Heterogeneity in Consumer Demands and the Income Effect: Evidence from Panel Data

Mette Christensen

THE INSTITUTE FOR FISCAL STUDIES
WP16/07

Heterogeneity in Consumer Demands and the Income Effect: Evidence from Panel Data*

Mette Christensen†

Abstract

All micro studies of demand are based on using time series cross sectional data. Because in such data each household is only observed once, it is only under strong identifying restrictions that one can interpret the coefficients on consumer behavior. For example, if tastes are correlated with income, the usual estimates of income elasticities from cross sectional data are biased. In contrast, panel data allows identification of the coefficients on consumer behavior in the presence of unobservable correlated heterogeneity. In this paper we make use of a unique Spanish panel data set on household expenditures to test whether unobservable heterogeneity in household demands (taste) is correlated with total expenditures (income). We find that tastes are indeed correlated with income for half of the goods considered.

JEL: C33, D12

Keywords: Engel curves, unobservable heterogeneity, panel data

*I am grateful to Richard Blundell, Martin Browning, Lola Collado, Tom Crossley, Mette Ejrnaes, Birgit Grodal, Frank Windmeijer, Allan Würtz and participants in seminars at Centre of Applied Microeconometrics, University of Copenhagen, The QED Workshop, University of Alicante and The 10th International Conference on Panel Data and the CHILD workshop for many helpful comments and discussions. I would also like to thank Lola Collado for providing the data set, and José M. Labeaga for help with the data. Financial assistance from Centre of Applied Microeconometrics, from Institute of Economics, University of Copenhagen and from The European Community's Human Potenetial Programme under contract HPRN-CT-2002-00235 [AGE] is gratefully acknowledged. All errors are mine.

†Department of Economics, University of Manchester, and The Institute for Fiscal Studies, London. E-mail: mette.christensen@manchester.ac.uk.

1 Introduction

The usual way of estimating demand systems is to use cross sectional household level data on expenditures. Because data is cross sectional, each household is observed only once. This means that the influence of differences in income, demographic composition, and prices faced (that is, observables) can be modeled, but that it is only under extremely restrictive (and not well-studied) assumptions it is possible to model the effect of unobservable differences between households in their tastes or propensities to spend in particular ways. Therefore demand system estimation usually assumes that all consumers have identical preferences. Now, if this is not the case and consumers *do* have different preferences (which is widely believed to be the case), we will still get unbiased estimates of income repsonses if this preference heterogeneity is random in nature, that is, if it does not vary with income in any systematic way. But if this assumption is not met and so preferences are correlated with income, then estimates of income responses will be biased (because differences in income will be confounded by differences in tastes). This issue is of great importance because the magnitudes of income responses are important inputs into many policy analyses. For example, the distributional impact of tax reforms (the extent to which tax changes differentially impact rich and poor households) depends on the way spending patterns vary by income levels.

In contrast to the usual way of estimating demand systems, this paper uses a unique panel data set, the Encuesta Continua de Presupuestos Familiares (hereafter the ECPF) on household expenditures. The ECPF is a rotating panel data set covering the period 1985-97 with households staying in the survey for between 5 and 8 quaters. It has information on a wide range of goods, information on several income categories as well as information on demographic variables. Because each household is observed multiple times, we can allow for time-invariant unobservable heterogeneity in demands. We estimate a demand system for a full range of non-durable consumption goods and test whether unobservable heterogeneity (tastes) is with total expenditure (income). We find that total expenditure and tastes are correlated for almost half of the goods considered.

Panel data on household expenditures for a full range of commodities is scarce, and as a consequence, only a few other studies have considered the question of correlation between preferences and incomes. Aasness, Biørn and Skjerpen (1993) estimate a Linear Expenditure System on a Norwegian panel of length two years with yearly information. They find that preferences are correlated with income. However, their study has some limitations. Firstly, they only have two observations per household, where we have between five and eight observations per household. Secondly, they use the Linear Expenditure System, which has some unfortunate functional form implications. One implication of the Linear Expenditure System is that all goods must be substitutes (if the system is to be integrable, which they impose). Another, and more important implication for the research question under consideration is that imposing linear Engel curves if data is in fact not linear may be problematic when testing for correlated heterogeneity. To see this, suppose that data is really quadratic. Fitting a linear specification to this data while allowing for different intercepts for different households may then imply that the household-specific intercepts seem to be significantly different when in fact it is the misspecification that forces them to be different. We will return to this point in the model section. Duncan, Gardes, Gaubert and Strarzec (2002) also examine possible biases in budget elasticities arising from not taking account of correlated heterogeneity, using the Panel Study of Income Dynamics (the PSID) and a Polish panel data survey. They also find evidence of bias caused by correlated heterogeneity. However, their study is limited in that their data only has expenditure information on food eaten at home and food eaten outside home. In contrast, the ECPF has expenditure information on a wide range of goods. Carrasco, Labeaga and Lopez-Salido (2004) use the ECPF to test for the presence of habit formation, employing a test proposed in Meghir and Weber (1996). They model preferences as flexible direct translog preferences allowing for time non-separabilities and preference shocks. They find that it is important to control for correlated heterogeneity ("fixed effects") and doing that they find evidence of habit formation in all three goods. However, they do not consider a full demand system, but only the three goods, Food at home, Transport and Services [1]. Labeaga and Puig (2002) also use the ECPF and find evidence of correlated heterogeneity. They also emphasize the linear model and consider a different sample than we do. Another paper that is closely related to ours is Browning and Collado (2004). They take as starting point our finding of correlated heterogeneity and ask whether there in addition to this correlated heterogneity also is evidence of habit formation. Their answer is affirmative. Finally, Calvet and Comon (2003) test for correlated heterogeneity, using the Family Expenditure Survey (the FES). The FES is a cross sectional data set and thus Calvet and Comon (2003) are forced to employ a restrictive specification of unobservable heterogeneity in order to be able to identify their model. Firstly, they need that budget shares are linear in the expenditure term, and secondly, they can only allow for a one-dimensional heterogeneity term, which is a product of a one-dimensional household-specific term and a one-dimensional commodity-specific term. In other words: Heterogeneity across consumers is modelled by a single parameter for each consumer. With panel data it is possible to allow for heterogeneity terms that are both household - and commodity-specific, thus allowing for heterogeneity across consumers to be modelled by an N-dimensional parameter, where N is the number of goods. The empirical results of Calvet and Comon (2003) are stunning: They conclude that basically all the variation in budget shares is due to correlated heterogeneity, i.e. that income and tastes are completely confounded. This means that there is no income effect, so changes in household income have no impact on budget shares. In this paper, we also test if there is an income effect. Our findings contradict those of Calvet and Comon (2003) in that, even though we also find evidence of correlated heterogeneity, we still find evidence of significant income effects for the vast majority of the commodities considered.

In order to illustrate the issue of income effect versus correlated heterogeneity more clearly, consider the following figure:

[1] They choose these three goods only in order to be able to compare their findings with those of Meghir and Weber (1996), who consider exactly these three goods.

	identical tastes (identical preferences)	non-identical tastes (non-identical preferences)
income independent budget shares (homothetic preferences)	*	pure taste effect
income dependent budget shares (non-homothetic preferences)	pure income effect	combination

If all the variation in budget shares we see in the data is due to heterogeneity and there is no impact of income on budget shares, it means that consumers have different homothethic preferences [2]. In this case, income and tastes are completely confounded and what we would usually attribute to changes in income when estimating demand systems on cross sectional data should, if this is true, in fact be attributed to heterogeneity only. This is the top right box in the figure, and this is the conclusion of Calvet and Comon (2003). The usual way of formulating and estimating demand systems is represented in the lower left box of the figure: Usually, it is assumed that consumers have identical non-homothetic preferences. But it could also be the case that tastes and income are only partly confounded and there is in fact an effect from both tastes and income. This is the lower right box of the figure.

The research question of this paper can thus be formulated in terms of this figure: Which box is the correct way to model and estimate demand systems? Is the variation in budget shares caused by an income effect, by a taste effect or by a combination of the two? That is the question we seek to answer in this paper.

Another motivation for considering the question of whether total expenditure is correlated with unobserved heterogeneity can be found in the theoretical literature that deals with nonparametric identification of demand models with preference heterogeneity from repeated cross sections. Every identification result in this literature presupposes that unobservables are independent of observables (see e.g. Roehrig (1988), Brown and Walker (1989), Brown and Matzkin (1998), Blundell, Browning and Crawford (2003), Matzkin (2003), Beckert and Blundell (2004)). This paper can thus, aside from the pure applied implications of bias or no bias in the usual estimates of income responses, also be seen as reaching into the broader research agenda on "How to formulate preference heterogeneity in demand systems". Obviously, the issue of whether observables and unobservables are correlated becomes less of a problem when panel data is available, because with panel data we have possibilities for allowing for such correlation. But panel data on consumer expenditures is scarce. And moreover, some of the existing cross sectional data sets on consumer expenditures are of very high quality and ongoingly collected (like for example the FES). The research question of how to identify consumer demand models with unobservable heterogeneity from

[2] When preferences are homothetic, budget shares are independent of income. One example is Cobb-Douglas preferences, $u(x_1, ..., x_N) = \Pi_{i=1}^{N} x_i^{\alpha_i}$ where the budget shares are given by the fixed coefficients $\alpha_1, ..., \alpha_N$.

cross sections is therefore also for this reason an important one.

In this paper, we formulate a demand system as a linearized QUAID system with unobservable heterogeneity representing differences in tastes in consumers' demands. Because we have panel data available we can let the additive unobservable heterogeneity terms be both commodity - and household-specific. Further, we can test whether total expenditure is correlated with tastes by an instrumental variables approach, because the time series of observations for each household provides us with (potential) instruments. This means that we are able to identify the income responses in the case of correlated heterogeneity. We carefully examine validity and relevance of the potential instruments. We carry out the test for correlated heterogeneity as the regression-based version of a Wu-Hausman test for endogeneity. We compare the estimated marginal effects when taking account of correlated heterogeneity with the marginal effects resulting from the usual way of estimating demand systems.

This rest of this paper is organised as follows. In Section 2 we specify the model. In Section 3 we describe the sample of the ECPF we use for this paper. Section 4 contains the empirical analysis and results and Section 5 discusses and concludes.

2 The Model

We will base our model on the linearized version of the Quadratic Almost Ideal Demand system (the QUAID) used in Blundell, Pashardes and Weber (1993). The original QUAID system is the quadratic extension of the Almost Ideal Demand system (the AID) and was introduced by Banks, Blundell and Lewbel (1997).

Let $i = 1, ..., N$ index the N commodities, let $h = 1, ..., H$ index households and let $t = 1, ..., T_h$ index time periods for households h. Let w_{iht} denote the budget share for commodity i for household h at time t, let x_{ht} denote total expenditure for household h at time t and let p_{it} denote the price of commodity i at time t and let p_t denote the vector of all prices at time t. The QUAID system is then given by

$$w_{iht} = \alpha_i + \sum_{j=1}^{N} \widetilde{\gamma}_{ij} \ln p_{jt} + \beta_i \left[\ln x_{ht} - \ln P(p_t)\right] + \frac{\lambda_i}{b(p_t)} \left[\ln x_{ht} - \ln P(p_t)\right]^2, \tag{1}$$

$i = 1, ..., N$, $h = 1, ..., H$ and $t = 1, ..., T_h$, where the price index is given by

$$\ln P(p_t) = \alpha_0 + \sum_{j=1}^{N} \alpha_j \ln p_{jt} + \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{N} \widetilde{\gamma}_{jk} \ln p_{jt} \ln p_{kt}, \qquad t = 1, ..., T$$

and the price-dependent part of the coefficient on the quadratic expenditure term is given by

$$b(p_t) = \prod_{i=1}^{N} p_{it}^{\beta_i}.$$

This demand system is integrable, that is, there exists a utility function such that the demand underlying the budget shares in (1) is given as the maximisation of this utility function subject to a linear budget constraint. A feature of the QUAID that has made it very popular is that it allows different income responses at different income levels: The marginal effect of the budget share for commodity i is given as

$$\frac{\partial w_i}{\partial x} = \beta_i + 2\frac{\lambda_i}{b(p)} \ln x,$$

that is, a household with a high level of total expenditure (high x) can have a different behavioral response to a change in its budget than a household with a low level of total expenditure (low x). Moreover, it allows for commodities to be a necessary good at some levels of total expenditure and a luxury good at other levels of total expenditure. To see this, consider the income elasticity for this model. In general, the income elasticity for commodity i can be written as[3]

$$e_i = \frac{1}{w_i}\frac{\partial w_i}{\partial(\ln x)} + 1,$$

and commodity i is defined to be a luxury good if and only if $e_i > 1$. We have that

$$e_i = \frac{1}{w_i}\left(\beta_i + 2\frac{\lambda_i}{b(p)} \ln x\right) + 1,$$

so commodity i is a luxury for a household with total expenditure level x if and only if $\beta_i + 2\frac{\lambda_i}{b(p)} \ln x > 0$. Therefore, a commodity can be a luxury at some levels of total expenditure and a necessity at other levels of total expenditure.

The QUAID system is clearly nonlinear in parameters, since both the price index $\ln P(p_t)$ and the coefficient on the quadratic expenditure term contain parameters. By replacing the parametric price index with a Stone price index and by making the coefficient on the quadratic expenditure term price-independent, the system becomes linear in parameters. We will replace the parametric price index $\ln P_t(p)$ with the Stone price index given by

$$\ln P^*(p_t) = \sum_{i=1}^{N} \overline{w}_{it} \ln p_{it}, \quad t = 1, ..., T$$

where $\overline{w}_{it}$ is the mean of the budget shares across households,

$$\overline{w}_{it} = \frac{1}{H}\sum_{h=1}^{H} w_{iht}.$$

[3]Letting q_i denote the quantity of commodity i, the income elasticity for commodity i is by definition $e_i = \frac{x}{q_i}\left(\frac{\partial q_i}{\partial x}\right)$, and since $\frac{\partial w_i}{\partial \ln x} = \frac{\partial(p_i q_i/x)}{(\partial x)/x} = \frac{\partial(p_i q_i/x)x}{(\partial x)} = \frac{(p_i(\partial q_i/\partial x)x - p_i q_i)x}{x^2} = p_i\left(\frac{\partial q_i}{\partial x}\right) - \frac{p_i q_i}{x}$, then $\frac{1}{w_i}\frac{\partial w_i}{\partial \ln x} = \left(\frac{x}{p_i q_i}\right)\left(p_i\left(\frac{\partial q_i}{\partial x}\right) - \frac{p_i q_i}{x}\right) = \frac{x}{q_i}\left(\frac{\partial q_i}{\partial x}\right) - 1$, i.e. $e_i = \frac{1}{w_i}\frac{\partial w_i}{\partial \ln x} + 1$.

Thus, we will use the following specification of budget shares:

$$w_{iht} = \alpha_i + \sum_{j=1}^{N} \widetilde{\gamma}_{ij} \ln p_{jt} + \beta_i \left[\ln x_{ht} - \ln P^*(p_t)\right] + \gamma_i \left[\ln x_{ht} - \ln P^*(p_t)\right]^2 , \qquad (2)$$

$i = 1, ..., N$, $h = 1, ..., H$, $t = 1, ..., T_h$. By linearizing the demand system in this way we lose two features as compared to the original QUAID system. One feature is integrability: The system in (2) is not integrable. We will return to this point later in this section. The other potential problem concerning this linearisation is that it could introduce bias in estimates of price elasticities. Pashardes (1993) shows that estimating an AID system (which is given as (1) without the quadratic expenditure term) with the original price index replaced by a Stone price index may introduce bias in estimates of the price effects. However, this is not a problem for our analysis, since we are not interested in the price effects; our focus is on the income effects.

We introduce unobserved heterogeneity (taste heterogeneity) into the model as an additive commodity - and household specific term. Let η_{ih} denote this commodity - and household specific term for commodity i, household h, $i = 1, ..., N$, $h = 1, ..., H$. Then $\eta_h = (\eta_{1h}, ..., \eta_{Nh})$ is the vector of "taste parameters" for household h, $h = 1, ..., H$ and we have

$$w_{iht} = \alpha_i + \beta_i \left[\ln x_{ht} - \ln P^*(p_t)\right] + \gamma_i \left[\ln x_{ht} - \ln P^*(p_t)\right]^2 + \eta_{ih}.$$

$\eta = (\eta_1, ..., \eta_H)$ thus represents unobservable heterogeneity in consumers' taste. We think of η as being a random variable. The value of η_h is unknown to the researcher, but is known to household h for each $h = 1, ..., H$.

As mentioned earlier, the imposition of linear Engel curves (as is the case when using the AID system as in Calvet and Comon (2003) or the Linear Expenditure System as in Aasness, Bioern and Skjerpen (1993)) can be problematic when testing for correlated heterogeneity. This is illustrated in Figure 1 (see Appendix): Suppose we have two households, a richer one (with high log total expenditure) and a poorer one (with low log total expenditure). Suppose that the data is like the points in the figure, that is, that the richer household has great variability in total expenditure and the poorer one does not. It is clear that in this case the true specificaion is quadratic in log total expenditure. Think for a moment about testing for correlated heterogeneity by a Hausman test, that is, basically testing whether usual pooled OLS is significantly different from the within groups estimator [4]. The pooled OLS estimator weighs all data points equally and therefore results in a prediction like the one shown in the figure. The within groups estimator uses differences from the mean in log total expenditure within households and thus gives more weight to households with high variation in total expendtiures. Therefore the within groups estimator results in a prediction like the one shown in the figure. Now, as a result of the linear specification of the Engel curves, OLS and Within groups seem very different, when in fact, the only reason they are different is misspecification. This is the main reason that we use the QUAID system as our base demand system.

[4] To be precise, in order to employ a Hausman test for correlated heterogeneity, one needs to estimate the model under the null of no correlated heterogeneity efficiently, i.e. by a random effects estimator. But since pooled OLS is consistent, though not efficient under the null, it suffices for our illustration here, since we only look at the estimators.

We are now in a position to return to the subject of integrability. As mentioned, the QUAID system given in (1) is integrable, and the linearized QUAID system in (2) is not. One could be under the misperception that if we made our demand system nonlinear in parameters in the same way as (1) (i.e. if we had added $\eta_{ih} + \varepsilon_{iht}$ to (1) instead of to (2)) then our system would be integrable. The point we want to make here is that this is not enough to ensure integrability. What we have is that individual demands are given as

$$\alpha_i + \beta_i \left[\ln x_{ht} - \ln P^*(p_t)\right] + \delta_i \left[\ln x_{ht} - \ln P^*(p_t)\right]^2 + \eta_{ih}$$

and it is known from BW and Lewbel that in order to make a demand system with additive error terms (additive preference heterogeneity) integrable it is necessary that the error terms are functionally dependent on prices and/or incomes, where functionally dependent means that the partial derivative of the unobservable term with respect to prices and/or income must be nonzero. Another way to express this necessity condition is that a minimum requirement for obtaining a demand system with preference heterogeneity which is consistent with consumer theory is that there must be unobservable heterogeneity in the marginal effects. This condition means that simply adding a time invariant individual-specific term representing preference heterogeneity to an integrable system does not again result in an integrable system. How to formulate and identify a demand system with preference heterogeneity, where the preference heterogeneity is functionally dependent on prices and/or on total expenditure - i.e. such that preference heterogeneity can enter not just as a level effect, but also in the marginal effects - is still very much an open research question. For one example of a demand system with lots of preference heterogeneity in the marginal effects, see Christensen (2005). Again, of course, that demand system is identified on panel data, but not on cross sectional data.

To conclude this section, we compare our specification of heterogeneity to that of Calvet & Comon (2003). The reason we focus on Calvet and Comon (2003) is that firstly, they have formulated a model with preference heterogeneity which is identified from repeated cross sections, and secondly, that their empirical results are so dramamtic. Calvet & Comon (2003) introduce preference heterogeneity into the budget share equations by applying Grandmont's homothetic α-transformation to the AID system (see Grandmont (1992), Grandmont (1987) and Calvet and Comon (2003)). The AID system is given by

$$w_{iht} = \alpha_i + \sum_{j=1}^{N} \widetilde{\gamma}_{ij} \ln p_{jt} + \beta_i \left[\ln x_{ht} - \ln P(p_t)\right], \tag{3}$$

$i = 1, ..., N$, $h = 1, ..., H$ and $t = 1, ..., T_h$, where the price index is given by as before

$$\ln P(p_t) = \alpha_0 + \sum_{j=1}^{N} \alpha_j \ln p_{jt} + \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{N} \widetilde{\gamma}_{jk} \ln p_{jt} \ln p_{kt}, \qquad t = 1, ..., T.$$

Grandmont's homothetic α-transformation of the budget share functions is given by

$$w_h(p, x) = w(p, e^{\tau_h} x)$$

for $\tau_h \in \mathbb{R}$. This yields budget share equations

$$\begin{aligned} w_{ih} &= \alpha_i + \beta_i \ln x_h + \sum_{j=1}^{N} \gamma_{ij} \ln p_j \\ &\quad -\beta_i \left\{ \alpha_0 + \sum_{k=1}^{N} \alpha_k \ln p_k + \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{N} \gamma_{kj} \ln p_k \ln p_j \right\} + \beta_i \tau_h . \end{aligned}$$

Preference heterogeneity (unobserved individual taste-heterogeneity) is thus represented by the term $\beta_i \tau_h$. The heterogeneity scheme in this model is related straightforward to the heterogeneity scheme in our model: $\eta_{ih} = \beta_i \tau_h$, $i = 1, ..., N$, $h = 1, ...H$. To establish identification of this model, first sum the budget share equations over households for each commodity i. Because the heterogeneity term τ_h has unconditonal mean zero, the term $\beta_i \tau_h$ then cancels out, and what is left is the AID system which is then identified off the time series of the repeated cross sections. From this, the β_i's can thus be identified. Afterwards, sum the budget share equations over commodities for each h - because the β_i's are identical across consumers for each i and because they sum to one across commodities, τ_h is then identifed. The heterogeneity scheme specified in Calvet & Comon (2003) can be interpreted in the following way: If $\beta_i > 0$ then commodity i is a luxury, so $\tau_h > 0$ implies that household h likes all luxuries more than the average household and likes all necessities less than the average household. This heterogeneity scheme thus allows only two different kinds of taste: Either a household likes all luxuries more than the average household ($\beta_i > 0$), or a household likes all luxuries less than the average household ($\beta_i < 0$). Thus it restricts taste to have a common linear structure across commodities where the linear structure is given by the directions of the income effects for the commodities. A serious drawback of this heterogeneity scheme is that for each household the random component capturing preference heterogeneity is only one-dimensional: Beckert (2003) showsh that in order for a random demand model to generate a non-degenerate distribution of demands, given prices and total expenditure, it is necessary that the random variable capturing preference heterogeneity has at least the same dimension as the number of commodities [5].

3 Data

We select a sample of the ECPF consisting of the households where the husband is all the time employed as a wage earner with a permanent job and the wife is out of the labor force. For a description of the ECPF, see Chapter 2 in Christense (2005). We select the full-time employed wage earners with permanent jobs because by only modelling the demand (and not modelling the labor supply) we have implicitly assumed separability between the consumption of goods and the labor supply, and there seems to be empirical evidence against this separability assumption: The empirical findings of Browning & Meghir (1991) show rejections in the FES. By selecting out the unemployed, the part-time employed and the employed with temporary jobs, we increase the probability that none of the husbands in our sample are making labor

[5] Lemma 1 in Beckert (2003).

market descisions during the sample period. Obviously, by only selecting the households where the wife is out of the labor force, we ensure that the wives are not making any labor market descisions during the sample period [6]. This selection makes the assumption of separability between consumption and labor supply more plausible.

We expect the sample selection bias arising from this selection to be in the direction of less unobserved heterogeneity, since we have selected a sample which is more homogenous than the original data set. This will make a finding of correlated heterogeneity stronger: If we find evidence of correlated heterogeneity in this sample, we would expect to find even stronger evidence in a more heterogeneous sample.

We do two more selections. We select out the households that report a zero budget share for the commodity group food eaten at home (14 households). We also select out the households that report earned household income to be zero (37 households). This leaves us with a sample of 3083 households and 21.513 observations.

In our sample we have information on the occupational status and education level of the husband, housing tenure, the number of people in different age groups living in the household, the age of the husband and the wife; we have expenditure information on seventeen different non-durable commodity groups and income information on several income categories. We will only use the income category "Earned household income", which in our sample is the sum of the husband's earnings and the potential earnings of cohabiting adults (grown up children living at home and contributing to household income). The reason for using this income category is that the expenditure information is at the household level, so it seems reasonable also to consider income at the household level.

In the following we provide some summary statistics for our selected sample. Firstly, the distribution of households according to number of quarters they stay in the survey:

Number of interviews	5	6	7	8	Total
Number of households	574	482	465	1562	3083
Percent	18,62	15,63	15,08	50,66	100

These percentages are very similar to the analogue percentages for the full data set; thus it does not seem to be the case that the households we selected stay longer or shorter in the survey than is the case for the full data set.

We construct a dummy variable for the husband's education level, which takes the value one if the husband has a high education (where we define high education to be secondary school or a university degree) and zero otherwise. We also construct a dummy for being a home owner, i.e. the dummy for housing tenure takes the value one if the household is a home owner and zero otherwise. The following table provides summary statistics for the husband's education level, the husband's occupational status and the household's housing tenure:

[6] The vast majority of the wives in the ECPf work as housewives and are thus out of the labor force, see Chapter 2 in Christensen (2005).

Husband's education level	**Percent**	**Frequency**
High (secondary school or university degree)	36,29	7806
Low	63,71	13.707
Total	100	21.513
Husband's occupational status		
Workers with university degree	9,77	2.101
Specialized workers without university degree	75,03	16.142
Unspecialized workers	15,20	3.270
Total	100	21.513
Housing tenure		
Home owner	78,13	16.808
Renter or other	21,87	4705
Total	100	21.513

Comparing these numbers to the numbers in the full data set, we find that the percentage of home owners is roughly the same: 78,13 percent are home owners in our sample and 80,87 percent are home owners in the full data set. With respect to the husband's education, there is a difference in that only 25,07 percent in the full data set have a high education level whereas 36,29 percent have high education in our sample. This could reflect two things. One explanation is that we have selected out the unemployed, and since there may be higher unemployment rates amongst the ones with lower education levels this could explain the difference. Another explanation could be that since we have also selected out the retired, we have selected out the oldest men, which have lower education levels than the younger men. The next table shows sample means and standard deviations for the household composition:

	Mean	Std.dev.
Husband's age	44,01	9,48
Wife's age	41,92	9,62
Number of children	1,44	1,14
Number of adults	2,80	1,03
Total household size	4,23	1,24

Again comparing with full data set we get that the couples in our sample are on average younger than the couples in the full data set (average age is 51 for men in the full data set and 48 for women) and that household size in our sample on average is larger due to the number of children being larger (average number of children in the household is 1,05 in the full data set). But these differences were to be expected, since one of the groups we have selected out is the group of retired households (which are older, and for which the number of children in the household is lower). In summary, the only real difference our sample and the full data set is that the husband's education level in our sample is higher than in the full data set [7].

Next, we turn to the commodities we want to examine. We define ten commodity groups from the non-durable commodity groups in the ECPF (see Chapter 2 for the full list of commodity groups recorded in the ECPF):

Commodity	**Description**
Food at home	Food and non-alcoholic drinks eaten at home
Foodout	Food eaten in restaurants, cafeterias and bars
Alcohol&Tobacco	Alcoholic drinks and tobacco
Clothing	Clothing and footwear
Transportation	Transportation (excluding petrol)
Energy	Energy at home (heating by electricity) and petrol
Services	Services at home (water, furniture repair), Non-durables at home (cleaning products), Personal services and Personal non-durables
Medication	Non-durable medicines and medical services
Leisure	Cinema, theatre, clubs for sports
Education	Education

Sample averages and sample standard deviations of the budget shares are presented below:

[7] We also tabulated or summarized the demographic variables by the number of periods in the survey does not reveal that any particular group should be more likely stay, say, longer in the survey than other groups. For example, one could have expected that households with few children were more likely to complete the full 8 periods in the survey, but this does not seem to be the case.

Commodity	**Mean**	**Std.Dev**
Food at home	.3782	.1394
Foodout	.1107	.0977
Alcohol&Tobacco	.0351	.0373
Clothing	.1360	.1110
Transportation	.0830	.0878
Energy	.1040	.0668
Services	.0577	.0526
Medication	.0287	.0590
Leisure	.0305	.0386
Education	.0360	.0590
Total	1.000	

A histogram of log deflated total expenditure for this selection of goods is shown in Appendix. As expected, the distribution of log total expenditure is close to a normal. Since we will use within-household differences in log deflated total expenditure to identify the income effect, the variation in log deflated total expenditure within the household is important for our analysis. We therefore examine this variation more closely in the following. We calculate the standard deviation as well as the mean of log deflated total expenditure within each household. The distribution of the ratio between the two, the coefficient of variation, is presented in the table below and a histogram of the values of this ratio is presented in Appendix (both the table and the graph are based on 3083 numbers, one for each household):

Percentile	5	25	50	Mean	75	95
Value of the coefficient of variation	.0093	.0147	.0194	.0210	.0255	.0370

Thus, there is on average a two percent variation in log deflated total expenditure within households in this sample, which seems adequate. Next, we examine if the variation in log deflated total expenditure within households correlates with the levels of log deflated total expenditure within households. To this end, we regress the within-household variance in log deflated total expenditure on the within-household mean in log deflated total expenditure. We also regress the within-household intertquartile range in log deflated total expenditure on the within-household median in log deflated total expenditure in order to control for outliers. The estimated coefficients, with t-statistics in parantheses, are reported below and graphs of the predicted regressions are presented in Appendix:

	Estimated coefficient and t-statistics
Regression of variance on mean	$-.0203$ (-5.70)
Regression of inter-quartile range on median	$-.0275$ (-3.63)

This shows that there is a small negative correlation between the within-household level of log deflated total expenditure and the within-household variation of log deflated total expenditure: The higher the level, the less is the variation. However, in the graphs of the predictions we also depicted the data points, and the correlation does not seem to be crucial [8].

A histogram of log deflated earned household income is also presented in Appendix. Also the distribution of log earned income is close to a normal, though not as close as log total expenditure is.

Finally, in order to compare the overall levels of total expenditure and household earnings in our sample with the full data set, we report sample means (standard deviations in parantheses) together with the corresponding figures for the full data set:

	Our sample	The full data set
Total non-durable expenditure	453.981 (255.734)	379.634 (234.837)
Household earned income	480.058 (258.972)	309.345 (337.385)

This shows that our selected sample on average spend more than what is on average spent in the full data set and that the households in our selected sample earn a higher income (though, the large difference between the income figures for the two sample is somewhat misleading, since the figure for the full data set is a sample mean and is thus based on all observations in the full data set, including the retired, unemployed and self-employed which only contribute with zero earned income). This was also what could be expected.

4 Empirical Analysis

Let w_{iht} denote the budget share for commodity i for household h at time t, let x_{ht} denote total expenditure for household h at time t, let p_{it} denote the price of commodity i at time t and let p_t denote the vector of all prices at time t, let W_{ht}^1 denote the vector of demographics that vary across households and within time, let W_t^2 denote the vector of time dummies and let u_{iht} denote the unobservable term. The

[8] Moreover, the R^2 of the regression of the variance on the mean is 0.01 and the R^2 of the regression of the interquartile range on the median is 0.0043, so not much of the variance (interquartile range) is explained by the mean (median).

empirical specification is then given by

$$w_{iht} = \alpha_i + \beta_i \left[\ln x_{ht} - \ln P^*(p_t)\right] + \gamma_i \left[\ln x_{ht} - \ln P^*(p_t)\right]^2 + \delta'_{1i} W^1_{ht} + \delta'_{2i} W^2_t + u_{iht},$$

$i = 1, ..., N-1,\ h = 1, ..., H,\ t = 1, ..., T_h$, where $\ln P^*$ is the Stone price index given by

$$\ln P^*(p_t) = \sum_{i=1}^{N} \overline{w}_{it} \ln p_{it}, \quad t = 1985, ..., 1997$$

and where

$$u_{iht} = \eta_{ih} + \varepsilon_{iht},$$

where η_{ih} denotes the time-invariant part of the unobservables and ε_{iht} denotes the idiosyncratic error term. Letting $\eta = (\eta_1, ..., \eta_H)$, $\varepsilon_{ih} = (\varepsilon_{ih1}, ..., \varepsilon_{ihT_h})$, $h = 1, ..., H$, $i = 1, ..., N-1$ and $\varepsilon_i = (\varepsilon_{i1}, ..., \varepsilon_{iH})$ and $T = \sum_{h=1}^{H} T_h$:

$$\begin{aligned} E[\eta] &= 0 \\ E[\varepsilon_i \varepsilon'_i] &= \sigma_{\varepsilon_i} I_T, \quad i = 1, ..., N-1 \\ E[\eta_{ih} \left| (\ln x_{ht}, \ln x^2_{ht}) \right.] &= \mu \neq 0, \quad E[\eta_{ih} \left| (W^1_{ht}, W^2_{ht})' \right.] = 0 \end{aligned}$$

where we allow for contemporaneous measurement error in total expenditures:

$$\begin{aligned} E[\varepsilon_{iht} \left| (\ln x_{hs}, \ln x^2_{hs}) \right.] &\neq 0, \quad t = s, \\ E[\varepsilon_{iht} \left| (\ln x_{hs}, \ln x^2_{hs}) \right.] &= 0, \quad t \neq s, \end{aligned}$$

$h = 1, ..., H,\ i = 1, ..., N-1.$

The demographic variables in W^1_{ht} are household size, the dummy for the husband's education level, the husband's occupational status, the dummy for housing tenure as well as time dummies for which "interview-week" the household is interviewed in. Because the survey design of the ECPF is such that the interviews of the households are spread out across the year such that a given household is always interview in the same week in each quarter (for example, if a household is interviewed in the second week of the first quarter, it is also interviewed in the second week of the other quarters), we have this "interview-week" variable. In order to control for this systematic variation, we include these interview-week dummies [9]. Since there are 32 interview-weeks in each year, $W^1 = (W^1_{ht})^{H,T}_{h,t=}$, thus consists of 36 demographic variables. $W^2 = (W^2_t)^T_{t=1}$ is the vector of time dummies that are identical for all households. We have chosen the most general specification of seasonality and macroeconomics effects and thus W^2

[9] The "interview-week-dummies" are always significant. In earlier specifications we also included age of the husband and age of the wife in various functional forms, but these terms were never significant, so we finally left them out.

includes a dummy for quarter q, $q = 1, ..., 4$, in year s, $s = 1985, 1986, ..., 1997$. Because we have 13 years of data with only one quarter of data in the last year, we have a total of 49 quarter-and-year dummies. Because the prices recorded in the data set are quarterly price indeces and are identical for all households, the quarter-and-year dummies are colinear with the prices, so we leave out the prices. Since we are not interested in estimating price effects, we thought it best to make the most general structure to control for seasonal variation and macroeconomic effects. The last thing is especially important, since the ECPF is a rotating panel. This means that one household may be in the survey for, say, the years 1985 and 1986 while another household is in the survey for 1995 and the first half of 1996. Obviously, these two households will be exposed to very different macroeconmic environments [10], which we control for by including the quarter-and-year dummies. Also, the quarter-and-year dummies control for seasonal variation, that is, that expenditure patterns vary across seasons.

In this section we first present our test for correlated heterogeneity. Secondly, we estimate the model both by the usual estimation method employed on cross sectional data as well as by an instrumental variables estimation and graph the marginal effects in both cases. We test for whether there is a significant income effect: For the commodities where we find evidence of correlated heterogeneity, we perform the test in the instrumental variables specification and for the commodities where we find no evidence of correlated heterogeneity we perform the test in the usual cross sectional specification. The left out good is Education.

4.1 Empirical Strategy: Testing

The first test that comes to mind when speaking about testing for correlated unobservable heterogeneity is a Hausman test for random effects versus fixed effects. However, the Hausman test is a test for whether *all* the regressors are simultaneously correlated with the household-specific term. This means that in our case we would be testing whether total expenditure, demographics and time dummies (and prices) are jointly correlated with the household-specific term, which is not quite what we are interested in. One could also argue that potential correlations between prices and tastes could exist and should be taken account of (for example, some people may have a high taste for sales and other people not), but that is not the focus of this paper. But even though we would find evidence of correlated heterogeneity, it would not specify what was the cause of this correlation; it could be that the correlation was driven entirely by, say, correlation between household size and the household-specific term. This problem could be overcome by employing the test described in Wooldridge (2002) pp. 290-291: If one is interested in testing whether only one regressor is correlated with the household-specific term, one can employ a t-statistic version of the Hausman test, using the estimated parameters under fixed effects and random effects, respectively, of this regressor together with appropriate standard errors from estimated variance-covariance matrix resulting from the fixed effects and random effects estimations. If one is interested in testing for correlation with more than one regressor, one can use an F statistic version of the Hausman test (this statistic turns out to be equivalent to the one suggested in Mundlak (1978) which consists

[10] In the years 1992-1994 Spain suffered a recession, while in the years 1985-1991 the Spanish economy was booming (M. D. Collado (1998) p. 229).

in including the within-means of the regressors in question). However, this test does not, as it stands in Wooldridge (2002), allow for measurement error in total expenditure [11]. Another drawback of the Hausman test is that the estimator under the alternative uses within-household differences in the data (most commonly the Hausman test uses the within groups estimator under the alternative, that is, one of the estimators in the test is calculated on the basis of within-household differences from the within-household mean). This means that doing a Hausman test would employ within-household differences of all variables, also demographics and time dummies. Because there is only little variation within households for most demographics this would entail losing a lot of information (quite simply, for many households the contribution of demographic variables to the within groups estimator would be a number of zeros) [12].

For these reasons, we choose instead to view the problem as an endogeneity problem. Because we have a panel data set, we have potential instruments in the case where endogeneity of total expenditure is due to time-invariant unobservable heterogeneity: Differences in total expenditure and in income will be uncorrelated with η. Hence we take an instrumental variables approach and carry out our test for correlated heterogeneity between tastes and total expenditure as a classic test of endogeneity of explanatory variables.

We will use an endogeneity test that utilizes the augmented regression (the reduced form) [13]. To illustrate the idea, suppose for simplicity that the equation of interest is (the structural model)

$$y_{ht} = X_{ht}\beta + u_{ht}, \qquad h = 1, ..., H, \; t = 1, ..., T,$$

We want to test if u is correlated with X and if so, we want to estimate β in a way that takes this correlation into account. Suppose we have an instrument Z for X, then a way to test for correlation is to do a residual augmented regression: First regress X on its instruments:

$$X_{ht} = \Pi Z_{ht} + v_{ht} \qquad h = 1, ..., H, \; t = 1, ..., T$$

and calculate the residuals from this regression:

$$\widehat{e}_{ht} = X_{ht} - \widehat{\Pi} Z_{ht}.$$

Second, regress y on X, including in addition the residuals $\widehat{e}_{ht}$ from the first stage. The t-test of the coefficient on the residuals being zero is then a test of exogeneity of X : The coefficient on the residuals should be insignificant if X is exogenous, i.e. if u and X are uncorrelated.

4.2 Choice of Instruments

An instrumental variable is a variable which is uncorrelated with the error term in the equation of interest (*validity*) and correlated with the endogenous variable in the equation of interest (*relevance*).I

[11] And it was not obvious how to adjust it to allow for measurement error.

[12] This also means that even if we could adjust the F statistic version of the Hausman test to allow for measurement error, we would still face this problem.

[13] See Wooldridge (2002) p. 118-121 and Browning (2002).

4.2.1 Potential Instruments

The panel data provides us with time series observations of total expenditures and incomes for each household. This provides us with potential instruments for the potential correlation between total expenditure and the household-specific term, when this correlation is assumed to be constant over time. Under the assumption that the correlation is time-invariant and that its variance is also constant over time, differences in log total expenditure and in log income are uncorrelated with the household-specific term. To show this, suppose that

$$\text{Cov}\left(\ln x_{ht}, \eta_h\right) = \rho \neq 0.$$

Let $\Delta \ln x_{ht} = \ln x_{ht} - \ln x_{ht-\tau}$, $\tau \in \{1, ...T\}$. Then

$$\begin{aligned}
\text{Cov}\left(\Delta \ln x_{ht}, \eta_h\right) &= E(\Delta \ln x_{ht} \eta_h) - E(\Delta \ln x_{ht}) E(\eta_h) \\
&= E((\ln x_{ht} - \ln x_{ht-\tau})\eta_h) - E(\ln x_{ht} - \ln x_{ht-\tau}) E(\eta_h) \\
&= E(\ln x_{ht}\eta_h) - E(\ln x_{ht-\tau}\eta_h) - E(\ln x_{ht})E(\eta_h) + E(\ln x_{ht-\tau})E(\eta_h) \\
&= Cov(\ln x_{ht}\eta_h) - Cov(\ln x_{ht-\tau}\eta_h) \\
&= \rho - \rho \\
&= 0,
\end{aligned}$$

i.e. the difference (for any time period) in log total expenditure is uncorrelated with the household-specific term. Applying the same logic yields that lagged differences in log total expenditure (i.e. $\Delta \ln x_{ht-1} = \ln x_{ht-1} - \ln x_{ht-\tau-1}$) and differences in log income are both uncorrelated with the entire error term $u_{ht} = \eta_h + \varepsilon_{ht}$ (assuming that the measurement error in the ε's is uncorrelated over time within a household).

We consider the following sets of potential instruments:

	Description of instrument set	Definition of instrument set
Z_1	Yearly differences of log total expenditure and log income	$(\ln x_t - \ln x_{t-4},\ \ln x_t^2 - \ln x_{t-4}^2,$ $\ln x_t^3 - \ln x_{t-4}^3,$ $\ln y_t - \ln y_{t-4},\ \ln y_t^2 - \ln y_{t-4}^2,$ $\ln y_t^3 - \ln y_{t-4}^3)$
Z_2	Lagged quarterly differences of log total expenditure and quarterly differences of log income, where lag is one quarter	$(\ln x_{t-1} - \ln x_{t-2},\ \ln x_{t-1}^2 - \ln x_{t-2}^2,$ $\ln x_{t-1}^3 - \ln x_{t-2}^3,$ $\ln y_t - \ln y_{t-1},\ \ln y_t^2 - \ln y_{t-1}^2,$ $\ln y_t^3 - \ln y_{t-1}^3)$
Z_3	Lagged yearly differences of log total expenditure and yearly differences of log income, where lag is one quarter	$(\ln x_{t-1} - \ln x_{t-5},\ \ln x_{t-1}^2 - \ln x_{t-5}^2,$ $\ln x_{t-1}^3 - \ln x_{t-5}^3,$ $\ln y_t - \ln y_{t-4},\ \ln y_t^2 - \ln y_{t-4}^2,$ $\ln y_t^3 - \ln y_{t-4}^3)$
Z_4	Lagged yearly differences of log total expenditure, yearly differences of log income, income and lagged log total expenditure, where lag is one quarter	$(\ln x_{t-1} - \ln x_{t-5},\ \ln x_{t-1}^2 - \ln x_{t-5}^2,$ $\ln x_{t-1}^3 - \ln x_{t-5}^3,$ $\ln y_t - \ln y_{t-4},\ \ln y_t^2 - \ln y_{t-4}^2,$ $\ln y_t^3 - \ln y_{t-4}^3,$ $\ln y_t,\ \ln y_t^2\ \ln y_t^3,$ $\ln x_{t-1},\ \ln x_{t-1}^2, \ln x_{t-1}^3)$
Z_5	Lagged log total expenditure and log income, where lag is one quarter	$(\ln x_{t-1},\ \ln x_{t-1}^2, \ln x_{t-1}^3$ $\ln y_t,\ \ln y_t^2\ \ln y_t^3)$

The set Z_1 is a potential set of instruments in case there is no measurement error, but there is correlated heterogeneity. The sets Z_2 and Z_3 are potential sets of instruments in case there is both correlated heterogeneity and measurement error. The sets Z_4 and Z_5 are potential sets of instruments in case there is no correlated heterogeneity, but there is measurement error. Whether any of these sets of instrumental variables are valid is an empirical matter. We therefore now turn to empirically testing validity.

4.2.2 Validity of Instruments

When the number of instruments exceeds the number of endogenous variables, one can employ a Sargan test for overidentifying restrictions to test for validity of the "extra" instruments. The Sargan test is a test for whether the overidentifying restrictions implied by the extra moment restrictions hold; i.e. it takes the moment restrictions needed in order to identify the model hold as given and then test whether the "extra" (i.e. the overidentifying) restrictions hold [14].

We carry out a Sargan test for each good separately, so in the following we suppress the commodity index. The model is

$$\begin{aligned} w_{ht} &= X_{ht}\beta + \upsilon_{ht}, \\ \upsilon_{ht} &= \eta_h + \varepsilon_{ht}, \qquad t = 1, ..., T_h, \quad h = 1, ..., H, \end{aligned}$$

where we specify the error term structure by (i.e. allowing for heteroskedasticity)

$$\begin{aligned} E\left[\upsilon_{ht}\upsilon_{hs}\right] &= \begin{cases} \sigma_{h\eta}^2 + \sigma_{h\varepsilon}^2, & t = s \\ \sigma_{h\eta}^2, & t \neq s \end{cases} \\ E\left[\upsilon_{ht}\upsilon_{ks}\right] &= 0, \qquad h \neq k, \text{ for all } t, s. \end{aligned}$$

Define $\upsilon_h = (\upsilon_{h1}, ..., \upsilon_{hT_h})'$ and $\upsilon = (\upsilon_1, ..., \upsilon_H)'$. Then

$$E\left[\upsilon_h \upsilon_h'\right] = \begin{bmatrix} \sigma_{h\eta}^2 + \sigma_{h\varepsilon}^2 & \sigma_{h\eta}^2 & \cdots & \sigma_{h\eta}^2 \\ & \cdot & & \cdot \\ & & & \cdot \\ & & \cdot & \sigma_{h\eta}^2 \\ & & & \sigma_{h\eta}^2 + \sigma_{h\varepsilon}^2 \end{bmatrix} \equiv \Omega_h.$$

$E\left[\upsilon\upsilon'\right]$ is then a block-diagonal matrix of dimensions $(\sum_h T_h)$ by $(\sum_h T_h)$ with $\Omega_1, ..., \Omega_H$ in the diagonal.

Let $Z = (Z_1, ..., Z_H)'$ denote the matrix of instrumental variables,

$$Z_h = \begin{bmatrix} z_{h1}^1 & .. & .. & z_{h1}^L \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ z_{hT_h}^1 & .. & .. & z_{hT_h}^L \end{bmatrix},$$

where L is the number of instruments plus the number of exogenous variables (i.e. the exogenous variables are included in Z). The moment conditions implied by validity of the instruments Z are

$$E\left[Z'(\eta + \varepsilon)\right] = 0.$$

The Sargan test statistic is given by

$$T_{\text{Sargan}} = \left[Z'\left(w - X\widehat{\beta}_{\text{GMM}}\right)\right]' \left[Z'\widehat{\Omega}Z\right]^{-1} \left[Z'\left(w - X\widehat{\beta}_{\text{GMM}}\right)\right] \sim_{as} \chi^2_{(L-K)},$$

[14] Thus, the Sargan test is not a test for whether all the instruments are valid, but rather a test for whether the "extra" instruments are valid. Note that the test gives no information about which of the instruments are valid.

where

$$\widehat{\beta}_{\text{GMM}} = \left[X'Z\left(Z'\widehat{\Omega}Z\right)^{-1}Z'X\right]^{-1}\left[X'Z\left(Z'\widehat{\Omega}Z\right)^{-1}Z'w\right]$$

and $\widehat{\Omega}$ is a White estimate of Ω. We thus calculate the Sargan test statistic in two steps. First, we get a consistent estimate of β by estimating β by Two Stage Least Squares (2SLS):

$$\widehat{\beta}_{\text{2SLS}} = \left[X'Z\left(Z'Z\right)^{-1}Z'X\right]^{-1}\left[X'Z\left(Z'Z\right)^{-1}Z'w\right].$$

From this regression, we compute the residuals $e = w - X\widehat{\beta}_{\text{2SLS}}$ and use them to estimate Ω, as suggested by White (1982). Denoting the residuals for household h by $e_h = (e_{h1}, ..., e_{hT_h})'$, we get

$$\begin{aligned}\widehat{\sigma^2_{h\eta} + \sigma^2_{h\varepsilon}} &= \frac{1}{T_h}\sum_{t=1}^{T_h} e^2_{ht}, \qquad h = 1, ..., H \\ \widehat{\sigma^2_{h\eta}} &= \frac{1}{T_h^2 - T_h}\sum_{\substack{t,s=1 \\ t\neq s}}^{T_h} e_{ht}e_{hs}, \qquad h = 1, ..., H\end{aligned}$$

and

$$\widehat{\Omega} = \begin{bmatrix} \widehat{\Omega}_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \widehat{\Omega}_H \end{bmatrix}$$

With this estimate of Ω, we calculate $\widehat{\beta}_{\text{GMM}}$. Since households are independent, Ω is block-diagonal so we can calculate the weighting matrix $Z'\widehat{\Omega}Z$ as

$$Z'\widehat{\Omega}Z = \sum_{h=1}^{H} Z_h'\widehat{\Omega}_h Z_h.$$

We always have two endogeneous variables ($\ln x$ and $\ln x^2$), so the number of degrees of freedom in the χ^2-distribution of the Sargan test statistic is always the number of instruments less 2. When taking yearly differences rather than quarterly differences and when using lagged values, we lose observations. The table below shows how many instruments and how many observations we have for each set of potential instruments:

Description of instrument set	Number of instruments	Number of households	Number of observations
Z_1	6	3083	9181
Z_2	6	2509	15.347
Z_3	6	2509	6098
Z_4	12	2509	6098
Z_5	6	2509	18.430

The critical value of the χ^2 distribution with 4 degrees of freedom is 9.488 at the 5% level and 13.277 at the 1% level. The critical value of the χ^2 distribution with 10 degrees of freedom is 18.307 at the 5% level and 23.209 at the 1% level. The critical value of the χ^2 distribution with 1 degree of freedom is 3.841 at the 5% level and 6.635 at the 1% level. The estimated Sargan test statistics are shown in the table below (bold numbers are significant on a 5% significance level, starred numbers significant on a 1% significance level):

	Z_1	Z_2	Z_3	Z_4	Z_5
Food at home	**6,622**	485	10,045 (*)	34,664	50,487
Foodout	12,742 (*)	300	**4,803**	**12,258**	**1,332**
Alcohol&Tobacco	**1,474**	234	**2,861**	24,809	21,311
Clothing	27,580	263	**3,150**	**16,932**	10,697 (*)
Transportation	**5,096**	1256	**7,886**	45,168	56,360
Energy	11,866	266	**1,051**	**5,934**	13,402 (*)
Services	**9,740**	160	**3,995**	**7,100**	**7,121**
Medication	13,340	12538	10,425 (*)	**14,295**	**7,596**
Leisure	14,615	469	**6,060**	23,653	22,347
Education	12,814 (*)	12.317	**3,833**	**10,641**	**4,087**

This table shows that the only set of instruments that are valid at a 1% significance level is Z_3, which consists of lagged yearly difference in log total expenditure and yearly differences in log income. The results for the set Z_2 are remarkable: Quarterly differences in log total expenditure and log total income are very far indeed from being valid. The instrument sets that would be valid in case of correlated heterogeneity and no measurement error, or in case of measurement error and no correlated heterogeneity (Z_1, Z_4 and Z_5) are valid for some, but not for all commodities. We therefore continue our analysis using only the instrument set Z_3.

Because Z_3 entails yearly differences and lagged values of yearly differences, where the lag is quarterly, everything that follows is now based on the sample that is observed long enough to provide lagged yearly differences. In other words, only households observed for 6,7 or 8 quarters can be used. This leaves us with a sample of 2509 households and 6098 observations. We thus have between 1 and 3 quarters of observations per household left. More than 60 percent of the households that are left are observed for 3 quarters.

4.2.3 Relevance of Instruments

Since only one set of instruments turned out to be valid for all commodities, namely Z_3, we only use that set of instruments from now on. The reduced form (the agumented regression) is thus

$$\ln x = \pi_1^0 + \Pi_1 Z_3 + \kappa_1 [W^1 \; W^2]' + \nu_1$$

$$\ln x^2 = \pi_2^0 + \Pi_2 Z_3 + \kappa_2 [W^1 \; W^2]' + \nu_2,$$

where Π_j is a six by one vector, $j = 1, 2$. The instruments are indeed relevant in that they are significantly correlated with each of the endogeneous variables:

	F(6,6019) 5%: 2,80	$\chi^2(6)$ (5%: 12.592)
Equation where dependent variable is $\ln x$	42,72	256,32
Equation where dependent variable is $\ln x^2$	8,31	49,86

We therefore carry on with the instrument set Z_3. Whether there is enough independent variation in the variables in Z_3 to identify both the β and γ parameters of the demand equations is not evident from this test of relevance, and thus Z_3 may be a weak instrument in the sense that it only identifies one of the two parameters. But since we perform our endogeneity test in an ordinary linear Two Stage Least Squares estimation, weak instruments will bias the estimates towards the OLS estimates. This means that if we have a weak instrument problem we are more likely not to be able to reject exogeneity [15].

4.3 Empirical Results

First we test for correlated heterogeneity by testing for exogeneity of total expenditure. We follow Banks, Blundell and Lewbel (1997) and include only the residual from the first equation in the reduced form. This is exactly correct when log total expenditure and the error terms in the Engel curves are jointly normal [16]. As mentioned earlier in Section 3, log total expenditure is very close to being normal in our sample. Denoting the residual for household h at time t by $\widehat{e}_{ht}$, the estimating equation for the endogeneity test thus is

$$w_{iht} = \alpha_i + \beta_i \left[\ln x_{ht} - \ln P^*(p_t)\right] + \gamma_i \left[\ln x_{ht} - \ln P^*(p_t)\right]^2 + \delta_{1i}' W_{ht}^1 + \delta_{2i}' W_t^2 + \widehat{e}_{ht} + v_{iht},$$

[15] Stock, Wright and Yogo (2002).
[16] Banks, Blundell and Lewbel (1997) p. 530.

$i = 1, ..., N-1$, $h = 1, ..., H$ and $t = 1, ..., T$ and our endogeneity test is whether the residual is significant. We estimate this equation by ordinary OLS. The results are:

	correlated heterogeneity: t-ratio on the residual	Bias in usual OLS estimates
Food at home	-0.08	no bias
Foodout	-2.14	downwards
Alcohol&Tobacco	2.28	upwards
Clothing	1.72	no bias
Transportation	2.39	upwards
Energy	-2.00	downwards
Services	-1.03	no bias
Medication	-0.01	no bias
Leisure	-1.36	no bias

From this we conclude that there is correlated heterogeneity in the commodities Foodout, Alcohol&Tobacco, Transportation and Energy. For these four commodities, it is thus appropriate to estimate the budget share equation by instrumental variables. For the remaining five commodities we estimate the budget share equation in the usual way as if data was cross sectional; we estimate the budget shares by pooled OLS, allowing for heteroskedasticity across households in these cases [17].

Next we examine whether there is a significant income effect. For Foodout, Alcohol&Tobacco, Transportation and Energy we will test this with the IV estimates and for the remaining budget share equations we test this in the GLS estimation. The table of the test results displays the F-ratio for the test of joint signifcance of $\ln x$ and $\ln x^2$ [18]:

[17] One could argue that we should instrument for measurement error for these five commodities. But the possible instruments in the case of measurement error (and no correlated heterogeneity) are represented by the instrument set Z_5, which is not valid. We therefore chose to use ordinary OLS.

[18] The critical value is is 2.99. Standard errors were clustered by household.

	income effect: F-ratio on $(\ln x, \ln x^2)$
Food at home	228.65
Foodout	12.36
Alcohol&Tobacco	4.14
Clothing	111.38
Transportation	.80
Energy	1.19
Services	5.06
Medication	6.33
Leisure	18.66

Only for the commodities Transportation and Energy do we find no evidence of a significant income effect. That is, only for those two commodities can we conclude that "it is all heterogeneity". For all the remaining commodities there is, even in the cases of the presence of correlated heterogeneity, still a signifcant effect from income.

In order to compare our results with other studies that have based their model on the AID system, and also in order to do some sort of robustness check of our endogeneity test results in the QUAID model, where we may have a weak instruments problem, we also did the exogeneity test and the test for signifcance of income effects in the AID model. The table of results is in Appendix. As can be seen, there is no qualitative difference in the exogeneity test between the QUAID and the AID specifications: In the AID model we find evidence of correlated heterogeneity for exactly the same commodities as with the QUAID specification, and the bias is in the same direction. This makes us feel more confident in our results, especially since the quadratic log expenditure term is only significant for 2 commodities:

	Significance of quadratic expenditure term: t-ratio on $\ln x^2$
Food at home	0.27
Foodout	1.44
Alcohol&Tobacco	-0.32
Clothing	-2.07
Transportation	-1.19
Energy	0.61
Services	-1.71
Medication	1.17
Leisure	-4.82

Finally, we present the estimated income repsonses (the estimates of the coefficients in the total expenditure terms, β_i and γ_i with standard errors in parentheses, standard errors are clustered by household):

	OLS: β	OLS: γ	IV: β	IV: γ
Food at home	$-.1579$ (.1837)	.0019 (.0070)	.5683 (1.5888)	$-.0260$ (.0610)
Foodout	$-.0091$ (.1258)	.0023 (.0049)	-1.5294 (1.1174)	.0619 (.0430)
Alcohol&Tobacco	$-.0307$ (.0522)	.0009 (.0019)	.1755 (.6052)	$-.0075$ (.0231)
Clothing	.2535 (.1093)	$-.0078$ (.0042)	1.1254 (1.2841)	$-.0423$ (.0493)
Transportation	$-.0748$ (.1163)	.0037 (.0045)	1.2330 (1.0424)	$-.0478$ (.0401)
Energy	$-.1714$ (.1035)	.0053 (.0039)	$-.6904$ (1.1070)	.0260 (.0426)
Services	.0958 (.0533)	$-.0035$ (.0020)	.0005 (.4864)	.0004 (.0187)
Medication	$-.0962$ (.0890)	.0040 (.0034)	$-.9265$ (.6401)	.0360 (.0246)
Leisure	.1731 (.0349)	$-.0065$ (.0013)	$-.1807$ (.3504)	.0074 (.0134)

We are now finally able to present the two interesting sets of estimates of income elasticities: The income elasticities estimates wihtout taking correlated heterogeneity into account (i.e. the demand system is estimated by OLS), and the income elasticities estimated taking account of correlated heterogeneity (i.e. the demand system is estimated by IV,using Z_3 as instruments). The resulting income elasticity estimates are presented below:

Commodity group	No heterogeneity (OLS)	Correlated heterogeneity (IV)
Food at home	.7031	.7087
Foodout	1.5571	1.8498
Alcohol & Tobacco	.7030	.2247
Clothing	1.2831	1.2509
Transportation	1.3830	.8625
Energy	.6335	.9574
Services	1.1087	1.2384
Medication	1.2626	1.027
Leisure	1.2253	1.5824

As can be seen from the table, the estimates are smaller in magnitude in the model with no heterogenity (OLS) for all the goods in which we found a downward bias, and reversely for the goods in which we found upward bias. The estimates are quite different in magnitude for exactly the goods in which there is bias. This suggests that policy analysis involving demand for those goods could give misleading results.

5 Conclusions

In this paper we have tested whether tastes are correlated with income, exploiting a unique panel data set on household expenditures. Our results suggest that this is inded the case for some, but not all, of the nondurable commodities usually considered in demand system analysis. Further, our results also suggest that there is, even in the cases of correlated heterogeneity, still a significant income effect for the vast majority of commodities. This means that our study supports what the few other studies of demand systems on panel data have found, namely that there seems to be evidence of correlated heterogeneity, which in turn introduces bias in estimates of income responses, but tastes and income are not completely confounded as is suggested by Calvet and Comon (2003). The identifying assumption of the heterogeneity scheme in Calvet and Comon (2003) could formally be tested with our data with a minimum chi square test.

The condition of unobservables being independent of observables, among then income, is assumed in all studies of nonparametric identification of consumer demand models with preference heterogeneity from cross sectional data. These findings suggest that this assumption is too strong, but also that it may not need to be abandoned for all commodities.

References

[1] Aasness, J., E. Biørn and T. Skjerpen (1993): "Engel Functions, Panel Data and Latent Variables", *Econometrica*, Vol. 61, No. 6, 1395-1422.

[2] Arellano, M., L. P. Hansen and E. Sentana (1999): "Underidentification?", mimeo.

[3] Banks, J., R. Blundell and A. Lewbel (1997): "Quadratic Engel Curves and Consumer Demand", *The Review of Economics and Statistics*, Nov, Issue 4, 527-539.

[4] Beckert, W., (2003): "On Specification and Identification of Stochastic Demand Models", Department of Economics, University of Florida.

[5] Beckert, W., and R. Blundell (2004): "Invertibility of Nonparametric Stochastic Demand Functions", working papaer, Birkbeck Working Papers in Economics & Finance, BWPEF 0406.

[6] Blundell, R., M. Browning and I. Crawford (2003): Nonparametric Engel Curves and Revealed Preference, *Econometrica*, Vol. 71, No. 1, 205-240.

[7] Blundell, R., A. Duncan and C. Meghir (1998): "Estimating Labor Supply Responses Using Tax Reforms", *Econometrica*, Vol. 66, Issue 4, 827-861.

[8] Blundell, R., P. Pashardes and G. Weber (1993): "What do we Learn About Consumer Demand Patterns from Micro Data?", *American Economic Review*, Vol. 83, Issue 3, 570-595.

[9] Brown, D., and R. Matzkin (1998): "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand", working paper, Northwestern University.

[10] Brown, B., and M. Walker (1989): "The Random Utility Hypothesis and Inference in Demand Systems", *Econometrica*, Vol. 57, No. 4, 815-829.

[11] Browning, M. (2002): "Notes on Instrumental Variables Estimation", teaching note for econometrics course, University of Copenhagen.

[12] Browning, M., and M. D. Collado (2001): "The Response of Expenditures to Anticipated Income Changes: Panel Data Estimates", *American Economic Review,* Vol. 91(3), 681-692.

[13] Browning, M., and M. D. Collado (2004): "Habits and Heterogeneity in Demands: A Panel Data Analysis", CAM working paper 2004-18.

[14] Browning, M., and C. Meghir (1991): "The Effects of Male and Female Labor Supply on Commodity Demands", *Econometrica*, Vol. 59, No. 4, 925-951.

[15] Calvet, L., and E. Comon (2000): "Behavioral Heterogeneity and the Income Effect", Harvard Institute of Economic Research Working Paper No. 1892, forthcoming in *Review of Economics and Statistics* (2003)

[16] Carrasco, R., J. M. Labeaga and J. D. Lopez-Salido (2004): "Consumption and Habits: Evidence from Panel Data", *The Economic Journal*, forthcoming.

[17] Deaton, A., and J. Muellbauer (1980a): Economics and Consumer Behavior, *Cambridge University Press*

[18] Deaton, A., and J. Muellbauer (1980b): "An Almost Ideal Demand System", *American Economic Review*, Vol. 70, No.3, 312-326.

[19] Duncan, G. J., F. Gardes, P. Gaubert and C. Strarzec (2002): "Panel and Pseudo-Panel Estimation of Cross-sectional and Time Series Estimates of Food Consumption: The case of American and Polish Data", working paper no. 2002-02, CREST.

[20] Grandmont, J-M. (1987): "Distributions of Preferences and the Law of Demand", *Econometrica*, Vol. 55, No. 1, 155-161.

[21] Grandmont, J-M. (1992): "Transformations of the Commodity Space, Behavioral Heterogeneity, and the Aggregation Problem", *Journal of Economic Theory*, Vol. 57, No. 1, 1-35.

[22] Horowitz, J., (2000): "The Bootstrap", forthcoming in Handbook of Econometrics.

[23] Hsiao, C. (1986): Analysis of Panel Data, *Econometric Society Monographs, Cambridge University Press*

[24] Johnston, J., and J. DiNardo (1997): Econometric Methods, fourth edition, McGraw-Hill.

[25] Labeaga, J. M., and J. Puig (2003): "Demand system estimation with individual heterogeneity: an analysis using panel data on households", mimeo.

[26] Matzkin, R. L., (2003): "Nonparametric Estimation of Nonadditive Random Functions", *Econometrica*, Vol. 71, No. 5, 1339-1375.

[27] Pashardes, P., (1993): "Bias in Estimating the Almost Ideal Demand System with the Stone Price Index Approximation", *The Economic Journal*, Vol. 103, 908-915.

[28] Robin, J-M., and R. J. Smith (2000): "Tests of Rank", *Econometric Theory*, 16, 151-175.

[29] Roehrig, C. S., (1988): "Conditions for Identification in Nonparametric and Parametric Models", *Econometrica*, Vol. 56, Issue 2, 433-447.

[30] Stock, J. H., J. H. Wright and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments", *Journal of Business & Economics Statistics*, Vol. 20, No. 4, 518-529.

[31] Verbeek, M., (1995): "Alternative Transformations to Eliminate Fixed Effects", *Econometric Reviews*, 14(2), 205-211.

[32] Wooldridge, J. M., (2002): Econometric Analysis of Cross Section and Panel Data, The MIT Press.

A Appendix

The results of the tests for correlated heterogeneity and significant income effects in the AID model:

	correlated heterogeneity: t-ratio on the residual	income effect: t-ratio on $\ln x$
Food at home	-0.09	-29.55
Foodout	-2.15	5.61
Alcohol&Tobacco	2.27	-3.66
Clothing	1.75	16.76
Transportation	2.37	-0.60
Energy	-2.03	-1.61
Services	-1.00	3.09
Medication	-0.04	4.90
Leisure	-1.29	4.72