Working papers SIET 2018 – ISSN 1973-3208

Households' expenditure elasticities for transport products and services: a geo-demographic approach

Angela S. Bergantino [1*], Mauro Capurso[2], Jeremy Toner[2]
[1] *Department of Economics, Management, and Business Law, University of Bari, Italy*
[2] *Institute for Transport Studies, University of Leeds, United Kingdom*

Abstract

In this paper we move forward the traditional approach to the estimation of a single, aggregate, expenditure elasticity for the commodity transport. In particular, we obtain separate elasticity measures for motoring expenditures and transport services, while also taking into account geo-demographic differences between households. To this end, we create a pseudo-panel for the United Kingdom (UK), using an exogenous clustering of households (Output Area Classification, OAC), thanks to which we are able to identify groups of geographically distant households, sharing the same characteristics, lifestyles, and needs. We then test whether the perception of transport services as inferior, necessity, or luxury goods significantly differ within clusters. Results show that, overall, total transport expenditure and motoring expenditure are weak luxuries, whereas bus and rail expenditures are strongly luxuries. Sample segmentation by OAC clusters shows some evidence of different income elasticities for different groups of consumers: transport in general, and motoring expenditure in particular, are necessities rather than luxuries for higher income groups.

Keywords: Transport, household expenditure elasticities, pseudo-panel data, geo-demographics, Output Area Classification.

1. Introduction

Individuals and households' spending behaviour for specific commodities has received notable attention in the literature since the seminal work of Engel (1857). The so-called Engel curves describe the relationship between households' expenditure on particular goods or services and their total expenditure or income. These relationships have attracted a considerable amount of attention given the important role they play in various models of income distribution (Bewley 1982; 1986). Either the income or the expenditure elasticity of demand are useful metrics for both governments and private firms, to help them decide on the goods to produce, and to understand how a change in the overall income in the economy would affect the demand for their products (i.e., whether it is inelastic or elastic, or even inferior for some consumers).

It is acknowledged that the classification of commodities as inferior, necessity, or luxury might change at different levels of total income or expenditure. Researchers, however, rarely investigated possible differences in expenditure behaviour within commodity groups. For example, if we consider the commodity "transport" – for which previous research estimated an aggregated elasticity close to (or slightly greater than) unity (Deaton and Muellbauer, 1980; Bewley, 1982; Giles and Hampton, 1985; Haque, 1988) –, it seems reasonable to expect different sensitivities with respect to the purchase

* Corresponding Author: Angela S. Bergantino (angelastefania.bergantino@uniba.it)

of a new car, which is an occasional and infrequent investment in durable equipment, and the expenditure on fuel or public transport tickets, which are purchased in a more regular way. This might be true also for airfares, given the so-called "democratisation" of air travelling.

At the same time, previous studies privileged the estimation of aggregated measures at a broad geographical level; however, people living in the same region/country not necessarily share the same expenditure behaviour, while geographically-distant individuals/households, with similar socio-demographic traits, lifestyles, and needs might show comparable expenditure patterns.

This paper aims at overcoming these limitations of previous research, proposing a novel approach to the estimation of disaggregated elasticity measures for transport products and services; in particular, we analyse the households expenditure behaviour in the United Kingdom (UK) taking into account intra-commodity as well as geo-demographic differences. To this extent, we first estimate an aggregate elasticity measure for transport expenditure, which is comparable with previous research. Then, we estimate separate measures for motoring expenditure and for public transport services, namely rail, bus, and taxis services. Moreover, we also introduce a geo-demographic dimension in the analysis, and we test whether the perception of transport products and services as inferior, necessity, or luxury goods significantly differ between clusters of households living in similar but geographically distant places (e.g. student neighbourhoods, city centres, countryside), and characterised by specific socio-demographic traits.

To this extent, a pseudo-panel was created pooling a series of independent cross-sectional datasets from the UK "Living Costs and Food" (LCF) surveys over the period 2008-2013, for which the UK Office for National Statistic also provided information on households' Output Area Classification (OAC, Vickers and Rees, 2006). This is a three-level geo-demographic classification (based on 2001 Census data), which goes beyond the standard urban-rural segmentation, grouping households into prototypical clusters. Statistical neighbourhoods are defined using socio-economic and residential information of the household reference person and of the household itself (on age, ethnicity, education, employment, type of housing), under the assumption that people with similar characteristics live in similar places, and vice versa. The huge potential of geo-demographic indicators in policy development has been recognised since the late '80s (Birkin and Clarke, 1989); these emerged as a very powerful tool for handling highly dimensional census data (Singleton and Spielman, 2013), although mainly used for marketing purposes.

Results partially confirm our expectations, i.e. that the perception of different transport products and services differ from a single measure, aggregated at the commodity level, and between clusters of household, exogenously identified using a geo-demographic classification. The elasticity measures obtained can be easily transposed into maps, and are supposed to be useful indicators for long-term transport planning investment decisions. For example, urban planners might take into consideration the geo-demographic perception for transport services to better ensure transport affordability, and to avoid transport-related social exclusion and poverty. Similarly, these outcomes might also be useful for private transport operators' (at either local or national level) in driving decisions on supply and pricing of transport services.

The remainder of the paper is as follows. The literature on household expenditure analysis is surveyed in Section 2. Section 3 describes the data and the pseudo-panel generating process, and Section 4 reports the empirical strategy and the proposed model. A discussion of the main features of the OAC can be found in Section 5. In Section 6 results are shown and commented upon, and Section 7 draws conclusions.

2. Literature review

The vast majority of authors who deal with the estimation of complete systems of demand equations use cross-sectional data, with only few authors employing other approaches (e. g. pseudo-panels, Berri et al., 1998), or concentrating their analysis specifically on a certain commodity (e.g. transport).

Deaton and Muellbauer (1980) propose, in their seminal work, an Almost Ideal Demand System (AIDS) of demand equations where commodities' budget shares depend linearly on the logarithm of real total expenditure and on the logarithm of relative prices. According to the authors, this model could be considered superior to other models previously proposed simply because it simultaneously possesses all properties typically desirable in demand analysis. They estimate total expenditure and price elasticities for eight groups of non-durable commodities using annual British data over the period 1954-1974, and they find the expenditure elasticity for transport and communication products to be greater than one (1.23).

Bewley (1982) derives a full system of Engel curves using data from the Australian Household Expenditure Survey 1975/6, with households grouped by average weekly income. Goods are divided into ten categories, and again transport and communication are grouped together. In this case, the income elasticity is estimated, and this is also found to be greater than one (1.22).

Giles and Hampton (1985) estimate a complete demand model with eight commodity groups using the 1981/2 New Zealand Household Survey. They employ the Full Information Maximum Likelihood (FIML) estimator, reporting transport as a luxury expenditure group. Using six different model specifications, the authors do not find large differences in the estimation of total expenditure elasticities. They report, however, that the AIDS and the Bewley's (1982) Addilog class of models perform slightly better, according to the Akaike Information Criterion (AIC). They also report similarities in the magnitude of the estimated total expenditure elasticities between the FIML and the OLS estimator, at least when computed at the sample mean.

Working with data from the Australian Household Expenditure Survey 1975/6, Haque (1988) estimates the total expenditure elasticity for food and nine non-food commodities. According to the author, none of the commonly used functional forms properly fitted the data; hence, a more flexible Box-Cox Engel function was needed to better specify the non-linearity of the relationships between the dependent and the independent variables. Using the Maximum-Likelihood estimator, he calculates a total expenditure elasticity of 1.30 for the commodity transport (again, this was grouped with communication). A few years later (1992), the same author obtains better results for this commodity group using the Double Semi-Log functional form. Even though he obtains a greater-than-one expenditure elasticity (i.e., 1.16), such value is not significantly greater than unity, meaning that transport and communication services could therefore be considered as necessities rather than luxuries.

Banks et al. (1997) propose a class of quadratic Engel curves (the so-called Quadratic Almost Ideal Demand System – QUAIDS) to nest the AIDS model by Deaton and Muellbauer (1980) and the Translog model by Jorgenson et al. (1982). According to them, some goods could be luxuries at certain income levels and necessities at others. They estimate a complete demand model for food, fuel, clothing, and alcohol on a pooled dataset using the UK Family Expenditure Survey from 1970 to 1986. Their results appear to be coherent and plausible descriptions of consumers' behaviour, thereby allowing the prediction of welfare effects of prices and tax changes.

Bergantino (1997) links together the most used models through a Box-Cox transformation, the Generalised Almost Ideal System, following an approach previously developed by Neves (1992). She concentrates exclusively on the commodity transport, employing the 1993 UK Family Expenditure

Survey; according to her, transport products and services should be considered necessities at higher levels of income and luxuries at lower ones. The author then ranks the nine functional specifications used, favouring the Banks et al.'s (1997) QUAIDS. Bergantino also calculates different elasticity values for public and private transport. She concludes that, while it is straightforward for the former category to be considered a necessity, for the latter, the estimated elasticity is not statistically greater than unity, thereby rejecting the hypothesis of considering private transport as a luxury category.

Berri et al. (1998) analyse car ownership and public and private transport expenditure elasticities using the pseudo-panel approach for Poland, France, Canada, and USA. They create a pseudo-panel dataset using different grouping criteria for each country (age and level of education of the householder, geographic location and income classes of the household). In terms of modelling, they employ the AIDS and the QUAIDS functional forms and the GLS estimator. The authors classify public transport as a normal good for Canada and as a luxury good for France, while they classify private transport as a luxury category in both countries. Furthermore, in all countries they find that private transport expenditure is less sensitive to income variations for rich households compared to poor ones. The opposite is found for public transport, where richer households show a higher elasticity.

Finally, Beneito (2003) chooses the most appropriate functional form starting with the estimation of a non-linear system and then deriving a linear one. Using the 1991 Spanish Household Expenditure Survey, he classifies transport as a luxury good, reporting an income elasticity of 1.6 and reinforcing the need to run separate regressions at different levels of income. The author concludes that an increase in the income of poor households favours the consumption of necessities, while the consumption of luxuries increases only after an increase in the income of richer households.

Table 1 summarises the main findings of the articles reviewed in this section. From this analysis, it seems that, at higher levels of aggregation, the estimated elasticities do not substantially differ from each other in absolute value. However, at further levels of disaggregation (e.g. public and private) transport services shift from being luxuries to necessities.

Table 1. A literature review on transport elasticity

Authors	*Year*	*Survey*	*Functional form*	*Elasticities*

Deaton and Muellbauer	1980	British HES 1954-74	AIDS	1.23
Bewley	1982	Australian HES 1975-76	Addilog	1.22
Giles and Hampton	1985	New Zealand HES 1981-82	AIDS - Addilog	1.23 (AIDS) 1.24 (Addilog)
Haque	1988	Australian HES 1975-76	Box-Cox	1.30
Haque	1992	Australian HES 1975-76	Double Semi-Log	1.16[a]
Bergantino	1997	British HES 1993	QUAIDS	1.00 (all) 1.02[a] (private) 0.99 (public)
Berri et al.	1998	Canadian HES 1969, 1978, 1982, 1986, 1992; French HES 1979, 1984, 1989, 1994; Polish HES 1987-90; US HES 1980-89	AIDS - QUAIDS	**Canada**: 1.17 (private) 1.16 (public); **France**: 1.19 (private) 1.36 (public); **Poland**: 2.20 (total transport); **US**: 1.16 (private) 1.84 (long-distance public) 0.91 (short-distance public)
Beneito	2003	Spanish HES 1991	Box-Cox	1.60[b]

a Not statistically greater than unity.

b This refers to the market income elasticity, which considers also the share that each commodity group represents over the total demand, as well as the elasticity of each commodity group with respect to total income.

Source: Author's elaboration.

A key development in our work is further disaggregation of transport expenditures looking at expenditure on: motoring; rail transport; bus transport; taxi. Existing modal evidence for the UK derived from direct demand models, suggests bus transport always has a negative income elasticity whatever the nature of the trip: that is, bus travel is an inferior good; rail transport has income elasticities varying between 0.1 and 2.0 depending on trip length, location and purpose (Balcombe et al., 2004). Thus rail travel can be either a necessity or a luxury. The evidence on car use income elasticities is somewhat sparse. Goodwin et al. (2004) summarise the recent evidence to be that short-run (one year) income elasticity of demand for car travel is 0.2 and the long-run elasticity 0.5, which numbers clearly imply car travel to be a necessity rather than a luxury. However, Graham and Glaister (2004) in a companion piece to Goodwin et al. (2004) suggest the long-run income elasticity of demand for car travel to be in the range 1.1 to 1.8, hence clearly a luxury good. We have found no evidence on income elasticity of demand for taxis. Thus, we present the first attempt to model expenditure elasticities at the level of transport mode using a demand system approach. This is clearly an advance on previous work and, given the evidence suggests very different income elasticities for bus and for rail, the previous disaggregation into "motoring" and "public transport" is clearly flawed. We develop the approach further by adopting a geo-demographic approach, thus permitting the possibility of income elasticity variation between different types of consumers. These developments are presented in the following sections.

3. Data and variables

Genuine panel datasets allow the researcher to track the behaviour of the same individual over a certain length of time; however very few panel datasets are available because it is too costly to collect them, and researchers interested in analysing individuals and households' expenditure behaviour typically employed cross-sectional data. However, a serious issue with cross-sectional data concerns the so-called zero-expenditure problem. This is the case when either the optimal level of consumption for a specific group is below the minimum expenditure required to purchase it (censoring problem), or when households do not actually purchase any items in that specific group in the period they are asked to keep the diary (purchase-infrequency problem; Beneito, 2003). Moreover, misreporting can also cause zero-expenditure to be recorded. With particular reference to the commodity transport (especially when data are disaggregated at mode level) zero values are more likely due to purchase infrequency or misreporting.

The zero-expenditure problem of cross-sectional surveys, and the unavailability of genuine panel datasets, can be solved by using the pseudo-panel approach, which allows to track cohorts of individuals. Here the units of analysis are subgroups of the population, clustered by roughly time-invariant characteristics (such as year of birth or sex), rather than single individuals/households as with genuine panel data. The pioneer of this approach is Deaton (1985), who states that although this methodology has been mainly formulated as a response to the absence of panel datasets, this does not necessarily mean that it gives inferior results. This technique has the advantage that indicators' means for the cohorts are error-ridden estimates of the unobservable population means, which allows the possibility of recognising measurement errors (hence controlling for them) even better than in genuine panel datasets, where these play a central role (Deaton, 1985).

Following the methodology described by Deaton (1985) and Verbeek (2008), we generate a pseudo-panel using micro-data from the Living Costs and Food Survey (LCF) - a national cross-sectional survey released by the Office for National Statistics - for the years 2008-2013. In the LCF Survey, every individual aged 16 and over in the surveyed household is asked to keep a diary for two weeks in which to record daily expenditures. Even children aged between 7 and 15 are asked to fill a simplified version of the diary. These surveys are conducted throughout the year in order to ensure the coverage of seasonality effects, and have been conducted each year since 1957, although under different names. The surveys are carried out every year with a new population sample, and geographic and demographic representativeness is constantly maintained. Two main datasets are available. One contains information at the individual level, and the other one contains information at the household level. In the latter case – which has been used for this work – only the "household reference person" answered the survey in place of all other members.

The definition of "cohort" is rather important in the creation of a pseudo-panel dataset. Some authors group individuals into cohorts on the basis of the year-of-birth only, while others also on the sex of the household reference person (Browning et al., 1985; Propper et al., 2001; Dargay, 2007; Meng et al., 2014). The way in which cohorts' characteristics are chosen is of strategic importance, as well as the number of households in each cohort, i.e., the size of the cohort. Households' characteristics that are used to define the cohorts need to be observed for all households in the sample at different points in time. Moreover, it is important that each household belongs to exactly one cohort.

For this paper, pseudo-panel cohorts are defined according to the year of birth of the household reference person, after segmenting the sample on the basis of households' membership into a specific OAC. This is a three-level geo-demographic classification by which households are clustered into statistical neighbourhoods using socio-economic and residential information of the household reference person, and of the household (age, ethnicity, education, employment, type of housing) obtained from the 2001 Census. The underlying assumption is that people with similar characteristics live in places with similar characteristics, even if those places are spatially apart. The highest level of aggregation has been used in this work, where UK households have been clustered into 7 "super groups" (SG hereafter; see Section 5 for more details).

The asymptotic behaviour of the pseudo-panel estimators depends on the number of households, N, and time periods, T (as in genuine panel data), but also on the number of cohorts, C, and the number of observations per cohort, nc. However, it is an empirical matter whether such behaviour provides a reasonable approximation of finite sample properties. According to Verbeek (2008), three possibilities are available:

1) $N \rightarrow \infty$, with C fixed, so that $nc \rightarrow \infty$;
2) $N \rightarrow \infty$ and $C \rightarrow \infty$, with nc fixed;
3) $T \rightarrow \infty$, with N and C fixed (nc is thus also fixed).

As Meng et al. (2014) suggest, there is a trade-off between the number of cohorts and the number of households in each cohort. When grouping households into cohorts some information related to the heterogeneity of households within each group is lost, and less efficient estimates are obtained (Dargay, 2007). However, this efficiency loss is minimised if the variation within cohorts is small, and the one between cohorts is, by contrast, great. This suggests a large number of small cohorts to preserve within-cohort homogeneity. But, if cohorts are identified using fewer households, however, the value of the indicators at the cohort level (e.g. expenditure, income) will be less representative of the actual population one; hence, there is also a conflicting need to seek to maximise the number of households to include in each cohort.

In this paper, we use a year-of-birth band smaller than usual (3 years instead of the more commonly-used 5 or 7 years). As an example, householders who were born in 1948, 1949, and 1950 were grouped together, and tracked for up to six years. Hence, we assume that their expenditure behaviour is consistent over a six-to-nine-years time span. We also assume that householders older than 75 behave similarly to each other. Moreover, a lower-bound constraint on the number of households in each cohort is enforced. According to Verbeek and Nijman (1993) the pseudo-panel observations need to be made up of at least 100 individuals, although this threshold can be reduced if individuals in each group are sufficiently homogeneous, as they are in our case. Given the large number of cohorts created (21 three-year birth cohorts times 7 SGs) we lower this threshold (15 households for SGs 2 and 7, which are the least populated ones, and 30 for the others). We create an un-balanced pseudo-panel (Table 2), given that cohorts not fulfilling the size requirement are removed from the dataset. On the one hand, we want to ensure robust estimates of each variable group means; on the other hand, we need a minimum number of pseudo-panel households to successfully estimate the models.

Table 2. Number of pseudo-panel IDs by LFC Survey

	Super group (SG)						
LFC Survey	**1**	**2***	**3**	**4**	**5**	**6**	**7***
2008	18	9	15	16	15	18	17
2009	20	14	14	16	12	18	6
2010	17	8	14	15	10	16	11
2011	17	9	12	17	15	18	18
2012	17	9	15	16	11	19	17
2013	17	6	13	17	10	17	16

Note: *The minimum number of households required (cohort size) for SGs 2 and 7 was lowered to 15.
Source: Authors' elaboration.

We choose 2013 as the base year, and we use the Retail Price Index (RPI) to obtain real expenditure and income data for each SG for each year. Since the focus is on expenditure and not quantities demanded (which, in any case, are not measured in commensurable units or, for some elements, not measured in meaningful units), prices can be assumed to be constant in real terms over the time series.

Another important issue relates to the variables to be included in the model. Table 3 shows the list of all variables used in this work, together with their definitions. The value of transport expenditure and, in turn, of its sub-components, is the dependent variable1. Total household expenditure (instrumented by the anonymised weekly income plus allowances), together with a series of dummies accounting for the characteristics of the households or of the household reference person (car availability, child presence, sex, and residential location – London vs the rest of the UK), are the independent variables.

[1] This is a derived variable containing motoring expenditure (which in turn includes petrol, repair, servicing and other motoring costs, but excludes the cost of the purchase of vehicles) as well as public transport services expenditure (which includes rail, bus, and air fares, and taxis expenditure).

Table 3. List of variables

	Dependent Variables
TransportExp	Cohort weekly total transport expenditure
In the other models Motoring Expenditure, Rail, Bus, and Taxis were used as dependent variables.	

	Explanatory Variables
TotExp OR Income	Cohort weekly total expenditure OR anonymised weekly income plus allowances
ln(TotExp) OR ln(Income)	Logarithm of *TotExp* OR logarithm of *Income*[2]
Age	Average age of the household reference persons grouped in the cohort
Age2	Square of *Age*
Sex	Proportion of males among the household's reference persons in the cohort
Children	Average number of children (under 18) in the households grouped in the cohort
CarAvailability	Average number of cars in the households grouped in the cohort
London	Average number of household based in the London Region in the cohort

The vast majority of empirical studies employ total household expenditure as the main independent variable, while only a few authors use total income as the instrument. Both variables, total expenditure and income, have their drawbacks. On the one hand, income might be affected by measurement errors, given that individuals may be reluctant to reveal how much they earn. Poorer individuals might feel uncomfortable when revealing such information to the surveyor and overstate this value. At the opposite, richer individuals might understate it for the same reasons, or for the fear that this information is passed to the fiscal authority. On the other hand, the use of income rather than total expenditure is justified by the endogeneity of the latter. In this regard, both Summers (1959) and Liviatan (1961) propose estimating Engel curves by adopting the instrumental variables technique. Although the use of pseudo-panel cohorts has been proven to potentially mitigate the endogeneity problem (Gardes et al., 2005), testing reveals that this was not the case with the data used for this analysis. For this reason the instrumental variables approach has been preferred to the use of fixed-effect models with total expenditure as the main explanatory variable, as explained further in the next section.

[2] Moreover, when employing a model specification in which the total outlay variable is in the logarithm form, such variable should be calculated as the mean of the logarithms rather than the logarithm of the means (Deaton, 1985, p.113).

4. The empirical strategy

When pseudo-panel data are used in the estimation of demand models, Deaton (1985) proposes using fixed-effect models, so as to account for time-invariant unobserved heterogeneity at the cross-sectional level (the individual effect) that is correlated with the set of explanatory variables. The linear fixed-effect model when genuine panel data are available can be expressed as (equations 1-3):

$$y_{ht} = x'_{ht}\beta + \alpha_h + u_{ht}, \ \mathrm{t} = 1, \dots, \mathrm{T} \tag{1}$$

$$\text{with } \mathrm{E}\{x_{ht}\, u_{ht}\} = 0 \text{ for each t,} \tag{2}$$

$$\text{and } \mathrm{E}\{\alpha_h x_{ht}\} \neq 0, \tag{3}$$

where x'_{ht} stands for a matrix of regressors and β is the relative vector of parameters of interest. If condition (3) is not met, model (1) "can be consistently estimated by pooling all observations and treating $\alpha_h + u_{ht}$ as a composite error term" (Verbeek, 2008).
When only repeated cross-sections are available, Deaton (1985) proposes aggregating all observations into cohorts C in order to obtain consistent estimators for β, and to estimate the following model (equation 4):

$$\bar{y}_{ct} = \bar{x}'_{ct}\beta + \bar{\alpha}_{ct} + \bar{u}_{ct}, \ \mathrm{c} = 1, \dots, \mathrm{C}; \mathrm{t} = 1, \dots, \mathrm{T} \tag{4}$$

where $\bar{y}_{ct}$ is the mean value of the observed y_{ht} in each cohort c in period t (and the same goes for all other variables – Verbeek, 2008, p.371). $\bar{\alpha}_{ct}$ depends on t (because individual households in each cohort are not the same every year) and is likely to be correlated with $\bar{x}_{ct}$ if α_h is correlated with x_{ht} in model (1). Treating $\bar{\alpha}_{ct}$ as part of the error term, will lead to inconsistent estimators (Verbeek, 2008).

However, because the cohort observations reflect the average values for all considered variables of all households included in the cohort, it is possible to consider these as error-ridden measurements of the unobserved population cohort's means, suggesting the use of an error-in-variables estimator (Deaton 1985; Dargay, 2007; Verbeek, 2008).

But if the cohort size is large as well as the time variation in the cohort's means, the measurement error can be ignored and the within-transformation estimator can be employed, since the bias is very small (Verbeek and Nijman, 1992). When this is the case, it is possible to consider $\bar{\alpha}_{ct}$ as a fixed unknown parameter and ignore any variation over time, so that $\bar{\alpha}_{ct} = \alpha_c$ (which simply means estimating a single intercept for each cohort), and estimating the following model (equation 5):

$$\bar{y}_{ct} = \bar{x}'_{ct}\beta + \alpha_c + \bar{u}_{ct}, \ \mathrm{c} = 1, \dots, \mathrm{C}; \mathrm{t} = 1, \dots, \mathrm{T} \tag{5}$$

In this work, type (1) asymptotic behaviour is preferred. The number of cohorts, C, is kept fixed, letting the number of households in each cohort, N, to tend to infinity. According to Verbeek (2008), under such circumstances, the fixed-effect (within-transformation) estimator, $\hat{\beta}_w$, will be consistent for β, provided that:

$$\lim_{n_c \to \infty} \frac{1}{CT} \sum_{c=1}^{C} \sum_{t=1}^{T} (\bar{x}_{ct} - \bar{x}_c)(\bar{x}_{ct} - \bar{x}_c)' \text{ is finite and invertible,}$$

and, that:

$$\lim_{n_c \to \infty} \frac{1}{CT} \sum_{c=1}^{C} \sum_{t=1}^{T} (\bar{x}_{ct} - \bar{x}_c)\bar{\alpha}_{ct} = 0.$$

Whether the first condition (which states that the cohort averages show some time variations even if they are composed of a large number of households – Verbeek, 2008) is satisfied or not depends on the way cohorts are identified. According to Verbeek and Nijman (1992), the bias in the standard within-transformation estimator $\hat{\beta}_w$ may be substantial even if the cohort size is large.

Another relevant issue related to the estimation of demand equations concerns the specification of the most appropriate functional form. Some authors propose to select a different functional form for each commodity group, which is clearly inconsistent with the fulfilment of the adding-up restriction in Engel curves analysis (Bewley, 1982). However, there is no need to simultaneously estimate a complete set of demand equations if every equation contains the same set of independent variables. All major estimators (OLS, GLS, IV) can be employed to estimate each equation separately, and the obtained estimates will be as efficient as those deriving from the simultaneous estimation.

With respect to transport, the double semi-log (which derives from the AIDS - Haque, 1992) is favoured in the context of cross-sectional analysis, and the linear functional form (Dargay, 2007) is favoured when using the pseudo-panel data. This motivates the use of these functional forms in this study.

The following models have been estimated (linear, equation 6 - double semi-log, equation 7):

$$\bar{y}_{(TransportExp)ct} = \widehat{\bar{x}'_{(TotExp)ct}}\beta_1 + \bar{z}'_{ct}\beta_i + \alpha_c + \bar{u}_{ct},\ \mathrm{c} = 1, \dots, \mathrm{C};\ \mathrm{t} = 1, \dots, \mathrm{T} \qquad (6)$$

$$\bar{y}_{(TransportExp)ct} = \widehat{\bar{x}'_{(TotExp)ct}}\beta_1 + \widehat{ln[\bar{x}']_{(TotExp)ct}}\,\beta_2 + \bar{z}'_{ct}\beta_i + \alpha_c + \bar{u}_{ct},$$
$$\mathrm{c} = 1, \dots, \mathrm{C};\ \mathrm{t} = 1, \dots, \mathrm{T} \qquad (7)$$

where $\widehat{\bar{x}'_{(TotExp)ct}}$ and $\widehat{ln[\bar{x}']_{(TotExp)ct}}$ are the predicted total household expenditure and the logarithm of the total household expenditure (estimated in the first stage by the average cohorts' anonymised weekly income plus allowances and by its logarithm). Standard errors are obtained using the clustered sandwich estimator to control for the dependency within cohorts. We then use the Kleibergen-Paap (2006) rk-statistic to test for the weakness of the instruments. This is a generalisation of the Cragg-Donald Wald statistic to the case of non-independently and identically distributed errors. We test endogeneity as the difference of two Sargan-Hansen statistics, where the null hypothesis is that the specified endogenous variable (e.g. total household expenditure and its logarithm) can be treated as exogenous.

Table 4 presents the formulas used to calculate the expenditure elasticities at mean values for the linear and for the double-semi-log functional forms.

Table 4. Elasticity formulas

Functional form	Elasticity formula
Linear	$\eta_j = \frac{\beta_1}{\overline{w^J}}$
Double-Semi-Log	$\eta_j = \frac{\beta_1}{\overline{w^J}} + \frac{\beta_2}{\overline{y^J}}$

According to the elasticity formula for the linear functional form, the elasticity only depends on the estimated coefficient β1 associated with the total expenditure and on $\overline{w^J}$, the share of the expenditure for the commodity of interest over total expenditure $\overline{y^J}$,, meaning that it is constant given $\overline{w^J}$. In the double semi-log functional form, instead, it also depends on the estimated coefficient β2.

5. The Output Area Classification

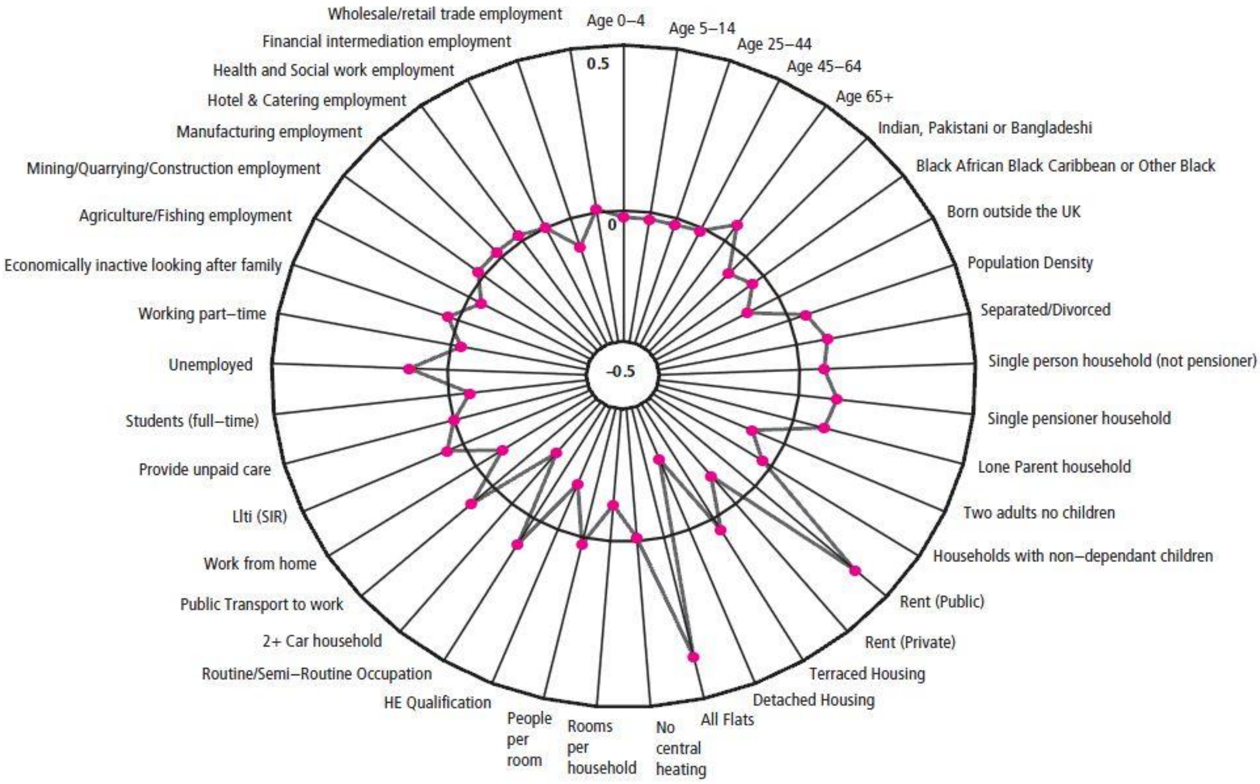
Geo-demographics assume that individuals who live close to each other are also behaviourally close. Hence, it seems reasonable to also assume that there might be individuals (and, in this case, households) who are not spatially close, but have similar socio-economic traits and lifestyles.

Using a carefully selected set of variables from five domains (demographic structure, household composition, housing, socio-economic and employment status) taken from the 2001 Census, Vickers and Rees (2006) group the UK population into clusters, minimising the within-cluster variability and maximising variation between them. In particular, they employ a K-means algorithm (which minimises the Euclidean sums of squared deviations from the cluster mean) to identify the groups, proposing a three-tier hierarchy classification made of 7 “super groups”, 21 “groups”, and 52 “sub groups”. Among the variables used, some have relevance with respect to transport: the number of households with two or more cars and the number of individuals going to work using public transport means.

The authors also provide names and profiles for the clusters with radial plots (Figure 1). Three circles are embedded within each other. The midway circle shows the average value across the UK for each variable, while the inner and outer circles represent 50% (less and more) variation with respect to the average. The aim of the names and profiles is “to create a short description, using text and visuals, which only takes a few seconds to read but significantly expands the user’s understanding of the group” (Vickers and Rees, 2006). However, they also stress the importance, besides the clusters’ names, of the statistical traits of the clusters with respect to each variable. A short description by Williams and Botterill (2006) of the main features of each super group is presented in Table 5.3.

[3] Households in the LFC surveys from 2008 to 2013 have been clustered by ONS using to the 2001 Census Data, while a new classification, which is based on the 2011 Census Data, has been used to cluster households in the LFC surveys from 2014 onwards. However, the two classifications are not directly comparable, and this is the reason why the analysis dates up to 2013, even though more recent LFC surveys were available (up to 2015 at the time of writing). We preferred the former classification given that this enabled us to create a pseudo-panel using six consecutive surveys rather than just two.

Figure 1. Radial plot for super group 5



Source: Vickers and Rees, 2006

Table 5. The main features of each super group

Super group number and name	*Where those people mainly live*	*% of the population clustered in each group**	*Variables far above the UK average*	*Variables far below the UK average*
1 - Blue collar communities	High concentration in the North East, South Wales, and cities around Scotland and Midlands	16.7	Terraced housing Public renting	Higher education qualifications Flats
2 - City living	High concentration within city areas, especially London	6.1	Single person households (not pensioner) Private Rents Flats Higher education qualifications People born outside the UK	Detached housing Households with non-dependent children age 5 to 14
3 - Countryside	All across the UK, especially in more rural areas	12.5	Detached housing Homeworkers People working in agriculture Two or more car households	Public transport to work Population density Flats
4 - Prospering suburbs	The most common area type in the UK	23.1	Detached housing Two or more car households	Public renting Private renting Terraced housing Flats No central heating
5 - Constrained by circumstances	Around cities, especially Scotland	10.9	Public renting Flats	Two or more car households Higher education qualifications Detached housing
6 - Typical traits	All across the UK	19.3	Terraced housing	Public renting
7 - Multicultural	High concentrations around major cities such as London and Birmingham	11.5	Minority ethnic population People born abroad Flats Public renting Private renting Use of public transport to work	Detached housing

Note: *Based on 2001 UK Census.
Source: Williams and Botterill (2006).

In the remaining of the paper we will consider the SG 6 to be "average". SGs 1, 5 and 7 are lower income groups than average, whereas SGs 2, 3, and 4 are higher income groups. Low income groups are expected to spend less (on average) on motoring but also SG2, as high density city living reduces the need for car travel relative to suburban or periurban or rural areas. These latter areas (SG3, SG4) are expected to be lesser users of bus because they are using car instead and also typically face less frequent bus services than the cores of large cities.

6. Results

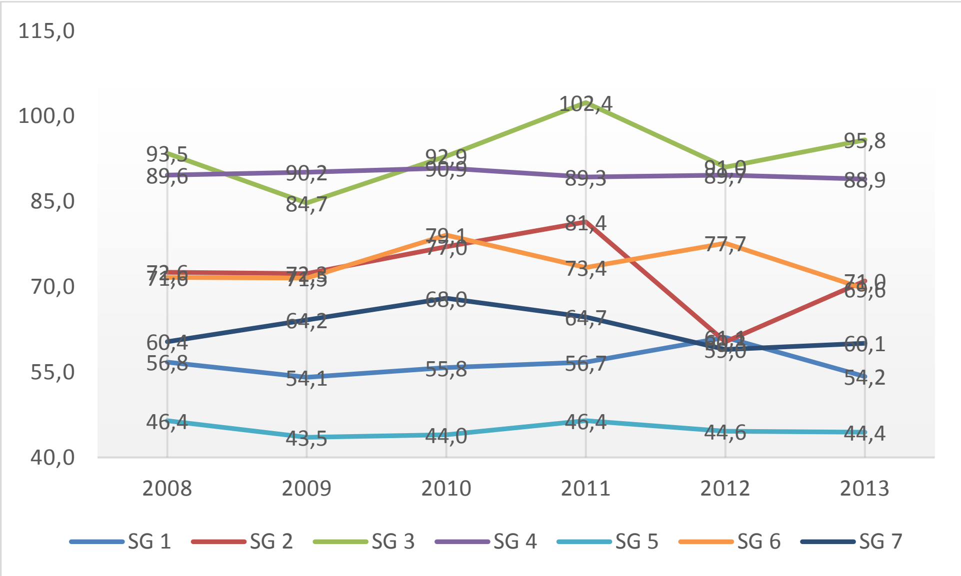
The elasticities for the commodity transport and its main sub-components are separately presented and compared in the following sub-sections. The elasticities calculated over the full sample are presented first, and then compared with those obtained using sub-samples according to the OACs. A series of t-tests then checks whether the elasticities are statistically different from zero, one, and from the full-sample values, respectively (at 90% significance level).

6.1.The commodity transport

In the UK, "transport" represents the highest-expenditure category.4 Note that in the UK official statistics, the purchase of housing is not included in expenditure surveys although rental of housing is included. If house purchase were included, housing would be the biggest expenditure category. That said, transport is clearly a substantial household expenditure, greater than food and drink, and recreation, and communications.

Figure 2 shows the pattern of the average weekly households' expenditure for each super group over the period 2008-2013. SGs 1, 4, 5, and 6 do not significantly modify their consumption habits over the considered period, while the remaining SGs show a greater variation.

Figure 2. Average weekly transport expenditure (in GBP) by super group and year



Source: Authors' elaboration based on pseudo-panel dataset.

The share of total expenditure that households clustered in each SG devoted to transport expenditure fluctuates from 9.5 to 14.5% in the considered period. Interestingly, while for SGs 2, 3, and 5, a drop in this share is observed for 2012, there is an increase for SGs 1 and 6 in the same year (Table 6).

[4] With an average of £74.80 per week in 2014, which represents 14% of total expenditure.

Table 6. Share of total expenditure devoted to transport expenditure by year and SG

Year	Super group 1	2	3	4	5	6	7
2008	12.3%	10.2%	10.2%	12.1%	12.0%	12.5%	12.3%
2009	12.6%	10.5%	10.5%	13.1%	11.2%	12.8%	11.4%
2010	12.7%	11.4%	11.4%	13.2%	11.6%	13.7%	12.4%
2011	13.5%	11.2%	11.2%	13.6%	12.8%	13.4%	13.3%
2012	14.5%	9.5%	9.5%	13.4%	11.6%	14.5%	12.1%
2013	12.9%	10.4%	10.4%	13.3%	12.1%	12.5%	12.7%

Source: Author's elaboration based on the pseudo-panel dataset.

Table 7 reports the results of the estimation of the elasticities for the full sample and for each SG separately.

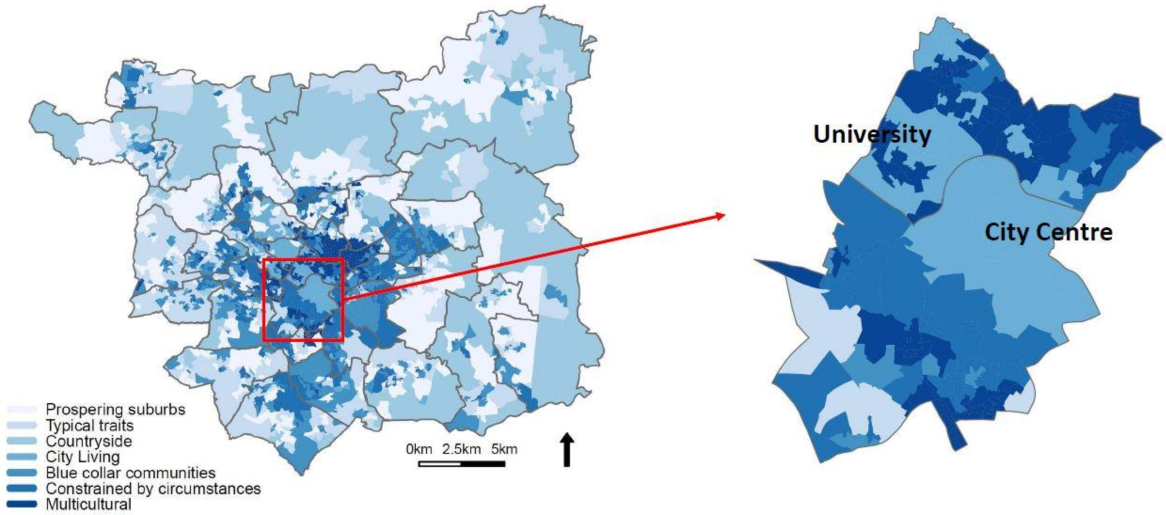
Table 7. Expenditure (income) elasticity for the commodity transport

	Total Exependiture (£)	% (total)		Elasticity	p-value (0)	p-value (1)	p-value (fs)
Full sample	544.37	70.74	Linear	1.267	0.000	0.378	na
			Dsl	1.333	0.000	0.300	na
SG 1	421.9	56.4	Linear	1.155	0.002	0.678	0.757
			Dsl	2.697	0.124	0.333	0.436
SG 2	689.0	72.4	Linear	0.654	0.174	0.472	0.200
			Dsl	9.897	0.898	0.908	0.912
SG 3	654.1	93.1	Linear	0.842	0.068	0.733	0.355
			Dsl	1.033	0.019	0.941	0.501
SG 4	679.5	89.7	Linear	0.449	0.311	0.213	0.064
			Dsl	0.711	0.287	0.665	0.354
SG 5	367.0	45.1	Linear	2.136	0.018	0.209	0.338
			Dsl	1.138	0.001	0.697	0.590
SG 6	547.7	73.8	Linear	1.208	0.000	0.529	0.850
			Dsl	1.172	0.000	0.530	0.566
SG 7	490.37	62.2	Linear	2.387	0.354	0.590	0.664
			Dsl	1.399	0.000	0.316	0.862

Results from Table 7 reveal that, over the full sample, the estimated elasticity for the commodity transport is 1.267 (when the linear functional form is used, or 1.333 with the double semi-log – DSL). This value is statistically different from zero, but not from one, confirming what was previously found in the literature. While earlier studies – among those reviewed in Section 2 – report an elasticity measure for transport services significantly larger than unity, in the latest works, the estimated elasticity tends to be greater but not significantly different from unity. For overall transport expenditure over the full sample, there is no substantive difference between the linear and the DSL elasticity. Next, we estimate the within-group elasticities. The elasticities for SGs 2, 3, and 4 slightly differ from the full-sample one in terms of absolute values, being less than the unity. However, this difference is statistically relevant for SG 4 only, which groups together households owning two or more cars in a proportion larger than the UK average. Moreover, only the value for SG 3 is directly interpretable, being statistically different from zero but not from one. The confidence interval for SG 2 is very large, making the estimate not precise.5 Considering the linear elasticities for these SGs to be more robust than the DSL elasticities, we conclude that higher income SGs have transport as a necessity rather than a luxury. At the opposite, we obtain elasticity values greater than two for SGs 5 and 7, even though these values are not statistically different from one and from the value obtained over the full sample. Note that in these cases, the DSL elasticities are more precisely estimated than the linear elasticities. Recalling that these are the poorest consumers, we find transport to be a luxury good for the lowest income groups. Interestingly, the estimated elasticity for households clustered in SG 6 is 1.20. This SG groups together typical households in the UK, and it confirms the value obtained over the full sample. To sum up, estimation results reveal that the commodity transport should be considered a luxury good for average income and below average income UK households. Nevertheless, a more in-depth investigation through segmentation of the sample by means of the households' membership into clusters allows the identification of slightly different sensitivities towards transport expenditure. The introduction of a geo-demographic dimension in the analysis allows transposing the different elasticity measures into maps. As an example, we choose a city in the UK, and Figure 3 represents the map of the Leeds County with a snapshot of the University and City Centre areas. Different degrees of expenditure elasticity for transport products and services are highlighted using shades of blue. It is evident how households living very close to each other might show completely different elasticities.

[5] In the remainder of the paper we will refer to elasticity estimates which are not "precise" when the p-value is greater than 0.1.

Figure 3. Elasticity map for Leeds county and selected city areas



Source: Authors' elaboration based on results presented in Table 7.

This information might be very valuable for any planner, policy maker or transport operator, given that it gives a better idea of how preferences are spatially organised within the population. For example, it might be useful to ex-ante infer the outcome of policies aimed at reducing the car usage for residents in specific areas, or of fare schemes designed to encourage the use of public transport.

6.2.Motoring expenditures

Motoring expenditures only account for operational costs of private vehicles, excluding the costs relative to the purchase of vehicles. Columns 1-2, Table 8, suggest that motoring expenditures contribute to an average of 8.8% of households' total expenditure. This percentage corresponds to slightly less than £50 (£48.3) every week, which are devoted to petrol and cars' repairs/maintenance.

Table 8. Expenditure (income) elasticity motoring expenditures

	Motoring Expenditure (£)	% (total)	Average car/household		Elasticity	p-value (0)	p-value (1)	p-value (fs)
Full sample	48.3	8.8%	0.99	Linear	1.246	0.000	0.480	na
				Dsl	1.240	0.000	0.443	na
SG 1	40.1	9.3%	0.85	Linear	1.384	0.004	0.419	0.778
				Dsl	2.932	0.107	0.288	0.353
SG 2	40.8	10.5%	0.78	Linear	1.293	0.293	0.812	0.972
				Dsl	-20.32	0.895	0.890	0.889
SG 3	69.9	10.7%	1.38	Linear	0.836	0.089	0.738	0.399
				Dsl	1.078	0.020	0.867	0.727
SG 4	64.5	9.5%	1.36	Linear	0.466	0.289	0.224	0.074
				Dsl	0.519	0.448	0.481	0.291
SG 5	30.3	8.1%	0.67	Linear	2.195	0.030	0.236	0.349
				Dsl	1.132	0.002	0.711	0.763
SG 6	51.2	9.2%	1.08	Linear	0.714	0.002	0.219	0.021
				Dsl	0.819	0.002	0.494	0.111
SG 7	36.0	7.3%	0.67	Linear	2.252	0.399	0.639	0.708
				Dsl	0.824	0.059	0.685	0.339

The linear elasticity for motoring expenditures (calculated) over the full sample is 1.246 (1.240 with DSL). This value is statistically different from zero, but not from one. Hence, this sub-commodity can be classified as a luxury, consistently with what previous research has found. Among the works reviewed in Section 2, Berri et al. (1998) classify motoring expenditure as a luxury good for Poland, France, US, and Canada. In particular, they find richer households to be also less sensitive to income variations than poorer families, which is exactly what we observe in this work. Households that spend the most (those in SGs 3, 4, and 6) show an expenditure (income) elasticity lower than, but not statistically different from, one (SGs 3 and 6), or lower than one (SG 4), although this estimate is not precise. Additionally, the estimated value seems to be statistically different from the value obtained over the full sample (at least for SGs 4 and 6). In this regard, it is worth noting that motoring expenditure is strictly related to the average number of cars per household. This value is much greater than the one for SGs 3 and 4 (where households with two or more cars are concentrated), and slightly more than the one for SG 6. At the same time, the average number of cars per household is lower for those SGs where, according to the classification, there is a lower proportion of households with two or more cars (SG 5), and a higher propensity to use public transport to go to work (SG 7). Interestingly, these households also show the largest elasticity for motoring expenditure, even though these estimates are either not statistically different from one (SG 5), or not very precise.

6.3.Rail transport expenditure

Households' expenditure for rail transport services accounts for an average of £3.3 per week (0.6% of total expenditure). Due to data restrictions, it was not possible to further distinguish between underground and long-transport rail services. Hence, rail transport expenditure here contains both (Table 9).

Table 9. Expenditure (income) elasticity for rail services

	Rail Transport Expenditure (£)	% (total)		Elasticity	p-value (0)	p-value (1)	p-value (fs)
Full sample	3.3	0.6%	Linear	3.253	0.000	0.000	na
			Dsl	5.402	0.001	0.005	na
SG 1	1.6	0.3%	Linear	0.377	0.570	0.347	0.000
			Dsl	-1.831	0.850	0.770	0.456
SG 2	8.8	1.3%	Linear	1.496	0.516	0.829	0.447
			Dsl	-21.44	0.907	0.903	0.884
SG 3	2.4	0.4%	Linear	4.978	0.018	0.060	0.413
			Dsl	3.572	0.026	0.108	0.253
SG 4	4.1	0.5%	Linear	4.923	0.119	0.214	0.596
			Dsl	7.623	0.051	0.090	0.569
SG 5	1.4	0.3%	Linear	7.533	0.007	0.018	0.122
			Dsl	4.938	0.013	0.047	0.816
SG 6	3.4	0.6%	Linear	6.981	0.000	0.000	0.013
			Dsl	5.598	0.000	0.001	0.884
SG 7	4.1	0.8%	Linear	5.084	0.126	0.219	0.581
			Dsl	4.208	0.021	0.078	0.513

The full sample linear elasticity for rail transport services is 3.253 (5.402 with DSL), a value which is significantly and statistically greater than one. This provides evidence that rail transport services can be classified as luxuries. At the SG level, this result seems to be confirmed also. Only in two cases statistically different elasticities have been obtained. SG 1 has an elasticity of 0.377, which implies considering rail services as necessities for these households, even though this estimate is not very precise. SG 6 shows, instead, an elasticity of 6.981, a value almost twice the one obtained over the full sample, which is also statistically different from one.

6.4.Bus transport expenditure

British households spend an average of 1.52 pounds per week for bus transport services (0.30% of their total expenditure). Also in this case, due to data restrictions, it was not possible to distinguish between local and long-distance bus services.

Table 10. Expenditure (income) elasticity for bus services

	Bus Transport Expenditure (£)	% (total)		Elasticity	p-value (0)	p-value (1)	p-value (fs)
Full sample	1.52	0.30%	Linear	3.253	0.000	0.001	na
			Dsl	0.185	0.855	0.419	na
SG 1	1.76	0.39%	Linear	1.332	0.663	0.914	0.754
			Dsl	1.753	0.661	0.851	0.695
SG 2	2.17	0.33%	Linear	3.391	0.058	0.181	0.538
			Dsl	17.63	0.899	0.905	0.900
SG 3	0.76	0.11%	Linear	6.198	0.005	0.019	0.078
			Dsl	6.202	0.001	0.005	0.001
SG 4	0.99	0.13%	Linear	2.898	0.596	0.728	0.911
			Dsl	2.714	0.480	0.655	0.510
SG 5	1.92	0.49%	Linear	7.755	0.034	0.065	0.136
			Dsl	4.256	0.064	0.157	0.076
SG 6	1.26	0.22%	Linear	-2.435	0.065	0.009	0.000
			Dsl	-2.285	0.054	0.006	0.037
SG 7	2.42	0.48%	Linear	3.043	0.437	0.602	0.847
			Dsl	1.091	0.265	0.926	0.353

Table 10 reports the elasticity values for bus services. The elasticity over the full sample is 3.253. This value is statistically different from both zero and one, and very similar to the value obtained for rail services. Hence, bus services also share the characteristics of luxuries.

When looking at the elasticities obtained at the SG level, two estimates are statistically different from the one obtained over the full sample. Households clustered in SG 3, which are more concentrated in rural areas according to the description provided in Section 5, show an elasticity twice as large as the full sample one (6.198). Interestingly, these households use very little public transportation to get to work, while also owning more cars than the UK average. At the opposite, households clustered in SG 6 (the "typical traits" SG) have a negative elasticity, hence providing evidence that, for the "average UK household", bus services should be considered inferior services. This value is both statistically different from zero, one, and the full sample value. Unfortunately, it is not possible to infer meaningful considerations for the other SGs given that the confidence intervals for the estimates are sufficiently large to not reject any hypothesis regarding the exact value of the elasticities.

6.5.Expenditure for taxis

Table 11 reports the estimated elasticities for taxi services, which account for an average of 0.25% of total weekly households' expenditure (£1.27).

Table 11. Expenditure (income) elasticity for taxi services

	Taxis Expenditure (£)	% (total)		Elasticity	p-value (0)	p-value (1)	p-value (fs)
Full sample	1.27	0.25%	Linear	2.646	0.624	0.760	na
			Dsl	-1.854	0.528	0.332	na
SG 1	1.25	0.30%	Linear	3.862	0.212	0.355	0.199
			Dsl	5.987	0.333	0.420	0.420
SG 2	2.66	0.40%	Linear	3.869	0.062	0.166	0.054
			Dsl	-5.798	0.924	0.910	0.910
SG 3	0.93	0.14%	Linear	2.656	0.208	0.432	0.190
			Dsl	2.931	0.190	0.388	0.388
SG 4	1.08	0.16%	Linear	-0.660	0.242	0.003	0.329
			Dsl	3.360	0.584	0.701	0.701
SG 5	1.09	0.31%	Linear	1.867	0.439	0.720	0.413
			Dsl	1.038	0.428	0.977	0.977
SG 6	1.14	0.21%	Linear	4.134	0.007	0.039	0.005
			Dsl	3.514	0.054	0.167	0.167
SG 7	1.27	0.28%	Linear	7.007	0.516	0.577	0.509
			Dsl	0.841	0.587	0.918	0.918

Large confidence intervals for the estimated elasticities (over the full sample) do not allow to reject any hypothesis regarding the classification of taxi services as necessity or luxury services. At the SG level, instead, it is possible to provide meaningful interpretations in a few cases. For example, for SG 4 (which is the most common area type in the UK) an elasticity of -0.660 is estimated and it is statistically different from one but not from zero. On the contrary, for SG 6 (which groups together households with the average UK traits) the estimated elasticity is 4.134, a value that is statistically different from both zero and one. A similar result is also obtained for SG 2 (3.869), that groups together households that are within city areas of major cities such as London or Birmingham. These households show the largest expenditure for taxi services, both in absolute and percentage terms. Finally, for the other SGs, greater-than-one elasticities are obtained, even though these measures are not very precise.

6.6. Summary of results

In the following Table 12 we provide a summary of main results presented and discussed in the previous subsections.

Table 12. Expenditure (income) elasticity measures - Linear functional form

	Full Sample	Typical traits (SG 6)	Lower Income SGs			Higher Income SGs		
			1	5	7	2	3	4
Commodity transport	1.267[b] weak luxury	1.208[b] weak luxury	1.155[b] weak luxury	2.136[b] weak luxury	2.387[d]	0.654[d]	0.842[b] weak necessity	0.449[d]
Motoring Expenditure	1.246[b] weak luxury	0.714[b] weak necessity	1.384[b] weak luxury	2.195[b] weak luxury	2.252[d]	1.293[d]	0.836[b] weak necessity	0.466[d]
Rail Expenditure	3.253[a] luxury	6.981[a] luxury	0.377[d]	7.533[a] luxury	5.084[d]	1.496[d]	4.978[a] luxury	4.293[d]
Bus Expenditure	3.253[a] luxury	-2.435[a] inferior	1.322[d]	7.755[a] luxury	3.043[d]	3.391[b] weak luxury	6.198[a] luxury	2.898[d]
Taxis Expenditure	2.646[d]	4.134[a] luxury	3.862[d]	1.867[d]	7.007[d]	3.869[b] weak luxury	2.656[d]	-0.660[c] weak inferior

Note: a - statistically different from both 0 and 1; b - statistically different from 0 but not from 1; c - not statistically different from 0 but statistically different from 1; d - not statistically different from either 0 or 1.

To sum up, the perception of different transport products and services is found to differ from a single measure, aggregated at the commodity level. Moreover, there are interesting differences also between clusters of household, exogenously identified using a geo-demographic classification. Overall, total transport expenditure and motoring expenditure can be considered weak luxuries, whereas bus and rail expenditures are definitely strongly luxuries. Looking then at the characteristics of the SGs, transport in general, and motoring expenditure in particular, are necessities rather than luxuries for higher income groups.

7. Conclusion

The aim of this study is twofold. On the one hand, it contributes to the literature on expenditure data analysis with an in-depth (and updated) investigation on households' transport expenditure in the UK, breaking down this broad commodity class into its main components. With few exceptions (Bergantino, 1997) previous research only obtained a single aggregated elasticity measure for the commodity transport, not distinguishing between car operating costs and the use of transport services (e.g. bus, train, taxis), and occasional and infrequent expenditures on durable equipment, such as the purchase of vehicles. These metrics need a constant update, given that consumers' behaviour and habits, especially those related to transport, change over time, with all implications this might have on both short- long-term planning.

On the other hand, we introduce a geo-demographical dimension in the estimation of elasticity measures; clusters of households living in similar places but geographically distant (e.g. student neighbourhoods, city centres, countryside) have been identified within the UK population, and we estimate separate elasticity measures for each cluster to reveal possible differences in the perception of transport products and services. The development of such indicators, and their use in the estimation of elasticity measures, have a huge potential in marketing as well as in policy development, going beyond the standard urban-rural or purely geographical segmentations.

Summing up our results, the aggregate expenditure (income) elasticity for the commodity transport is found to be greater than, but not statistically different from, one. This result is comparable with previous research, which also found total transport expenditure and motoring expenditure to be weak luxuries. However, further segmentation of the sample shows some evidence of different income elasticities for different clusters of consumers: transport in general, and motoring expenditure in particular, are necessities rather than luxuries for higher income groups. A similar pattern is found for motoring expenditure, which constitutes the largest share of households' weekly transport expenditure.

Rail and bus expenditures are found to be strongly luxuries. However, this result is again not confirmed for the first cluster (blue collar communities), for which we find rail transport services to be considered necessities. With respect to bus transport services, households that share the traits of the average UK households show a negative elasticity. This means that a discrete proportion of households considers bus services as inferior services. Finally, taxi services are found to share the characteristics of luxury services, even though SGs' estimates for sub-categories are not very robust.

To conclude, these results are encouraging, suggesting the incorporation of further layers of heterogeneity into the analysis, thanks to geo-demographics. These can be easily transposed into maps, and provide useful insights to policymakers, urban planners, and transport operators. The results obtained can not only increase the probability of success of transport policies, such as those aimed at reducing the car usage for residents in specific areas, but also drive better decisions on both supply and pricing of transport services, proposing, for example, tailor-made fare schemes to encourage the use of public transport for segments of the population.

Bibliography

Balcombe, R., Mackett, R., Paully, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., and White, P. (2004) The demand for public transport – a practical guide. TRL Report 593.

Banks, J., Blundell, R., Lewbel, A. (1997) Quadratic Engel curves and consumer demand, Review of Economics and Statistics, 79(4), pp. 527-539.

Beneito, P. (2003) A complete system of Engel curves in the Spanish economy, Applied Economics, 35(7), pp. 803-816.

Bergantino, A. S. (1997) Estimating Engel curves for transport expenditures: Evidence from UK household budget data, International Journal of Transport Economics/Rivista internazionale di economia dei trasporti, XXIV (2), pp. 279-305.

Berri, A., Gardes, F., Madre, J.-L., Starzec, C. (1998) Income elasticities for transport expenditures in Canada, France, Poland and U.S.A. An analysis on pseudo-panels of expenditure surveys. Paper presented at the World conference on transport research, 12-17 July, Antwerp.

Bewley, R. A. (1982) On the functional form of Engel curves: The Australian household expenditure survey 1975–76, Economic Record, 58(1), pp. 82-91.

Birkin, M., and Clarke, M. (1989) The generation of individual and household incomes at the small area level using synthesis, Regional Studies, 23(6), pp. 535-548.

Browning, M., Deaton, A., Irish, M. (1985) A profitable approach to labor supply and commodity demands over the life-cycle, Econometrica: journal of the econometric society, 53(3), pp. 503-543.

Dargay, J. M. (2007) The effect of prices and income on car travel in the UK, Transportation Research Part A: Policy and Practice, 41(10), pp. 949-960.

Deaton, A. (1985) Panel data from time series of cross-sections, Journal of econometrics, 30(1), pp. 109-126.

Deaton, A. and Muellbauer, J. (1980) An almost ideal demand system, The American economic review, 70(3), pp. 312-326.

Engel, E. (1857) Die productions-und consumtions-verhältnisse des königreichs Sachsen, Bulletin de l'Institut International de la Statistique, 9.

Gardes, F., Duncan, G. J., Gaubert, P., Gurgand, M., Starzec, C. (2005) Panel and pseudo-panel estimation of cross-sectional and time series elasticities of food consumption: The case of US and polish data, Journal of Business & Economic Statistics, 23(2), pp. 242-253.

Giles, D. E. and Hampton, P. (1985) An Engel curve analysis of household expenditure in New Zealand, Economic Record, 61(1), pp. 450-462.

Goodwin, P. B., Dargay, J.M. and Hanly, M. (2004) Elasticities of road traffic and fuel consumption with respect to price and income: a review, Transport Reviews, 24(3), pp. 275-292.

Graham, D.J., and Glaister, S. (2004) Road traffic demand elasticity estimates: a review, Transport Reviews, 24(3), pp. 261-274Haque, M. O. (1988) Estimation of Engel elasticities from the Box-Cox Engel function, Metroeconomica, 39(3), pp. 317-335.

Haque, M. O. (1992) Aggregate expenditure elasticity for transport and communication in Australia, Transportation, 19(1), pp. 43-57.

Jorgenson, D. W., L. J. Lau, and T. M. Stoker (1982) The transcendental logarithmic model of aggregate consumer behavior. In: Basmann, R. and Rhodes, G. eds. Advances in Econometrics, vol. 1., JAI Press, Greenwich.

Kleibergen, F. and Paap, R. (2006) Generalized reduced rank tests using the singular value decomposition, Journal of Econometrics, 133(1), pp. 97-126.

Liviatan, N. (1961) Errors in variables and Engel curve analysis, Econometrica: Journal of the Econometric Society, 29(3) , pp.336-362.

Meng, Y., Brennan, A., Purshouse, R., Hill-McManus, D., Angus, C., Holmes, J., Meier, P. S. (2014) Estimation of own and cross price elasticities of alcohol demand in the UK — A pseudo-panel approach using the Living Costs and Food Survey 2001–2009, Journal of health economics, 34, pp. 96-103.

Neves, P. D. (1992) Towards a more general version of the almost ideal demand system, Economics Letters, 38(3), pp. 305-308.

Office for National Statistics and Department for Environment, Food and Rural Affairs. (2014) Living Costs and Food Survey, 2001 - 2012 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor], SN: 7472, http://dx.doi.org/10.5255/UKDA-SN-7472-2.

Propper, C., Rees, H., Green, K. (2001) The demand for private medical insurance in the UK: a cohort analysis, Economic Journal, 111(471), pp. 180-200.

Singleton, A. D., and Spielman, S. E. (2014) The past, present, and future of geodemographic research in the United States and United Kingdom, The Professional Geographer, 66(4), pp. 558-567.

Summers. R. (1959) A note on least squares bias in household expenditure analysis, Econometrica: Journal of the Econometric Society, 27(1), pp. 121-126.

Verbeek. M. (2008) Pseudo-panels and repeated cross-sections. In: The Econometrics of Panel Data, Springer Berlin Heidelberg, Berlin, pp. 369-383.

Verbeek, M., and Nijman, T. (1992) Can cohort data be treated as genuine panel data?, Empirical economics, 17(1), pp. 9-23.

Verbeek, M., and Vella, F. (2005) Estimating dynamic models from repeated cross-sections, Journal of econometrics, 127(1), pp. 83-102.

Vickers, D., and Rees, P. (2006) Introducing the area classification of output areas, Population trends-London, 125, 15.

Williams, S., and Botterill, A. (2006). Profiling areas using the output area classification. Regional Trends, 39, 11-18.