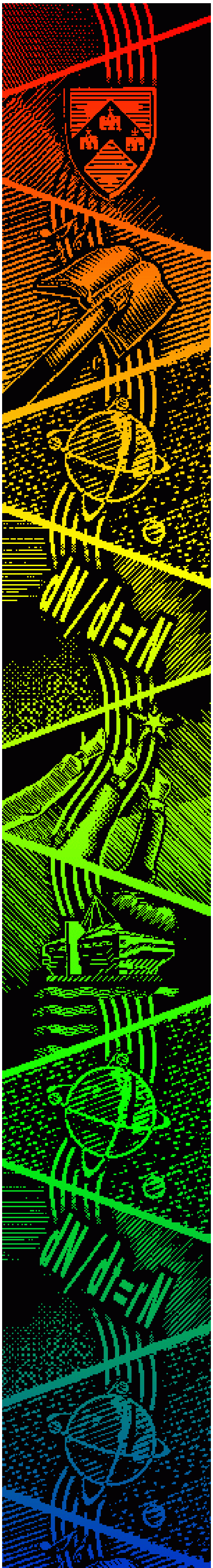

THE UNIVERSITY *of York*

Discussion Papers in Economics

No. 2004/08

Measuring Nash Equilibrium Consumption Externalities

by

Emanuela Randon

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Measuring Nash Equilibrium Consumption Externalities

Emanuela Randon*

Dept. of Economics and Related Studies, University of York, UK

Dip. di Economia Politica, Università di Milano-Bicocca, Italy

May 6, 2004

Abstract

We estimate Nash equilibrium consumption externalities in household petrol budget shares. The reaction curves are obtained from an AIDS with petrol consumption externality. Using a continuous set of ten year cross sections from FES (1991-2000), we analyse the externality generated by households living in Newcastle area (UK). In each year, income decile cohorts are created. Panel techniques are used after pooling cross section estimates have been discussed. Using non nested procedures, two restricted models are compared: the cohort specific externality effect and the single popular case. The single popular is the model accepted by the data.

Keywords: Household Economics, Nash Equilibrium, Externalities, Cross-Sectional Models, Models with Panel Data.

JEL: C31, C33, C72, D12, D62.

*This paper has been extracted from my Ph.D. Thesis at the University of York. I wish to thank my supervisor P. J. Simmons. I am also grateful to Martina Menon.

1 Introduction

Estimation and measurement of consumption externalities are still challenging problems in applied research. Particularly, estimation of Nash equilibrium consumption externalities has not been the object of applied economic works. This paper provides a first contribution in this direction. We estimate the externalities generated by a particular good, petrol, on the consumption patterns of families living in a metropolitan area in North England (Newcastle area). Since we want to analyse how the family expenditure for petrol is influenced by the petrol expenditure decision of the other families, a Nash setting seems the most appropriate to study this case. The estimation of the Nash equilibrium externalities caused by petrol is interpreted as a proxy of the externality caused by traffic congestion in this area.

Traffic congestion has been theoretically recognised as a complex phenomenon either for its physical representation or for modelling the individual decisions that cause it. For its physical description, improvements have been done from microscopic to macroscopic analysis, from static to dynamic models. Microscopic models have the aim of tracking the behavior of individual vehicles: queuing theory is an example of this kind of analysis as well as the car following theory (see May (1990)). Macroscopic models, instead, treat traffic as vehicle streams, described in term of density, speed and flow. The hydrodynamics theory is applied to analyse stationary and non stationary conditions of these variables. The analysis of the *fundamental diagram of traffic flow* introduced by Haight (1963) is a central issue in the time independent literature on this field (see May (1990) for a survey on this topic). In a time dependent contest, instead, the hydrodynamic model developed by Lighthill and Whitham (1955) and Richards (1956) has been extensively used (see Daganzo (1997) for a survey). But it is reductive modelling traffic congestion only as a physical phenomenon without considering it as the result of consumer decisions. Traffic congestion in consumer theory has been essentially analysed as a negative externality on consumer welfare. As with other negative externality phenomena, the literature on this topic focuses on the equilibrium inefficiency and on possible tolls or pricing mechanisms to correct it (see for example Small (1992)). The positive and normative results of interdependent negative externalities can be applied to the congestion problem. A more specific congestion Nash equilibrium has been examined in bottleneck models. Vickrey (1969) studied the optimal departure time decision of a fixed number of identical drivers going to work using a road with a bottleneck of fixed capacity. Each driver is facing a trade-off problem: leave home inconveniently early to avoid the queue or postpone the departure time incurring a schedule delay cost because of the bottleneck. The driver can incur an additional trip cost if a time-varying toll is applied. The equilibrium is reached when no driver has an incentive to change his optimal departure time, given the departure time decision of the other individuals. This model has been extended with different contributions: heterogeneous individuals have been considered in Cohen (1987) and in Arnott et al. (1993); different pricing regimes (uniform toll and step toll) have been the object of Arnott et al. (1993); Ramsey prices have been analysed in Arnott and Marvin (1990). In these works only numerical simulations are presented, no empirical application has been performed. Only Small (1982) presented an estimation of the optimal scheduling of work trips subject to a peak load demand, but without considering a Nash equilibrium setting.

In our work we consider a different traffic congestion Nash equilibrium. We are not interested in analysing how the daily individual departure time for work trip is influenced by the departure time decisions of other individuals, but how the daily individual decisions of using the car is affected by the decision of the car users living in the same area. For this purpose we want to estimate if the annual household expenditure of petrol is interdependently influenced by the annual household expenditure of petrol of the other families living in the same metropolitan area. We have chosen to analyse the sample of households living in Newcastle. A continuous set of ten year cross sections from the Family Expenditure Surveys (FES) is used to estimate the petrol externality. In each year we have created ten income decile cohorts. This particular data treatment allow us to observe the consumption behaviour of ten representative households through time and to analyse how the externalities can affect the Engel Curves of these ten representative households. Data can thus be treated as panel data.

Firstly, we estimate the congestion Nash equilibrium assuming specific cohort externalities: petrol total expenditure of each cohort affects petrol budget shares of other cohorts with the same intensity. A restricted form of this model is accepted. Secondly, a special case of the original model is considered: the single popular externality effect. In this case, the cohort petrol budget share is affected by the total petrol expenditure of other representative cohorts. Using nonnested test procedures the two restricted models are compared. The single popular externality effect is the most suitable to describe the Nash equilibrium household behavior. This result has an important welfare implication. In order to internalise the negative externality effect, households should be taxed independently of household income. Both models are first analysed assuming pooled cross sections. Next we estimate them within a panel framework. Fixed effect model and random effect model are compared and tested. The cross section estimates give evidence that the household petrol consumption pattern is explained by income and externality variables. Particularly, income matters for the six poorest households. For the richest households only externality variables and the intercept are significant. The fixed effect model is rejected. The random effect model confirms the results obtained by cross section techniques.

The structure of this paper is as follows. In the second section we analyse the theoretical model. The stochastic framework is introduced in Section 3. After a section of data description (Section 4), we present the Nash equilibrium estimate (Section 5). The specific cohort externality model is estimated in Section 5.1. In Section 5.2 we show that the single popular externality effect model is accepted by the data. The same result is obtained also for larger data set (see Section 6).

2 Almost Ideal Demand System with externalities

In this section we provide a theoretical economic foundation of our work. A partial equilibrium framework is assumed, in which prices, demographic variables, incomes are the exogenous variables of the model. The reaction curves are derived from an Almost Ideal Demand System (AIDS) with externalities (2). They can be interpreted as Engel curves with externalities. They are the structural forms of our demand system. Solving for the system of reaction curves, the Nash equilibrium is obtained.

2.1 Reaction curves

Suppose that petrol is good 1. Each household has a AIDS preferences with an expenditure function of this type:

$$\text{log} g_{ht}(p, z_{ht}, u_{ht}) = a_{ht}(p, z_{ht}, \overline{W}_{1kt}) + b_{ht}(p) u_{ht} \tag{1}$$

where p is the vector of prices, z_{ht} is the vector of demographic variables such as number of children, age, number of workers in family h at time t, $\overline{W}_{1kt} = \sum_{k \neq h} W_{1kt}/(H-1)$ is the deflated average total expenditure by each family other than h at time t, u_{ht} is the utility function of family h. This term represents the externality of our model. Essentially, petrol consumption is proportional to car travel which is the source of congestion.

To develop a model consistent with the standard conditions required by the consumer demand theory (satisfaction of the budget constraint and homogeneity of degree 0 in prices of the income budget share constraint), particular restrictions on the parameters of the function specified in (1) should be imposed:

$$a_{ht}(p, z_{ht}, \overline{W}_{1kt}) = a_{0h} + \sum_i a_{ih} z_{iht} + \sum_j a_{0hj} \text{log} p_j + c_{0jh} \overline{W}_{1kt}$$

$$log b_{ht}(p) = \sum_j b_{jh} \text{log} p_j$$

with:

$$\sum_j a_{0jh} = 1; \sum_j b_{jh} = 0$$

Applying Hotelling rule, the budget share of good 1 (petrol) of household h is given by[1]:

$$w_{1ht} = a_{01h} + b_h(p)b_{1h}u_{ht} =$$

$$= a_{01h} + b_{1h}(\log M_{ht} - a_h(p, z_{ht}, \overline{W}_{1kt}))$$

Finally, the reaction curves of the model are:

$$w_{1ht} = a_{01h} + b_{1h}[\text{log}M_{ht} - a_{0h} - \sum_i a_{ih}z_{iht} - \sum_j a_{0jh}\text{log}p_j - c_{01h}\overline{W}_{1kt}]$$

or in a compact form:

$$w_{1ht} = A_{01h} - b_{1h}\sum_i a_{ih}z_{hit} + C_{1h}\overline{W}_{1kt} + b_{1h}\text{log}m_{ht} \qquad (2)$$

where:

$$\begin{aligned} A_{01h} &= a_{01h} - b_{1h}a_{0h} \\ m_{ht} &= (M_{ht}/P) \\ P &\sim \prod_j p_j^{a_{0jh}} \\ C_{1h} &= -c_{01h}b_{1h} \end{aligned}$$

In the sequel we approximate P by the deflator.

Differences in petrol budget share expenditures among households are imputed to differences in income level, demographics, externality effects, type of preferences. These reaction curves can be interpreted as petrol Engel curves for each household with externality effect. The applied microeconomic analysis of the relationship between commodity consumption and income has a long tradition in the literature.

[1] $\frac{\partial logg_i}{\partial logp_i} = a_{0hj} + c_{0hj}\sum_{k\neq h} w_{jk} + \sum_i a_{ijh}z_{hi} + \frac{\partial b_h(\cdot)}{\partial logp_i} =$
$a_{0hj} + c_{0hj}\sum_{k\neq h} w_{jk} + \sum_i a_{ijh}z_{hi} + b_h(\cdot)b_{jh}$

Engel (1895), Working (1943), Leser (1963) are recognised as the seminal works in this area. Muellbauer (1976), Deaton and Muellbauer (1980) and Jorgeson et all (1982) specify all the requirements to make Engel curves compatible with integrability consumer theory. The most recent studies of the Engel Curves (see for example Hildenbrand (1994), Hausman et al. (1995), Banks et al. (1997)) indicate that the standard linearity assumption in the logarithm of expenditure doesn't provide a satisfactory explanation of the budget share Engel curves for particular goods (i.e. clothing, alcohol, non durable goods), but additional terms are required. The non parametric analysis suggests that higher order income terms should be added (Banks et al. (1997) restrict these higher income terms to being quadratic in order to satisfy the integrability requirements of the demand system). In our work, instead, we suggest that for the petrol case, the term to be added is the externality effect. To estimate and test the significance of the externality, particular restrictions on the externality effect are imposed in the stochastic framework.

3 Stochastic framework

Let us assume that the stochastic term satisfies the condition $E \in_{jh}= 0$ and Var $(\in_{jh})=\sigma_j^2$ for each h. The disturbance is normally distributed in the population as a whole with the implication that the standard normal variable $\in_{jh} / \sigma_j^2$ has a mean of zero and a variance of 1. Since we are looking only at the single equation for petrol, we are also dropping the subscript specification of the good. The stochastic framework of our structural model coincides with the reaction curves previously defined, corrected by the error term:

$$w_{ht} = A_{0h} + b_h \sum_i a_{ih} z_{iht} + C_h \overline{W}_{1kt} + b_h \log m_{ht} + \in_{ht}$$

To avoid identification problems (numbers of parameters greater than the number of observations), the data are reorganised as follow:

- the households are grouped in 10 different cohorts according to the income decile of belonging;

- in each cohort, each household has identical parameters, i.e. identical preferences ($b_h = b$; $a_{ih} = a_i$; $C_h = C$ for each h);
- for each decile we compute the mean of each variable in every year, which is represented by a bar symbol over the letter;
- for any decile h, any other decile k has the same externality effect C_k for all decile h (with $h \neq k$);
- the coefficient of the externality is constant through time.

Since we are dealing with ten cross sections, the stochastic framework of the structural form of the model becomes under these assumptions:

$$\begin{aligned}
\overline{w}_{1t} &= \overline{A} + \sum_i A_{i1}\overline{z}_{i1t} + C_2\overline{W}_{2t} + C_3\overline{W}_{3t} + ... + C_{10}\overline{W}_{10t} + b_1\overline{\log m}_{1t} + e_{1t} \\
&\vdots \\
\overline{w}_{1t+9} &= \overline{A} + \sum_i A_{i1}\overline{z}_{i1t+9} + C_2\overline{W}_{2t+9} + C_3\overline{W}_{3t+9} + ... + C_{10}\overline{W}_{10t+9} + b_1\overline{\log m}_{1t+9} + e_{1t+9}
\end{aligned}$$

$$\begin{aligned}
\overline{w}_{2t} &= \overline{A} + \sum_i A_{i2}\overline{z}_{i2t} + C_1\overline{W}_{1t} + C_3\overline{W}_{3t} + ... + C_{10}\overline{W}_{10t} + b_2\overline{\log m}_{2t} + e_{2t} \\
&\vdots \\
\overline{w}_{2t+9} &= \overline{A} + \sum_i A_{i2}\overline{z}_{i2t+9} + C_1\overline{W}_{1t+9} + C_3\overline{W}_{2t+9} + ... + C_{10}\overline{W}_{10t+9} + b_2\overline{\log m}_{2t+9} + e_{2t+9}
\end{aligned}$$

$$\vdots$$

$$\begin{aligned}
\overline{w}_{10t} &= \overline{A} + \sum_i A_{i10}\overline{z}_{i10t} + C_1\overline{W}_{1t} + C_2\overline{W}_{2t} + ... + C_9\overline{W}_{9t} + b_{10}\overline{\log m}_{10t} + e_{10t} \\
&\vdots \\
\overline{w}_{10t+9} &= \overline{A} + \sum_i A_{i10}\overline{z}_{i10t+9} + C_1\overline{W}_{1t+9} + C_2\overline{W}_{2t+9} + ... + C_9\overline{W}_{9t+9} + b_{10}\overline{\log m}_{10t+9} + e_{10t+9}
\end{aligned}$$

C_h is the parameter of the externality variable of cohort h. It indicates the effect on average of the externality caused by the petrol total expenditure of decile h on the petrol budget share of the remaining deciles. A_{ih}, b_h measure respectively the mean effect of the demographic variable i and of the log of income of decile h of each cohort on the dependent variable. Ten representative decile cohorts have been thus created, each of them representing the mean consumption behaviour through time of a household with particular income condition. The first decile cohort is tracking the consumption pattern of the poorest representative household through time, the last decile of the richest one. Previous literature on cohort analysis has been used to test life-cycle theory. For example, Browning et al. (1981) estimate the individual life-cycle of hours and wages. They simulate a panel data using cohort means of a continuum of household cross sections. The cohorts have been created according to the age of the head of the household. The age cohort mean of each variable can be thus essentially interpreted as an individual panel observation, reproducing the behaviour of the representative consumer in a particular period of his life. In our case, the information on consumption patterns are at a more aggregate level since the households have not been grouped on the basis of an individual variable but considering a household variable. The representative consumer interpretation seems appropriate only assuming that the head of household income is the main source of household total expenditures. It seems difficult, also, to interpret this model in term of life cycle theory. We can't test if the financial conditions of the representative household in time t are stationary or are evolving in the following time periods.

To estimate the model, the externality effect can be approximated to $\overline{W}_{ht} = \overline{w}_{ht} * \overline{m}_{ht}$ for each $h = 1, ..10$ and $t = 1, ..10$. Considering the low variability of the income in each cohort decile, this approximation is acceptable. In a vector form the model becomes :

$$\begin{aligned}
\overline{w}_1 &= A + \sum_i A_{i1}\overline{z}_{i1} + C_2\overline{w}_2 * \overline{m}_2 + C_3\overline{w}_3 * \overline{m}_3 + ... + C_{10}\overline{w}_{10} * \overline{m}_{10} + b_1\overline{\log m}_1 + e_1 \\
\overline{w}_2 &= A + \sum_i A_{i2}\overline{z}_{i2} + C_1\overline{w}_1 * \overline{m}_1 + C_3\overline{w}_3 * \overline{m}_3 + ... + C_{10}\overline{w}_{10} * \overline{m}_{10} + b_2\overline{\log m}_2 + e_2 \\
& \qquad (3) \\
\overline{w}_{10} &= A + \sum_i A_{i10}\overline{z}_{i10} + C_1\overline{w}_1 * \overline{m}_1 + C_2\overline{w}_2 * \overline{m}_2 + ... + C_9\overline{w}_9 * \overline{m}_9 + b_{10}\overline{\log m}_{10} + e_{10}
\end{aligned}$$

where $\overline{w}_h$, $\overline{z}_{ih}, \overline{m}_h$, $\overline{\log m}_h$ are the ten observation vectors respectively for the petrol budget share, demographic variable i, income and log of income of the decile cohort h (with $h = 1, ..10$).

The structural form is a non linear simultaneous equation model with 10 endogenous variables represented by the total petrol budget expenditure means of each decile cohort. The estimation of each equation independently of the others is equivalent to the estimation of each single reaction curve. But the estimation of the overall system requires the imposition of an equilibrium condition. In brief, the estimation of the equations simultaneously is equivalent to the estimation of the Nash equilibrium behaviour in the petrol pattern consumption of each representative cohort. To have a unique numerical estimation of the structural coefficients of our model, the system of simultaneous equations should be identified. Let us define M as the number of exogenous variables in the system; M_i as the number of exogenous variables appearing in equation i; G the number of endogenous variables. A necessary condition for a system to be identified is that for each equation the number of excluded exogenous variables ($M - M_i$) should be at least as great as the number of included right-hand-side endogenous variables ($G-1$). This is the "order condition" for identification of each equation (see, for example, Wooldridge (2001) p. 215). In our case this condition is satisfied since $M - M_i > G - 1$ in each equation. The satisfaction of the order condition is not sufficient for the identification of the parameters of the structural equations. Another condition should be satisfied, the so called "rank condition". Let us define R_i as the $J_i \times (G + M)$ matrix of known constant where J_i is the number of restrictions on the vector of structural parameters of equation i. In a *linear* system of endogenous variables, the coefficients of the structural equations are identified if and only if the *rank* $R_iB = G - 1$, where B is the $(G + M) \times G$ matrix of structural parameters in the system (see, for example, Wooldridge (2001) p. 218). In our case, the system is *non linear* in the endogenous variables since the right hand side endogenous variables are multiplying the exogenous income variables. To simplify the estimation, one method extensively used is to relabel the non linear function of the endogenous variables as new variables. In our model, thus, we have ten additional endogenous variables. It has been shown (see Wooldridge (2001) pp. 230- 234) that using this method the rank condition should be applied "without increasing the number of equations" (Wooldridge (2001) p. 234): in our case *rank* $R_iB = 9$ for the identification of the system. In the appendix B, we show that this condition is satisfied. The model is, in principle, overidentified, since the number of excluded exogenous in each equation is greater that the number of right hand side endogenous variables. The estimates of the structural parameters are not unique. One way of ensuring uniqueness is to apply 2sls to a system of simultaneous equation model overidentified, the estimates of the parameters are uniquely identified. Before presenting the empirical results applying this technique, the data set is described in the next section.

4 Data description

We used data from the 1991 to 2000 Family Expenditure Surveys (FES) of the United Kingdom to estimate the externality effect. The ten cross sections of seven thousand British families have been divided into different subsamples according to the geographic location of each household. The geographic location has been specified by two spatial variables: standard region (i.e. North West, North, Welsh, Greater London,...) and administrative area (i.e. metropolitan-non metropolitan area classified according to the density of the population). Each subsample represents thus the expenditure record of families living in the same geographic region with the same population density. Among all the subsamples created, we have decided to focus on the samples of families living in metropolitan zones since only in these zones the public transportation system is extensively developed, offering a valid alternative to personal motoring. The decision between using the car or using the public transportation system can be thus affected by the traffic congestion of the area. In non-metropolitan or rural area this alternative is not effectively present due to the relative lack of the local transportation. Personal motoring may represent the only option for households, despite of the level of traffic congestion reached in the area: it should be difficult to test the presence of negative interdependent externalities in the petrol expenditures of resident households. We have also excluded London since it has always represented an exception for the complexity of its urban structure and transportation systems. From the five samples created in each year (respectively for Glasgow, Newcastle, Leeds, Manchester, Birmingham), in this paper we have chosen to present the results of the estimates of the petrol consumption externalities of households located in Newcastle and surroundings. Table 1 shows the number of observations available for our study in each year. Each cross section represents the random sample of households living in Newcastle each year. We cannot follow the individual household behaviour through time, as in the panel data. But since we are dealing with a continuum of cross sections, in each year we can create representative households according to particular criteria and follow the consumption of these representative households through time. In Browning et al. (1985) the households have been grouped according to the age of the head of the households. In our case the representative households are created in relation to the income decile to which they belong. We have originated ten income cohorts in each cross sections. Looking at the mean behaviour of these cohorts of households through the surveys, we can track the mean behaviour of households placed in the same income decile through time. The cohorts means can be interpreted thus as a panel data.

Table 2 indicates the number of observations in each random sample once the missing values have been dropped. As expected, the missing values are present in the less recent cross sections (1991-1993).

Table 1: Number of households in each cross section for Newcastle

Year	N. of observations
1991	152
1992	157
1993	159
1994	150
1995	155
1996	125
1997	143
1998	132
1999	135
2000	116
total	1424

Table 2: Number of households in each cross section without missing values

Year	N. of observations
1991	124
1992	146
1993	138
1994	150
1995	155
1996	125
1997	143
1998	132
1999	135
2000	116
total	1364

Table 3 and Table 4 examine respectively the means over time for each decile and the overall means of each sample without missing values. The means of the demographic variables are not significantly affected once the missing values have been eliminated. Only the sample overall means of the variables in the three cross sections 1991-1993 are influenced. In principle, we could have left the missing values: they are irrelevant for the robustness of our estimations. For methodological accuracy, we have decided to drop them. The number of workers and the number of cars are on average increasing in each decile. The age of the head of the household tends to decrease. Since our aim is to measure the traffic congestion of the Newcastle area, the samples have been further selected considering only households owning at least

Table 3: Mean by decile of number of workers, children, cars and age without missing values

Decile	workers	children	cars	age
1	.0991736	.1570248	.0495868	61.87603
2	.0967742	.2903226	.1209677	58.77419
3	.2518519	.4740741	.2222222	51.57037
4	.4744526	.4525547	.4525547	54.31387
5	.7769784	.7338129	.5467626	47.47482
6	1.06993	.6013986	.7202797	47.97203
7	1.340278	.7916667	.7916667	45.34722
8	1.438849	.7625899	.9208633	45.57554
9	1.819444	.6388889	1.326389	43.94444
10	1.934783	.9202899	1.456522	43.7971
tot	.957478	.5923754	.6788856	49.761

a car and with positive petrol expenditure. We are not considering corner solution equilibria.

Table 4: Mean by year of number of workers, children, cars and age without missing values

Year	workers	children	cars	age
1991	1.266129	.6048387	.7016129	46.79839
1992	1.212329	.6986301	.8424658	48.13014
1993	1.130435	.5869565	.7101449	47.25362
1994	.82	.5066667	.6333333	50.68
1995	.8322581	.5870968	.6322581	51.65161
1996	.872	.672	.632	50.44
1997	.8741259	.5804196	.6433566	50.32867
1998	.9848485	.5227273	.9208633	45.57554
1999	.6444444	.5185185	.7424242	49.80303
2000	.9741379	.6637931	.7068966	50.42241
tot	.957478	.5923754	.6788856	49.761

Table 5 shows the number of households in each random sample once households without cars or with a null annual petrol expenditure have been eliminated. Since the random samples selected are now more restricted, the households are less heterogeneously distributed as before. In Table 6 it is possible to observe that the means of the demographic variables for each year sample don't vary significantly over time: the factor of randomness of the samples has been reduced. Looking at Table 7, the average number of workers increases in each decile while the average age is signifi-

Table 5: Number of households in each cross section with positive budget shares

Year	N. of observations
1991	59
1992	80
1993	66
1994	67
1995	72
1996	47
1997	57
1998	60
1999	53
2000	46
total	607

Table 6: Mean by year of number of workers, children, cars and age with positive budget share

Year	workers	children	cars	age
1991	1.355932	.5084746	1.237288	47.20339
1992	1.5625	.6875	1.3625	46.725
1993	1.424242	.6060606	1.242424	46.06061
1994	1.238806	.7014925	1.208955	46.40299
1995	1.152778	.5694444	1.263889	48.93056
1996	1.425532	.5744681	1.297872	46.48936
1997	1.350877	.5438596	1.245614	46.50877
1998	1.3	.4166667	1.366667	49.63333
1999	1.132075	.5660377	1.188679	49.73585
2000	1.413043	.8478261	1.391304	45.73913
tot	1.337727	.601318	1.280066	47.36244

cantly reduced in each cohort, sensitively in the lowest deciles. This suggests that the number of observations dropped refer principally to single unit households of pensioners or unemployed.

As expected, the means over time for workers and cars are increasing in each decile. The decile cohorts in the middle of the income distribution have the higher number of children on average. The oldest households belong to the lowest decile, characterised principally by pensioners. Let us now consider the means within decile and over time of total petrol expenditure deflated[2] and petrol budget shares (see

[2]To deflate the variable used in this Table, we have used the OECD 1999 Deflator Serie and the petrol price per litre from "Transport Statistics Great Britain 27th edition (2001)" by the Department for Transport, Local Government and the Regions, p. 38.

Table 7: Mean by decile of number of workers, children, cars and age with positive budget share

Decile	workers	children	cars	age
1	.3606557	.0655738	1	59.2459
2	.6229508	.3934426	1.032787	49.95082
3	.9672131	.5245902	1.147541	48.96721
4	1.183333	.6666667	1.166667	49.4
5	1.311475	.7704918	1.163934	45.72131
6	1.622951	.8688525	1.229508	43.01639
7	1.516667	.5333333	1.216667	45.81667
8	1.852459	.6885246	1.442623	43.98361
9	1.983607	.7540984	1.704918	44.88525
10	1.966667	.75	1.7	42.56667
tot	1.337727	.601318	1.280066	47.36244

Table 8: Average total expenditure, petrol expenditure, and petrol budget shares over time

Decile	totexp.	petrol exp.	petrol budget shares
1	53.6121	8.129355	.1481841
2	92.60303	6.606084	.0766885
3	126.3519	9.729563	.0764968
4	158.8654	11.39887	.0720053
5	201.4903	12.12694	.0608531
6	250.3391	14.43917	.057878
7	300.2375	15.61528	.0522787
8	361.0419	16.59269	.0461743
9	464.1758	20.92153	.0453375
10	733.3229	23.82057	.0346958
tot	374.1571	16.87412	.0521555

Table 8 and Table 9).

From Table 8 it is evident that a representation of the externality in term of petrol budget shares can not be suitable to measure the influence of traffic congestion on the household petrol consumption decision since will underestimate the impact of this phenomenon. This is due to the fact that the increase in total expenditure in each decile cohort is more than proportional to the increase in petrol expenditure. Even if the total petrol expenditure is increasing in each decile cohort, the petrol budget share results diminishing. In Table 9, instead, we can observe that the year mean of total expenditure deflated and petrol expenditure deflated tend to increase

Table 9: Average total expenditure, petrol expenditure, and petrol budget shares in each cross section

Year	totexp.	petrol exp.	petrol budget shares
1991	346.2454	15.02674	.0554029
1992	372.5229	15.2792	.0471209
1993	357.5089	16.43975	.0537472
1994	346.2005	15.75768	.0515714
1995	379.7666	17.51672	.0491743
1996	375.5325	15.65215	.0502422
1997	407.7989	16.46249	.0484754
1998	404.9928	18.74149	.0553383
1999	344.2831	18.42868	.0588379
2000	419.7327	20.79272	.0546436
tot	374.1571	16.87412	.0521555

proportionally through time: the petrol budget share means are almost constant. In the following sections we present the estimates of our work.

5 Estimating Nash equilibrium

In this section the Nash equilibrium estimates are presented. Firstly, we show that a restricted version of model specified in (3) is accepted (the accepted restrictions are $C_1 = C_2 = C_4 = C_5 = C_6 = C_9 = 0$ (see Section 5.1)). Secondly, another restricted model is considered: the single popular externality effect. The restrictions imposed are: $C_1 = ... = C_{10} = C$ (see Section 5.2). The two restricted models are compared and tested. For both of them, after the independent pooling cross section estimates we present the results of the panel data estimators.

5.1 The restricted specific cohort model

As previously explained in section 3, the structural form of our model is a non linear system of simultaneous equations. To simplify the method of estimation, the system is rearranged as follows:

$$
\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{10} \end{bmatrix} = \begin{bmatrix} 1 & z_1 & \cdots & 0 & \log m_1 & \cdots & 0 & 0 & \cdots & E_{10} \\ 1 & 0 & & 0 & 0 & & 0 & E_1 & & E_{10} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & z_{10} & 0 & \cdots & \log m_{10} & E_1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} A \\ a_1 \\ \vdots \\ a_{10} \\ b_1 \\ \vdots \\ b_{10} \\ C_1 \\ \vdots \\ C_{10} \end{bmatrix}
$$

We have transformed the original non linear simultaneous equation model in a linear system with endogenous variables. The dependent variable is represented by the petrol budget share vector of each decile (w_h with $h = 1, ..10$), the included exogenous variable i is a variable with demographic or income decile h observations if the decile= h and null observations if decile $h \neq k$ (with h and $k = 1....10$). The externality variables are the endogenous variables of the system. Externality variable h is characterised by null observations if decile=h, and petrol total expenditure for the remaining deciles. We are assuming that the petrol total expenditure of each cohort is affecting the petrol budget shares of the other cohorts with the same intensity. Since the endogenous variables are correlated with the error terms, to provide consistent and efficient estimations an appropriate estimation method is required. We should find a set of instrumental variables (excluded exogenous) correlated with the endogenous variables but uncorrelated with the error terms. The first stage of the estimation consists in regressing each endogenous variable on all the exogenous variables. The second stage consists in running the structural equation regression, replacing each endogenous variable with its own fitted values stored in the first stage regressions. To ensure valid standard errors and t statistics, we have applied 2sls Stata econometric package. The estimates of the pooled independently cross sections

Table 10: Pooling cross section estimates

	Coefficients	Standard error	t-value	p-value
constant	0.6834	0.274	2.77	0.007
logincome2	-.0909****	0.0374	-2.43	0.018
logincome3	-.0889***	0.0387	-2.30	0.024
logincome4	-0.094****	0.0375	-2.51	0.014
logincome5	-0.0760***	0.0360	-2.11	0.038
logincome6	-0.0756***	0.0350	-2.16	0.034
logincome7	-0.0795***	0.0350	-2.28	0.026
logincome8	-0.0850***	0.0367	-2.31	0.023
logincome9	-0.070***	0.0322	-2.19	0.032
logincome0	-0.0726***	0.0312	-2.33	0.022
car1	-.41201***	0.2056	-2.00	0.048
car2	-0.4628	0.1075	0.43	0.668
car4	0.0828	0.0552	1.50	0.138
car9	0.0350	0.0280	1.25	0.213
workt9	-0.0308	0.0310	-1.00	0.322
ext.income3	-0.0038****	0.0015	-2.54	0.013
ext.income7	-0.0018***	0.0009	-2.02	0.046
ext.income8	-0.0037***	0.0017	-2.23	0.029
ext.income10	-0.0012***	0.0005	-2.05	0.044
R^2	0.010			
$R^2 - adjusted$	-0.2095			
n	100			

NOTES: (1)***indicates significance at the 2.5 per cent level
(2)****indicates significance at the 1 per cent level

are shown in Table 10. Instead of adding year dummy variables to take into account the effect of different variables distribution through time, we have made the observations of each random sample identically distributed through time deflating all the variables dependent on prices (i.e. total expenditure, log of total expenditure, total petrol expenditure).

These estimations are the final results of the 2sls stepwise regression, once the less significant variables have been dropped out in each step on the base of a t test. The unrestricted model has been thus rejected in favour of a more restricted one, assuming that $C_1 = C_2 = C_4 = C_5 = C_6 = C_9 = 0$. The demographic variables seem almost irrelevant in the explanation of the consumption pattern of household petrol budget expenditures (only the number of cars in decile 1 is significant at the 2.5 percent of the confidence interval). The significant variables that explain the household

Table 11: Diagnostic tests of the pooling cross sections

Sargan Test:	$\chi^2(15)$=9.7948	[0.8324]
Test for endogeneity:	$F(4,77)$=13.60	[0.0000]
Heteroskedasticity Test:	$F(35,45)$=1.1459	[0.3303]
Testing beta=0:	$\chi^2(18)$=69.444	[0.0000]
Normality Test:	$\chi^2(2)$=0.2921	[0.8628]

consumption behaviour of this good are the log of deflated total expenditure and externality variables. All the diagnostic test requirements are satisfied and there is evidence of endogeneity (see Table 11).

The set of instruments satisfies the requirement of exogeneity (see Sargan Test)[3]. The test for endogeneity is checking that the residuals obtained from the reduced form for the endogenous variables using OLS are significant in the structural equations. In this case we reject the null (coefficients of the residuals null) at the 1 per cent significance level. We can not reject the null hypothesis that the error terms are homoskedastic (Heteroskedasticity test)[4], independently and identically distributed as a normal $N \sim (0, \sigma^2)$ (Normality test). This is evident also from the Fig. 1.

We reject also the hypothesis that the coefficients of the structural model are null (testing beta=0).

In Figure 2 we graph the actual and fitted values of our model. The observations are distributed along the intercept of the quadrant, but with a significative dispersion.

Our analysis can be further extended considering Panel data econometric approaches. Particularly, in Table 12 the results of the fixed effect model are presented.

The externality variables have still a significative explanatory power, but not surprisingly the log of income variables are no longer significant in this case. Probably because of the lack of variability of these variables through time in each decile. This is evident from Figure 3 and 4 where the log of deflated total expenditure is almost

[3]In details, the set of instruments used are: the variables for the number of cars in real terms of decile 1,3,0; total number of workers of decile 7; total number of workers in real term for decile 1,2,4,6,7,8,10; log of income squared for decile 1; 3 additional income variables for decile 3,4,8 in which the own income decile observations are repeated for each decile and 3 additional car variables for decile 2,3,7 having the own car decile observations repeated in each decile.

[4]Consider model (6.3). In the equilibrium, the budget share of each equation is a function also of the error term of all the remaining equations. In principle, we should expect Heteroskedasticity. There is not evidence of it.

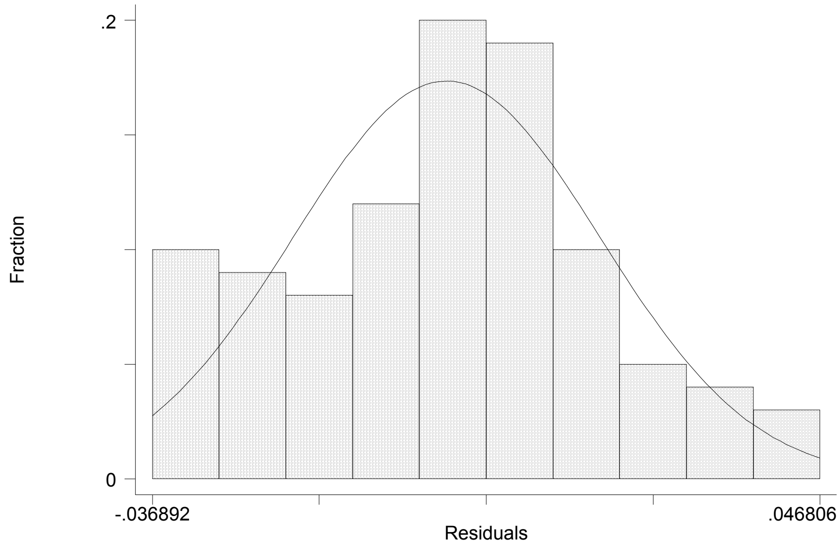

Figure 1: Standard errors are independently and identically distributed as a normal $N \sim (0, \sigma^2)$

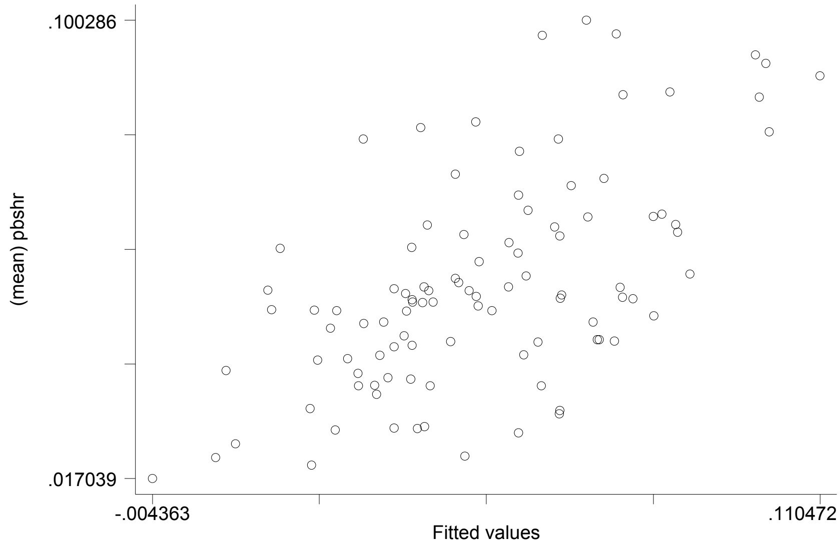

Figure 2: Plot of actual and fitted values of petrol budget shares with pooling cross sections.

Table 12: Panel data: fixed effect estimates

	Coefficients	Standard error	t-value	p-value
constant	0.6349	0.2725	2.33	0.020
logincome2	-.0791	0.0584	-1.35	0.175
logincome3	-.0740	0.0885	-0.84	0.403
logincome4	-0.0670	0.0881	-0.76	0.448
logincome5	-0.1242	0.1682	-0.74	0.460
logincome6	-0.1299	0.1209	-1.07	0.282
logincome7	-0.0182	0.0770	-0.24	0.813
logincome8	-0.0575	0.0650	-0.88	0.377
logincome9	-0.0779	0.0745	-1.05	0.296
logincome0	-0.1251	0.0851	-1.47	0.142
car1	-.41201***	0.2056	-2.00	0.048
car2	-0.0625	0.1058	0.59	0.555
car4	0.0843*	0.0523	1.61	0.107
car9	0.0388	0.0275	1.41	0.158
workt9	-0.0320	0.0304	-1.05	0.293
ext.income3	-0.0036***	0.0015	-2.33	0.020
ext.income7	-0.0015**	0.0008	-1.75	0.080
ext.income8	-0.0032**	0.0017	-1.88	0.059
ext.income10	-0.0010**	0.0005	-1.79	0.074
R^2	0.1577			
n	100			
$F^{\dagger}(9,73)$	0.16 [0.9975]			

NOTES: (1)*indicates significance at the 10 per cent level

(2)**indicates significance at the 5 per cent level

(3)***indicates significance at the 2.5 per cent level

(4)† This is the F test on the individual dummy variables

constant for each decile through time.

The fixed effect model cannot be a suitable framework to explain our problem, since the hypothesis of null individual specific dummy variables is accepted on the base of the F test specified in Table 12. The unobserved fixed individual effects are thus irrelevant to explain the household externality behaviour. The random effect model is inestimable due to the fact that the rank of the variance covariance matrix is null. In this framework, when few cohort specific externality variables are significant, the pooling independent cross section across time is the model that better describes the data set. Let us consider another specific case of our original model: the single popular externality effect.

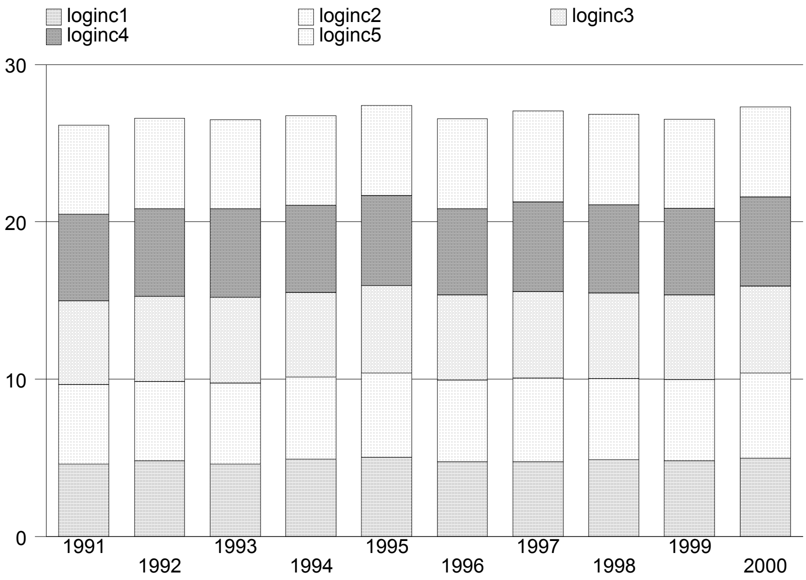

Figure 3: Log of income of the first five cohort income deciles through time

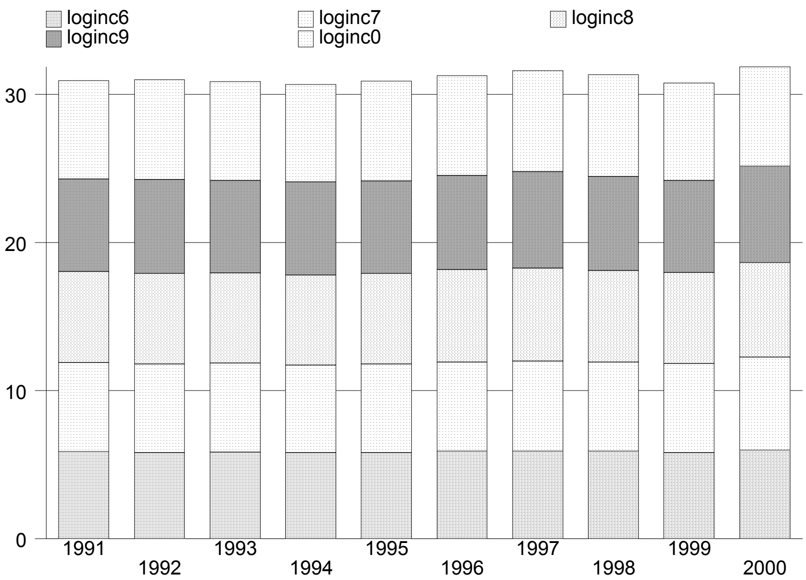

Figure 4: Log of income of the five richest income decile cohort through time

Table 13: Pooling cross section estimates of the single popular externality effect

	Coefficients	Standard error	t-value	p-value
constant	0.1017	0.0294	3.45	0.001
logincome1	-0.0522*	0.0364	-1.43	0.156
logincome2	-0.0258**	0.0148	-1.74	0.085
logincome3	0.0044*****	0.0010	4.73	0.000
logincome4	-0.0205*****	0.0074	-2.76	0.007
logincome5	0.0032******	0.0009	3.47	0.001
logincome6	0.0014**	0.0009	1.62	0.109
car1	0.2948**	0.1746	1.69	0.095
car2	0.1511***	0.0692	2.18	0.032
car4	0.1194*****	0.0364	3.27	0.002
car9	0.0481*****	0.0173	2.78	0.007
workt2	0.0122	0.0221	0.55	0.584
wortk9	-0.0401*****	0.0147	-2.72	0.008
sumext	-0.0004***	0.00019	-2.12	0.037
R^2	0.5177			
$R^2 - adjusted$	0.4448			
n	100			

NOTES: (1)*indicates significance at the 10 per cent level
(2)**indicates significance at the 5 per cent level
(3)***indicates significance at the 2.5 per cent level
(4)****indicates significance at the 1 per cent level
(5)*****indicates significance at the 0.5 per cent level

5.2 The single common popular channel

In the previous section a restricted form of the original model has been estimated. Another special case of the original model can be considered: the common popular single channel effect, in which the cohort petrol budget share is affected by the petrol expenditure of the other representative cohorts. The restrictions imposed are thus: $C_1 = ... = C_{10} = C$.

The result of the pooling independent cross section estimates are shown in Table 13. As in the previous case only income and externality regressors are relevant in the explanation of the household petrol consumption: demographic variables result almost irrelevant, numbers of car are significant only for the less wealthy households. Income is important for the 6 poorest deciles. For the richest households only externality variables and the constant are significant in explaining their petrol consumption behaviour.

The independent pooling cross section regressions satisfy all the diagnostic test requirements (see Table 14).

Since we are dealing only with one right hand side endogenous variable, the test of endogeneity is a *t test* on the residuals of the reduced form of the externality variable in the structural equation. Since the residual variable is significant, we can not reject the hypothesis of endogeneity. Also in this case the standard errors are independently and identically distributed as a normal function (see Fig. 5).

Figure 6 plots the actual and fitted values of petrol budget shares assuming the single popular externality hypothesis. Comparing to Figure 2, the observations show less dispersion along the 45 degree line.

This observation can be sufficient to reject the model estimated in the previous section. For a more rigorous procedure we use the Davidson-MacKinnon test for non nested model[5].

In Table 15 we show the results of testing the single popular externality model against the restricted cohort specific externality model previously estimated. Since the fitted values of the first model presented in this section are not statistically significant we can not reject the single popular externality model.

In Table 16 instead the cohort specific externality effect is tested against the single popular channel. Also in this case the fitted values of the second model are not significant: it is not possible even to reject the first model tested. Since both the models are accepted according to the MacKinnon-Davidson test procedure, an additional selection criterium should be considered. There is common agreement that the comparison of the adjusted R^2 can be used to discriminate between the models (see, for example, Wooldridge (2003) p.295). In this case the common popular model appears to be the most appropriate to describe households consumption behaviour (its adjusted R^2 value is higher that the one of the other model).

Since the single common popular is the accepted model, we can extend further the

[5]As previously argued, both the models estimated are specific cases of the original model presented in section 6.3. In the previous case we have imposed that $C_1 = C_2 = C_4 = C_5 = C_6 = C_9 = 0$. In the single popular model, instead, that $C_1 = ... = C_{10} = C$. None of them can be obtained as a special case from the other.

Table 14: Diagnostic tests for the single popular externality case with pooling cross sections

Sargan Test:	$\chi^2(6) = 3.0015$	[0.8087]
Test for endogeneity:	$t(84) = 2.74$	[0.007]
Heteroskedasticity Test:	$F(25, 60) = 0.42031$	[0.9904]
Testing beta=0:	$\chi^2(13) = 110.52$	[0.000]
Normality Test:	$\chi^2(2) = 1.1666$	[0.5581]

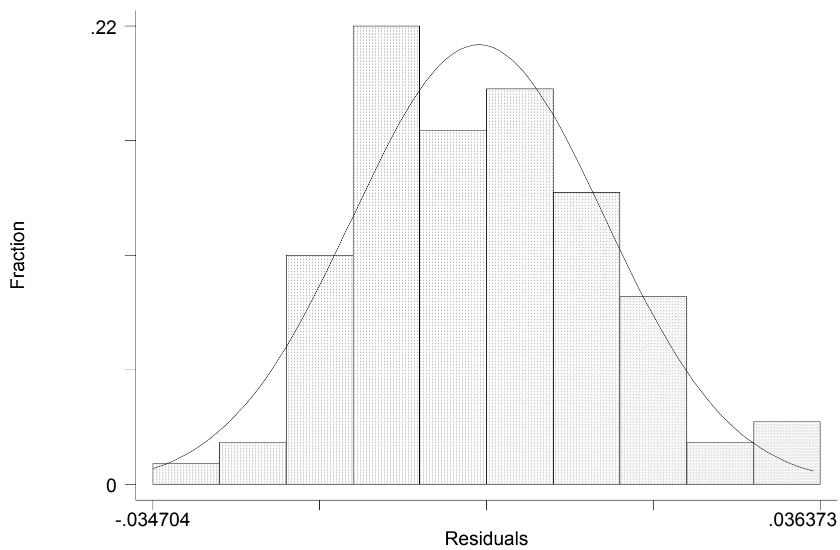

Figure 5: Standard errors are indipendentely and identically distributed as a normal $N \sim (0, \sigma^2)$ in the case of the single popular externality effect

Table 15: Davidson-MacKinnon Test for the single popular externality model

	Coefficients	t-value
constant	.1033961	3.28
logincome1	-.0486236	-1.33
logincome2	-.0233554	-1.57
logincome3	.0041952	3.70
logincome4	-.0193712	-2.47
logincome5	.0030664	3.04
logincome6	.0014314	1.52
car1	.2756243	1.58
car2	.1383472	1.96
car4	.1127416	2.88
car9	.046157	2.51
workt2	.0105295	0.47
wortk9	-.0385542	-2.46
sumext	-.0004431	-2.28
fitted values model 1	.0656315	0.68

Table 16: Davidson-MacKinnon test for the specific cohort externality model

	Coefficients	t-value
constant	0.6834	2.77
logincome2	-.0608709	-2.43
logincome3	-.052442	-2.14
logincome4	-.0612951	-2.21
logincome5	-.0458087	-1.96
logincome6	-.0458607	-2.00
logincome7	-.0456112	-1.98
logincome8	-.0481004	-2.08
logincome9	-.0456158	-2.08
logincome0	-.043574	-2.12
car1	-.24524	-1.87
car2	.0597396	0.74
car4	.0718665	1.50
car9	.0293898	1.75
workt9	-.0170048	-0.69
ext.income3	-.0019103	-2.61
ext.income7	-.0003032	-0.57
ext.income8	-.00129	-2.06
ext.income10	-.0003799	-1.05
fitted values II model	.2512712	0.79
n	100	

NOTES: (1)***indicates significance at the 2.5 per cent level
(2)****indicates significance at the 1 per cent level

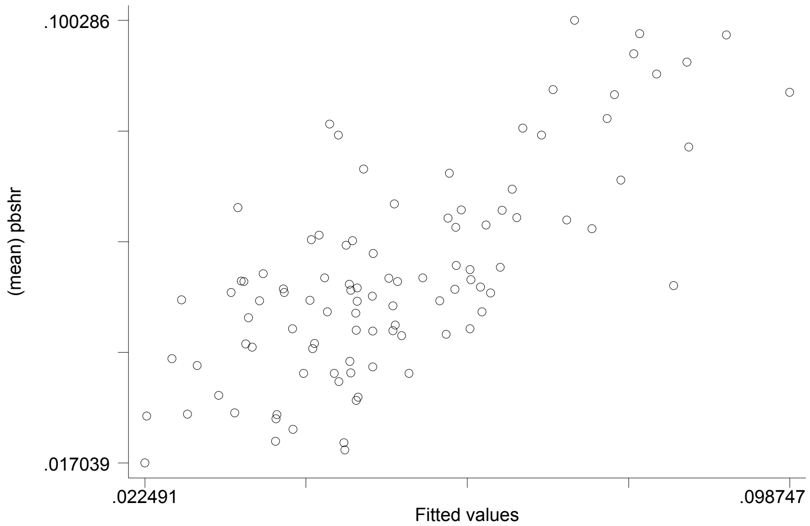

Figure 6: Actual and fitted values of petrol budget shares in the case of a single common popular externality effect with pooling cross sections.

analysis of its interpretative power considering a panel data setting. In Table 17 and Table 18 the results of the fixed effect model and of the random effect are respectively presented. As in the previous case, the fixed effect is rejected on the base of a F *test*. The Hausman test favours to the random effect model: $\chi^2(12)$=1.19 [1.0000].

The null hypothesis that the difference between the fixed and the random effect coefficient is not systematic is accepted. It is not possible to reject the hypothesis at the base of the Random effect model that the unobserved individual effect is uncorrelated with each explanatory variable.

Comparing the random effect model with the independent pooling cross section model, it is interesting to notice the similarity of the estimate results.

6 Other metropolitan area data sets

This chapter focuses on the measuring of congestion Nash equilibrium in Newcastle area. In Uk, there are other four metropolitan areas: Birmingham, Leeds, Glasgow, Manchester. Glasgow data set has similar number of observations to the Newcastle

Table 17: Panel Data: fixed effect estimates of the single popular externality effect

	Coefficients	Standard error	t-value	p-value
constant	0.1063	0.1000	1.06	0.288
logincome1	-0.0538*	0.0366	-1.47	0.141
logincome2	-0.034	0.0490	-0.70	0.486
logincome3	-0.012	0.0661	-0.18	0.854
logincome4	-0.0520	0.0613	-0.85	0.397
logincome5	0.0804	0.1020	0.79	0.430
logincome6	0.01513	0.08533	0.18	0.859
car1	0.2948**	0.1746	1.69	0.095
car2	0.1442**	0.0817	1.76	0.078
car4	0.1199*****	0.03605	3.33	0.001
car9	0.0497*****	0.0176	2.83	0.005
workt2	0.0106	0.0251	0.42	0.671
wortk9	-0.0327*	0.0216	-1.52	0.129
sumext	-0.0004**	0.0002	-1.85	0.065
R^2	0.041			
n	100			
$F^{\dagger}(9, 78)$	0.93 [0.5056]			

NOTES: (1)*indicates significance at the 10 per cent level
(2)**indicates significance at the 5 per cent level
(3)***indicates significance at the 2.5 per cent level
(4)****indicates significance at the 1 per cent level
(5)*****indicates significance at the 0.5 per cent level
(6)†This is the F test on the individual dummy variables

data set (795 observations without missing values and with positive petrol budget shares). The data set of families living in Birmingham with a positive petrol budget expenditure is characterised by 1543 observations. The number of observations for households in Leeds is 2155. The biggest data set is characterised by the year random samples of households living in Manchester (2292 observations once missing value and null petrol expenditure have been eliminated). The results obtained in the relatively small data set representing the Newcastle area are still confirmed in the other data sets, despite of the increase in the number of observations. For example in Table 19 we estimate the single popular externality model with independent pooling cross sections for Manchester (the sample with the largest number of observations). There is still evidence of negative externality effects in the households petrol expenditures.

Also in this case all the diagnostic requirements are satisfied (see Table 20) and the externality variables is endogenous.

Table 18: Panel data: random effect estimates of the single popular externality effect

	Coefficients	Standard error	t-value	p-value
constant	0.1029	0.0290	3.54	0.000
logincome1	-0.0515*	0.0354	-1.45	0.147
logincome2	-0.0256**	0.0145	-1.76	0.078
logincome3	0.0044***	0.0019	2.32	0.000
logincome4	-0.0205*****	0.0074	-2.78	0.007
logincome5	0.0032**	0.0018	1.82	0.069
logincome6	0.0015	0.0017	0.87	0.386
car1	0.2918**	0.1700	1.72	0.086
car2	0.1505***	0.0673	2.24	0.025
car4	0.1195*****	0.0354	3.37	0.001
car9	0.0484*****	0.0168	2.88	0.004
workt2	0.0117	0.0215	0.55	0.585
wortk9	-0.0399*****	0.0147	-2.70	0.007
sumext	-0.0004***	0.00019	-2.23	0.026
R^2	0.5177			
n	100			

NOTES: (1)*indicates significance at the 10 per cent level
(2)**indicates significance at the 5 per cent level
(3)***indicates significance at the 2.5 per cent level
(4)****indicates significance at the 1 per cent level
(5)*****indicates significance at the 0.5 per cent level

The fixed effect model is not accepted: the F test on the individual dummies doesn't reject the null hypothesis: $F(9, 77) = 0.60$. The results obtained in relatively small samples are still confirmed once larger number of observations are considered.

7 Conclusion

We provide empirical estimates of Nash equilibrium negative externalities in household petrol budget shares. For this purpose we have used a continuum of ten cross sections from FES (1991-2000), selecting random samples of households living in Newcastle. In each year the household observations have been grouped according to the household income decile. Ten representative income decile cohorts have been created. We evaluate two restricted models, namely the cohort specific externality effect and the single popular case, using cross section and panel data techniques. The single popular model is the one most suitable to describe the consumption behaviour of our samples. This result is confirmed also for data sets with larger observations.

Table 19: Pooled cross section estimates for Birmingham

	Coefficients	Standard error	t-value	p-value
constant	0.0788	0.0151	5.21	0.000
logincome1	-0.0320**	0.0167	-1.92	0.058
logincome3	0.0011**	0.0007	1.69	0.095
logincome5	-0.0075*	0.0051	-1.49	0.140
logincome6	-0.0011**	0.0007	-1.71	0.091
logincome8	-0.0089**	0.0052	-1.72	0.090
logincome9	-0.0020*****	0.0006	-3.31	0.001
car1	0.1765***	0.0784	2.25	0.027
car2	0.0095*****	0.0034	2.73	0.008
car5	0.0331*	0.0228	1.45	0.149
car7	-0.0053**	0.0027	-1.95	0.054
car8	0.0291*	0.0200	1.45	0.150
car0	-0.0108*****	0.0020	-5.30	0.000
sumext	-0.0002**	0.0001	-1.77	0.080
R^2	0.6795			
$R^2 - adjusted$	0.6311			
n	100			

NOTES: (1)* indicates significance at the 10 per cent level
(2)** indicates significance at the 5 per cent level
(3)*** indicates significance at the 2.5 per cent level
(4)***** indicates significance at the 0.5 per cent level

Table 20: Diagnostic tests for the single popular externality case with independent pooling cross sections for Birmingham

Sargan Test[6]:	$\chi^2(3)$=0.74148	[0.8634]
Test for endogeneity:	$t(16, 83)$=4.11	[0.0000]
Heteroskedasticity Test:	$F(25, 59)$=1.1457	[0.1166]
Testing beta=0:	$\chi^2(13)$=205.97	[0.0000]
Normality Test:	$\chi^2(2)$=2.7475	[0.2532]

We suggest that, in order to internalise the externality effect, a tax independent of household income should be implemented.

Appendix A: Identification of the system

Consider the system (3):

$$
\begin{aligned}
\overline{w}_1 &= A + \sum_i A_{i1}\overline{z}_{i1} + C_2\overline{w}_2 * \overline{m}_2 + C_3\overline{w}_3 * \overline{m}_3 + ... + C_{10}\overline{w}_{10} * \overline{m}_{10} + b_1\overline{\log m}_1 + e_1 \\
\overline{w}_2 &= A + \sum_i A_{i2}\overline{z}_{i2} + C_1\overline{w}_1 * \overline{m}_1 + C_3\overline{w}_3 * \overline{m}_3 + ... + C_{10}\overline{w}_{10} * \overline{m}_{10} + b_2\overline{\log m}_2 + e_2 \\
\\
\overline{w}_{10} &= A + \sum_i A_{i10}\overline{z}_{i10} + C_1\overline{w}_1 * \overline{m}_1 + C_2\overline{w}_2 * \overline{m}_2 + ... + C_9\overline{w}_9 * \overline{m}_9 + b_{10}\overline{\log m}_{10} + e_{10}
\end{aligned}
$$

To simplify the analysis suppose that the system is organised in three decile cohorts and that only one demographic variable is present in each equation:

$$
\begin{aligned}
\overline{w}_1 &= A + A_{11}\overline{z}_1 + C_2\overline{w}_2 * \overline{m}_2 + C_3\overline{w}_3 * \overline{m}_3 + b_1\overline{\log m}_1 + e_1 \\
\overline{w}_2 &= A + A_{22}\overline{z}_2 + C_1\overline{w}_1 * \overline{m}_1 + C_3\overline{w}_3 * \overline{m}_3 + b_2\overline{\log m}_2 + e_2 \\
\overline{w}_3 &= A + A_{23}\overline{z}_3 + C_1\overline{w}_1 * \overline{m}_1 + C_2\overline{w}_2 * \overline{m}_2 + b_3\overline{\log m}_3 + e_{10}
\end{aligned}
$$

The model is a non linear simultaneous system in three endogenous variables $\overline{w}_1, \overline{w}_2$, $\overline{w}_3$ ($G = 3$) with six predetermined exogenous variables ($M = 6$) and two predetermined variables for each equation ($M_i = 2$). According to Fisher (1965), for identification it is sufficient to relabel $\overline{w}_i * \overline{m}_i$ as new variables (with $i = 1, .., 3$) and to obtain that $rank\ R_iB = 2$ (the number of the *original* endogenous variables minus 1).

The "extended" form of our model becomes:

$$
\begin{aligned}
\overline{w}_1 &= e_{12}\overline{w}_2 + e_{13}\overline{w}_3 + A_{11}\overline{z}_1 + A_{12}\overline{z}_2 + A_{13}\overline{z}_3 + C_{11}\overline{w}_1 * \overline{m}_1 + C_{12}\overline{w}_2 * \overline{m}_2 + \\
&\quad C_{13}\overline{w}_3 * \overline{m}_3 + b_{11}\overline{\log m}_1 + b_{12}\overline{\log m}_2 + b_{13}\overline{\log m}_3 + A + e_1 \\
\overline{w}_2 &= e_{21}\overline{w}_1 + e_{23}\overline{w}_3 + A_{21}\overline{z}_1 + A_{22}\overline{z}_2 + A_{23}\overline{z}_3 + C_{21}\overline{w}_1 * \overline{m}_1 + C_{22}\overline{w}_2 * \overline{m}_2 + \\
&\quad C_{23}\overline{w}_3 * \overline{m}_3 + b_{21}\overline{\log m}_1 + b_{22}\overline{\log m}_2 + b_{23}\overline{\log m}_3 + A + e_2 \\
\overline{w}_3 &= e_{31}\overline{w}_1 + e_{32}\overline{w}_2 + A_{31}\overline{z}_1 + A_{32}\overline{z}_2 + A_{33}\overline{z}_3 + C_{31}\overline{w}_1 * \overline{m}_1 + C_{32}\overline{w}_2 * \overline{m}_2 + \\
&\quad C_{33}\overline{w}_3 * \overline{m}_3 + b_{31}\overline{\log m}_1 + b_{32}\overline{\log m}_2 + b_{33}\overline{\log m}_3 + A + e_3
\end{aligned}
$$

with $e_{ij} = A_{ij} = b_{ij} = 0$ if $i \neq j$ with $i = 1, 2, 3$ and $j = 1, 2, 3$ and $C_{ij} = 0$ if $i = j$.

Consider the rank condition for equation 1. We should prove that rank $R_1 B = 2$. The matrix R_1 and B (previously defined in section III) are respectively:

$$R_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} \text{-1} & e_{21} & e_{31} \\ e_{12} & \text{-1} & e_{32} \\ e_{13} & e_{23} & \text{-1} \\ b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \\ a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \\ 1 & 1 & 1 \end{bmatrix}$$

The matrix $R_1 B$ is thus:

$$R_1B = \begin{bmatrix} e_{12} & -1 & e_{32} \\ e_{13} & e_{23} & -1 \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ c_{11} & c_{21} & c_{31} \end{bmatrix}$$

Imposing the restrictions previously specified ($e_{ij} = A_{ij} = b_{ij} = 0$ if $i \neq j$ with $i = 1, 2, 3$ and $j = 1, 2, 3$ and $C_{ij} = 0$ if $i = j$), we get finally:

$$R_1B = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \\ 0 & c_{21} & c_{31} \end{bmatrix}$$

The rank $R_1B = 2$: the first equation satisfies the rank order condition for identifiability. Applying the same method, we can prove that the order condition is satisfied also for the other equations. We have riorganised the system in three cohorts, for analytically convenience. This proof can be easily extended to the system of 10 cohorts. Using the same method, the condition $rank R_1B = 9$ is proved.

Appendix B: Price series used in this work

Year	Price Deflator*	Petrol Price per litre**
1991	80.2	49.0
1992	83.4	47.8
1993	85.7	54.8
1994	87	56.4
1995	89.3	60.1
1996	92.2	60.4
1997	94.9	64.6
1998	97.9	72.4
1999	100	77.8
2000	102.4	84.5

NOTES: *OECD 1999 price deflator series

**Transport Statistics Great Britain 27th edition (2001)

References

[1] Arnott, R. de Palma, A. and Lindsey, R. (1993), "A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand", *American Economic Review*, 38, 161-79.

[2] Arnott, R. and Kraus, M. (1995), "Self-Financing of Congestible Facilities in a Growing Economy", *Boston College Working Paper*, 304.

[3] Arnott, R. and Marvin, K. (1993), "The Ramsey Problem for Congestible Facilities", *Journal of Public Economics*, 50, 371-96.

[4] Banks, J. W., Blundell, R. W. and Lewbel, A. (1996), "Tax Reform and Welfare Measurement: Do We Need Demand System Estimation", *Economic Journal*, 100, 1227-41.

[5] Banks, J., Blundell, R. and Lewbel, A. (1997), "Quadratic Engel Curves and Consumer Demand", *Review of Economics and Statistics*, 79, 527-39.

[6] Browning, M., Deaton, F. and Irish, M. (1985), "A Profitable Approach to Labor Supply and Commodity Demands over the Life-Cycle", *Econometrica*, 53, 503-44.

[7] Cohen, Y. (1987), "Commuter Welfare under Peak-Period Congestion Tolls: Who Gains and Who Loses?", *International Journal of Transport Economics*, 14, 238-66.

[8] Daganzo, C. F. (1997), *Fundamentals of Transportation and Traffic Operations*, Elsevier Science (New York).

[9] Deaton, A. S. and Muellbauer, J. (1980), "An Almost Ideal Demand System", *American Economic Review,* 70, 321-36.

[10] Engel, E. (1895), "Die Lebenskosten Belgischer Arbeiter-Familien Fruher and Jetzt", *International Statistical Institute*, 9, 1-74.

[11] Fisher, F. M. (1965), "Identifiability Criteria in Nonlinear Systems: A Further Note", *Econometrica*, 33, 197-205.

[12] Haight, F. A. (1963), *Mathematical Theories of Traffic Flow*, Academic Press (New York).

[13] Hausman, J. A., Newey, W. K. and Powell, J. L. (1995), "Nonlinear Errors in Variables: Estimation of Some Engel Curves", *Journal of Econometrics*, 65, 205-33.

[14] Hildenbrand, W. (1994), *Market Demand: Theory and Empirical Evidence*, Princeton University Press (Princeton, NJ).

[15] Jorgenson, D. W., Lau, L. J. and Stoker, T. M. (1982), "The Trascendal Logarithmic Model of Aggregate Consumer Behavior", in *Advances in Econometrics*, 1, edited by R. Basmann and G. Rhodes, JAI Press, (Greenwich, CT).

[16] Leser, C. E. V. (1963), "Forms of Engel Curves", *Econometrica*, 31, 694-703.

[17] Lighthill, M. J. and Whitham, G. B. (1955), "On Kinematic Waves II: A Theory of Traffic Flow on Long Crowded Roads", *Proceedings of the Royal Society* (London), 229A, 317-45.

[18] May, A. D. (1990), *Traffic Flow Fundamentals*, Prentice Hall, Englewood Cliffs, (New Jersey).

[19] Muellbauer, J. (1976), "Community Preferences and the Representative Consumer", *Econometrica*, 44, 525-43.

[20] Richards, P. I. (1956), Shock Waves on the Highway, *Operations Research*, 4, 42-51.

[21] Small, K. A. (1982), The Scheduling of Consumer Activities: Work Trips, *American Economic Review*, 72, 467-79.

[22] Vickrey, W. S. (1969), "Congestion Theory and Transport Investment", *American Economic Review*, 59, 251-60.

[23] Wooldridge, J. M. (2001), *Econometric Analysis of Cross Section and Panel Data*, MIT Press (Cambridge, MA).

[24] Wooldridge, J. M. (2003), *Introductory Econometrics. A Modern Approach (2nd edition)*, South-Western (Ohio).

[25] Working, H. (1943), "Statistical Laws of Family Expenditure", *Journal of the American Statistical Association*, 38, 43-56.