How Revealing is Revealed Preference?*

Richard Blundell†
University College London and Institute for Fiscal Studies

May 2005

Abstract

This paper addresses the following two key criticisms of the empirical application of revealed preference theory: When it does not reject, it doesn't provide precise predictions; and when it does reject, it doesn't help us characterise the nature of irrationality or the degree/direction of changing tastes. Recent developments in the application of RP theory are shown to have rendered these criticisms unfounded. A powerful test of rationality is available that also provides a natural characterisation of changing tastes. Tight bounds on demand responses and on the welfare costs of relative price and tax changes are also available and are shown to work well in practice.

*This review paper is based on the results of joint research with Martin Browning and Ian Crawford. I would also like to acknowledge the input of joint research and discussions with Xiaohong Chen, Dennis Kristensen, Rosa Matzkin, and James Powell. This study is part of the program of research of the ESRC Centre for the Microeconomic Analysis of Public Policy at IFS. The financial support of the ESRC and of the Leverhulme Foundation is gratefully acknowledged. The usual disclaimer applies.

†Department of Economics, University College London, Gower Street, London, WC1E 6BT, and Institute for Fiscal Studies, r.blundell@ucl.ac.uk, http://www.ifs.org.uk

1 Introduction

Measuring the responses of consumers to variation in prices and income is at the centre of applied welfare economics: it is a vital ingredient of tax policy reform analysis and is also key to the measurement of market power in modern empirical industrial economics. Parametric models have dominated applications in this field but, I will argue, this is both unwise and unnecessary. To quote Dan McFadden in his presidential address to the Econometric Society: "[parametric regression] interposes an untidy veil between econometric analysis and the propositions of economic theory". Popular parametric models place strong assumptions on both income and price responses. The objective of the research reported here is to accomplish all that is required from parametric models of consumer behaviour using only nonparametric regression and revealed preference theory. The idea is to fully exploit micro data on consumer expenditures and incomes across a finite set of discrete relative price or tax regimes. This is achieved by combining the theory of revealed preference (RP) with the nonparametric estimation of consumer expansion paths (Engel curves).

Central to the criticism of nonparametric revealed preference theory is that it has no bite. That is, it cannot really discriminate between rational and irrational behaviour. There are two main concerns: When it doesn't reject, it doesn't provide us with precise predictions. When it does reject, it doesn't help us characterise either the nature of irrationality or the degree or direction of changing tastes.

I will argue that recent developments in the application of RP theory have rendered these criticisms unfounded. It is relatively easy to construct a powerful test of rationality for both experimental and observational data. Moreover, we can consider rationality over groups of decisions, over types of individuals and over periods of time. This allows a characterisation of changing tastes. Where we do not reject we can also provide tight bounds on welfare costs of relative price and tax changes as well as tight bounds on demand responses (and elasticities).

At the heart of this analysis are three key aims. First, to provide a powerful test of integrability conditions on individual household data without the need for parametric models of consumer behavior. Second, to provide tight bounds on welfare costs of relative price and tax changes. Third, to provide tight bounds on demand responses (and elasticities) to relative price and tax changes. The overall aim is to lift McFadden's untidy veil!

Historically, parametric specifications in the analysis of consumer behaviour have been based on the Working-Leser or Piglog form of preferences in which budget shares are linear in the log of total expenditure (see Muellbauer 1976). These underlie the popular Almost Ideal and Translog demand models of Deaton and Muellbauer (1980) and Jorgenson, Lau and Stoker (1982). However, more recent empirical studies on individual data have suggested that further nonlinear terms, in particular quadratic logarithmic income terms, provide a much more reliable specification (see, for example, Hausman, Newey, Ichimura and Powell (1995), Lewbel (1991), Blundell, Pashardes and Weber (1993)). This was brought together in the Quadratic Almost Ideal Demand System (QUAIDS) by Banks, Blundell and Lewbel (1997) which provided a fully integrable system consistent with the quadratic logarithmic Engel curve specification and allowing second-order flexibility of relative price responses. Nonetheless, relative price effects remain constrained in an unnatural way across individuals with different incomes. In this paper, this line of research is taken one important step further allowing the fully nonparametric estimation of Engel curves and, by using revealed preference restrictions alone at each point in the income distribution, also allowing price responses to be quite unrestricted across individuals with different incomes.

With a relatively small number of price regimes: across locations or points in time or both, nonparametric revealed preference theory provides a natural setting for the study observed behaviour. The attraction of RP theory is that it allows an assessment of the empirical validity of the usual integrability conditions without the need to impose particular functional forms on preferences. Although developed to describe individual demands by Afriat (1973) and Diewert (1973) following the seminal work of Samuelson (1938) and Houthakker (1950), it has usually been applied to aggregate data even though this presents a number of problems[1]. First, on aggregate data, 'outward' movements of the budget line are often large enough, and relative price changes are typically small enough, that budget lines rarely cross (see Varian 1982; and Bronars 1995). This means that aggregate data may lack power to reject revealed preference conditions. Second, if we do reject revealed preference conditions on aggregate data then we have no way of assessing whether this is due to a failure at the micro level or rather to inappropriate aggregation across households that do satisfy the integrability conditions yet have different nonhomothetic preferences. By combining nonparametric statistical methods with a revealed preference analysis of micro data, we can overcome these problems.

The central contribution of the Blundell, Browning and Crawford (2003)

[1] See Manser and McDonald (1988), and the references therein.

study was to develop a method for choosing a sequence of total expenditures that maximise the power of tests of generalised axiom of revealed preference (GARP) with respect to a given preference ordering. They term this the sequential maximum power (SMP) path and present some simulation evidence showing that these GARP tests have considerable power against some key alternatives. From this idea it is possible to develop a method of bounding true cost-of-living indices. In particular, extending the insights in Arrow (1958) and Wald (1939) we may thus obtain the tightest upper and lower bounds for indifference curves passing through any chosen point in the commodity space. Blundell, Browning and Crawford (2004) turn their attention to demand responses and show that these methods can be used to calculate best nonparametric bounds on demand responses and on responses to tax reforms. The tightness of these bounds depends on the closeness of the new prices to the sets of previously observed prices and the restrictions placed on cross-price effects.

In this paper I review these advances in RP application and point to the direction of further research.

2 Data: Observational or Experimental?

2.1 Observational Data

For most interesting problems in consumer economics we must rely on observational data. In particular, we typically have consumer budget surveys in which there is continuous micro data on incomes and expenditures, a finite set of observed price and/or tax regimes, and discrete demographic differences across households. The aim of the empirical RP research reported here is to use this information alone, together with revealed preference theory, to assess consumer rationality and to place tight bounds on behavioural responses and welfare changes.

However, as in much of applied microeconomics, working through the optimal design for an experiment is a useful lead-in to observational design.

2.2 Experimental Data: Is There a Best Design?

Suppose we were running a lab experiment: What would be the best design to test restrictions from RP theory? In order to answer this first consider a sequence of demands that constitutes a rejection of RP. The demands $\mathbf{q}(x_3)$ in the sequence $\mathbf{q}(x_1)$, $\mathbf{q}(x_2)$, $\mathbf{q}(x_3)$ described in Figure 1 display such a rejection. In Figure 1; there are two goods and three price regimes; $\mathbf{q}(x_2)$ is revealed to be at least as good as $\mathbf{q}(x_1)$ and $\mathbf{q}(x_3)$ at least as good as $\mathbf{q}(x_2)$. The transitive closure shows that $\mathbf{q}(x_3)$ is at least as good as $\mathbf{q}(x_1)$. On the other hand, a direct comparison of $\mathbf{q}(x_3)$ and $\mathbf{q}(x_1)$ shows that $\mathbf{q}(x_3)$ is strictly preferred to $\mathbf{q}(x_1)$. Thus we have a rejection. Were the budget line x_3 moved further out, the possible region of rejection would shrink. Moreover, if it were to move in it would be uninformative in deciding upon a RP rejection. Thus, the particular path of budgets for any sequence of prices influences the chance of finding a rejection. So is there an optimal path?

Figure 1 here

To investigate this further, consider the following assumptions on preferences and the resulting proposition.

Assumption 1. For each agent, there exists a set of demand functions $q(p,x)$: $\Re^{J+1}_{++} \to \Re^{J}_{++}$. We define $q_t(x)$ to be an expansion path for given prices in time and location t.

Assumption 2. (Weak normality). If $x > x'$ then $q_t^j(x) \geq q_t^j(x')$ for all j and all p_t.

Now define the sequential maximum power (SMP) path

$$\{\tilde{x}_s, \tilde{x}_t, \tilde{x}_u, ..., \tilde{x}_v, x_w\} = \{\mathbf{p}_s'\mathbf{q}_t(\tilde{x}_t), \mathbf{p}_t'\mathbf{q}_u(\tilde{x}_u), \mathbf{p}_v'\mathbf{q}_w(\tilde{x}_w), x_w\}$$

Proposition 1: (Blundell, Browning and Crawford (2003)). *Suppose that the sequence*

$$\{\mathbf{q}_s(x_s), \mathbf{q}_t(x_t), \mathbf{q}_u(x_u), ..., \mathbf{q}_v(x_v), \mathbf{q}_w(x_w)\}$$

rejects RP. Then the SMP path also rejects RP.

If there is an RP rejection to be found on any budget path along a particular sequence of relative prices, the SMP will find it. This result is great for experimental design, but (as argued above), for the most part we will want to work with individual observational data. In observational data we typically have a given finite sequence of relative price regimes and cannot experimentally vary the budget along that sequence. However, individual data will allow us to estimate local expansion paths: nonparametric Engel curves. If we knew the expansion paths $q_t(x)$, as in Figure 2, then we can improve the RP test as indicated by the proposition. The aim therefore is to develop nonparametric expansion paths that mimic the experimental design.

Figure 2 here

3 What Does the Observational Data Look Like?

The data were drawn from the repeated cross-sections of household-level data in the British Family Expenditure Survey (1974 to 2001). The FES is a random sample of about 7,000 households per year. From this we used a sub sample of all the two-adult households, including those with and without children.[2] For the purposes of this discussion, each year of data is treated as a separate price regime. The commodity groups are non durable expenditures grouped into: beer, wine, spirits, tobacco, meat, dairy, vegetables, bread, other foods, food consumed outside the home, electricity, gas, adult clothing, children's clothing and footwear, household services, personal goods and services, leisure goods, entertainment, leisure services, fares, motoring, and petrol.[3] We turn first to the total budget variable in our Engel curve analysis. This is typically transformed by the log transformation because total outlay is often supposed to have a normal cross-section distribution. To see the power of the kernel method, Figure 3 presents the (Gaussian) kernel density estimation using a group of about 1,000 household, from the U.K. Family Expenditure Survey. These are married couples with no children (so as to keep a reasonable degree of homogeneity in the demographic structure).

Figures 3 and 4 here

The results are interesting and show that it is relatively difficult to distinguish the nonparametric density from the fitted normal curve, which is also shown. The bivariate kernel density plot in Figure 4 indicates that the joint density of food expenditure share and log total expenditure seems close to bivariate normal, with strong negative correlation. The line through the bivariate

[2] A further selection of households with cars was made in order to include motoring expenditures and, in particular, petrol as commodity groups.

[3] More precise descriptions of components of the commodity groups are provided in Blundell, Browning and Crawford (2003).

distribution is the local conditional mean or the kernel regression, to which we next turn.

To provide an idea of the importance of allowing flexibility in the shape of the Engel curve relationship, Figures 5 and 6 present kernel regressions for the Engel curves of two commodity groups in the FES together with a quadratic polynomial regression. These curves are presented for a relatively homogeneous group of married women without children, although the next section will discuss how socio demographic heterogeneity might be accommodated in kernel-based regression techniques.

Figures 5 and 6 here

4 From Statistical Engel Curves to Structural Expansion Paths

4.1 Improving RP Tests

In order to utilise these statistical Engel curves in the analysis of RP conditions, we need to develop structural expansion paths. From these we can then ascertain the demands as we move the budget line and in turn mimic the experimental design. Structural expansion paths need to be based on an empirically acceptable and theoretically sound method for pooling over demographic types. They must also allow for endogeneity of total expenditure within the nonparametric regression. The nonparametric nature of these estimates means that they provide flexible expansion paths that differ across markets, i.e. by time period and location.

4.2 Pooling Expansion Paths over Demographic Types

Let $\{(\ln x_i, w_{ij})\}_{i=1}^n$ represent a sequence of n household observations on the log of total expenditure $\ln x_i$ and on the jth budget share w_{ij}. Also, let $\mathbf{z}$ represent a vector of discrete demographic variables. The Engel share curve is given by:

$$E(w_{ij}|x, \mathbf{z}) = G_j(\ln x_i, \mathbf{z}_i).$$

A popular semiparametric specification in the partially linear model (see Robinson 1988):

$$E(w_{ij}|x, \mathbf{z}) = g_j(\ln x_i) + \mathbf{z}_i' \boldsymbol{\gamma}_j.$$

However, the following proposition shows this to be a particularly restrictive choice once RP conditions are imposed.

Proposition 2 (Blundell, Browning and Crawford 2003). *Suppose that budget shares have the following form that is additive in functions of* $\ln x$ *and demographics* $\mathbf{z}$

$$w_j(\ln \mathbf{p}, \ln x, \mathbf{z}) = m_j(\ln \mathbf{p}, \mathbf{z}) + g_j(\ln \mathbf{p}, \ln x)$$

If Slutsky symmetry holds and if the effects of demographics on budget shares are unrestricted, then $g_j(.)$ *is linear in* $\ln x$.

Thus, if the simple partially linear form is used then, to make it generally consistent with RP, preferences are restricted to the semi-log budget share class know as Piglog (Muellbauer 1976).

An attractive alternative is the shape-invariant or shape-similar specification. Härdle and Marron (1990) and Pinske and Robinson (1995) propose such a generalisation of the partially linear model:

$$E(w_{ij}|x, \mathbf{z}) = g_j(\ln x_i - \phi(\mathbf{z}_i' \boldsymbol{\theta})) + \mathbf{z}_i' \boldsymbol{\alpha}_j.$$

Defining $s = 0, 1, .., S$ distinct demographic groups of size n_s say, different family sizes, these shape invariant restrictions have the form

$$g_j^s(\ln x_i) = g_j^0(\ln x_i - \phi(\mathbf{z}^{s\prime}\boldsymbol{\theta})) + \mathbf{z}^{s\prime}\alpha_j.$$

In practice this transformation has been found to work well, see Blundell, Duncan and Pendakur (1998), for example. This semiparametric method of pooling across household types is adopted in the work that follows.

4.3 Endogeneity of Total Expenditure

Endogeneity has to be a key concern in giving a structural interpretation to a statistical relationship. Consider the endogeneity of total expenditure x. It is very likely that the total budget and individual commodity demands are jointly determined. Instrumental variable estimates for nonparametric regression have been developed in a sequence of recent papers: Newey and Powell (2003), Darolles, Florens and Renault (2000) and Hall and Horowitz (2003); these are reviewed in Blundell and Powell (2003). Here we consider the estimation of the semiparametric model that includes the shape-invariant restrictions. For this we consider the semiparametric IV estimates under the shape-invariant restrictions as developed in Blundell, Chen and Kristensen (2003).

Figure 7 presents the estimates of for food shares for two adult families with and without children.[4] The plots offer a comparison of the fully nonparametric estimate vs. the semiparametric one, and the endogenous case vs. the exogenous one. Together with the estimated Engel curves, they also report 95% pointwise confidence bands of these. The bands were obtained using the non-

[4] The comparison is between households with no children and those with one or two children.

parametric bootstrap based on 1,000 resamples.[5] As noted in Blundell, Chen and Kristensen (2003), the nonparametric IV estimates using the subsample of households without children should be interpreted with care. The estimates are quite imprecise. Our main focus is on the lower RHS plot in each panel which represents the semiparametric IV estimates under the shape invariant restrictions. Several interesting features are present in the plots. As may be expected, the estimated shares of alcohol and food-out for households with children are everywhere below those for households without children. As family size increases, for any given total outlay the shares going to alcohol and food-out fall while the share going to food-in increases. So there is a shift in expenditure shares from one set of nondurables to another when families have children. The curvature also changes significantly as we allow for endogeneity. Therefore, neglecting potential endogeneity in the estimation can lead to incorrect estimates of the Engel curve shape. The Engel curve for food-in, for example, shows a much more pronounced reverse "S" shape under endogeneity, with a more dramatic shift to the right in the curve resulting from the presence of children.

Figure 7 here

The semiparametric efficient estimates of the θ and α parameters for the full set of goods are given in Table 1. The estimates are plausibly signed in both the endogenous and exogenous cases. However, the data supports the hypothesis that total expenditure is endogenous. The results show a strong impact on θ of allowing for endogeneity. This parameter measures the general log equivalence scale for the presence of children with a couple normalised to

[5] Further details of these plots and more plots for a range of goods are described in the Blundell, Chen and Kristensen paper.

unity. The LS estimate is implausibly low, whereas the IV estimate is very plausible and represents an equivalence scale of about 0.45, normalised to unity for a couple without children. This is also seen in the more dramatic shift in the plotted curves between the two groups as commented on previously. One can give an interpretation to the estimates of α; for example the negative value of α for alcohol shows the decline in the overall alcohol budget share, given total equalised expenditure, that occurs for larger households.

Table 1: Efficient estimates of θ in the exog. and endog. case

	Semiparametric IV		Semiparametric LS	
	coefficient	std. ($10^{-3}\times$)	coefficient	std. ($10^{-3}\times$)
θ	0.3698	57.4712	0.1058	34.3810
α - alcohol	-0.0216	4.5047	-0.0239	2.5322
α - fares	-0.0023	2.5089	-0.0092	1.4027
α - food-in	0.0213	6.5406	0.0461	4.8861
α - food-out	0.0006	3.6744	-0.0046	2.4182
α - fuel	-0.0035	2.7611	0.0054	1.9069
α - leisure	0.0388	10.9148	-0.0016	6.2392
α - travel	-0.0384	5.9912	-0.0226	3.9748

It is these semiparametric IV estimates of Engel curves that we use to construct structural expansion paths. With these in place, for each price regime we can now go on to test for periods that do not reject the RP conditions and provide bounds on demand responses and welfare measures.

5 Bounds on Demand Responses

With the nonparametric expansion paths in place, we can consider rejections of the RP restrictions. This plan is carried out in detail in Blundell, Browning and Crawford (2003). However, not only can we improve the power of the RP test, we can also address the first of the two main concerns raised in the Introduction.

Namely, where we don't reject, we can show how to improve the precision of the bounds on demand responses and on welfare costs of price regulation or tax reforms.

5.1 Bounding Demand Curves

Varian (1982) provided a comprehensive analysis on demand response bounds under RP conditions. In Figure 8 we present a two-good two-period Varian (1982) best support set for demand responses for new prices p_0.

Figure 8 here

Suppose we observe a set of demand vectors $\{\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_T\}$ that record the choices made by a consumer when faced by the set of prices $\{\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_T\}$. Varian (1982) poses the question of how, whilst respecting the standard requirements of rationality but without making any parametric assumptions about preferences, we can use such data to predict demands if we have a new price vector $\mathbf{p}_0$ with total outlay x_0. He suggests the notion of a *support set,* which is defined as:

$$S^V(\mathbf{p}_0, x_0) = \left\{ \mathbf{q}_0 : \begin{array}{c} \mathbf{p}_0'\mathbf{q}_0 = x_0, \ \mathbf{q}_0 \geq \mathbf{0} \text{ and} \\ \{\mathbf{p}_t, \mathbf{q}_t\}_{t=0,\ldots,T} \text{ satisfies RP} \end{array} \right\}.$$

That is, the set of demands on the budget surface that are nonparametrically consistent with the existing data. If the original set of demands does not satisfy the RP conditions then this set is, of course, empty.

Another possible resolution is to assume something more specific about the way in which demands vary with the total budget. In this case we generally have tighter bounds. To illustrate this, consider Figure 9, which simply adds linear expansion paths (denoted by $\mathbf{q}_1(x)$ and $\mathbf{q}_2(x)$) through the origin (homothetic).

Figure 9 here

As the figure shows, the only demands on the new budget line that are consistent with GARP and with the original data and expansion paths constitute a strict subset of the original bounds. The problem with this approach is, of course, that preferences may not be homothetic, and the assumed expansion paths may not be anything like the true expansion paths. Nevertheless, if the true expansion paths were available then this would provide a basis upon which to proceed with tightening the bounds.

Blundell, Browning and Crawford (2004) derive the following properties for the support set for demand responses $S(\mathbf{p}_0, x_0)$ derived from the intersection demands on each nonparametric expansion path. These support sets generate the E-bounds for demand curves.

Proposition 3 (Blundell, Browning and Crawford (2004)).

A. *For any* $(\mathbf{p}_0, x)$, *if the intersection demands* $(\mathbf{p}_t, \mathbf{q}_t(\tilde{x}_t))_{t=1...T}$ *satisfy GARP then the support set is non-empty.*

B. *The set* $S(\mathbf{p}_0, x_0)$ *is convex.*

C. *For any point on the new budget line that is not in* $S(\mathbf{p}_0, x_0)$, *we have that the intersection demands and this point fail GARP.*

A corollary of this proposition is that $S(\mathbf{p}_0, x_0)$ provides the best nonparametric bounds on demand responses that are local to each income percentile.

We label the bounds that correspond to this best support set 'E-bounds' because they are based on expansion paths. From Figure 10 it is straightforward to see how these bounds can be used to tighten the bounds on complete demand responses. Note that the resulting bounds on demand responses are local to each point in the income distribution.

Figure 10 here

5.2 What Features of the Data Narrow the Bounds?

Here we briefly investigate what conditions lead to narrow E-bounds. That is, when do we get tight bounds on behavioural responses to new prices?

To do this, consider the set of relative prices we observe in our U.K. data over the period 1975 to 2000. This is presented in Figure 11 and here we choose just three aggregate goods in order to enhance the visual analysis. The relative prices show the dramatic shift in the relative price of food in the late 1970s. The line with the crosses show one particular path for the change in the price of food, holding all others at their normalised value of unity. Suppose we are interested in bounding the demand curve for this relative price change. We can do this at different income levels, but where will the bounds be tightest?

Figure 11 here

In Figures 12 and 13 we consider a simple illustrative case. In the first figure we have two observed periods (with the price of good 3 normalised to unity); the observed relative prices are given by the stars in Figure 12. We can always find a starting hypothetical price that is a convex combination of observed prices; in the figure, one is shown by the circle. This should give good bounds. As can be seen, however, any variation of the price of good 1 leads to hypothetical prices outside the convex hull of observed prices (in this case simply the line between the observed prices) and this yields wide bounds. Yet giventhree observed prices (figure 13), we can find a starting value and variations that stay within the convex hull. Thus, as long as we have as many periods as goods, we can find a hypothetical starting value and variations that allow us to stay within the convex hull.

Figures 12 and 13 here

Blundell, Browning and Crawford (2004) provide a general result for this intuition and show that, as we expand the convex set hull, we generally also shrink the E-bounds on demand responses.

5.3 E-Bounds on Demand Responses

These ideas work in application extremely well. For example, consider applying these ideas to the estimated nonparametric expansion paths and the relative price path in Figure 11. Figure 14 shows the convex hull of prices over which there are no RP rejections for an individual defined by the median income in 1985; the resulting E-Bounds on demand responses for food at this income level are presented in Figure 15. These figures show that tight bounds are achievable inside the convex hull, where the relative price information is dense.

Figures 14 and 15 here

5.4 Separability

Separability and other dimension-reducing restrictions can also help enormously in improving the precision of the bounds on demand responses. With two goods we can achieve point identification of demand responses at each observed relative price. With more than two goods, support sets only collapse to a point at the base price. Varian (1986) carefully lays out the RP conditions for weak and homothetic separability. The imposition of these separability restrictions strictly narrows the support set and can be easily added to the set of RP conditions defining E-bounds.

6 Bounds on Welfare Measures and Cost-of-Living Indices

In addition to bounding demand responses, we can also show that using nonparametric expansion paths provides the best nonparametric bounds on welfare costs of price regulation or tax reforms. This approach was developed systematically in Blundell, Browning and Crawford (2003), and the intuition can be seen from Figure 16. In this figure the upper and lower E-bounds on the indifference surface passing through $\mathbf{q}_1$ are given by the piecewise linear lines that describe the revealed better and revealed worse sets relative to q_1.

The dashed lines marked 'upper' and 'lower' shows the bounds on the cost function for some new set of relative prices $\mathbf{p}_z$. An early reference for this insight can be found in Arrow (1958), where the 'tightest' upper and lower bounds for indifference curves passing through any chosen point in the commodity space.

Figure 16 here

The application of these bounds will be important in welfare economics and can be illustrated in the analysis of cost of living bounds. Figure 17 provides such an analysis using the British FES data. In this graph, taken from the Blundell, Browning and Crawford (2003) study, the E-bounds on cost of living are represented by the solid lines and the classical revealed preference bounds by the dashed line. The bounds from classical revealed preference restrictions of the type used by Varian (1982) and calculated using the demands in each period at median within-period total budget are also reported. Confirming the results in Varian (1982) and Manser and McDonald (1988), classical non-parametric/revealed preference bounds based on the median demand data give little additional information on the curvature of the indifference curve through commodity space, and hence the bounds on the true index are wide. However, by the use of expansion paths we can dramatically improve these bounds.

Figure 17 here

Blundell, Browning and Crawford (2003) also show that the chained Törnqvist index is the only one of the traditional parametric indices that lies almost everywhere within these bounds and gives some support to its use in practise. Moreover, the E-bounds are also shown to be tight and perhaps sufficient for most policy purposes without making untenable assumptions.

7 How Should we Characterise Changing Tastes?

The question arises of how we should react when the data are not exactly in line with RP conditions. In this approach we allow local perturbations to preferences to describe the degree of taste changes, or a shift in marginal utility. This will allow us to assess the direction of taste change and will enable an evaluation of how tastes change for rich and poor. Essentially we ask the question: Are relative price changes enough or do we require changes in tastes?

First we explore whether these estimated changes in tastes are statistically significant or simply reflect sampling variation in the estimated expansion paths. The details of this method are developed in Blundell, Browning and Crawford (2004). To implement the approach, a minimum distance method is used in order to estimate perturbed demands that are local to each income percentile and data period. The minimum is taken subject to the RP conditions.

For each good and time period and for each income level, a series of perturbations is calculated. For the three goods described before the sequence of perturbations is given in Figure 18.

Figure 18 here

The results of this empirical analysis show that tastes do appear to change. They also show that tastes change slowly and that this taste change differs across the income distribution. Blundell, Browning and Crawford document periods of taste stability for some types of consumers over certain groups of goods. For quite long contiguous periods of time, RP conditions are not rejected. Consequently the convex hull can be expanded and E-bounds further improved. This approach allows an increased set of goods and periods for which RP conditions are not rejected and consequently expands the convex hull. Using this idea the improvement in the bounds reported in Figure 15 are shown in Figure **??**.

Figure 19 here

8 What Has Been Achieved and What Is Next?

This paper has shown the attraction of recent new developments in empirical revealed preference analysis. There is now a powerful test of rationality that is achieved completely within a nonparametric framework and allows an assessment of rationality by point in the income distribution, by groups of goods, by time periods. This approach enables a characterization of changing tastes. We have derived tight nonparametric bounds (E-bounds) on cost-of-living indices and welfare measures as well as tight nonparametric bounds on demand responses while avoiding the specifications in traditional welfare and IO analysis that heavily restrict substitution effects and their variation across the income distribution.

To close this discussion I shall to highlight three further challenges for empirical RP analysis. The first concerns dimension reduction. The second relates to the case of continuous price data and concerns the imposition of economic shape restrictions in nonparametric regression. The third concerns the development of a theoretically consistent approach to unobserved heterogeneity. These are pressing, yet exciting, areas for research.

Empirical studies of differentiated products in industrial organisation have increased the need for flexible measures of substitution parameters. With the large number of products typically under study, some dimension reduction is required. The most natural approach is the Gorman-Lancaster style hedonic models of characteristics demand. This has become more prevalent with the advent of widely available consumer panels. Important steps have already been taken in the recent paper by Blow, Browning and Crawford (2004) which analyses the characteristics demand using a consumer panel data. Consumer panels provide a natural source of information for characteristics models and an exciting prospect for RP analysis.

In relation to the second concern, when price data is continuous, the RP algorithm used in the work reported in this paper is not available. We need to be able to test and impose economic shape restrictions in nonparametric regression. Here again there is important recent work. For example, Yatchew and Bos (1997) develop a procedure that can easily incorporate constraints on derivatives (such as the Slutsky conditions). Hall and Yatchew (2004) propose a class of bootstrap-based tests for a variety of hypotheses including additive separability, monotonicity, and convexity as well as radial symmetry in density estimation.

Finally there is the issue of unobserved heterogeneity. In nonparametric regression analysis, this is the equivalent of nonseparable errors; it is well known that additive preference errors presents a highly restrictive preference specification, see Brown and Walker (1989) and Lewbel (1996). In the absence of long panels on individual consumer choices, the identification of preferences with unobserved heterogeneity places restrictions on behaviour. McFadden (1973, 2004), McFadden and Richter (1991), and Matzkin (1994) were at the forefront of developing a stochastic revealed preference analysis, and in a path-breaking study Brown and Matzkin (1998) develop a relatively flexible and theoretically consistent specification in which marginal utilities are linear in preference heterogeneity. This establishes a powerful framework on which to build a fully stochastic revealed preference analysis.

References

[1] Afriat, S.N. (1973),"On a System of Inequalities in Demand Analysis: An Extension of the Classical Method", *International Economic Review,* 14, 460-472.

[2] Arrow, K. (1958), "The Measurement of Price Changes", Joint Economic Committee, US Congress, Washington, D.C. Reprinted in Collected Works of Kenneth J. Arrow, Vol. 3: Individual Choice under Certainty and Uncertainty, Harvard University Press, Cambridge.

[3] Banks, J. Blundell, R.W. and Lewbel, A. (1996), "Tax Reform and Welfare Measurement: do we need Demand System Estimation?", *Economic Journal*, 106, 1227-1241.

[4] Banks, J. Blundell, R.W. and Lewbel, A. (1997), "Quadratic Engel Curves, Indirect Tax Reform and Welfare Measurement", *Review of Economics and Statistics*, LXXIX, 4, 527-539.

[5] Becker, G. S. (1962), "Irrational behaviour and economic theory", *Journal of Political Economy*, 70, 1-13.

[6] Blow, L. M. Browning and I. Crawford (2003), "Non-parametric methods for the consumer characteristics model", mimeo, Centre for Applied Microeconometrics, University of Copenhagen, May.

[7] Blundell, R. (1998), "Consumer Behaviour: Theory and Empirical Evidence - A Survey", *Economic Journal*, Vol. 98 (389), March, 6-65.

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

[9] Blundell, R., M. Browning and I. Crawford (2004), "Best Nonparametric Bounds on Demand Responses", mimeo IFS, Walras-Bowley Lecture, presented at the North American Econometric Society Meetings, UCLA.

[10] Blundell, R., Chen, X. and D. Kristensen, (2003) "Semiparametric Engel Curves with Endogenous Expenditure", CeMMAP Working Paper, IFS, CWP 15/03.

[11] Blundell, R. and Duncan, A. (1998) "Kernel Regression in Empirical Microeconomics" *Journal of Human Resources*, 33, 62-87.

[12] Blundell, R. Duncan, A. and Pendakur, K.(1998), "Semiparametric Estimation and Consumer Demand", *Journal of Applied Econometrics*, 13, 435-461.

[13] Blundell, R. P. Pashardes and G. Weber, (1993), "What do we Learn about Consumer Demand Patterns from Micro Data?", *American Economic Review*, 83, No.3, 570-597

[14] Blundell, R. and J. Powell, (2003), "Endogeneity in Nonparametric and Semiparametric Regression Models", Chapter 8 in Dewatripont, M., Hansen, L. and S. J. Turnovsky (eds), *Advances in Economics and Econometrics*, Cambridge University Press, ESM 36, pp 312-357.

[15] Blow, L. and I. Crawford (2001), The cost of living with the RPI: Substitution bias in the UK Retail Prices Index", *Economic Journal*, 111.

[16] Bronars, S.G. (1995), "The Power of Nonparametric Tests of Preference Maximisation", *Econometrica*, 55, 693-698.

[17] Brown, D.J. & R.L. Matzkin (1998) "Estimation of Simultaneous Equations Models, with an Application to Consumer Demand" Cowles Foundation Discussion Paper.

[18] Brown, B.W. and Walker, M.B., (1989), "The random utility hypothesis and inference in demand systems" *Econometrica*, 57, 815-829.

[19] Browning, M. and C. Meghir, 1991, The effects of male and female labor supply on commodity demands, *Econometrica,* 59, 925-951.

[20] Darolles, S., J-P Florens and E. Renault, (2000), "Nonparametric Instrumental Regression", mimeo, GREMAQ, University of Toulouse, April.

[21] Deaton A. S. and J. Muellbauer (1980),"An Almost Ideal Demand System", *American Economic Review*, 70, 312-36.

[22] Diewert, W.E. (1973), "Afriat and Revealed Preference Theory", *Review of Economic Studies*, 40, 419-426.

[23] Diewert, W.E. and Parkin, C. (1985), "Tests for the Consistency of Consumer Data", *Journal of Econometrics*, 30, 127-148.

[24] Famulari, M. (1995),"A Household-based, Nonparametric Test of Demand Theory", *Review of Economics and Statistics*, 77, 372-382.

[25] Hall, P. and J. Horowitz (2003), "Nonparametric Methods for Inference in the Presence of Instrumental Variables", Working Paper, Northwestern University.

[26] Hall, P. and A. Yatchew (2004), "Unified Approach to Testing Functional Hypotheses In Semiparametric Contexts", forthcoming *Journal of Econometrics*

[27] Härdle W. and Jerison, M. (1988), "The Evolution of Engel Curves over time", Discussion paper no. A-178. SFB 303, University of Bonn.

[28] Härdle W. (1990) *Applied Nonparametric Regression*, Cambridge: Cambridge University Press.

[29] Härdle W. and Marron J. (1990), "Semiparametric Comparison of Regression Curves," *Annals of Statistics*, 18, 63-89.

[30] Hausman, J.A. and Newey, W.K. (1995), "Nonparametric Estimation of Exact Consumer Surplus and Deadweight Loss", *Econometrica* 63, 1445-1476.

[31] Houthakker, H.S. (1950), "Revealed Preference and the Utility Function", *Economica,* 17, 159-74.

[32] Jorgenson,D.W., L.J.Lau, and T.M.Stoker (1982),"The Transcendental Logarithmic Model of Aggregate Consumer Behavior", in *Advances in Econometrics*, Vol1, ed. by R. Basmann and G. Rhodes, Greenwich, Connecticut: JAI Press.

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

[34] Lewbel, A. (1991), "The rank of demand systems: theory and nonparametric estimation", *Econometrica*, 59, 711-730.

[35] Lewbel, A (1996), " Demand systems with and without errors: reconciling econometric, random utility and GARP models", *American Economic Review.*

[36] Manser,M. y R. MacDonald (1988), "An Analysis of Substitution Bias in Inflation Measurement", *Econometrica*, 56: 909-930.

[37] McFadden, D. (1973), “Estimation and Testing of Models of Individual Choice Behavior from Data on a Population of Subjects”, Berkeley, working paper, University of California.

[38] McFadden, D. (2004), “Revealed Stochastic Preference”, Mimeo, Berkeley, February, University of California.

[39] McFadden, D. and K. Richter (1991), “Stochastic Rationality and Revealed Stochastic Preference”, in J. Chipman, D.McFadden, and K. Richter (eds) *Preferences, Uncertainty and Rationality*, Westview Press.

[40] Matzkin, R.L. (1994), “Restrictions of Economic Theory in Nonparametric Methods”,in *Handbook of Econometrics: Vol IV*, edited by R.F. Engle and D.L. McFadden, Elsevier Science, North Holland.

[41] Matzkin, R.L. (2003) “Nonparametric Estimation of Nonadditive Random Functions”, *Econometrica*, 71, 1339 - 1375.

[42] Mossin, A. (1972),“A Mean Demand Function and Individual Demand Functions Confronted with the Weak and Strong Axioms of Revealed Preference”, *Econometrica*, 40, 177-192.

[43] Muellbauer, J. (1976), “Community Preferences and the Representative Consumer”, *Econometrica*, 44, 525-543.

[44] Newey, W.K. (1994), “Kernel Estimation of Partial Means and a General Variance Estimator”, *Econometric Theory*, 10, 233-253.

[45] Newey, W.K. and J. Powell (2003): “Instrumental Variables Estimation for Nonparametric Models,” *Econometrica*, 71, 1565-1578.

[46] Pinske, C. and Robinson, P. (1995), “Pooling Nonparametric Estimates of Regression Functions with a Similar Shape”, *Advances in Econometrics and Quantitative Economics*, eds. G. Maddala, P. Phillips and T.N. Srinivisan, 172-195.

[47] Robinson, P.M. (1988), “Root n-consistent semiparametric regression”, *Econometrica*, 56, 931-954.

[48] Samuelson, P.A. (1938),“A Note on the Pure Theory of Consumer Behavior”, *Econometrica*, 5, 61-71.

[49] Sippel, R. (1997), “An Experiment on the Pure Theory of Consumer’s Behaviour”, *Economic Journal*, 107, 1431-1444.

[50] Varian, H.(1982), “The Nonparametric Approach to Demand Analysis ”, *Econometrica*, 50, 945-974.

[51] Varian, H.(1983), “Nonparametric Tests of Consumer Behaviour”, *Review of Economic Studies*, 99-110.

[52] Varian, H.(1986), “Nonparametric Analysis of Optimizing Behavior with Measurement Error”, *Journal of Econometrics*30, 445-459.

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

[54] Yatchew, A. and L. Bos (1997), "Nonparametric Least Squares Estimation and Testing of Economic Models", *Journal of Quantitative Economics*, 13, 81-131.

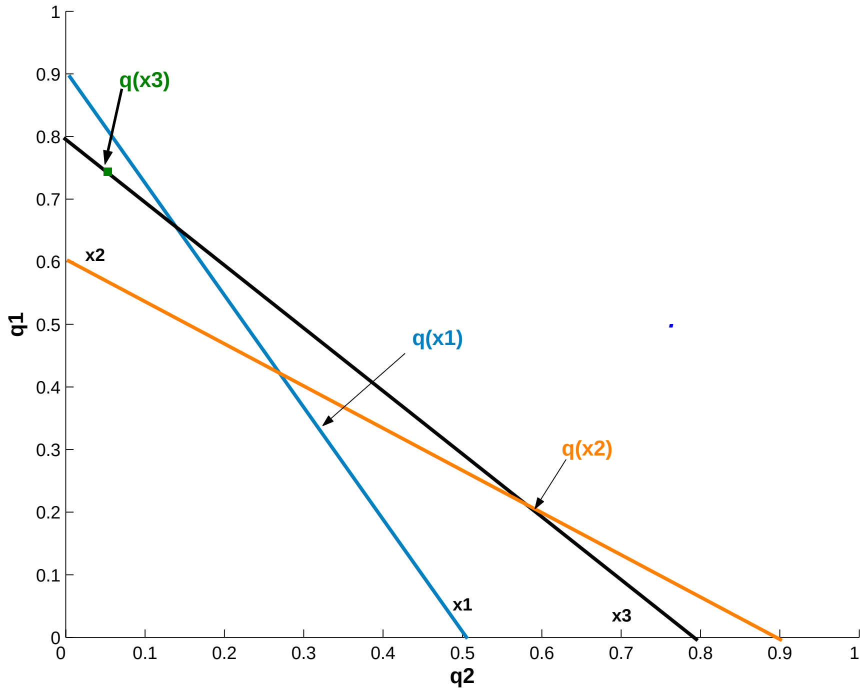


Figure 1: RP rejection.

Figure 2: Testing GARP with expansion paths.

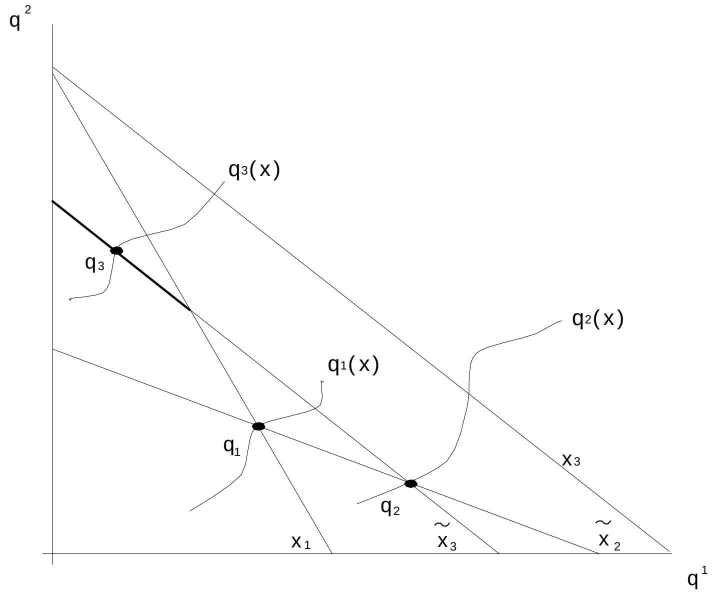


Figure 3: The density of log expenditure, FES 1980.

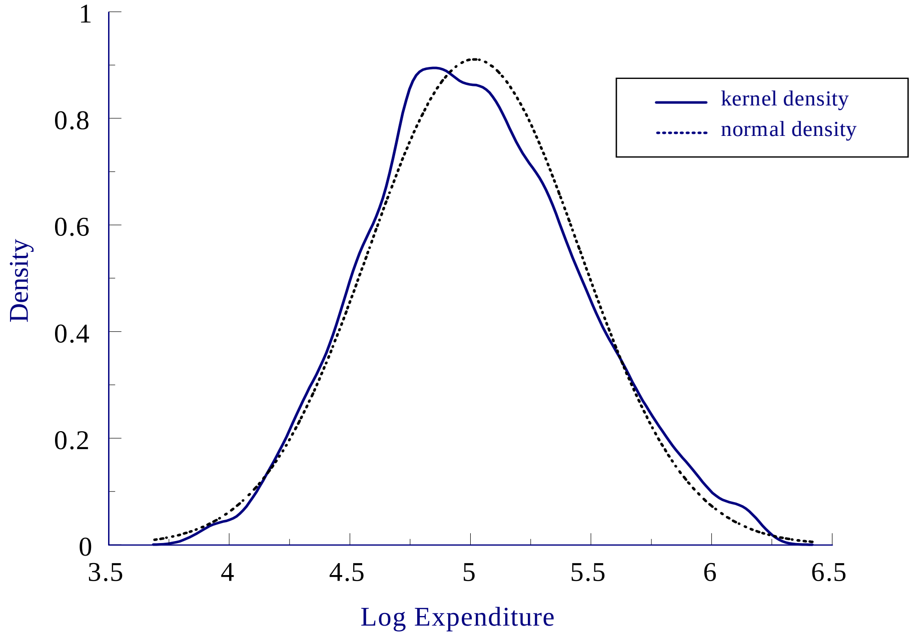


Figure 4: Bivariate kernel density: food share and log expenditure.

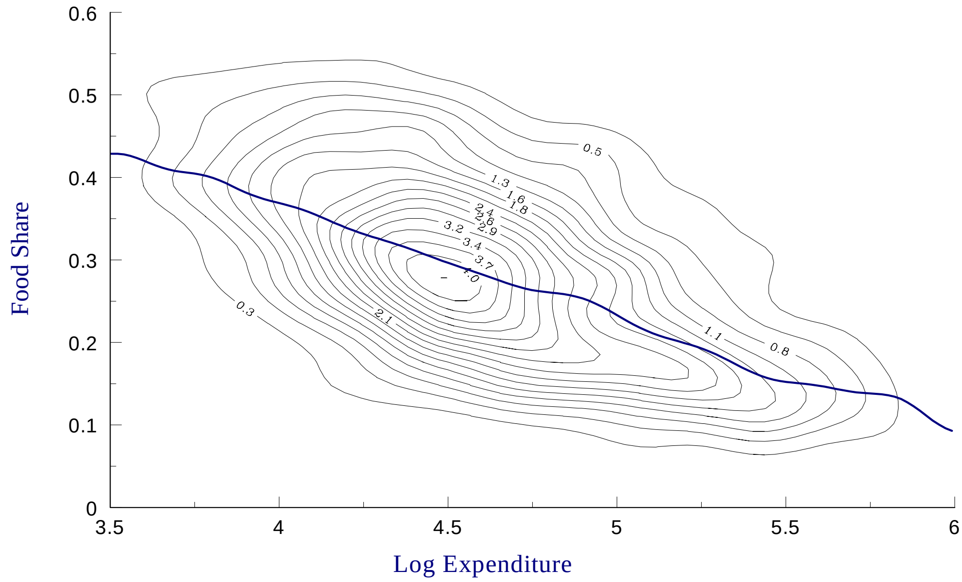


Figure 5: Nonparametric Engel curve: food share.

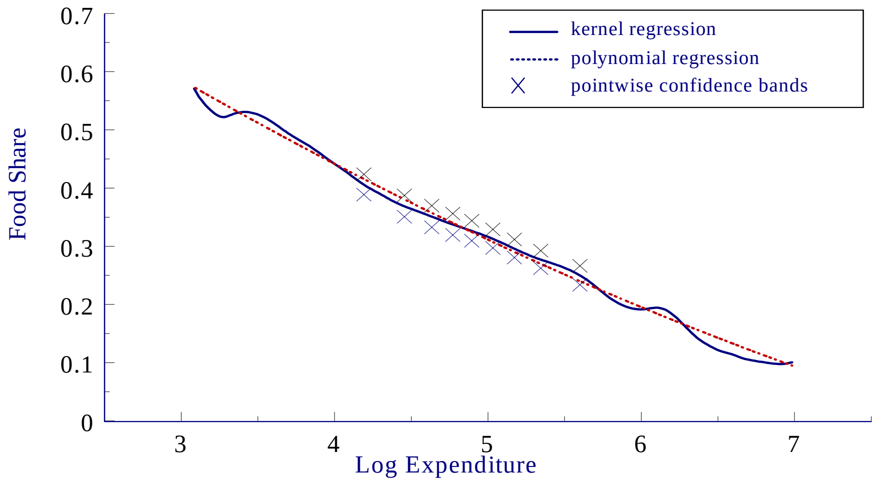
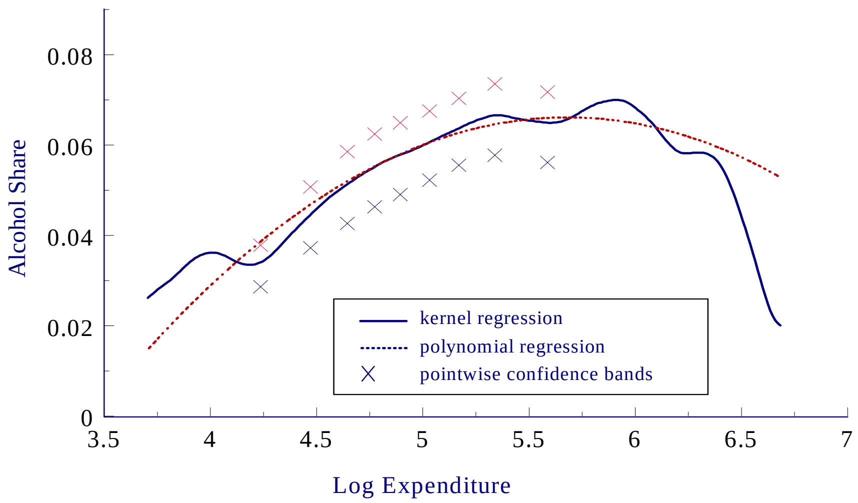


Figure 6: Nonparametric Engel curve: alcohol share.



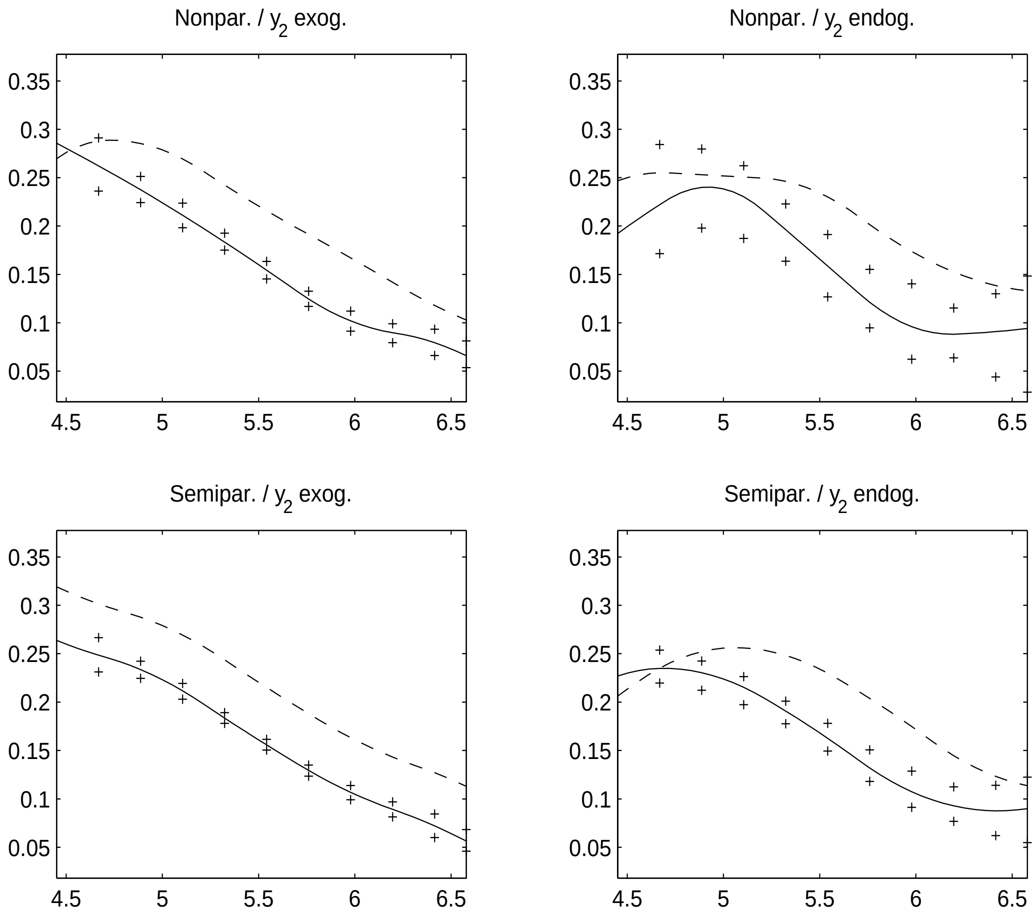


Figure 7: The estimated Engel curve for food:dashed curves with children; solid curves without children; plusses (+) 95% confidence bands.

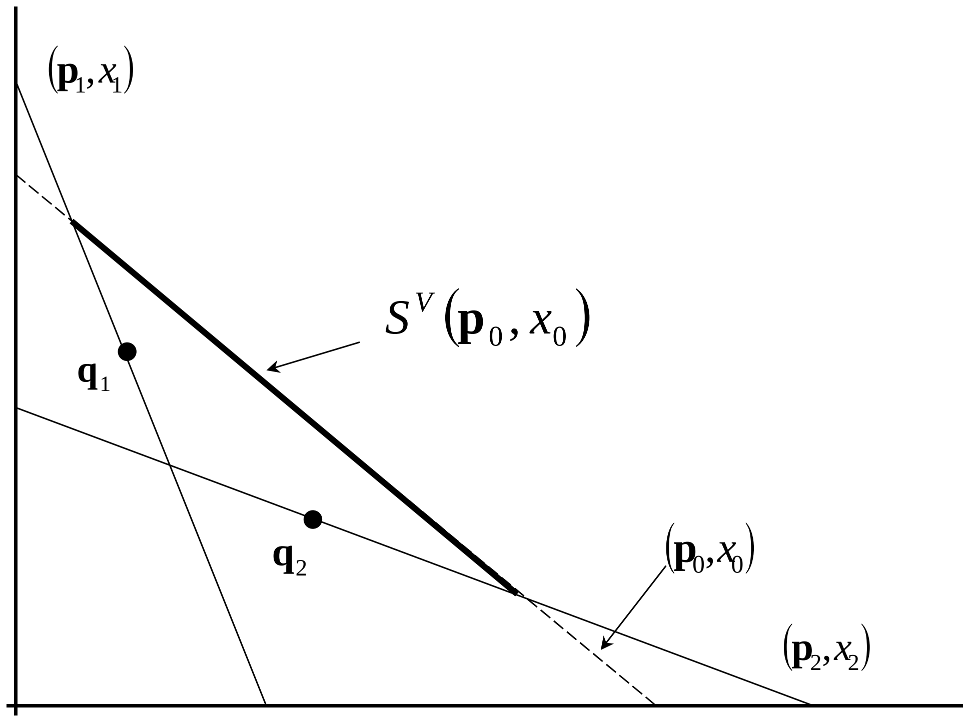


Figure 8: The Varian support Set.

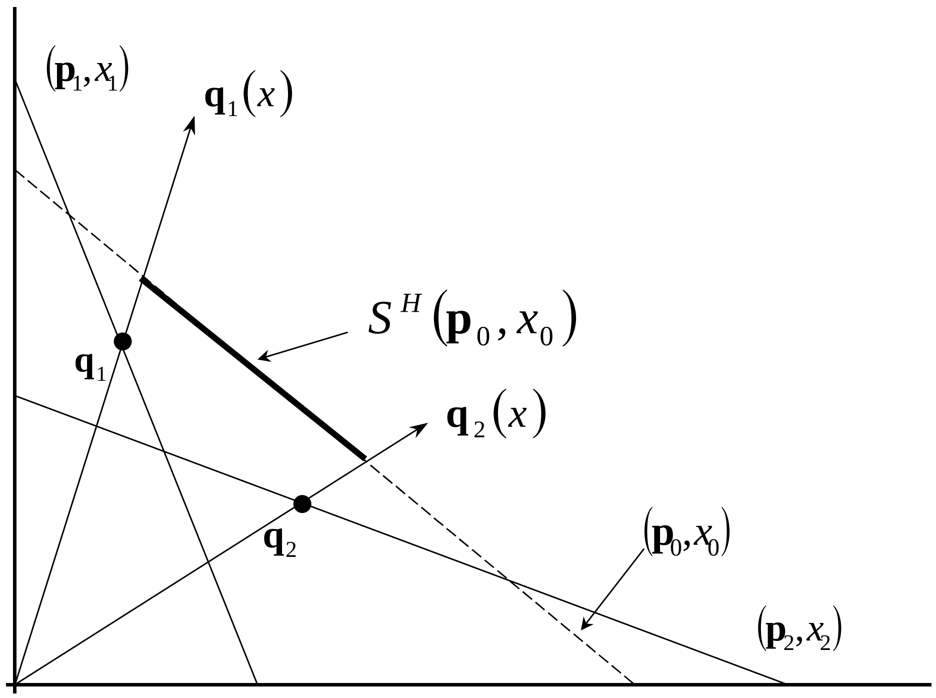


Figure 9: The support set under homotheticity.

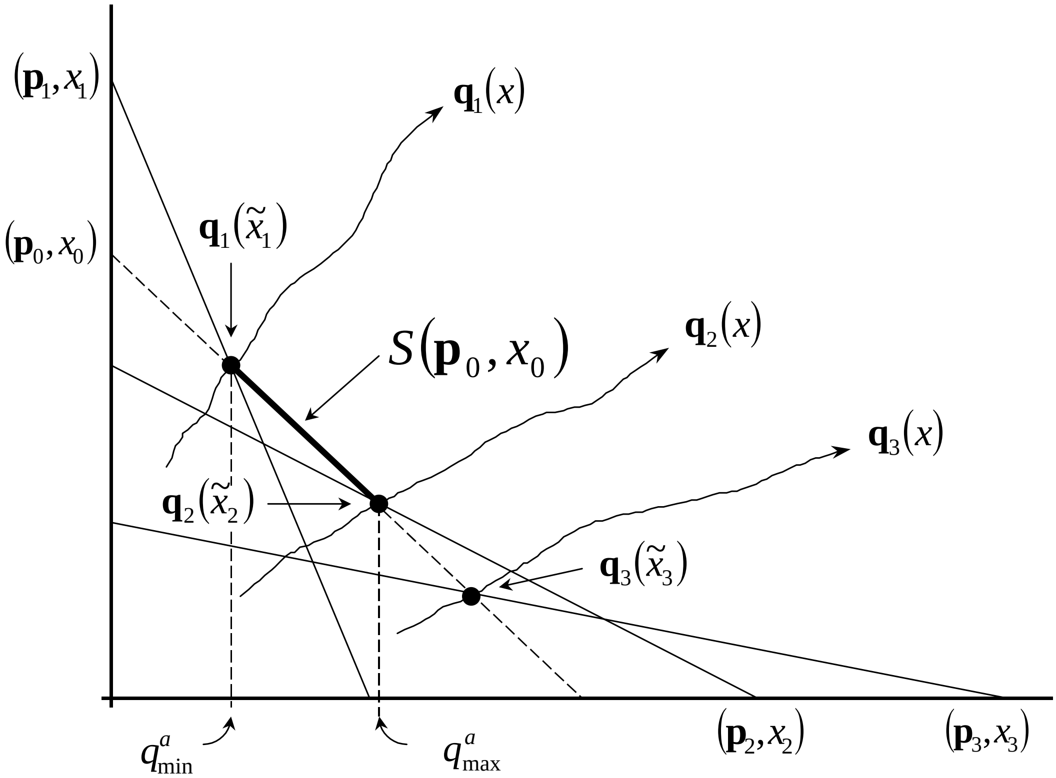


Figure 10: E-bounds on demand responses

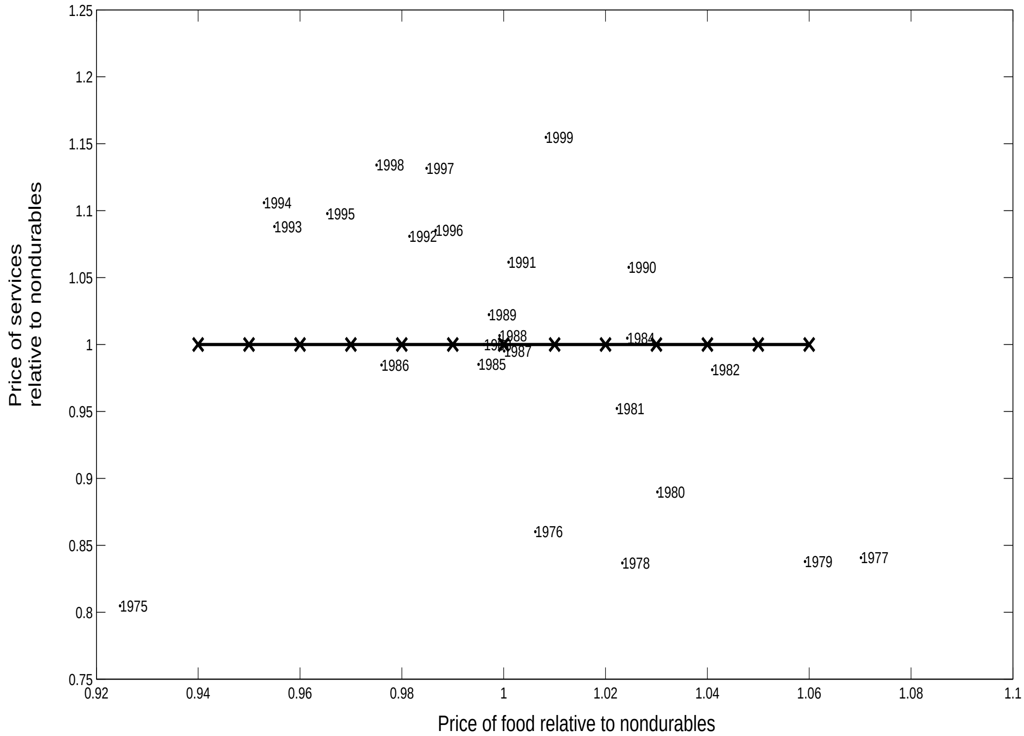


Figure 11: The times series of relative prices.

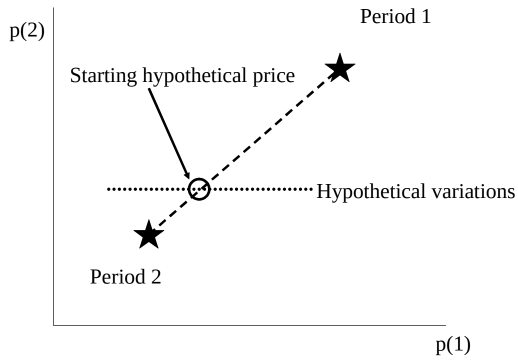


Figure 12: Convex combination.

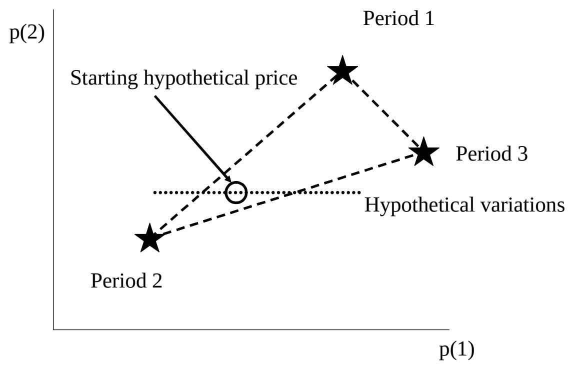


Figure 13: Convex hull.

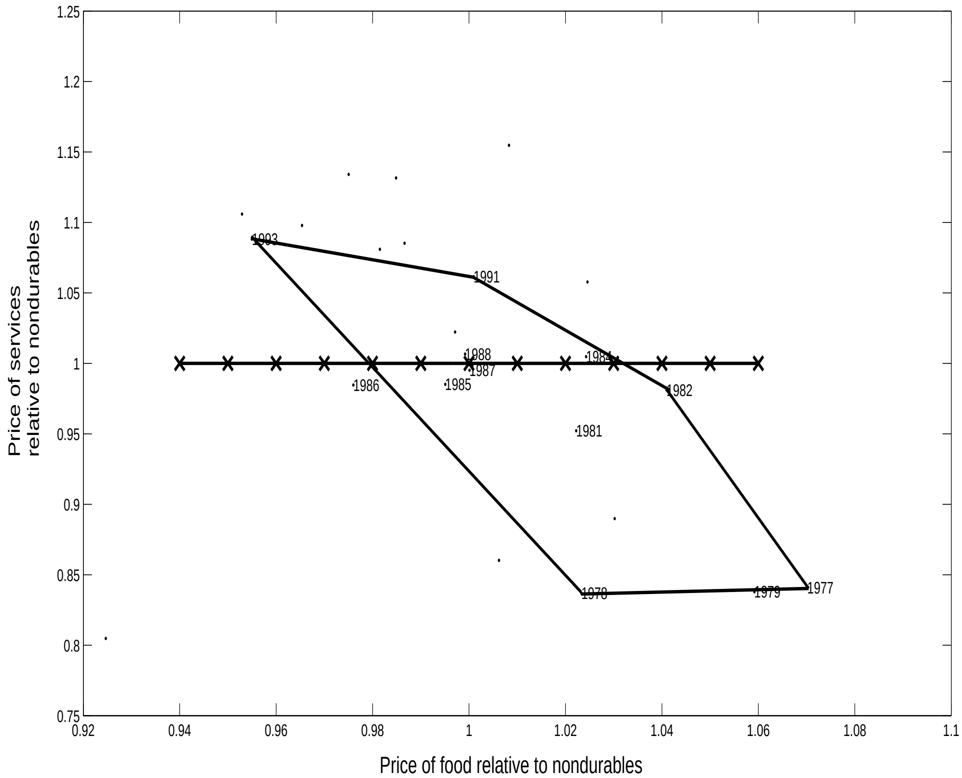


Figure 14: Convex hull for RP consistent comparisons.

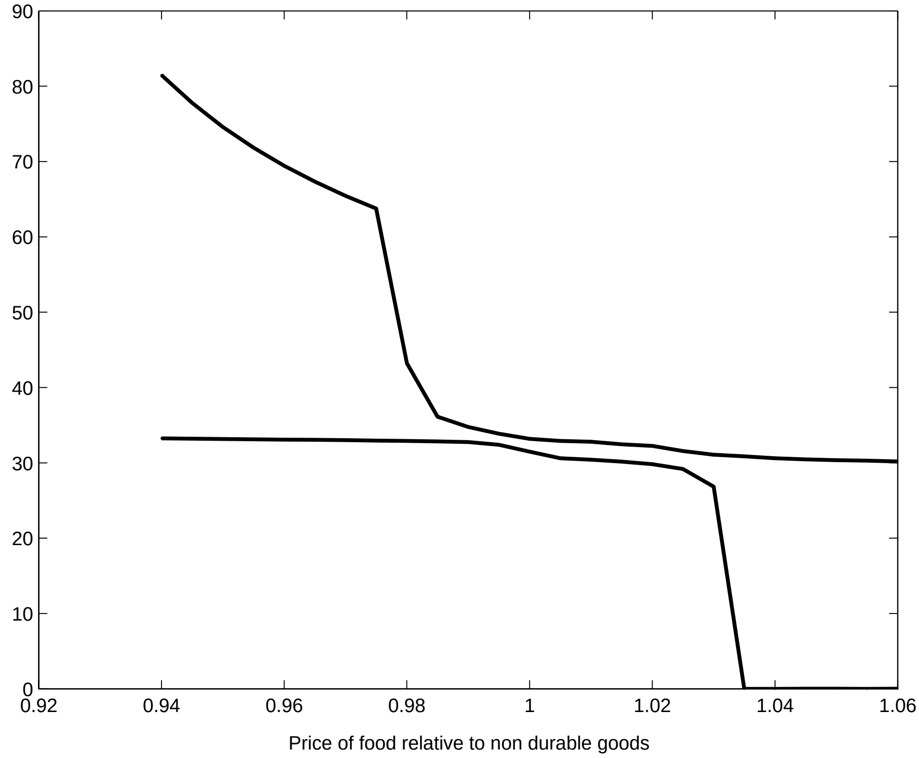


Figure 15: E-bounds to demand responses.

Figure 16: Bounds on the cost-of-living index using expansion paths.

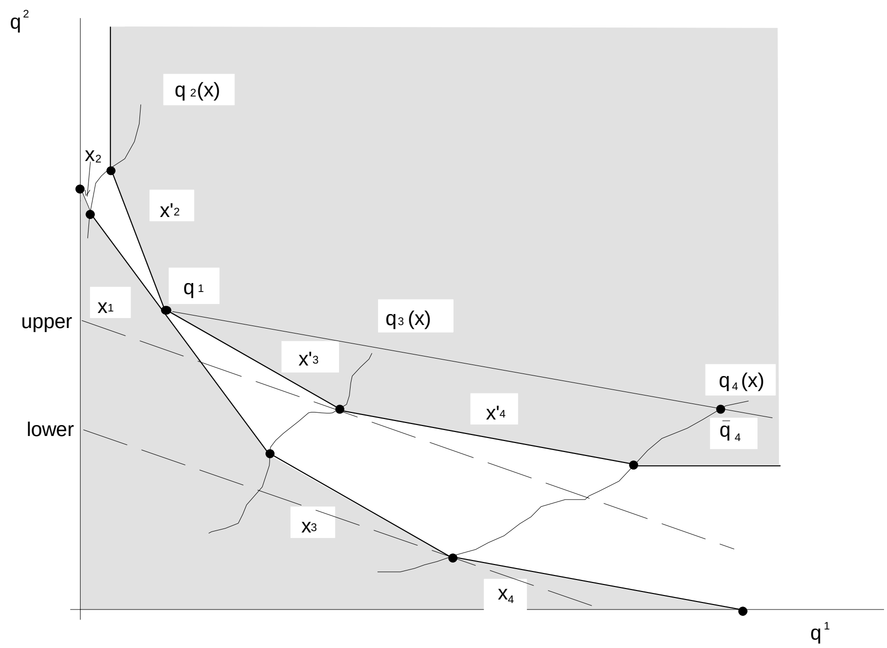
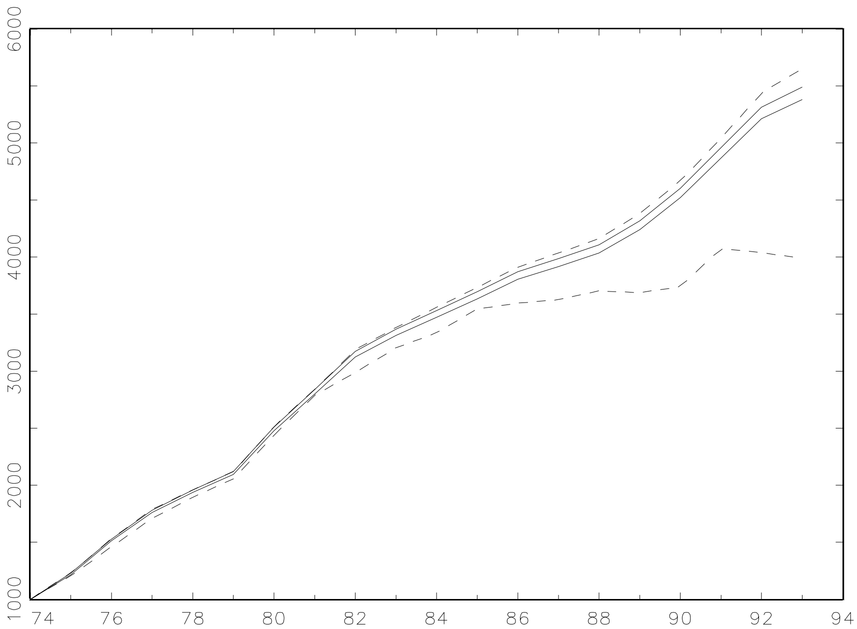


Figure 17: GARP bounds and classical RP bounds, 1974 to 1993.



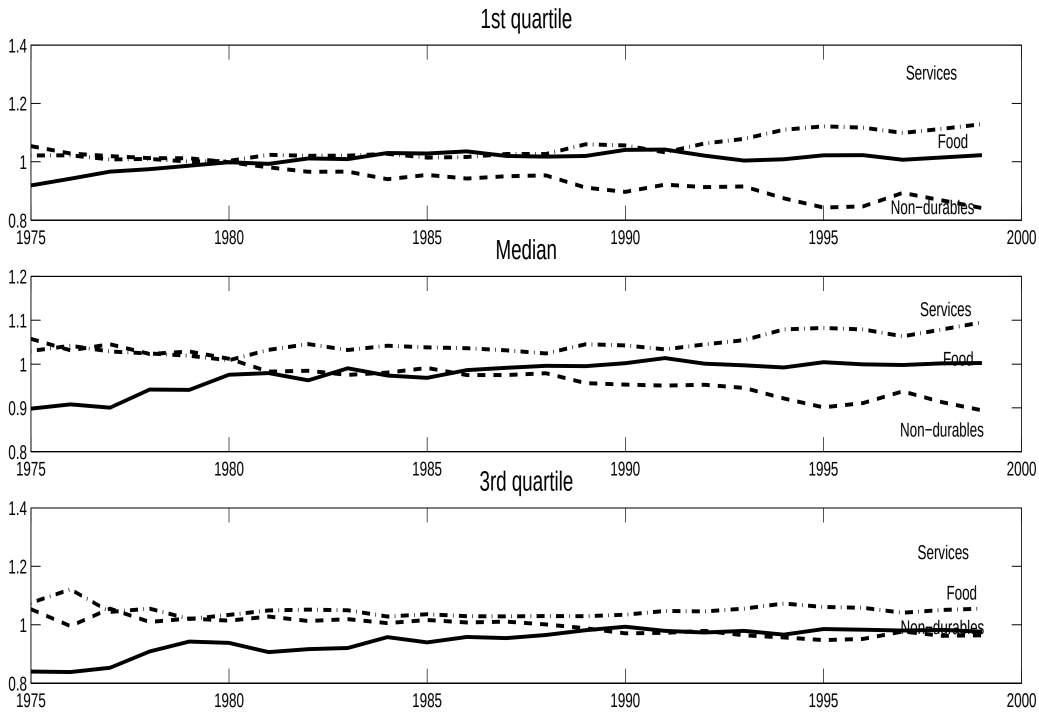


Figure 18: Taste changes across the income distribution.