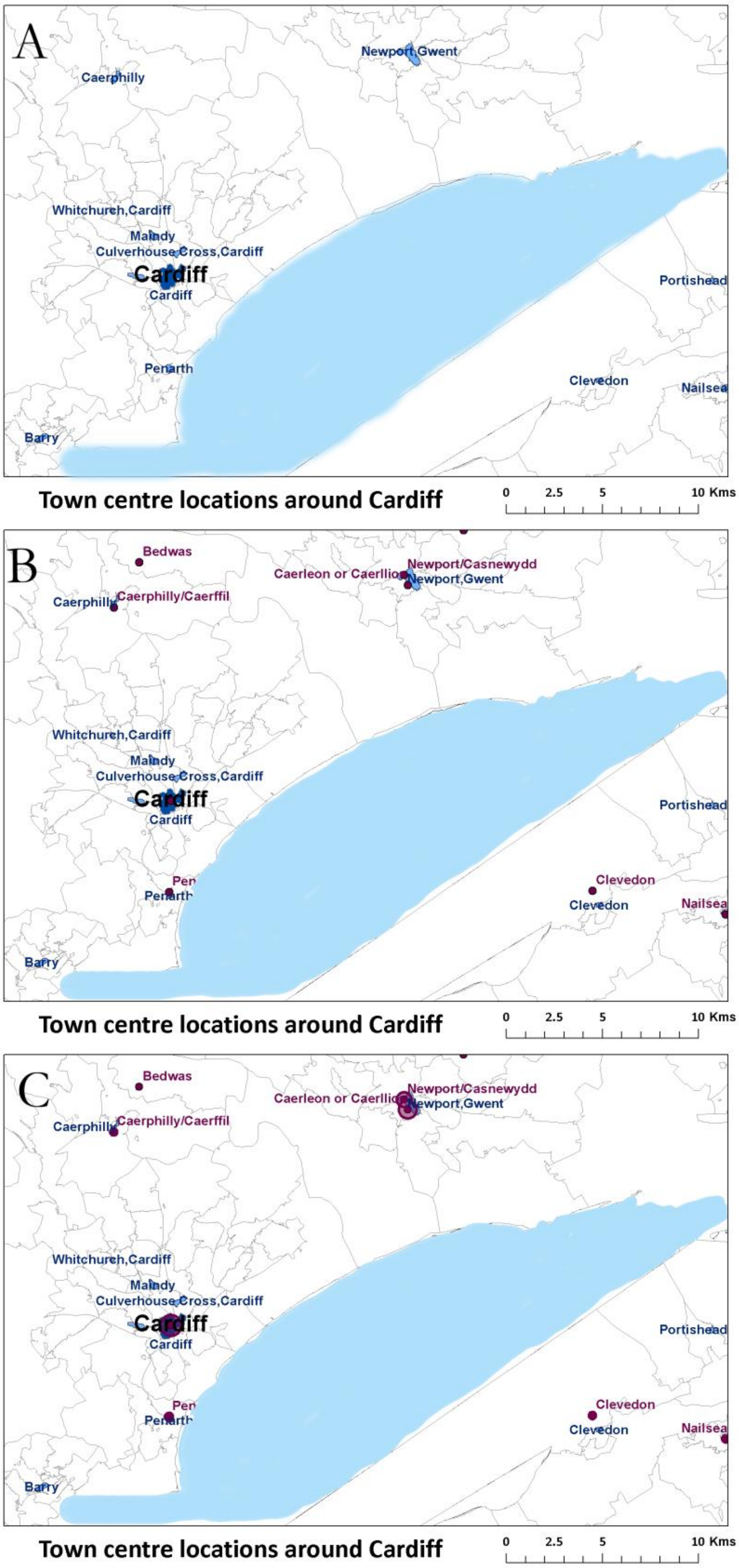


Figure 5:. TCs prediction, Edinburgh and Glasgow.

Appendix

Table A1: England and Wales TC radiuses, actual and predicted

ENGLAND

	TC Derived Radius in km					TC Area in km^2				
Variable	**Obs**	**Mean**	**StdD**	**Min**	**Max**	**Obs**	**Mean**	**StdD**	**Min**	**Max**
Observed value	944	0.248	0.156	0.113	3.187	944	0.270	1.077	0.040	31.905
Average all samples	944	0.263	0.168	0.107	2.973	944	0.316	1.021	0.036	28.018
Sample 1 betas	944	0.276	0.200	0.104	3.095	944	0.366	1.257	0.034	30.091
Sample 2 betas	944	0.274	0.199	0.106	3.076	944	0.360	1.237	0.035	29.718
Sample 3 betas	944	0.263	0.181	0.105	3.051	944	0.320	1.089	0.034	29.245
Sample 4 betas	944	0.259	0.150	0.105	2.526	944	0.282	0.738	0.034	20.040
Sample 5 betas	944	0.257	0.155	0.107	2.711	944	0.283	0.831	0.036	23.088
Sample 6 betas	944	0.252	0.166	0.106	3.382	944	0.285	1.218	0.035	35.929
Sub-sample1	768	0.258	0.160	0.104	3.095	768	0.289	1.120	0.034	30.091
Sub-sample2	820	0.252	0.156	0.106	3.076	820	0.276	1.075	0.035	29.718
Sub-sample3	876	0.246	0.150	0.111	3.051	876	0.260	1.025	0.039	29.245
Sub-sample4	780	0.253	0.121	0.105	1.040	780	0.247	0.311	0.034	3.397
Sub-sample5	832	0.248	0.119	0.107	1.070	832	0.237	0.308	0.036	3.600
Sub-sample6	894	0.241	0.113	0.108	1.030	894	0.222	0.287	0.036	3.336

WALES

	TC Derived Radius in km					TC Area in km^2				
Variable	**Obs**	**Mean**	**StdD**	**Min**	**Max**	**Obs**	**Mean**	**StdD**	**Min**	**Max**
Observed value	57	0.218	0.096	0.113	0.613	57	0.178	0.198	0.040	1.180
Average all samples	57	0.216	0.095	0.128	0.570	57	0.175	0.199	0.052	1.022
Sample 1 betas	57	0.219	0.097	0.125	0.585	57	0.180	0.206	0.049	1.077
Sample 2 betas	57	0.217	0.098	0.128	0.590	57	0.178	0.209	0.051	1.095
Sample 3 betas	57	0.213	0.091	0.122	0.534	57	0.168	0.182	0.047	0.895
Sample 4 betas	57	0.219	0.098	0.128	0.588	57	0.181	0.209	0.052	1.087
Sample 5 betas	57	0.217	0.098	0.127	0.589	57	0.178	0.210	0.050	1.089
Sample 6 betas	57	0.213	0.090	0.123	0.532	57	0.167	0.179	0.048	0.888
Sub-sample1	44	0.218	0.093	0.125	0.538	44	0.218	0.093	0.125	0.538
Sub-sample2	50	0.212	0.089	0.128	0.538	50	0.212	0.089	0.128	0.538
Sub-sample3	55	0.208	0.082	0.122	0.495	55	0.208	0.082	0.122	0.495
Sub-sample4	44	0.218	0.093	0.128	0.538	44	0.218	0.093	0.128	0.538
Sub-sample5	50	0.212	0.089	0.127	0.535	50	0.212	0.089	0.127	0.535
Sub-sample6	55	0.207	0.081	0.123	0.494	55	0.207	0.081	0.123	0.494

Figure A1: Bandwidth selection (modified from Everitt *et al*, 2011)

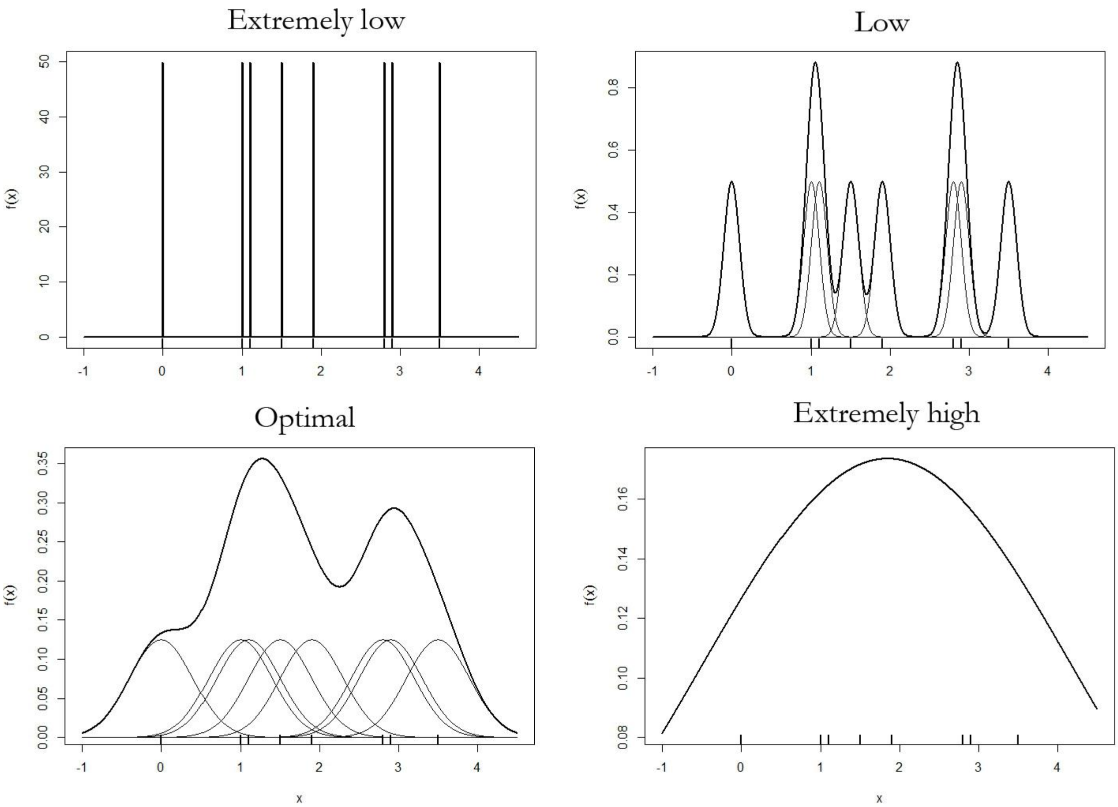
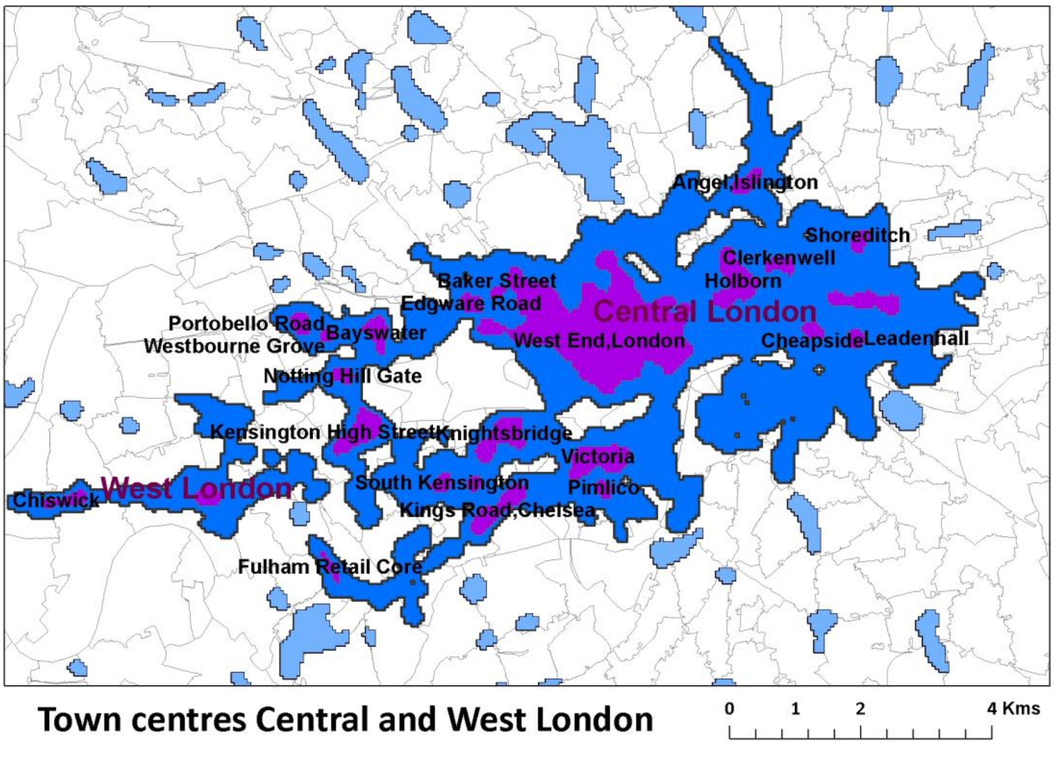


Figure A2: ACTA and Retail Cores London



Spatial Economics Research Centre (SERC)
London School of Economics
Houghton Street
London WC2A 2AE

Web: www.spatialeconomics.ac.uk