A national-scale application of the Huff gravity model for the estimation of town centre retail catchment area

Michail Pavlis[1], Les Dolega[1], Alex Singleton[1]

[1]Department of Geography and Planning, School of Environmental Sciences, University of Liverpool

November 7, 2014

Summary

This work presents the application of an unconstrained retail gravity model based on the Huff algorithm. The main objective of the analysis was to estimate the retail catchment areas for town centres in England as a function of relative attractiveness and distance from potential customers, while taking into account the effects of competition among the retail destinations. This was achieved by relaxing the areal constraint for spatial interactions, allowing thus the relative size and attractiveness of the town centres to define the scale at which competition might occur.

KEYWORDS: Retail gravity model, Huff model, National scale

1. Introduction

This study is concerned with the problem of estimating the extent and volume of potential customer patronage flows between retail centres, that is, forming an estimate of the catchment area of a retail centre. Perhaps the most commonly used techniques to tackle such a problem belong to a family of models known as gravity models. The fundamental hypothesis of gravity models is that the likelihood of a person from an origin location i patronising a retail centre in destination j is inversely proportional to distance and influenced by some measure of attractiveness of the retail centre (Wilson, 1974). In other words the closer a retail centre and the more attractive it is, the more likely it will be patronised by customers from a given location. Historically, the development of gravity models derived analogies from Newtonian physics.

The Huff gravity model is one of the most commonly used models for catchment estimation and can be defined as follows (Huff, 1964):

$$P_{ij} = A_j^a D_{ij}^{-b} / \Sigma(A_j^a D_{ij}^{-b}) \qquad (1)$$

Where P_{ij} the probability of a consumer from a location i patronising a retail centre j, A_j a measure of attractiveness of the retail centre j, D_{ij} a measure of distance between i and j, a (alpha) the exponent of attractiveness and b (beta) the exponent of distance. Thus, the numerator of the Huff algorithm is obtained by calculating the product between the store attractiveness raised to a power indicating the degree of attractiveness of that store, with the inverse distance raised to a power indicating how fast the attractiveness of that store decays, divided by the sum of all products between a given point of origin and every retail centre within the study area. Essentially, the denominator of the Huff model provides a way of standardising the numerator so that the sum of probabilities for each point of origin add up to 1 and, hence, allow consideration of the effect of competition between retail centres. Another important attribute of the Huff probabilities is that the probability of a location patronising a given destination is independent of the probability of another location patronising the same destination, and hence they could be calculated separately if required.

For the development of the distance decay parameter (D_{ij}) the time and money costs associated with travel is the most useful proxy, followed by the shortest road distance and the Euclidean (straight) distance (Wilson, 1974). Concerning the exponent of distance (b) it is more realistic to be disaggregated by some characteristic of the destination (e.g. retail centre size) (Drezner and Drezner, 2002) or by some characteristic of the customer (e.g. whether or not is a car owner) (Birkin et al., 2010). It has also been suggested that the distance decay parameter could be better approximated as an exponential function (rather than the power function presented in equation 1), with the advantage of the exponential function argued as providing more realistic results (Drezner and Drezner, 1996). The retail centre attractiveness score (A_j) is often derived from some measure of the relative size of a retail destination (Moryadas and Lowe, 1975), nevertheless additional data have also been used (when available) (De Beule et al., 2014) and could offer a more accurate indicator of attractiveness. Concerning the attractiveness exponent (a) it is not often used (set to a default value of 1), nevertheless, it could offer a way of weighting the attractiveness score either based on a characteristic of the retail centre or of the points of origin.

2. Developing the Huff model

A Huff gravity model was developed to estimate the retail catchment area of 1192 retail centres in England. The boundaries of the retail centres relate to those DCLG definitions defined in 2004. The centroids of 32.843 Lower Super Output Areas (LSOA) were used as the origin locations of customers patronising the retail centres. In addition, the UK national road network was collated from the Meridian 2 dataset provided by Ordnance Survey.

A composite indicator for the attractiveness of the retail centres was created using the following data:
1) Retail centre size (number of units),
2) Number of comparison retailers,
3) Number of leisure units,
4) Number of anchor stores,
5) Number of vacant outlets

Each variable was first standardised to the range between 1 and 100. Consequently the composite indicator was calculated by summing the values of the first four variables and subtracting from the result the vacant outlets units. Following this the retail centres were divided into four groups based on the attractiveness score and these groups formed the basis for the development of the beta coefficient. More specifically, town centres with score above 100 are all metropolitan retail centres (e.g. Manchester, Birmingham), while town centres with score between 50 and 100 are regionally important retail areas such as Bolton. Town centres with score between 20 and 50 serve as district centres such as Boston and Warwick. Finally, half of the town centres had a score bellow 20 and represent locations that have small and local retail catchments. The beta exponent was disaggregated into four different values, one for each group of town centres. These values were 1.2, 1.4, 1.6 and 2.0, in decreasing order of town centre size.

The shortest road distance from each LSOA centroid to the boundary of each retail centre was employed as measure of distance in this study. Given the variable shape and extent of retail centres, this was found to produce more realistic catchments than when simply using a centroid of a retail centre. This was calculated by extracting the coordinates of the points which defined the retail area boundaries and, consequently, applying the Dijkstra algorithm in order to calculate the shortest road distance. The analysis was all performed in R using the rgraph library. Following this, the minimum distance of each unique pair of LSOA – retail centres was obtained. This estimated distance also formed the basis for the development of an attractiveness exponent (a). More specifically, any retail centre within 0.5 kilometres from the centroid of a LSOA was assumed to be the primary retail destination and hence the attractiveness score for that pair was raised to the power of two. For all other distances, the default alpha value equal to 1 was used.

3. Results

The presentation of the results of the analysis focuses on the area and population of the primary and secondary catchments of the retail centres. As primary catchment was considered in this analysis the area that had a Huff probability equal to or larger than 0.5 of patronising a retail centre. A secondary catchment was considered to be the area that had a Huff probability equal to or greater than 0.2 (note that primary catchments are nested within secondary catchments). The results from the application of the Huff model are presented in Table 1 as the absolute and the percentage value of the area and population share for each of the four types of retail centres. Out of 1192 retail centres 1163 were assigned a primary and/or secondary catchment area while 29 were not (all were smaller retail centres). The effect of the exponents on the output of the model was investigated first by decreasing the beta values by 0.1 (Model 2 in Table 1) and then increasing them by 0.1 (Model 3 in Table 1). In addition, the effect of using only the default alpha coefficient (equal to 1) is also shown in Table 1 as Model 4. Finally, the output of a model with the same values as the Base Model when the distance is measured to the centroids of the retail centres is also shown in Table 1 as Model 5.

Table 1. Comparison of the output of 5 Huff models in terms of area and population assigned to the primary and secondary catchment, aggregated based on the size of the retail centres.

		Area		Population	
Total		130.537 Km^2		53.010.253	
Model	Rank	Primary (%)	Secondary (%)	Primary (%)	Secondary (%)
Base Model beta = 1.2, 1.4, 1.6, 2.0 alpha = 1, 2	1	1.233 (0.94)	11.357 (8.7)	4.229.864 (7.98)	11.571.711 (21.83)
	2	1.128 (0.86)	5.704 (4.37)	3.651.869 (6.88)	9.194.608 (17.34)
	3	1.308 (1.00)	4.259 (3.26)	3.752.775 (7.08)	8.347.951 (15.75)
	4	745 (0.57)	1.731 (1.32)	2.377.897 (4.48)	3.992.822 (7.53)
Model 2 beta = 1.1, 1.3, 1.5, 1.9	1	826 (0.63)	8.767 (6.72)	3.361.106 (6.34)	10.831.968 (20.43)
	2	724 (0.55)	4.087 (3.13)	2.820.000 (5.32)	7.833.338 (14.78)
	3	940 (0.72)	3.116 (2.39)	3.046.818 (5.75)	6.924.254 (13.06)
	4	608 (0.46)	1.362 (1.04)	2.143.857 (4.04)	3.422.134 (6.45)
Model 3 beta = 1.3, 1.5, 1.7, 2.1	1	1.793 (1.37)	14.025 (10.74)	5.118.533 (9.65)	12.050.600 (22.73)
	2	1.595 (1.22)	8.347 (6.39)	4.556.380 (8.59)	10.335.884 (19.50)
	3	1.792 (1.37)	5.905 (4.52)	4.683.244 (8.83)	9.611.359 (18.13)
	4	940 (0.72)	2.238 (1.71)	2.722.559 (5.13)	4.619.869 (8.71)
Model 4 alpha = 1	1	1.242 (0.95)	11.380 (8.72)	4.287.560 (8.09)	11.727.393 (22.12)
	2	1.125 (0.86)	5.710 (4.37)	3.610.125 (6.81)	9.243.030 (17.44)
	3	1.267 (0.97)	4.259 (3.26)	3.432.624 (6.47)	8.336.511 (15.73)
	4	642 (0.49)	1.695 (1.3)	1.768.890 (3.34)	3.724.097 (7.02)
Model 5 Distance to Centroids	1	767 (0.59)	10.207 (7.82)	2.841.208 (5.36)	11.297.676 (21.31)
	2	744 (0.57)	4.924 (3.77)	2.032.737 (3.83)	7.911.645 (14.92)
	3	754 (0.58)	3.285 (2.51)	1.814.316 (3.42)	6.941.614 (13.09)
	4	392 (0.30)	1.289 (0.99)	1.231.337 (2.32)	2.802.417 (5.29)

Based on the comparison of the five different models it is obvious that when decreasing the beta values of the Huff model, the area and population of the secondary and primary catchments, on average, are also decreasing across all groups of retail centres. This might be due to greater competition and cannibalisation of the Huff probabilities among the retail centres as the distance decay parameter gets smaller for all of the retail centres. On the other hand, Model 3 shows that when the beta values are increasing this results in greater market share, most likely due to reduced competition among the retail centres. As it was noted in Section 1, the alpha parameter is often overlooked, however, its use could be useful to model certain scenarios that could be quite difficult to model otherwise. For example, by increasing the alpha exponent for distances smaller than 0.5 km, it is possible to model the behaviour of customers preferring a retail centre simply due to proximity and convenience to reach. The comparison between the Base Model and Model 4 shows that it is possible to model such a scenario, given that it is mostly the primary catchment area of small retail centres which increases when the former model is applied. Finally, it can be seen from Table 1 that there is a striking difference between the Base Model that uses the distance to the boundary of the retail centres and Model 5 that uses the distance to the centroid of the retail centres. Across all retail centres the Base Model predicts a greater primary and secondary catchment area, which might indicate that using the centroids of the destinations could result in underestimating a catchment extent. In addition, it should be noted that the estimated catchment areas appear potentially more realistic for the Base Model as they follow the physical extent of the retail centres (Figure 1 A), compared to Model 5 which produces catchment areas that have more circular shapes (Figure 1 B).

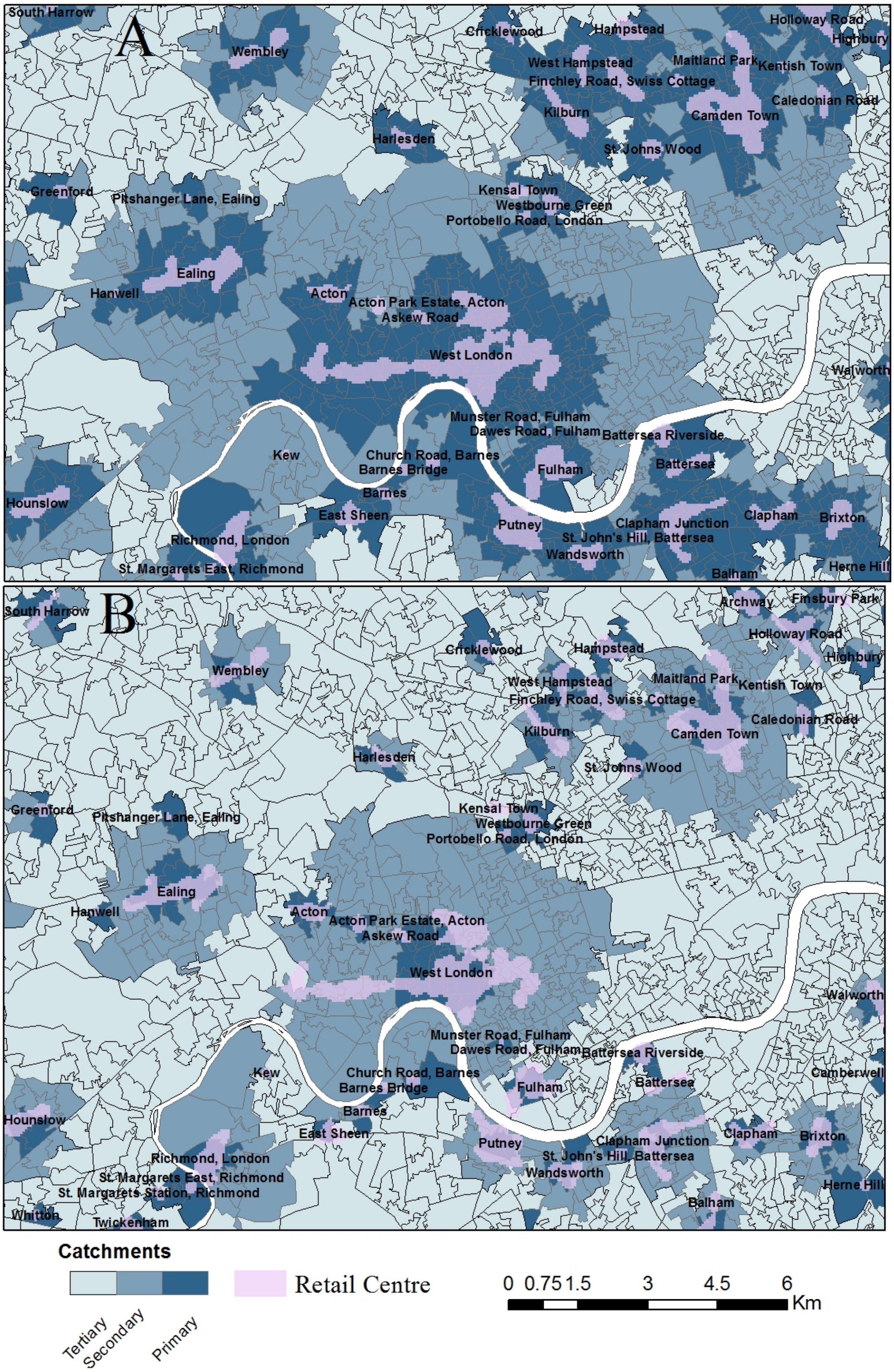

Figure 1. Comparison between the Base Model (distance to boundary) and Model 5 (distance to centroid).

Conclusion

The analysis presented here has shown that by removing the areal constraint of the Huff model it is possible to develop a national-scale model of catchment areas for the retail centres in England. The analysis also investigated the effect of the Huff model parameters on the estimated catchment area and suggested that a more accurate representation of the catchment areas might be obtained when the distance to the boundaries of the retail catchments is used instead of the distance to their centroids. The next step in the analysis would be to calibrate and test the model using data of customer behaviour.

References

Birkin M, Clarke G and Clarke M (2010). Refining and operationalizing Entropy-Maximizing model for business applications. Geographical Analysis, 42(4), 422-445.

De Beule M, Van den Poel D, Van de Weghe N (2014). An extended Huff-model for robustly benchmarking and predicting retail network performance. Applied Geography, 46, 80-89.

Drezner T and Drezner Z (1996). Competitive facilities: Market share and location with random utility. Journal of Regional Science, 36(1), 1-15.

Drezner T and Drezner Z (2002). Validating the gravity-based competitive location model using inferred attractiveness. Annals of Operations Research, 111, 227-237.

Huff DL (1964). Defining and estimating a trading area. Journal of Marketing, 28, p. 34-38.

Moryadas S and Lowe JC (1975). The geography of movement. Houghton Mifflin Company, p. 333.

Wilson AG (1974). Urban and regional models in geography and planning. John Wiley & Sons, p. 418.