Cost of Living Indices

Ian A. Crawford

November 1997

A Thesis Submitted for the Degree of Doctor of Philosophy to University College London, the University of London.

Contents

List of Tables

List of Figures

Abstract

The correct measurement of changes in the cost of living is important for many reasons. Indeed, the many uses to which measures of inflation are put make it difficult to overstate the importance of obtaining an accurate value. For example, the rate of inflation is important in wage settlements, the indexation of social security benefits, estimates of macroeconomic growth rates and, generally, all economic analyses which require knowledge of real, rather than nominal, magnitudes. Unfortunately, whilst a true cost of living index (which is defined as the ratio of the minimum cost of reaching a given level of economic welfare under alternative sets of prices) is a fairly straightforward theoretical concept, in practice its validity and the way it should be implemented are much less clear. The overall theme of this thesis is that of trying to address some of the fundamental issues which cost of living indices face without assuming either the existence or the form of the preferences which might underlie an index.

Chapter 1 introduces some of the ideas which appear later and describes some of the central issues. It defines the idea of a true cost of living index and discusses methods of determining whether a given dataset is compatible with the existence of a single, unique cost of living index constructed from it. The two main approaches to measuring cost of living indices; approximation and full parametric estimation are then discussed as are problems of

measurement, some of which arise when approximations are used, and some of which are encountered whatever the method adopted for computing the index. It concludes with summaries of the material contained in the following chapters.

Chapter 2 is an empirical investigation into the question of uniqueness: the extent to which different types of households have experienced differential changes in their cost of living in the period 1979 to 1992. It also considers the question of the influences which indirect tax changes over the period may have had, and looks at the effects which the method used to calculate housing costs has on the resulting indices. Overall, the fall in the relative price of necessities and the corresponding increase in the price of luxuries over the period studied, and the difference in expenditure patterns between rich and poor households have meant that the cost of living has increased slightly faster for richer households than it has for poorer households. The progressive nature of indirect tax reforms between 1979 and 1992 is shown to have contributed to this effect. The use of a user cost of housing approach to measuring shelter costs gives evidence of widely different inflation rates by date of birth cohorts as older cohorts reap large capital gains on housing over the period.

While Chapter 2 takes the idea of existence for granted, Chapter 3 considers the issue in more detail. The basic question asked is; is the empirical evidence consistent with the existence of a cost of living index ? So far this has proved difficult to answer. This is for two reasons. Firstly, many studies, particularly those which use microeconomic data, have found that parametric econometric models reject the hypothesis of utility maximisation or cost minimisation. But whether this is due to econometric misspecification or to a genuine rejection of the theory is hard to judge because under such circumstances the hypothesis under test is a joint one; namely that (a) the consumer has rational preferences and (b) that the functional form of these preferences is consistent with the econometric model. We have no easy way of telling whether a rejection is due to (a), or to (b) or to both (a) and (b). Secondly, while revealed preference tests present the possibility of avoiding this simultaneity (since they don't require functional form assumptions) the data used to carry out the tests often lack the power to reject the theory. This is because income growth often swamps relative price movements so that budget hyperplanes may not cross. If budget planes do not cross then it is impossible to reject the revealed preference conditions for rational choice. A further problem with revealed preference tests is that, so far, they have lacked a stochastic framework within which to judge the seriousness of any rejections. This chapter suggests a method which uses nonparametric Engel curves to simultaneously improve the power of revealed preference tests, and to make the problem amenable to statistical testing. It also looks at a method which, in the event of rejections, can be used to allow for changes in the quality of goods or equivalently changes in preferences for a particular good over the period of the data.

Chapter 4 concerns the question of the construction of true cost of living indices. Estimating true indices is computationally difficult and, as described above, the results may conflict with the theory. A literature has grown up around the issue of how to compute exact or approximate cost of living indices without estimating full, integrable demand systems. This approach uses a mixture of first and second order approximations, and indices which correspond exactly to certain forms for preferences but which are also simple to implement. If we do not wish to make functional assumptions, however, then there is only so far we can go with bounds to the true index because of limitations in the data. This chapter presents a method of using revealed preference restrictions in conjunction with nonparametric Engel curves to improve the bounds available without the need for assumptions on functional forms. The improvement is shown to be quite significant. This is used to examine the question of substitution bias, to revisit the question of uniqueness and to look at the question of whether substitution bias, which affects price indices with fixed quantity weights, varies with total expenditure.

Chapter 5 is concerned with another source of bias; specifically with the problem of new goods bias in price indices. This is an area in which functional form matters a great deal. A cost of living index needs to include the price of the new good for the period before the one in which it first existed in order to account properly for the welfare effects of its introduction. Functional form matters a great deal here as in order to do this the main method has been to use parametric demand systems to extrapolate the demand curve to the point at which demand becomes zero. The results of such an exercise are likely to be heavily dependent on functional form. This Chapter describes a nonparametric method which uses revealed preference restrictions to place a lower bounds on the virtual price of a new good without the need for functional form assumptions. This is the lowest bound consistent with the (partly testable) maintained hypothesis that the data are, on average, consistent with utility maximisation.

Acknowledgements

I would particularly like to thank Laura Blow, Richard Blundell and Martin Browning for their help. Chapter 2 is joint work with Richard Blundell and Martin Browning. Part of the theoretical work in chapter 4 is joint work with Laura Blow. I would also like to thank my supervisors Ian Preston and David Ulph, and all of my colleagues at the Institute for Fiscal Studies, past and present. In particular I would like to thank James Banks. Part finance for this thesis has been provided by the E.S.R.C. Centre for the Microeconometric Analysis of Fiscal Policy at I.F.S. and is gratefully acknowledged. The Family Expenditure Survey data used in this thesis was made available by the O.N.S. through the E.S.R.C. Data Archive. The author is solely responsible for any errors.

Declaration

No part of this thesis has previously been presented to any University for any degree.

Ian Crawford

Chapter 1

Introduction

This thesis is about aspects of cost of living indices: their existence, their uniqueness, their construction and problems encountered in the course of their measurement. The calculation and use of index numbers is essential to most aspects of social accounting, since without the aggregation of the great mass of information on quantities and prices which they perform, raw economic data would be difficult to comprehend. Further, the many uses to which measures of the rate of change of the cost of living (inflation) are put, make it difficult to overstate the importance of obtaining an accurate value. For example, the rate of inflation is important in wage settlements, the indexation of social security benefits, estimates of macroeconomic growth rates and, generally, all economic analyses which require knowledge of real, rather than nominal, magnitudes. Practicality is thus a prime concern in much of the literature on cost of living indices; as Deaton (1981) points out, cost of living indices combine "side-by-side some of the most difficult and abstruse theory with the most immediately practical issues of everyday measurement"[1]. Afriat (1977) questions whether, in most practical situations, this abstruse theory and its associated paraphernalia of utility functions and

[1] Deaton (ed.),(1981) p. 130.

cost functions really contributes very much of any use. He argues that "utility functions give service in theoretical discussions where they contribute structure which is an essential part of the matter"[2]. In fact they often confer exactly such an essential structure. Stone (1956), for example, suggested three reasons why the utility-based approach is useful in defining and constructing indices:

> "First, they give content to such concepts as real consumption which might otherwise be vague and obscure; second they bring out fundamental difficulties in establishing empirical correlates to concepts such as real consumption and so help to show what can and what cannot usefully be attempted in the present state of knowledge; finally they show the circumstances in which particular empirical correlates ... are likely to provide a good or a bad approximation to the concepts of the theory."[3]

These are precisely the concerns of this thesis. This chapter introduces some of the material which will appear later. The plan is as follows. Section 1.1 concerns the question of existence. It defines the idea of a true cost of living index, and discusses methods of determining whether a given dataset is compatible with the existence of a cost of living index constructed from it. It turns out that testing a dataset for consistency with the existence of a cost of living index is, at least in principle, a relatively straightforward matter. In practice, however, it may often be the case that the nature of the available data makes informative application of the theory difficult.

If we are unable to reject the hypothesis that a suitable index exists using, say, data on the average demands of a sample of households, the matter

[2]Afriat (1977), p. 3.

[3]Stone, (1956), pp. 18-19, also cited in Deaton, ed., (1981).

of how we can compute the index remains. The first issue is whether the data are consistent with the existence of a unique index which is identical for all households, or whether different indices should be calculated for different households. Generally speaking it has been found that cost of living indices vary with household characteristics and incomes and that a single index is therefore not universally applicable. This is considered in section 1.2. If this can be resolved, then the question of actually computing the index or indices arises. This is discussed in section 1.3 which summarises the two main approaches to measuring cost of living indices; approximation and full parametric estimation. In practice, parametric estimation may be both difficult and unreliable, and nonparametric bounds or approximations are often used. Section 1.4 discusses problems of measurement, some of which arise when approximations are used, and some of which are encountered whatever the method adopted for computing the index – even full parametric estimation. Section 1.5 concludes the introduction and summarises the contents of the next four chapters which seek to make a contribution in all of these areas.

Chapter 2 is an empirical investigation into the pattern and extent of variations in the rate of change of the cost of living between different types of households and households with different incomes. Chapter 3 presents a framework which improves the method whereby a finite dataset can be tested for consistency with the neoclassical model of consumer choice. It also looks at the issue of changes in quality or tastes, which can both lead to a rejection of the neo-classical model. Chapter 4 uses the techniques applied in Chapter 3 to develop a method of recovering two-sided bounds to true cost of living indices. It also revisits the question of differences in indices by total expenditure and provides an assessment of how well popular price indices approximate the true index using the improved yardstick which this chapter provides. Chapter 5 looks at another problem common to both price indices and true cost of living indices: the arrival of new goods. The normal parametric method of estimating the virtual price of a new good (for inclusion in an index) is both computationally expensive and heavily dependent upon a maintained hypothesis concerning the functional form of preferences. Chapter 5 presents a method of bounding this virtual price which is both computationally straightforward and does away with the need for functional form assumptions.

1.1 Existence

Change can only be recognised through reference to some standard which persists unchanged. In the case of the cost of living index the standard is a given level of economic welfare. The cost of living index measures the changing cost of reaching this standard as prices change. Traditionally in economics, the standard of living is measured by the welfare derived from consumption[4]. The concept of the true cost of living index was introduced by Konüs (1924) and is defined as the minimum cost of achieving the reference welfare level u_R when prices are $\mathbf{p}_t$, relative to the minimum cost of achieving the same welfare with prices $\mathbf{p}_s$. In notation this is written as

$$P_T = \frac{c\left(\mathbf{p}_t, u_R\right)}{c\left(\mathbf{p}_s, u_R\right)} \tag{1.1}$$

where $c\left(\mathbf{p}_t, u_R\right)$ is the consumer's cost or expenditure function evaluated at the period t price vector $\mathbf{p}_t$ and the reference welfare level u_R. The cost

[4] Sen (1985) argues persuasively against the usual approach of think about the standard of living as utility, income and wealth, suggesting a wider interpretation in which living standards are conceived of in terms of human functionings and capabilities (such as being warm and having the ability to be warm). Sen may be right but it is difficult to see how to implement his ideas with existing data.

function is central to the whole area of cost of living indices and is defined by

$$c(\mathbf{p}, u_R) \equiv \min_{\mathbf{q}} \{\mathbf{p}'\mathbf{q} : u(\mathbf{q}) \geq u_R\} \tag{1.2}$$

where the utility function $u(\mathbf{q})$ is usually assumed to be a real value function with the three properties (i) continuity, (ii) increasingness and (iii) quasi-concavity. The reference welfare level can be chosen more or less arbitrarily, however u_s and u_t are often popular and obvious choices since, by the assumption of cost minimising behaviour on the part of consumers, the cost functions evaluated at $(\mathbf{p}_t, u_t)$ and $(\mathbf{p}_s, u_s)$ are directly observable: $c(\mathbf{p}_t, u_t) \equiv \mathbf{p}_t'\mathbf{q}_t$ and $c(\mathbf{p}_s, u_s) \equiv \mathbf{p}_s'\mathbf{q}_s$. Note that as the index depends upon utility, comparing two indices with two different reference welfare levels (u_{R1} and u_{R2}) gives

$$\frac{c(\mathbf{p}_t, u_{R1})}{c(\mathbf{p}_s, u_{R1})} \neq \frac{c(\mathbf{p}_t, u_{R2})}{c(\mathbf{p}_s, u_{R2})}, \tag{1.3}$$

except under certain special circumstance which are discussed below.

If the consumer's preferences were known, and the price vectors in each period were observed, then the cost function and the index could be constructed. In many instances, however, all we ever observe are the consumer's demands $(\mathbf{q})$ and the prevailing price vector $(\mathbf{p})$, both of which are possibly measured with error. From these we need to infer the existence of $u(\mathbf{q})$ and the form and parameters of $c(\mathbf{p}, u)$ if we wish to compute the true index. In fact, it is often the case that it is aggregate or mean demands, rather than individual demands which are observed. The first, and most basic issue which this raises is whether, given a dataset of, say, mean demands and prices, can we determine the existence of a utility function with the appropriate properties outlined above which is also consistent with the data ? If we can, then Diewert (1978) shows that a corresponding cost function must also exist which is also a real value function, and which is (i) continuous, (ii) takes the value zero if $u = u(0_N)$ for all $\mathbf{p}$, (iii) is increasing in u, (iv) is concave and increasing (although not strictly so) in $\mathbf{p}$, (v) is linearly homogeneous. The question of how we might conduct such a test is considered next.

1.1.1 Revealed Preference Tests

The axioms of revealed preference were introduced by Samuelson (1947, 1948, 1950) and Houthakker (1950). Afriat (1965) and (1973), and Diewert (1973) pointed out the equivalence between counterparts of the axioms of revealed preference of Samuelson and Houthakker and the consistency of a system of homogeneous linear simultaneous inequalities, any solution to which will allow the construction of a finitely computable utility function which rationalises a finite set of demand data $(\mathbf{p}, \mathbf{q})$. The benefit of revealed preference restrictions (i.e. the restrictions which data must respect if they are not to violate the axioms of revealed preference) is that they do not presuppose the existence of utility functions and they can be used to test a dataset for consistency with the existence of a utility function with well behaved properties but an unspecified functional form.

Following Varian (1982) the definitions below set out revealed preference conditions, the notation to be used and the Generalised Axiom of Revealed Preference (GARP).

Definition 1. $\mathbf{q}_t$ *is directly revealed preferred to* $\mathbf{q}$*, written* $\mathbf{q}_t R^0 \mathbf{q}$*, if* $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$.

Definition 2. $\mathbf{q}_t$ *is directly revealed strictly preferred to* $\mathbf{q}$*, written* $\mathbf{q}_t P^0 \mathbf{q}$*, if* $\mathbf{p}_t'\mathbf{q}_t > \mathbf{p}_t'\mathbf{q}$

Definition 3. $\mathbf{q}_t$ *is revealed preferred to* $\mathbf{q}$*, written* $\mathbf{q}_t R \mathbf{q}$*, if* $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}_s$*,* $\mathbf{p}_s'\mathbf{q}_s \geq \mathbf{p}_s'\mathbf{q}_r$*,...,*$\mathbf{p}_m'\mathbf{q}_m \geq \mathbf{p}_m'\mathbf{q}$*, for some sequence of observations* $(\mathbf{q}_t, \mathbf{q}_s, ..., \mathbf{q}_m)$*. In this case, we say that the relation* R *is the transitive* *closure of the relation R^0.*

Definition 4. $\mathbf{q}_t$ *is revealed strictly preferred to* $\mathbf{q}$*, written* $\mathbf{q}_t P\mathbf{q}$*, if there exist observations* $\mathbf{q}_s$ *and* $\mathbf{q}_m$ *such that* $\mathbf{q}_t R\mathbf{q}_s$, $\mathbf{q}_s P^0\mathbf{q}_m$, $\mathbf{q}_m R\mathbf{q}$.

Definition 5. *Data can be said to satisfy the Generalised Axiom of Revealed Preference (GARP) if* $\mathbf{q}_t R\mathbf{q}_s \Rightarrow \mathbf{p}_s'\mathbf{q}_s \leq \mathbf{p}_s'\mathbf{q}_t$*. Equivalently, the data satisfy GARP if* $\mathbf{q}_t R\mathbf{q}_s$ *implies not* $\mathbf{q}_s P\mathbf{q}_t$.

Afriat's Theorem[5] summarises the correspondence between revealed preference (GARP), Afriat's inequalities and the existence of a utility function, with all of the desired properties.

Afriat's Theorem: *The following statements are equivalent:*

(1) There exists a non-satiated utility function which rationalises the dataset.

(2) There exist numbers $U_s, U_t, \lambda_t \geq 0$*, for* $s, t = 0, ..., T$ *such that*

$$U_s \leq U_t + \lambda_t \mathbf{p}_t' (\mathbf{q}_s - \mathbf{q}_t)$$

(3) The data satisfy the Generalised Axiom of Revealed Preference (GARP).

(4) There exists a non-satiated, continuous concave, monotonic utility function which rationalises the data.

Statements (1) and (4) show that, if a dataset is consistent with any non-trivial utility function, then it is also consistent with a well-behaved utility function[6]. Afriat's Theorem also gives two methods with which we can check for the existence of a set of preferences underlying a given finite dataset: either we can check the data for violations of GARP (statement (3)), or we can check to see if we can compute numbers $U_s, U_t, \lambda_t \geq 0$, which satisfy the

[5]The term seems to be Varian's (1982).

[6]Varian ((1982), p. 946) also points out that this means that violations of concavity, continuity and monotonicity cannot be rejected by a finite dataset since the violations can occur in the "gaps" between observations.

Afriat inequalities (statement (2)). Statement (3) is computationally somewhat arduous[7] since it requires the solution of a linear programming problem with $2n$ variables in n^2 constraints in which the number of constraints rises with the square of the number of variables. On the other hand, checking GARP is quite straightforward using Warshall's algorithm[8] to check for violations in transitive cycles.

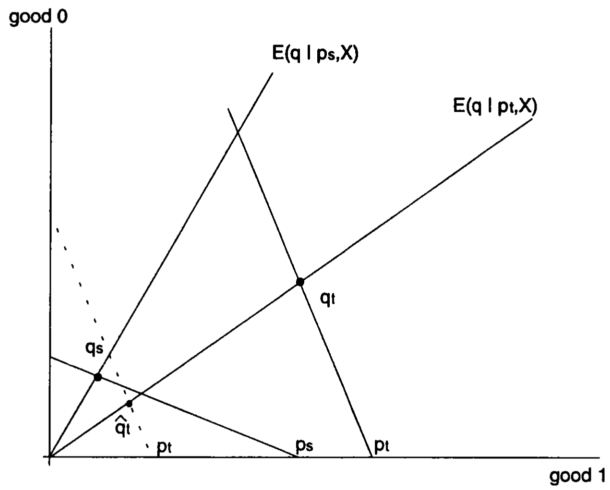
In practice the usefulness of revealed preference tests has been somewhat limited by the data to which these techniques have typically been applied. In general the problem seems to be that revealed preference restrictions often do not reveal very much about the preferences which might be consistent with a given dataset. Usually applications of tests of the axioms of revealed preference have been carried out with aggregate demand data. This has presented a number of problems. First, with aggregate data, 'outward' movements of the budget plane are often large enough and relative price changes are typically small enough that budget lines rarely cross (this has been pointed out by Varian (1982), Bronars (1987) and Russell (1992)). Thus aggregate data may lack power to reject revealed preference (GARP) conditions. This is illustrated in Figure 1.1.

In this example there are two goods and two price regimes (s, t). The (hypothetically homothetic) straight lines from the origins show the expansion paths for demands as total spending grows under each set of prices. Since the budget shares of each good is larger under circumstances in which their relative prices in higher, these demand patterns would be hard to rationalise with well behaved preferences. Suppose that growth in spending between the two price regimes (periods) is such that we observe the price/quantity com-

[7]Although Varian (1982) provides an algorithm which will compute numbers satisfying the Afriat Inequalities as part of his proof of Afriat's Theorem.

[8]Warshall (1962).

Figure 1.1: Spending growth and the power of revealed preference tests.



binations $(\mathbf{p}_s, \mathbf{q}_s)$ and $(\mathbf{p}_t, \mathbf{q}_t)$. Since $\mathbf{p}_t'\mathbf{q}_t > \mathbf{p}_t'\mathbf{q}_s$ the bundle $\mathbf{q}_t$ is revealed strictly preferred to $\mathbf{q}_s$. There is no violation of GARP simply because the $\mathbf{q}_t$ bundle contains more of both goods than does $\mathbf{q}_s$ and is therefore preferred through the monotonicity of utility. If instead we had observed less spending growth (say the bundles $(\mathbf{p}_s, \mathbf{q}_s)$ and $(\mathbf{p}_t, \widehat{\mathbf{q}}_t)$), then this would have revealed an inconsistency and a violation of GARP since $\mathbf{p}_t'\widehat{\mathbf{q}}_t > \mathbf{p}_t'\mathbf{q}_s$ and $\mathbf{p}_s'\mathbf{q}_s > \mathbf{p}_s'\widehat{\mathbf{q}}_t$. Growth in total spending can thus swamp relative price movements and reduce the ability of a dataset to reject GARP.

The second problem with empirical tests of GARP is that it has proven difficult to devise tests of the significance of rejections as the data used are usually aggregate or average demands with unknown variance. All of these issues are discussed and solutions are put forward in Chapter 3.

Revealed preference tests thus provide a potentially workable, albeit probably problematic, approach to testing for the existence of a set of preferences which are consistent with a given dataset. Suppose that a given dataset of, say, average demands has been tested for violations of GARP and none have been found. By Afriat's Theorem it is known that these data are consistent with a utility maximisation model and Diewert (1978), for example, shows that the cost function must therefore exist and have the desired properties outlined above. A true cost of living index must exist. But we are not much further forward; we still need to compute an index (and we don't know anything about $c(\cdot)$ other than its general characteristics). In particular we don't know whether any index which we do calculate will be unique or whether the cost of living index varies by household income and demographics and should be calculated separately for different household groups.

1.2 Uniqueness

A true cost of living index defined in equation (1.1) above usually depends upon utility. If the cost function is known then it can be evaluated under any utility level. But the resulting index evaluated at u_x will, in general, be different from the index evaluated at u_y (as discussed in relation to equation (1.3) above). Both indices are 'true', and neither is any truer than the other; they simply measure different things. Under what circumstances with there be a unique index which is independent of the reference welfare level ?

Malmquist (1953) first proved that homotheticity was both necessary and sufficient for the existence of a unique and unambiguous cost of living index[9]. Even if households all have the same preferences and face the same prices, price changes will affect their economic welfare differently if variations in income or total expenditure affect spending patterns. The only circumstance, then, under which one can speak accurately about *the* cost of living index is one in which this is not the case, i.e. when household expenditure patterns

[9] Afriat (1981) provides a revealed preference-based proof.

do not vary with total spending. If preferences are homothetic, then the cost function takes the form

$$c(\mathbf{p}_t, u_R) = a(u_R)\, b(\mathbf{p}_t)\,. \tag{1.4}$$

The true index evaluated at prices $\mathbf{p}_t$, $\mathbf{p}_{t+1}$ and welfare u_x then becomes

$$\frac{c(\mathbf{p}_{t+1}, u_x)}{c(\mathbf{p}_t, u_x)} = \frac{b(\mathbf{p}_{t+1})}{b(\mathbf{p}_t)}$$

and this is obviously identical to the index evaluated at the same prices and welfare level u_y;

$$\frac{c(\mathbf{p}_{t+1}, u_x)}{c(\mathbf{p}_t, u_x)} = \frac{b(\mathbf{p}_{t+1})}{b(\mathbf{p}_t)} = \frac{c(\mathbf{p}_{t+1}, u_y)}{c(\mathbf{p}_t, u_y)}$$

since the index is independent of utility. This means that even if two households have identical non-homothetic preferences the relative cost of one household achieving, say, its base period welfare level will be different to the relative cost of other household maintaining its different base period welfare; i.e. the welfare effects will differ. If relative prices vary and the homotheticity condition does not hold then a single index cannot be appropriate for every utility level and thus for every household. The answer is to abandon homotheticity and allow indices to vary by household income/spending and by demographic characteristics.

One of the earliest empirical studies of household budgeting was Engel's famous analysis of the consumption of poorer households[10] in which he rejected homotheticity. He concluded that the proportion of spending allocated to necessities declines as total expenditure and income increases. Homotheticity, on the other hand, implies unitary income elasticities and thus rules out the idea of luxuries and necessities defined in the Engel sense.

[10] Engel (1895).

The question of uniqueness may seem rather minor – surely an index based on average demand will be right on average ? However, official prices indices are put to many uses to which they may not always be appropriate if preferences are not homothetic. For example, the inflation adjustment to state benefits and pensions utilises the headline figure which is based upon average demands. However, the demands of those in receipt of the benefits and pensions which are being up-rated may be far from average, and therefore they may be either over or under-compensated by such an up-rating exercise. Deaton and Muellbauer (1980), cite the more drastic example of large fluctuations in the relative prices of staple foods in less developed countries. They cite Sen's (1977) description of the effects of such price variations on the Bengal Famine in which between three and five million people died and where the price system cause a dramatic change in the distribution of real consumption and welfare which would not have been apparent using measures based on average demands. Marshall (1890) also puts the point.

> "A perfectly exact measure of purchasing power is not only unattainable, but even unthinkable. The same change of prices affects the purchasing power of money to different persons in different ways. For him who can seldom afford to have meat, a fall of one-fourth in the price of meat accompanied by a rise of one-fourth in that of bread means a fall in the purchasing power of money ... While to his richer neighbour, who spends twice as much on meat as on bread, the change acts the other way."

To conclude, tests of revealed preference restrictions might show that a cost function compatible with a dataset exists. But, unless the cost function derives from homothetic preferences, then there is no unique true cost of living index; rather the index will vary with utility. This also means that the same price changes will affect the welfare of different households differently according to their total spending and their demographics, so allowing indices to vary by these characteristics may be sensible. The extent to which the index varies between household types is an empirical one which may, potentially, have important policy implications. This is subject matter of Chapter 2 and is revisted in Chapter 4. The question of how to compute the index or indices remains.

1.3 Construction: Estimation and Approximation

There are two main approaches to the computation of a cost of living index; parametric estimation and approximation[11].

1.3.1 Estimation

The parametric approach tackles the measurement issue head-on by attempting to estimate the cost function via a system of demand equations. Examples of this approach include Braithwait (1980) and Banks *et al* (1994). This is not easy and requires a great deal of data. Even when such data is available the results obtained may not live up to the restrictions which theory places upon them. The main problem is often the reliable estimation of price effects. This is because cross-sectional variation in prices is usually not observed[12]

[11] In fact there is also a third: this is based on the parametric estimation of not necessarily integrable demand systems upon which curvature is imposed locally. See Vartia (1983) for an early example and Ryan and Wales (1996) for a recent one.

[12] Exceptions may include those datasets which record both quantities and expenditures and which therefore allow for the calculation of unit values. These however, are not prices in the pure sense because they represent a combination of a price level and a quality-

and so price effects have to be recovered from the lesser amount of time series variation in published retail price sub-indices. Demographic effects and income effects, on the other hand, are usually fairly easy to estimate with precision since these parameters can exploit information in repeated cross-sections in which the dimension of the data is usually much greater than it is for price data. More importantly, since the hypothesis of (a) the existence of a utility function and (b) a specific functional form are jointly imposed in parametric models, we have no way of knowing whether violation of, for instance, Slutsky symmetry, is due to rejection of (a) or (b) or both. This is why revealed preference tests are used; they don't presuppose either the existence or the form of preferences[13].

1.3.2 Approximation

Overall, the practice of estimating full demand systems in order to recover the cost function has generally been thought to be best avoided, and methods with less stringent data and computational requirements have been sought. This has given rise to the literature on price indices. There are essentially three classes of price index: first order approximations, exact indices and superlative indices.

First order approximations (like the Laspeyres ($P_L = \mathbf{p}_1'\mathbf{q}_0 \;/\; \mathbf{p}_0'\mathbf{q}_0$) and

reflective element. Deaton (1987, 1988) suggests a method for estimating price elasticities from such data but in doing so makes a functional identification restriction and assumes that the amount of variation in price levels is zero within geographical clusters. Crawford, Laisney and Preston (1996) suggest a generalisation of Deaton's approach.

[13]Of course, while revealed preference tests do away with the jointness with respect to hypothesis (b), there are likely to be, in any empirical application of the test, a large number of supporting auxiliary hypotheses (mostly econometric) which make a 'crucial experiment' of this type impossible (see Cross (1982) for a discussion of the Duheme-Quine thesis). As Popper (1963) says " in these case it is sheer guesswork which of the ingredients should be held responsible for falsification". Nevertheless, in dropping one important auxiliary assumption, revealed preference tests may be a step in the right direction.

the Paasche ($P_P = \mathbf{p}_1'\mathbf{q}_1 \ / \ \mathbf{p}_0'\mathbf{q}_1$) indices) require no assumptions on functional forms and correspond to first order approximations to any cost function (based at u_0 and u_1 respectively).

For example, taking the first three terms of a Taylor series expansion of the cost function about $\mathbf{p}_0$ corresponding (i.e. based upon base period (period 0) welfare) to the Laspeyres index gives

$$c(\mathbf{p}_1, u_0) = \mathbf{p}_1'\mathbf{q}_0 + \frac{1}{2}\sum_i\sum_j \left.\frac{\partial q_0^i}{\partial p_0^j}\right|_{u=u_0} \left(p_1^i - p_0^i\right)\left(p_1^j - p_0^j\right) \tag{1.5}$$

If prices don't vary much between periods 0 and 1, or if prices move proportionally, or if there is little substitution, then the final term will tend toward zero so that

$$P_T = \frac{c(\mathbf{p}_1, u_0)}{c(\mathbf{p}_0, u_0)} \simeq \frac{\mathbf{p}_1'\mathbf{q}_0}{\mathbf{p}_0'\mathbf{q}_0} = P_L \tag{1.6}$$

That is, the Laspeyres approximates the true u_0-referenced cost of living index. The case of the Paasche end period-referenced index is entirely analogous. This property of these indices (that they provide first order approximations to their corresponding base or end period weighted true counterparts), and the ease with which they can be calculated (particularly the Laspeyres which can condition on currently available information) makes them popular choices. Unfortunately, empirical studies usually find ample evidence of substitution effects[14] (particularly for large relative price changes) rendering the approximations afforded by the Paasche and Laspeyres less acceptable.

Exact indices are simple indices which correspond to particular functional forms. An example is Fisher's Ideal index $\left(P_F = (P_L P_P)^{1/2}\right)$ which corresponds exactly to homogeneous quadratic preferences[15]. As with first order

[14]Blundell *et al* (1996), for example, find evidence of large own- and cross-price substitution effects in U.K. Family Expenditure Survey data.

[15]Proved by Frisch (1936) amongst others.

approximations, exact indices are chosen for the property that ratios constructed from their cost functions depend upon prices and quantities alone and so are easy to apply (i.e. they don't require the parameters of the cost function to be estimated). Simple first order price indices can also be exact. For example, as well as being first order approximations, the Paasche and Laspeyres are also exact indices if preferences are Leontief[16].

Superlative[17] indices are also exact for certain forms of preferences. These, however, have the added attractive feature that they can serve as good local *second* order approximations to any cost function; i.e. they are defined by the property that their cost functions have flexible functional forms. The Törnqvist index $\left(P_{TQ} = \prod_{i=1}^{n} (p_1^i/p_0^i) \exp\left(\frac{1}{2}(w_0^i + w_1^i)\right)\right)$ (where w_1^i is the budget share of good i in period 1) is an example of a superlative index which is the geometric mean of two true indices when the cost function is a general translog[18].

The main benefit of all of these approximations over full parametric estimation is that none require any more information than the usually readily observable price and demand vectors. The problem is that, unless underlying preferences are either exactly those assumed in the construction of the index (or a close approximation to them), then the resulting index may not be a good measure of the true effects of price changes on welfare. These issues are discussed further in Chapter 4 which sets out a method of improving bounds to a true index without making functional form assumptions.

[16]Proved by Pollak (1971) and Samuelson and Swamy (1974).
[17]The term is Diewert's (1976).
[18]Proved by Diewert (1976).

1.4 Problems of Measurement and Bias

So far, methods for determining the existence of a cost of living index (or rather the consistency of a dataset with the concept), its uniqueness and methods for its measurement or approximation have been discussed. Even if the existence of some cost of living index cannot be rejected, its uniqueness or otherwise has been established and a method for its estimation or approximation chosen and applied, a number of problems of measurement and bias still confront the resulting index. These possible sources of mismeasurement have become highly topical matters of debate recently, particularly so in the United States. This was largely stimulated by a calculation which showed that if indexed welfare programs and taxes in the U.S. were reduced by 1 percentage point in line with a recalculated Consumer Price Index (C.P.I.), the level of the budget deficit would be lower by $55 billion after five years[19].

There are two main types of bias: survey biases which relate to issues of price survey design and include outlet bias and formula bias, and economic biases. These include substitution bias, which concerns the extent to which official indices approximate a true index by accounting for behavioural changes in response to relative price changes, and the more complicated effects of quality change and the arrival of new goods which affect true indices as well as approximations. These sources of bias and mismeasurement are discussed in turn below, beginning with the survey biases.

[19] Alan Greenspan in a prepared statement to the U.S. Congress, January 1995. Cited in Moulton (1996).

1.4.1 Survey Biases

Formula Bias

This bias arises in prices indices in which a proportion of the sample prices are rotated each month or quarter. For new items brought into the price survey in this way (say in June) their base (January) weights are not observed. These are imputed by deflating June expenditures by the price increase of the goods which the new items are replacing. The way that this is done in the U.S. C.P.I. gives too much weight to goods which are on sale in the month of their introduction. Once these goods come off sale the price rise is likewise given too great a weight introducing an upwards bias. Similarly, too low a weight is given to items which are not on sale in the month in which they are rotated into the sample. When these items go on sale the price falls are not given their true weight.

For example[20], suppose that a fairly homogeneous item is sampled at three outlets each month. Suppose that in June two outlets are selling at £2 and the third at £1.25, that they each have base (January) prices of £1 and expenditure weights of £100 so that base quantity weights are 100 units at each outlet. Next month the prices are the same in the three outlets but one of the outlets selling at £2 now sells at £1.25, and the one selling at £1.25 now sells at £2. As the outlet weights are the same there is no inflation. In August the three new outlets which happen to sell at the same three price are sampled along side the orginal sample. Inflation is still zero. In September prices are the still same. The Bureau of Labor Statistics (B.L.S.) wish to calculate base prices for the new sample of outlets so that they can weight the prices in new outlets. For some reason, since there is no inflation between the new outlets and the old, the B.L.S. use the new outlet prices in August

[20] This example is from Moulton (1996).

(£2, £2, £1.25). This gives new base quantity weights for the outlets since, assuming that these outlets have the same £100 expenditure weights as the first three sampled, this means that the new quantity weights are 50 units (100/2) for two of the outlets and 80 units for the third (100/1.25). Between August and September the item comes off sale in the outlet selling at £1.25 (with weight 80) and goes on sale at one of others at the same price (with a weight of 50). There has not been any inflation but the new index is 1.075 simply through the change in the outlet weights. Formula bias such as this normally disappears within two months of re-sampling. This problem was noticed in the U.S. C.P.I. by Reinsdorf (1993).

Outlet Bias

This is another survey based problem which also centres on sample rotation. In this case the problem concerns the growth of discounted retail outlets, many of which price below established shops. Piachaud and Webb (1996) provide evidence of this for food retailers in the U.K., while trade estimates of growth in the market shares of warehouse stores in the U.S. are running at about 0.7% per year[21]. Given fixed sampling weights (even with periodic rotation) this means that price observations from such outlets will be under weighted compared to older, more established shops with possibly higher prices. This type of bias is similar to the more general problem of substitution bias which is discussed next.

[21] Moulton (1996).

1.4.2 Economic Biases

Substitution Bias

Substitution bias tends to be a feature of indices which are either first order approximations to true indices or exact indices which are based on the wrong functional form (i.e. one that is inconsistent with the data). Most official prices indices like the C.P.I. or the U.K. Retail Price Index (R.P.I.) are (chained) Laspeyres indices and the substitution bias inherent in such a fixed based weight index was set out in Konüs's famous inequality[22]

$$P_L = \frac{\mathbf{p}_1'\mathbf{q}_0}{\mathbf{p}_0'\mathbf{q}_0} \geq \frac{c(\mathbf{p}_1, u_0)}{c(\mathbf{p}_0, u_0)} \tag{1.7}$$

That is, the Laspeyres index is always greater than or equal to the corresponding base-referenced true cost of living index. The intuition behind this result (and the analogous one which says that the Paasche index is always less than or equal to the corresponding end-period reference true cost of living index) is straightforward. The fixed bundle $\mathbf{q}_0$ may have been the cheapest way of reaching u_0 under the original set of prices $\mathbf{p}_0$. But this is not necessarily the case once prices have changed to $\mathbf{p}_1$ (unless preferences actually *are* Leontief). This result is usually explained with recourse to diagrams with indifference curves (see for example Deaton and Muellbauer (1980) p. 171) but all that is required is the reflexivity of preferences as this ensures that $\mathbf{q}_0$ is at least as good as itself ($\mathbf{q}_0 R^0 \mathbf{q}_0$). One way in which to purchase $u_0(\mathbf{q}_0)$ is to purchase $\mathbf{q}_0$ itself at a cost of $\mathbf{p}_1'\mathbf{q}_0$, and hence the minimum cost of purchasing u_0 at $\mathbf{p}_0$ $(c(\mathbf{p}_1, u_0))$ cannot exceed $\mathbf{p}_1'\mathbf{q}_0$. The argument that the Laspeyres index overstates the true change in the cost of living follows immediately. Superlative indices like the Törnqvist do not suffer from substitution bias but require additional information on the current demand

[22]Konüs (1924).

vector as well as last period's.

Empirical comparisons between the Laspeyres and a full demand system are rare. One example is Braithwait (1980). More recently studies have compared Laspeyres to superlative indices to try to gauge the extent of substitution bias. In the U.S., Manser and McDonald (1988) found that the Laspeyres index tended to grow 0.14 to 0.22 percentage points faster per year than alternative measures which are free of substitution bias like the Törnqvist or Fisher's Ideal index. However, they found that their results were sensitive to the aggregation of their data. Cunningham (1996) sensibly suggests that substitution bias is sensitive to the frequency with which the index is rebased which up-dated commodity weights. In the U.S. the C.P.I. has been rebased every 10 years. In the U.K. the R.P.I. is rebased annually so the R.P.I. might be expected to exhibit less substitution bias. He suggests a plausible level of bias for the U.K. R.P.I. of around 0.05 percentage points per annum. This issue is further discussed in Chapter 4.

Quality Bias

Quality change poses a problem for all indices, even true cost of living indices. The problem is that even seemingly homogeneous goods may be subject to change over time. Gorman's (1956) original paper in this field, for example, highlighted this when he chose to look at eggs in his early work on adjusting prices for quality differentials. Price differentials between varieties of a good within a period are usually ascribed to quality variations. Price changes over time are a combination of changes in the characteristics of different varieties, and general shifts in the price level of all varieties. Disentangling these effects is a particular problem in the construction of cost of living indices. If quality is seen as increasing over time, failure to strip out the quality-related element of the time series price changes may lead to an over-estimate of the true index.

Most national statistical offices attempt to make some rudimentary adjustments to prices to account for quality changes when a data-collector can no longer find a price of an item with a given set of specified attributes. There are a number of methods and assumptions employed. For example, the prices of old and new items may be 'linked', i.e. all of the price difference between the old and new version of the good is ascribed to quality. Alternatively there may be a zero quality change assumption and the prices of the two goods are then directly compared. A third method is to assume that the price of the substitute good has changed at the same rate as the price of other similar goods. It is then omitted for the month of the specification change and the average price-change for similar goods of constant quality is used instead. Finally, some attempt may be made to measure quality change via changes in producer costs.

Then are two main approaches to measuring quality change which economists have put forward. The first is the characteristics model of Gorman (1956) later developed by Lancaster (1966), and the repackaging model of Fisher and Shell (1971)[23] later generalised by Muellbauer (1975). Both of these techniques take the view that the quality of a good is a function of its observable characteristics and that changes in price which are to do with quality can be estimated via the correlation of price changes with changes in product specifications. Each approach, however, lends itself more easily to one or other of the two main variants of applied hedonic regression techniques. The Gorman-Lancaster model fits the use of single years of repeated cross-section linear regressions of prices on characteristics, the Fisher and Shell model is more in sympathy with log-linear regressions using a pooled

[23]The idea goes back to Prais and Houthakker (1955).

time series of cross-sections. The Gorman-Lancaster approach is to regress price on characteristics for a number of years thereby seeking to estimate the shadow values of different characteristics. These prices are then combined into a Laspeyres index with fixed characteristics weights from the product as it existed in the first year of data. The Fisher and Shell approach is to introduce the quality of a good directly into the utility function in a way which pre-multiplies the quantity of the good in question, or alternatively deflates its price in the cost function. This is intuitively similar to getting more for your money over time (or paying less for the same amount of the good). In this case the ratio of prices of different varieties should be identical to the ratios of their qualities. This then lends itself to regressing log prices on time dummies (to pick up the quality constant price variation) and a (usually) linear function of characteristics data from pooled cross-sections of data. The Fisher and Shell approach to accounting for quality changes in discussed in Chapter 3 in which a revealed preference analogy is put forward.

Gordon (1990) estimated that the inflation rate of consumer durables in the U.S. C.P.I. was biased upwards by 1.5 % per year over the period 1947 to 1983 due to its current rule-of-thumb quality adjustment practices. Berndt, Griliches and Rappaport (1995) suggest that Gordon's estimates for the micro-computer component of the index was overly optimistic by 2% per year. Cunningham (1996) calculates that Gordon's estimates imply a 0.2% to 0.3% annual upward bias to the U.K. R.P.I. If cars are included this will increase bias to between 0.25% and 0.35% per year.

New Goods Bias

New goods bias concerns both the timely introduction of new goods to the data from which price indices are computed, and the problem of accounting for the welfare effects of their introduction. The complication caused by new goods arises because, when the number of goods changes across time periods, the full price vector will not be observed in all periods. For example, in order to compare two periods when a new good is introduced in the second period (using either a cost-of-living index[24] or an approximation which includes the new good in the reference bundle[25]) it is necessary to calculate what the (virtual) price of the new good was in the first period.

John Hicks (1940) discussed the question of how to value new goods and, more generally, the issue of how to deal with rationed goods when constructing index numbers. He showed that the price of a rationed good in an index number should not be the actual price, but the price which would make the rationed level of consumption consistent with an identical unrationed choice. New goods are viewed as a special case of rationing, in which the ration level in the period prior to the one in which they first exist is zero. Thus the virtual prices for new goods would be those which "just make the demand for these commodities (from the whole community) equal to zero"[26].This approach captures the benefit of the introduction of a new good by imagining that its price has reached its period t value from a level in period 0 which was marginally above the maximum value of the good to consumers.

The usual parametric approach to estimating virtual prices proceeds by assuming a particular functional form for demand which is consistent with maximisation of a particular form for the utility function. A system of demand equations is estimated using data from periods in which all goods are

[24]To deal with new goods such an index would, of course, need to be based upon preferences which are complete and stable over time.

[25]A Laspeyres price index, for example, would not include the new good.

[26]Hicks (1940), p.114.

available, and this is used to predict the lowest price which would result in zero demand for the new good (for given income levels and demographics) in the period immediately prior to the one in which it first exists. A recent example of this technique is Hausman (1997). Typically this procedure gives very big estimates of the price fall or welfare gain taking place between non-existence and existence.

There are a number of possible problems associated with this approach. In particular, the estimate of the virtual price will be heavily dependent on the maintained hypothesis concerning functional form. Furthermore, determining the best functional form is difficult when non-nested models are being compared. In addition, parametric methods are reliant upon (possibly suspect) out-of-sample predictions of the demand curve to solve for the virtual price. This is because it is usually necessary to extrapolate the demand curve across regions over which relative price variations have never been observed in the data. An alternative approach which does not require functional form assumptions or extrapolation is described in Chapter 5.

As with quality change, new goods are normally dealt with in practice by linking the price of the new good and the most close previously available substitutes. Cunningham (1996) suggests that this, and the fact that the U.K. R.P.I. is a chained Laspeyres index in which new goods are given zero weight initially, combine to give a bias of between 0.06% per year and 0.1% per year. Moulton's (1996) view is that empirical studies have done little to identify a plausible range of bias due to new goods because of the problems outlined above.

1.5 Conclusion and Summaries

This chapter has introduced some of the issues which will be addressed in greater detail in later chapters. It highlighted some of the problems with the current state of research in these areas, most of which stem from the problem of having to have a maintained hypothesis regarding functional form for preferences. It began with the question of establishing the existence of cost functions consistent with price and demand data. This can, in principle be dealt with by tests of revealed preference conditions or by checking for the existence of numbers which satisfy Afriat's Inequalities but such tests often lack a statistical framework and may also lack power in practice.

Uniqueness was then discussed. Even if a cost function which can rationalise a given finite dataset of average demand does exist, unless preferences are homothetic then there will not be a unique and unambiguous cost of living index. The disparity between the rates of increase in the cost of living of different population and income groups will be greater, the greater the rate of price increase and relative price movements.

If the issues of existence and uniqueness can be satisfactorily addressed, the next question of is that of construction and measurement (either of a single index or of indices for different household types). Two main approaches are available: estimation of the cost function by means of an integrable demand system, or approximation. These were discussed in turn. Generally, because of the extensive data requirements and the possibility that the final estimates may reject (for example) Slutsky symmetry, direct estimation has typically been avoided. Various approximations (first order, exact and superlative) have been suggested, each with different properties. The extent to which any of them give good approximations to a true index depends upon either the validity of functional form assumptions, limits to substitution effects or the size of relative price movements. Superlative indices which are based on preferences which give second order approximations to any cost functions suffer less ill effects from substitution in response to relative price variations than other indices assuming that substitution effects are not negligible in the first place. The measurement of the extent of any inaccuracy in the more popularly used indices like the Laspeyres, however, depends upon the estimation of the true index itself and so is hard to gauge.

Whatever the approach used to come up with an index, all indices suffer from problems of bias introduced in the survey process, by quality change in the goods in the demand vector, and by the arrival of new goods. This has become an important topic of research particularly in view of all of the uses to which price indices are put. The main survey biases are formula bias (which over weights sale goods) and outlet bias. Outlet bias involves the problem of keeping up-to-date the price survey weights when consumers switch their shopping to discount stores. This is an intrinsically similar problem to substitution bias which affects first order approximations to true indices. Quality bias is introduced when the price increases which are due to improvement in the quality of an item are not removed from the index. New goods bias is a somewhat different problem which concerns the estimation of the virtual price of a new good for the period before the one in which it first exists. The distinction between what is a new good and what is simply an improved version of an existing good is a difficult one to make. Parametric solutions to both of these problems may give results which are heavily determined by the functional form chosen, but in the case of the estimation of the virtual price of a new good in particular this may be an especially worrying aspect.

These themes are picked up in the following chapters which are summarised next.

Chapter 2: Disaggregating the Cost of Living Index: An Empirical Investigation for the U.K. 1979 to 1992.

This chapter is concerned with an empirical analysis of the uniqueness of cost of living indices: do cost of living indices vary with income and household characteristics ? It computes and compares cost of living indices for different household types and income levels to look at differences in their rate of increase. The approach adopted is to compute a series of chained Törnqvist indices. Particular attention is paid to variations with respect to income group, and major demographic characteristics. This chapter also looks at the effects which indirect tax changes over the period have had on differential inflation measures. Housing costs are a major component of most official price indices. Two different methods of computing housing costs are discussed (the current method applied in the UK Retail Prices Index and the user cost of housing approach which *inter alia* includes the expected annual capital gains) and their effects compared. The likely differential effects of quality and new goods bias on the results are also discussed.

Chapter 3: Revealed Preference Tests and Nonparametric Engel Curves

This chapter concerns the questions of existence and of correcting prices for quality or taste change. This chapter applies nonparametric statistical analysis to revealed preference tests of consumer behaviour. It exploits the idea that price-taking households in the same market face the same relative prices, in order to smooth across the demands of individual households for each common price regime. This is achieved by means of nonparametric regression techniques. This approach has two major benefits. Firstly, it is shown to provide a stochastic structure within which to examine the consistency of data and revealed preference theory. Secondly, knowledge of nonparametric Engel curves can be used to recover expansion paths which allow the power of revealed preference tests to be improved by accounting for period-to-period spending growth.

A method which maximises the power of the test of revealed preference for any given preference ordering is presented. An application is made to a long time series of repeated cross-sections from the 1974-1993 U.K. Family Expenditure Surveys. The consistency of this data with revealed preference theory is examined. Where rejections do occur, a method which allows for suitable adjustments to prices for quality or taste changes such that the data are made consistent with the existence of well-behaved preferences and hence the existence of a cost function is explored and applied.

Chapter 4: Revealed Preference Methods for Bounding True Cost of Living Indices

This chapter is concerned with the construction of approximations to true cost of living indices, the extent of substitution bias and the issue of uniqueness once more. This chapter suggests a revealed preference method which, without the need for functional and parametric assumptions, allows two-sided bounds to be placed on a given true index recovered from average demands. These bounds, are shown to be as least as tight as the classical nonparametric bound and are also the tightest bound obtainable given only the testable assumption that a given finite dataset of prices and average demands was generated by a well-behaved utility function. The algorithm which computes the bounds is shown to provide a powerful method of performing this test.

These ideas are applied to UK micro data from 1974 to 1993 (the same data used in Chapter 3) in order to identify sub-periods within which coherent indifference curves are recoverable at a number of quantile points in the total expenditure distribution. The improved bound for average demands is compared to other popular approximations and classical nonparametric bounds and is used to assess how closely these indices seem to approximate the true index. In particular the bounds on the true index are used to assess the extent of the substitution bias which is incurred by using simple approximations like the Laspeyres. The bounds on the true index are also compared to the Törnqvist. This procedure is also used to provide evidence on differences in cost of living indices by total expenditure level (i.e. non-homotheticity) of the household, and to estimate the extent to which substitution bias may vary with total expenditure.

Chapter 5: New Goods Bias in Cost of Living Indices: A Revealed Preference Approach

As outlined above, the calculation of cost-of-living or price indices is complicated by changes in the number of goods available between periods. This is because the full set of prices is not observable in every period. When, for example, a new good is introduced, it is necessary to impute its virtual price for the period before it existed. The usual approach is to set this virtual price at the level which would just have driven demand for the new good to zero in that period. It is shown that revealed preference restrictions can be used to calculate the lower bound on the virtual price of a new good which is consistent with the maintained hypothesis that the data were generated by the maximisation of a well-behaved utility function. It is also shown how this bound can be improved through use of nonparametrically estimated expansion paths. This approach has two principal merits compared to parametric estimation of virtual prices. Firstly, it does not require a maintained assumption regarding the form of the utility function. Secondly, it is computationally simple. These ideas are applied to U.K. household expenditure data to calculate the virtual price of the U.K. National Lottery at a point one year before its launch, and the effect of its inclusion in measures of annual inflation is examined. The variation of the virtual prices / welfare effects of the introduction of the good with respect to household income is also examined.

Chapter 2

Disaggregating the Cost of Living Index: An Empirical Investigation for the U.K. 1979 to 1992

2.1 Introduction

As discussed in the introductory chapter, even if all households have the same preferences, the only circumstance under which one can speak accurately about *the* cost of living index is one in which those preferences are homothetic. The empirical evidence ever since Engel (1895) is quite clear that this property does not hold in practice. As a result, relative price movements mean that, since even demographically similar households consume goods and services in different proportions according to the size of their total budget, then each household will have a different cost of living index. Further, if different households with different demography, for example, have different preferences then cost of living indices will vary by both household characteristics and income. This chapter concerns the pattern and extent of these variations in cost of living indices between households according to income and household composition.

This is not a new topic for research. The first study which explicitly considered the idea that different individuals experience different cost of living changes seems to date from 1707 when, in response to a student at Oxford, William Fleetwood calculated how much it would cost to buy a certain bundle of goods at current (1707) prices compared to the same bundle of goods at the prices prevailing 400 years earlier during the reign of Henry VI. The bundle of goods he chose was supposed to reflect the typical annual spending pattern of an 18th century student – 4 hogsheads of ale, etc. His conclusion was that "£5 in H VI 's days " would make a student "fall as rich as he who has now £20". Afriat (1977) points out that he might have taken another individual as an example – he gives the example of a civil servant - and gives the figures £25 and £50 as the costs of achieving the civil servant's standard of living in each period. On the basis of these data (5,20) and (25,50) a consumer price index for the average would be 2.3. This would be good news for the civil servant and bad news for the student since if their incomes were uprated according to the consumer price index the civil servant would receive £58.30 in 1707 compared to £11.70 for the student. One would be over-compensated, the other under-compensated. In the 1974 Reith Lectures, Dahrendorf argued that

> "at a time of general expansion (inflation), average figures gave some indication of a trend felt by most ... such averages are now losing much of their meaning. The end of the average means that much more attention will have to be given to differences."

This is particularly the case during periods of accelerated inflation: as Afriat (1977) puts it "(this) index number problem, which has been taken up with interest at times and dropped with exhaustion or boredom at others, is always present,and it has usually been taken up during periods of inflation"[1].

This forms the topic of this chapter which calculates and documents recent trends in inflation rates for different types of households and different income groups within the U.K. between 1979 and 1992. In particular, it looks at the differing effects which indirect tax reforms over the period, of which there have been many, may have had on the rate of increase in the cost of living for different groups. It also presents the effects of two alternative methods of calculating housing costs for owner occupiers: the user cost method, and the method currently used in the construction of the U.K. R.P.I.. The plan of this chapter is as follows. Section 2.2 presents a discussion of the properties of some alternative cost of living indices, the method and the data to be used in the study. Section 2.3 focuses on patterns of non-housing inflation for different income groups and demographic groups and section 2.4 looks at the influence of indirect tax reforms over the period on non-housing inflation. Section 2.5 examines the results of the inclusion of housing costs in the analysis. Two possible methods of calculating housing costs are discussed and alternative all-item cost of living indices are calculated using both measures. Section 2.6 concludes.

2.2 Household Cost of Living Indices

2.2.1 Non-Homotheticity and Differential Indices

True indices (defined in equation (1.1) above) will typically depend upon the reference welfare level which will itself depend upon preferences, prices and total income. As discussed in Section 1.2, even if households all have the same preferences *and* face the same prices, price changes will affect their welfare

[1] Afriat (1977), preface, p. xi.

differently if income or total expenditure affect spending patterns. The only circumstance under which one can speak accurately about *the* cost of living index is one in which this is not the case. In other words, homotheticity is both a necessary and sufficient condition for the existence of a single cost of living index, even within a demographically homogeneous group of households (this was first proved by Malmquist (1953)). If relative prices vary and this condition does not hold, and there is ample evidence that it does not, then a single index cannot be appropriate for every household. Further, the greater the extent of relative price changes and the greater the disparity in spending patterns between households, then the greater is the variation around the average measure. If we now allow preferences to vary with respect to household composition, then the welfare effects of price changes will vary with both income and demography.

One of the earliest serious studies of household budgeting was Engel's famous study of the consumption of poorer households[2] in which he found plenty of evidence that homotheticity does not hold. As discussed in section 1.2 of Chapter 1, Engel concluded that the proportion of spending allocated to necessities (like food) declines as total expenditure and income increases. Homotheticity, on the other hand, implies unitary income elasticities and thus rules out the idea of luxuries and necessities defined in the Engel sense.

To illustrate some *prima facie* evidence, consider the data on a typical necessity: domestic fuels. Figure 2.1 shows the Engel curve[3] for domestic fuel drawn nonparametrically using U.K. data from the 1992 Family Expenditure Survey (F.E.S.). The fuel share of total spending declines as the logarithm

[2] Engel, (1895).

[3] The proportion of the total household budget allocated to fuel against log total expenditure.

Figure 2.1: The Engel curve for domestic energy, FES, 1992

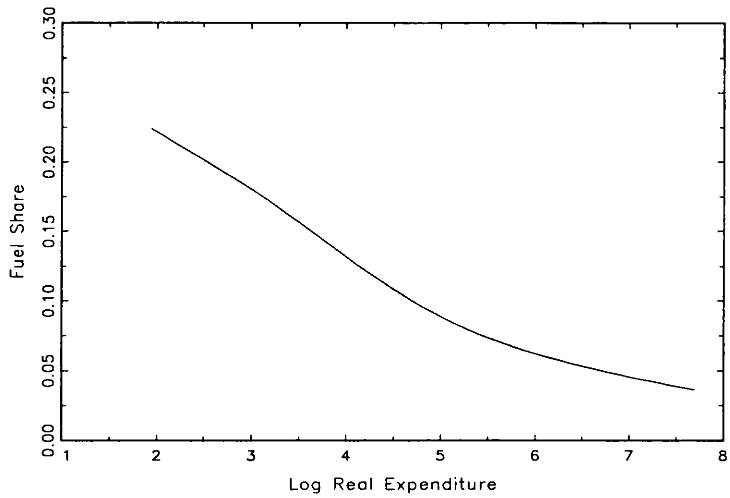


Figure 2.2: The relative price of domestic fuel, UK, 1978-92.

of total expenditure increases. This downward-sloping Engel curve is typical of goods that are usually thought of as necessities; poorer households with lower total expenditure spend a greater proportion of that total on necessities like fuel and food than do richer households[4]. Figure 2.2 shows the price of domestic fuels relative to the all-item retail price index from 1978 to 1992. Figures 2.1 and 2.2 are sufficient to show the existence of systematic differences between the cost of living of different households. The relative price movements illustrated will have a greater effect on the cost of living of households which consume more fuel as a proportion of their total budget than others. The data appendix to Chapter 3 which also uses U.K. F.E.S. data, shows that Engel curves are neither flat nor always linear for a wide range of commodities.

Table 2.1: The post-war experience in the UK, 1939-1980's.

1939-45	The prices of basic commodities were kept low favouring poorer households. The prices of luxuries were much increased (where they were available)
1945-61	The relative prices of necessities rose fast, introducing a substantial anti-poor bias
1961-66	Prices of luxuries caught up somewhat. Anti-poor bias in inflation reduced.
1966-71	The bias against the poor disappeared altogether by the beginning of the 1970's.
1970's	The prices of necessities rose despite food subsidies; later, with membership of the Common Market, dismantling of food subsidies and the oil crisis, the cost of necessities rose again.
Early 1980's	Some evidence of anti-poor bias.

[4]Luxuries are usually defined by upward-sloping Engel curves, necessities by downward-sloping Engel curves.

There have been a number of studies[5] which have analysed inter-household differences in the cost of living and in inflation since the Second World War. The general picture of relative inflation is presented in Table 2.1. As can be seen, the postwar period can be divided into fairly well defined phases in which increases in the relative prices on necessities pushed up the cost of living of poorer households faster than that of richer households. This chapter provides evidence on the 1980's and early 1990's, the run of data nests the papers of Bradshaw and Godfrey (1983) and Fry and Pashardes (1986) which found evidence that average inflation measures under-stated the rate of increase in the cost of living index for poorer households in the early 1980's.

2.2.2 Choosing an Index

Even given that cost of living indices will vary according to household income and composition, the calculation of true cost of living indices even for a homogeneous group of households requires that the cost function (describing the minimum cost of attaining a given standard of living / utility level) is known. As discussed in the introductory chapter, the parametric method for estimating the cost function may not be a desirable one to follow (because it is computationally arduous and the results often reject the usual integrability conditions). Furthermore, such an approach is only practicable for a low number of very broadly defined goods which incurs the cost of discarding information on variations in spending patterns within these groups. As Afriat (1977) puts it, "No simple and satisfactory correction of the C.P.I. can be obtained for particular population groups merely by adjusting the weights

[5] Allen (1958), Brittain (1960), Tipping (1970), Muellbauer (1974) and (1977), Piachaud (1978), Bradshaw and Godfrey (1983), Fry and Pashardes (1986).

on aggregate indices, like the food index, to correspond to their particular budgets. It is necessary to go back to the original detailed, disaggregated data and start again."[6] Rather than attempt to estimate cost functions from highly disaggregated data, economists have attempted to devise measures which avoid the need for explicit estimation of welfare effects and behavioural responses to price changes.

A useful measure, and the one used in this chapter, is one proposed by Törnqvist (1936) which Diewert (1976) shows to be equivalent to the true index under a relatively more plausible model of household consumption behaviour which allows for substitution effects.

Proposition 3.1. *The Törnqvist index is the geometric mean of two true cost of living indices when the cost function is a general translog.*

Proof.

Suppose that the functional form for the cost function is a general translog defined by

$$\begin{aligned}\ln c(\mathbf{p}, u) &= \alpha_0 + \sum_i \alpha^i \ln p^i + \frac{1}{2}\sum_i \sum_j \gamma^{ij} \ln p^i \ln p^j \\ &\quad + \beta_0 \ln u + \sum_i \beta^i \ln p^i \ln u + \frac{1}{2}\varepsilon_0 (\ln u)^2\end{aligned}$$

where the parameters satisfy the following restrictions

$\sum_i \alpha^i = 1, \;\; \sum_i \gamma^{ij} = \sum_i \beta^i = 0 \;\; \forall\, i$ for adding up;

$\sum_j \gamma^{ij} = 0 \;\; \forall\, i$ for homogeneity;

$\gamma^{ij} = \gamma^{ji} \; \forall\, ij$ for symmetry.

[6] Afriat (1977) p 15.

Rearranging the cost function gives

$$\ln c(\mathbf{p}, u) = \ln a(u) = \sum_i \left[\alpha^i + \frac{1}{2} \sum_j \gamma^{ij} \ln p^j + \beta^i \ln u \right] \ln p^i$$

Differentiating $\ln c(\mathbf{p}, u)$ with respect to $\ln \mathbf{p}$ gives the budget share vector $\mathbf{w}$ (by Shephard's lemma) hence

$$\frac{\partial \ln c(\mathbf{p}, u)}{\partial \ln p^i} = \sum_i \left[\alpha^i + \frac{1}{2} \sum_j \gamma^{ij} \ln p^j + \beta^i \ln u + \frac{1}{2} \sum_i \gamma^{ij} \ln p^i \right] = w^i$$

Substituting into the cost function gives

$$\ln c(\mathbf{p}, u) = \ln a(u) + \sum_i \left[w^i - \frac{1}{2} \sum_i \gamma^{ij} \ln p^j \right] \ln p^i$$

The translog true cost of living index (P_{TL}) is, then, the ratio of the cost function in period $t+1$ to that in period t, evaluated at u^*, so the log of the true translog index is

$$\ln P_{TL}(\mathbf{p}_t, \mathbf{p}_{t+1}, u^*) = \sum_i \left[w^i - \frac{1}{2} \sum_j \gamma^{ij} \ln p^j \right] \ln \left(\frac{p^i_{t+1}}{p^i_t} \right)$$

Evaluating this at u_t and u_{t+1} and taking the geometric mean gives the Törnqvist index (P_{TQ})

$$\ln P_{TQ}(\mathbf{p}_t, \mathbf{p}_{t+1}, u_T) = \frac{1}{2} \left(\ln P_{TL}(\mathbf{p}_t, \mathbf{p}_{t+1}, u_t) + \ln P_{TL}(\mathbf{p}_t, \mathbf{p}_{t+1}, u_{t+1}) \right)$$

$$\ln P_{TQ}(\mathbf{p}_t, \mathbf{p}_{t+1}, u_T) = \frac{1}{2} \sum_{s=t,t+1} \left(\sum_i \left[w^i_s - \frac{1}{2} \sum_j \gamma^{ij} \ln p^j_s \right] \ln \left(\frac{p^i_{t+1}}{p^i_t} \right) \right)$$

The γ^{ij} terms drop out because of symmetry giving

$$\ln P_{TQ}(\mathbf{p}_t, \mathbf{p}_{t+1}, u_T) = \sum_i \frac{1}{2} \left(w^i_t + w^i_{t+1} \right) \ln \left(\frac{p^i_{t+1}}{p^i_t} \right)$$

or equivalently

$$P_{TQ}(\mathbf{p}_t, \mathbf{p}_{t+1}, u_T) = \prod_i \left(\frac{p^i_{t+1}}{p^i_t} \right) \exp \left(\frac{1}{2} \left(w^i_t + w^i_{t+1} \right) \right)$$

and hence

$$P_{TQ}\left(\mathbf{p}_t, \mathbf{p}_{t+1}, u_T\right) = \exp\left[\frac{1}{2}\left(\frac{c\left(\mathbf{p}_{t+1}, u_t\right)}{c\left(\mathbf{p}_t, u_t\right)} + \frac{c\left(\mathbf{p}_{t+1}, u_{t+1}\right)}{c\left(\mathbf{p}_t, u_{t+1}\right)}\right)\right]$$

■

The Törnqvist index (which is also the discrete time analogue of the Divisia) is, therefore, based upon a preferred model of household behaviour, and although it avoids the need to estimate substitution effects it does not suffer the substitution bias inherent in the Paasche and Laspeyres indices. It also has the advantage that the model of preferences underlying it is fairly general[7] and performs relatively well in applied work on demand analysis[8].

The method adopted here is to calculate chained series of pairwise Törnqvist indices for each commodity. This will mean that each link in the chain refers to a different reference welfare level. Nevertheless, Diewert (1978) shows that these indices differentially approximate each other as well as the true index provided that variations in prices and expenditures between each period are small. He argues that this provides a strong justification for minimising period-to-period variations in prices and quantities by means of frequent rebasing and by chaining annual indices.

2.2.3 Data

The indices calculated in this chapter use information on price movements from the 74 sub-indices of the retail price index for the period 1978 to 1992, and correspondingly grouped household expenditure data from the Family Expenditure Survey for the same period.

The orginal aim of the F.E.S. was to provide the basis of an average basket of goods to be used in the calculation of the U.K. R.P.I. The F.E.S. is an

[7]See Christensen *et al* (1975).

[8]Deaton and Muellbauer (1980), Blundell *et al* (1996).

annual random cross section survey of around 7,000 households (this represents a response rate of around 70%). The F.E.S. records data on household structure, employment, income and the spending over the course of a two week diary period. All members of participating households over the age of 16 are asked to complete a spending diary. In the F.E.S. the information is aggregated to the household level and averaged across the two week period to give weekly expenditure figures for over 300 different goods and services.

The F.E.S. has much to recommend it as a data source on household spending; in particular the coverage of goods is comprehensive, and it excludes expenditures by businesses. Indeed it is heavily used by government statisticians and academics. However, it does have a number of drawbacks. For example, it does not measure spending by all households: it does not cover the institutional population of people living in retirement homes, military barracks, student hall of residence and the residents of hostels and temporary homes. Also, up until 1995 the F.E.S. ignores spending by household members under the age of 16. There may also be a problem of non-response as nearly one third of households which are initially approached do not respond to the survey, and these non-respondents may be different in a systematic way from households which take part. In particular, non-response is highest amongst richer households, among very young households and among the very old[9].

These problems may not be terribly serious, but there are other potential problems in the F.E.S. which might be more substantive. In particular, there may be problems of under or over-reporting of expenditures either through genuine forgetfulness (e.g. food consumed outside the home), or active concealment (e.g. flowers for the mistress) or a combination of forgetfulness and

[9]Tanner (1996).

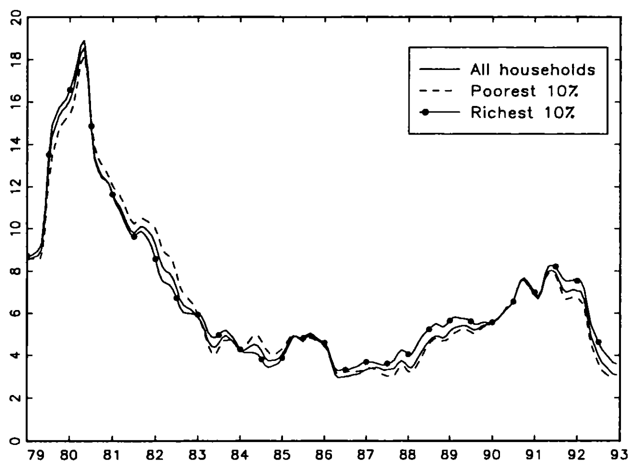
guilt (e.g. alcohol). Problems of under-reporting in relation to alcohol and tobacco are thought, by the Office for National Statistics, to be so severe that the F.E.S. data are supplemented with data from other sources (clearances from bonded warehouses, for example) for use in national accounting. Tanner (1996) shows that under reporting of alcohol spending compared to the National Accounts is of the order of 60% (i.e. 60% of the National Accounts total) and has been relatively stable over time (1978 to 1992). Tobacco under-reporting has increased with the F.E.S. capturing around two third of National Accounts spending in 1992, compared to three quarters in 1978. Another problem is the extent to which the two week diary period in the F.E.S. means that large infrequent purchases (of durables for example) may be under estimated. Data on durables from the F.E.S. are bolstered by data from other sources in the computation of the R.P.I..

The price data use here and throughout the following chapters are the 74 published sub-indices of the R.P.I. Because these data are collected from national sources, there is no regional variation and as a result this chapter ignores regional issues and issues to do with whether or not the prices actually paid by rich and poor households for ostensibly the same goods, may have changed differentially over the period. I return to these issues in the conclusions. Differences in cost of living indices between population groups are thus generated entirely by differences in their spending patterns, scaled by relative price movements. In the following sections, cost of living indices for specific population groups are calculated and compared to the all-household average measure.

2.3 Non-Housing Measures

There are several ways of illustrating group cost of living indices. Most previous studies (Fry and Pashardes (1986) and Bradshaw and Godfrey (1983), for example) present cost of living levels. However, most of the policy- relevant issues are to do with annual changes in the level, i.e. inflation. Benefit up-rating, for example, is designed to compensate households for year-to-year changes in their cost of living rather than the levels. Figure 2.3 illustrates the annual change (inflation) in the Törnqvist cost of living indices (exclusive of housing) for all households and for those in the top and bottom income deciles from 1979 to 1992.

Figure 2.3: Inflation rates, by income group, per cent, 1979-92



Non-housing inflation rates for households at the top and bottom of the income distribution follow the average closely. In general, the all-households average rate lies between the other two but the ranking changes; there are periods when poorer households are facing a higher rate of inflation and richer households a lower rate than the average, and there are also periods when this is reversed. Figure 2.4 emphasises the between-group differences by plotting the difference in inflation rates from the average at each point. The all-households average index is therefore normalised to zero and the differences for each income group are traced around it. For example, in early 1982 when the average all-households inflation measure is around 8 per cent (see Figure 2.3), Figure 2.4 shows that the richest 10 per cent of households saw their cost of living increasing at a rate approximately 0.8 percentage points lower than average (i.e. at around 7.2 per cent), while the cost of living of the poorest 10 per cent was increasing at a rate approximately 0.8 percentage points faster than average (i.e. at around 8.8 per cent). The difference in inflation rates between the richest and poorest households was thus about 1.6 percentage points at this time.

Figure 2.4: Difference in inflation rates, by income group, percentage points, 1979-92.

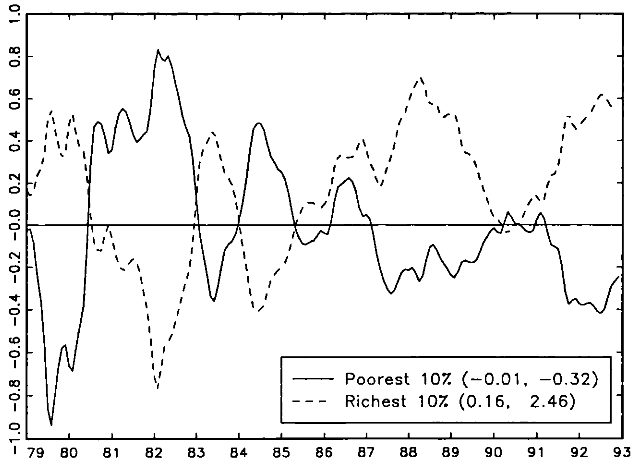
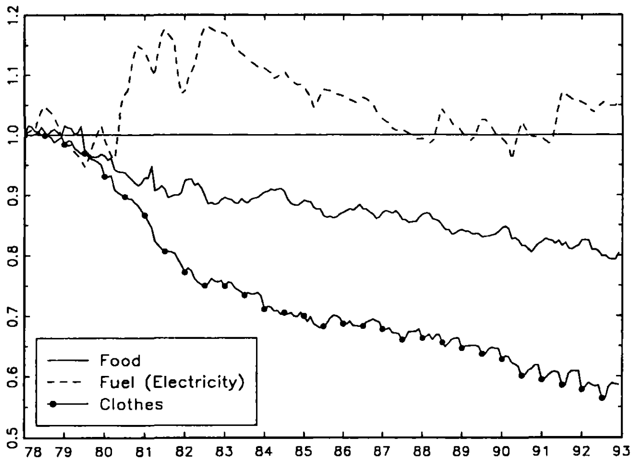


Figure 2.4 shows the cycling nature of the indices more clearly than Figure 2.3. The first number in parentheses in the legend for richer households

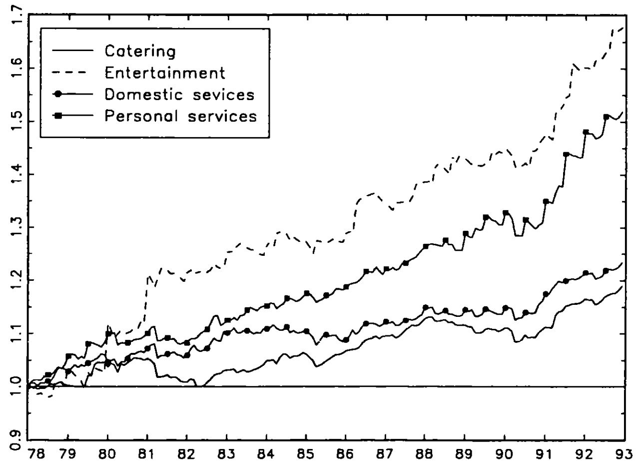
is the average difference from the all-households index for the whole period. This says that on average, inflation for richer households was 0.16 percentage points higher than the average for all households between 1979 and the end of 1992. The second number in parentheses shows the difference in the level of the cost of living index at the end of the period expressed as a percentage of the all-households average index level. This shows that at the end of the period, the cost of living of richer households had grown 2.46 per cent faster than average, and follows directly from their higher-than-average inflation rate. The corresponding numbers for poorer households show that, on average, their inflation rate was 0.01 percentage points lower than the average, and that by the end of the period their cost of living had grown 0.32 per cent less than the average.

Figure 2.5: The relative price of some necessities: food, electricity, clothing, 1978-92



The overall downward effect on relative inflation for poorer households is largely a product of falls in the relative price of necessities such as food and

Figure 2.6: The relative price of some luxuries: catering, entertainment and services, 1978-92.



clothing and (since the early 1980s) domestic fuels (which form a relatively large part of their total spending), and increases in the prices of many luxuries such as eating out, entertainment and other services (which form a relatively small part). Figures 2.5 and 2.6 illustrate these trends in relative prices, and Table 2.2 reports the average expenditure shares for each group at the beginning and end of the period.

The table shows that the average share of spending allocated to necessities (food, fuel, clothing) for all households has fallen from 0.41 to 0.30 over the sample period. The downward-sloping Engel curve relationship for necessities is apparent at both ends of the period. Richer households spend less on necessities than average (0.31 falling to 0.20), and poorer households spend more (0.52 falling to 0.40). The corresponding share increases have been in luxury goods such as entertainment and the 'other' category which is mostly services. One of the largest differences between the two groups over time

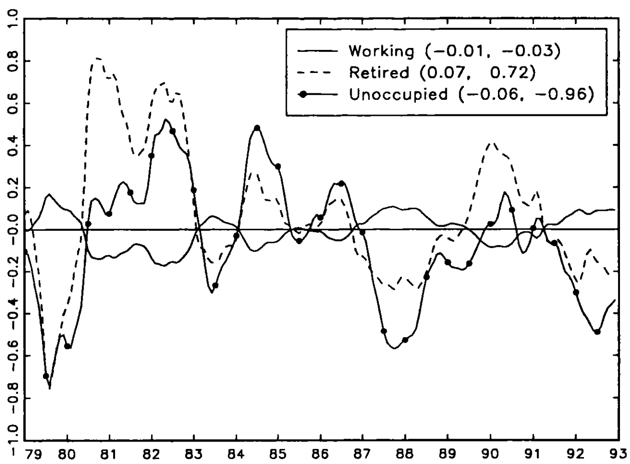
Table 2.2: Grouped average non-housing budget shares, FES, 1978 and 1992.

Group	Year	All	1st Decile	10th Decile
Food	1978	0.24	0.34	0.16
	1992	0.17	0.23	0.10
Catering	1978	0.04	0.03	0.05
	1992	0.04	0.03	0.05
Alcohol	1978	0.06	0.04	0.06
	1992	0.05	0.05	0.04
Tobacco	1978	0.04	0.06	0.02
	1992	0.02	0.05	0.01
Fuel	1978	0.07	0.10	0.05
	1992	0.06	0.09	0.04
Durables	1978	0.07	0.05	0.10
	1992	0.07	0.05	0.08
Clothing	1978	0.10	0.08	0.10
	1992	0.07	0.08	0.06
Motoring	1978	0.13	0.09	0.16
	1992	0.15	0.11	0.15
Fares	1978	0.03	0.03	0.03
	1992	0.04	0.04	0.03
Entertainment	1978	0.05	0.03	0.08
	1992	0.11	0.05	0.22
Other	1978	0.17	0.15	0.18
	1992	0.21	0.21	0.23

is spending on entertainment, which has grown much faster among richer households. The expenditure patterns shown in the table, coupled with the relative price movements illustrated in Figures 2.5 and 2.6, largely explain why the non-housing cost of living of richer households increased by more over this period than that of poorer households did.

Figure 2.7 illustrates the difference from the all-households inflation index by employment status of the head of household. Employment status and income are closely related and therefore it is not surprising that the cycles of the retired and unoccupied groups are similar to those of the poorer households in Figure 2.4. The main differences lie in the period 1989-90 when inflation for these groups was above the average to a greater extent than it was for the poorer households shown in Figure 2.4. As with the poorer households, the average difference for the unoccupied group is negative (-0.06 percentage points) as is the percentage difference in cost of living growth levels at the end of the period (-0.96 per cent). However, longer periods above the average for retired households in the early 1980s and in 1989-90 mean the retired households have done, on average, slightly worse with a positive average difference over the period (+0.07 percentage points) and corresponding higher cost of living growth level at the end (+0.72 per cent).

Figure 2.7: Difference in inflation rates, by employment status, percentage points, 1979-92.



Taking the poorest 10 per cent of the population and calculating changes in their average cost of living gave Figure 2.4. Variations in income and total expenditure are naturally small within the group and consequently differences in spending patterns due to households' positions along the Engel curve are also small. However, differences in household demographics within this section of the population may entail differences between Engel curves defined on these characteristics. There are, for example, poor households with children and poor households without children, young poor households and old poor households. These other factors will contribute to within-group variations in budget shares which may also be well determined.

A major demographic characteristic which influences households' expenditure patterns is the presence of children. However, the differences in relative inflation rates for households with and without children are small, no more than 0.2 percentage points at the most in the very early 1980s. The presence of children makes a household take on some of the spending characteristics of poorer households, indeed their equivalised income falls and adults forgo spending on luxuries like entertainment for more spending on necessities like food and clothing. This sort of spending pattern reduces the incidence of inflation over the period on households which consume these goods. The presence of children within a household results in an average inflation rate which is 0.07 percentage points below the population average over the period, and a cost of living 1 per cent below average at the end. Households without children, like richer households, are able to spend more on luxuries and over the period had a higher-than-average inflation rate.

Households in the bottom decile group with children have experienced an average rate of inflation over the period 0.04 percentage points less than the decile group average (0.05 percentage points less than the all-households average). Poor households without children, with a little more money to spend on luxuries, had an average rate of inflation which was 0.05 percentage points above the decile group average (0.04 percentage points above the population average).

2.4 Indirect Taxation

Since 1979, there have been various reforms to the structure and rate of V.A.T. and excise duties. This section removes the influence of tax changes from the cost of living indices presented in Section 2.3. In the U.K., V.A.T. is a broadly progressive tax, in the sense that richer households pay more V.A.T. as a proportion of total spending. This progressivity is entirely due to the base upon which V.A.T. is levied and the spending patterns shown in Table 2.2. During the period 1979 to 1992, food, domestic fuels, passenger transport and children's clothing, *inter alia*, were zero-rated for V.A.T. (i.e. entirely untaxed). Given that these types of goods are more important elements of total expenditure for poorer households, zero-rating means that the burden of V.A.T. falls most heavily on better-off households.

The incidence of excise duties is more mixed. The main dutiable goods are tobacco, alcohol and petrol. In general, petrol expenditure is higher for richer than for poorer households because of wider car-ownership amongst wealthier households. As a result, petrol excise duties are progressive when looked at across the whole population[10]. Tobacco duties, however, are regressive. Table 2.2 shows that poorer households spend proportionately more than richer households on tobacco. This is due to higher rates of smoking in the bottom income decile group rather than higher consumption by smoking households. Patterns of alcohol consumption and the incidence of duties, however, are more complex.

The Engel curve for alcohol has an upside-down U shape. Alcohol expenditure therefore has the characteristic of a luxury for poorer households (the

[10]Amongst car-owners, however, petrol duties are regressive and fall particularly hard on poorer rural households for which car-ownership, and therefore petrol expenditure, are more of a necessity.

upward-sloping portion of the curve), and of a necessity for richer households (downward-sloping portion of the Engel curve)[11]. Within the alcohol commodity group, there are further differences, with richer households spending more on wines and spirits than poorer households, with a general shift from beer to wines and spirits over the period across all households. Because of their higher alcohol expenditure shares, the overall incidence of alcohol taxation is upon poorer households. A shift in the balance of alcohol taxation away from wines and spirits also impacts more upon poorer households.

To illustrate the effects of indirect tax changes on the cost of living of different income groups, price increases due to V.A.T and excise duty changes have been removed from the price indices from 1978 onward and the cost of living indices recalculated[12]. Figures 2.8 and 2.9 show the differences from the average inflation index for the poorest and richest households. The solid lines correspond to the lines in Figure 2.4; however, here the indices are calculated using the Laspeyres formulation and not the Törnqvist.

The problem with the Törnqvist index in this application lies in the use of the end- period weight. The end-period weight depends on the end-period price vector, so when the counter-factual tax-exclusive price series is used, the correct end-period weights are not observed. Instead, only the base-period weights are observed and therefore the Laspeyres index is calculated.

The first major difference between the taxed and untaxed series occurs in mid-1979. This corresponds to the V.A.T. reforms in Sir Geoffrey Howe's first Budget. The amalgamation of the two V.A.T. rates to a single, higher, 15 per cent rate caused the faster increase in the cost of living of richer households and the slower-than-average increase for poorer households illustrated. One

[11]Banks, Blundell and Lewbel (1996).

[12]It is assumed that indirect taxes are passed on in full.

Figure 2.8: Differences in inflation rates for the 1st income decile,with and without taxes, percentage points, 1979-92.

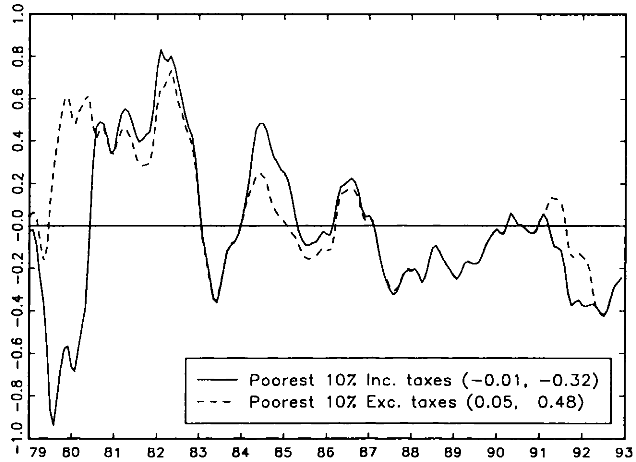
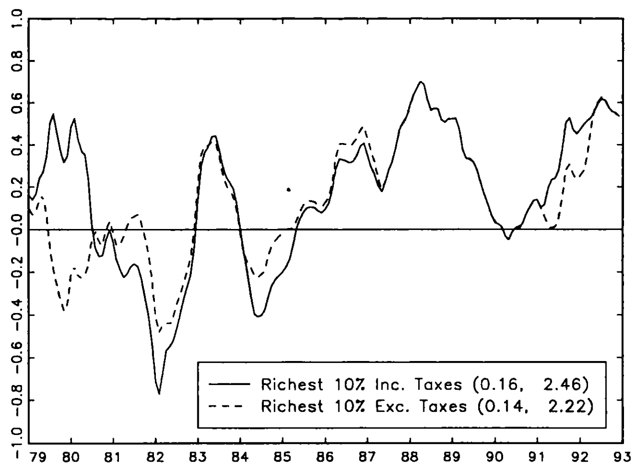


Figure 2.9: Differences in inflation rates for the 10th income decile,with and without taxes, percentage points, 1979-92.



year later, the effects of the V.A.T. increase drop out of the inflation rates for both groups, and return the tax-inclusive series to close to the tax-exclusive path.

Increases in excise duties, particularly on beer and cigarettes, and later the cut in wine duties are shown to push up inflation for poorer households between mid-1980 and 1987. The next period was one in which most excise duties were simply up-rated in line with inflation in each Budget. The final feature of note comes with the increase in the V.A.T. rate to 17.5 per cent in 1991 by Norman Lamont. Just as it did in 1979, the V.A.T. increase pushed up cost of living inflation for richer households faster than for poorer households. Again, the effects only last one year.

Overall, the effects of indirect taxes have been to slow cost of living inflation for poorer households relative to the average. In the absence of V.A.T. and indirect taxes, the poorest 10 per cent of households in the income distribution would have had an average increase in their cost of living which was 0.05 percentage points higher than average instead of 0.01 percentage points lower. Richer households' cost of living increases would have remained higher than average due to increases in the relative price of luxuries, but by a lesser amount (0.14 percentage points rather than 0.16 percentage points).

2.5 Housing

Housing costs form one of the largest components of total household expenditure. Not only are the weights relatively large, but the contribution of mortgage payments in particular has been quite volatile. These factors together make the cost of living indices extremely sensitive to fluctuations in mortgage interest rates; on average, a 1 per cent increase in mortgage interest rates raises the R.P.I. by 0.5 per cent. There is no reason to suppose that this increase in living costs would be distributed evenly across the population. Instead, it will impact on home-owners, with some rents possibly increasing after some time-lag. These different effects across tenure groups may add substantially to the differences in the non-housing cost of living for different population groups illustrated in Section 2.3.

The treatment of shelter costs for home-owners is practically and conceptually difficult. At present, shelter costs for home-owners are represented in the R.P.I. by nominal mortgage interest payments. Essentially, the current approach is to multiply the average outstanding mortgage debt (calculated as a weighted average of the value of mortgages taken out over the previous 25 years) by the current interest rate.

The use of the interest charge measures current expenditure by the household, but does not reflect the price of the shelter service which the house provides. In the same way as the price of a new consumer durable is unaffected by the monthly payments made to the finance company when it is bought on hire-purchase, there is a clear and obvious distinction between the price of shelter services and the borrowing costs of the household[13]. Mortgage costs go up and down with interest rates and fall to zero at the end of the term, but this is not related to the price of the flow of shelter services which the house provides.

While the current approach entails a high degree of sensitivity to interest rate changes, large variations in house prices hardly affect it at all due to the 25-year moving average. Current expenditure on shelter by incumbent home-owners will be unaffected, but if the price of shelter services is the imputed rent then this should rise with house prices. In the U.K., however, the imputed rent approach is difficult to apply because the house rental

[13]See Robinson and Skinner (1989).

market is heavily influenced by the provision of public housing. The use of imputed rents in the R.P.I. was abandoned in 1975.

There is a particular problem with the measurement of shelter costs for owner-occupiers (households which own their homes outright). These households do not make mortgage payments and so the use of mortgage interest payments for them would give a zero cost. Nevertheless, there must be some cost to owner-occupation; after all, the capital invested in the house may be more profitably invested elsewhere. Furthermore, these households own an asset which is slowly deteriorating physically and technologically. It is also an asset with a capital value which fluctuates. The concept of the user cost approach is an alternative designed to deal with this.

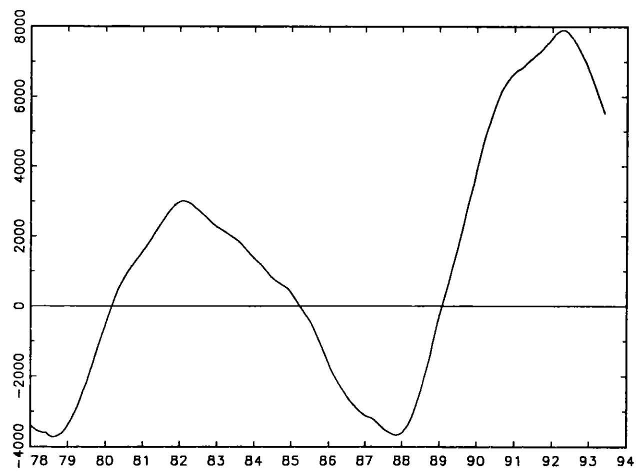
If a household were to borrow in order to buy a house at the beginning of the year, and sell it at the end, the *ex-ante* costs to the household would be given by:

$$uc_t = \left[m_t \left(1 + r_t^m\right) + \left(1 - m_t\right) r_t^e + d_t - E\left(\frac{p_{t+1}^h - p_t^h}{p_t^h}\right)\right] p_t^h$$

where m_t is the ratio of the amount borrowed to the purchase price, r_t^m is the tax-adjusted mortgage interest rate, r_t^e is the interest rate on alternative investments, d_t is the depreciation rate and transactions costs, p_t^h is the purchase price of the house and the final term in the square brackets reflects the expected capital gain (or loss) made on the house over the year. Dougherty and van Order (1982) show that in a competitive market, user costs equal imputed rents.

Under some, not particularly uncommon, circumstances (rapid house-price inflation and relatively low real interest rates), the expected capital gain on housing can outweigh the cost of borrowing and as a result the user

Figure 2.10: Nominal user cost of shelter, £ per annum, 1978-93.



cost can be negative. This is illustrated in Figure 2.10 which shows the user cost of shelter in the U.K. from 1978[14]. The house-price increases in the mid to late 1980s show up clearly as the expected capital gains on housing sent the user cost negative at less than £3,500 p.a. at its lowest point. This simply reflects the fact that, at the time, housing represented a good investment, the returns to which substantially outweighed the costs. The house-price falls at the beginning of the 1990s and the high interest rates at the time combine to push the user cost up to a peak of around £8,000 p.a. The beginnings of the recovery in the housing market and the falls in interest rates towards the end of the period pull the index down again in 1993.

The idea of a negative price raises obvious problems for the definition of a

[14]User costs were calculated using average monthly house-price data supplied by the Department of the Environment. Expected capital gains were estimated using nonparametric (kernel) methods. Essentially, this process applied a weighted moving average around each data point. Following the Bank of England's treatment of user costs in its housing-adjusted retail price index, depreciation was set at 0.5 per cent, transactions costs at 2 per cent and average proportion of the price borrowed at 65 per cent. The mortgage interest rate is from Table 7.1L in Financial Statistics (HMSO). The opportunity cost calculations are based on the Treasury Bill yield from Table 38 in Economic Trends (HMSO).

cost of living index. Presumably, given rational preferences and no other constraints, households would consume an infinite amount of mortgaged housing. Rationing and transactions costs clearly stop this from happening in practice but this raises the question of what is the relevant price to use in the index. Hicks (1940) shows that the relevant price is the virtual price consistent with the observed (rationed) level of consumption. Given that the ration level is positive but finite (indeed most households only have a single home), that would imply a positive virtual price. The issue of rationing and the estimation of such virtual prices is the topic of Chapter 5. For the present this chapter assumes that the welfare costs of rationing and transactions costs are roughly constant over the period and so the virtual price is just a constant positive translation of the user cost. This allows an indexed version of the user cost itself to serve as the price in the index as the index reflects variations from a given base price and movements of the whole level of the price series (i.e. rebasing) do not matter.

Figure 2.11: Shelter costs; rents, mortgage payments, RPI method and the User Cost, 1978-93.

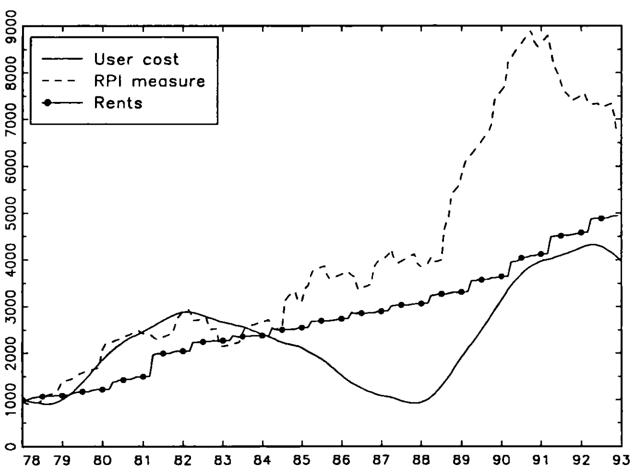


Figure 2.11 shows the user cost index, the mortgage payment index used in the R.P.I. and the price series for rents. The fact that the influence of house-price movements on the R.P.I. measure is negligible is illustrated quite clearly as the R.P.I. measure continues to rise gently in the mid-1980s when expected capital gains cause the user cost to fall. The R.P.I. measure also peaks earlier than the user cost when interest rates first started to fall. Because the lowest point in the house-price cycle was not reached for a few months after interest rates fell, the user cost measure continues to rise, although at a slower rate. The pattern of steps in the rents series is due to the influence of annual changes in rents charged on public housing.

The issue of the appropriate weight for the user cost series is difficult to resolve. The concept of a user cost is notional. The cost is incurred by the household but accrued rather than actually paid. Mortgagors, for example, accrue capital gains and losses but only pay their monthly mortgage bills. The usual weight applied to changes in the price of a good is the expenditure share where expenditure is price multiplied by quantity. In the case of housing, the implicit quantity is one. The expenditure is therefore the current price. This implies that the weight to apply to the user cost price series is the average nominal user cost itself.

One problem with this is that the size of the weight is both large and extremely volatile, as can be seen from Figure 2.11. In 1978, for example, average total weekly non-housing expenditure was £68. The average weekly user cost was around £70. In 1992, the average weekly user cost was around 150, while average total non-housing expenditure was £224. At other times (early 1980, 1985 and 1989), the user cost is zero. Annual increases in the user cost series reach around 100 per cent in early 1979 and in 1988, while they are negative at other times. Including the user cost price series with the nominal user cost weight would result in an unacceptably volatile index which was completely dominated by shelter costs. For the same reasons discussed in relation to a negative price it is also difficult to know how to deal with a negative weight. The approach adopted here is a compromise aimed at focusing on the different effects of the two price series. The weight used under the user cost approach is mortgage payments for households with mortgages, and average mortgage payments for households that own their houses outright. This has the benefit of using similar weights to those used under the R.P.I. method for mortgagors, but also gives a positive weight to owner-occupiers. Section 2.5.1 presents results based on housing-inclusive cost of living indices calculated using the R.P.I. method. Section 2.5.2 discusses and compares the effects of using the user cost approach.

2.5.1 The Mortgage Interest Approach

Figure 2.12 shows the Törnqvist all-households average inflation rate calculated with and without housing costs using the mortgage interest payment method used in construction of the R.P.I.. The effects of rents and mortgage payments are clear, particularly in the late 1980s when increases in interest rates pushed inflation in the all-items index above inflation in the non-housing index. The differential effects on renters versus mortgagors are shown in Figure 2.13.

The first major point of departure is 1981 when local authority rents were increased sharply as grants from central government were cut, and in the following year mortgage interest rates fell. The main differences, however, are apparent from 1988 onward as increases in interest rates pushed the cost of living of home-owners up faster while rents lagged. The interest rate cuts

Figure 2.12: Inflation with and without housing costs; all households, RPI measure, per cent, 1979-92.

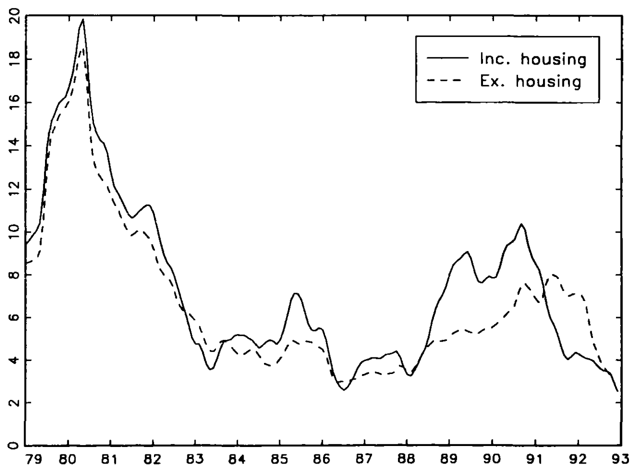
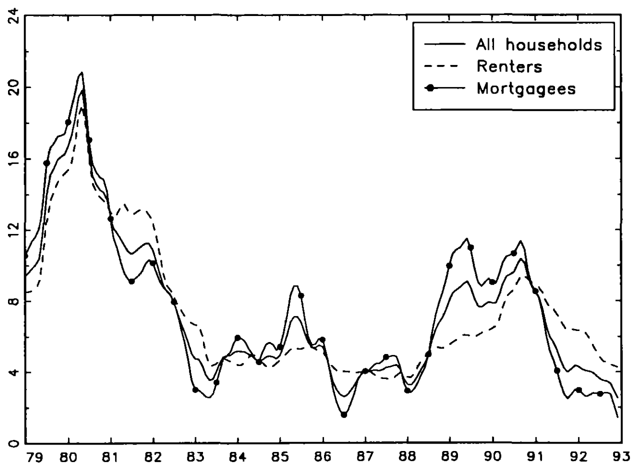


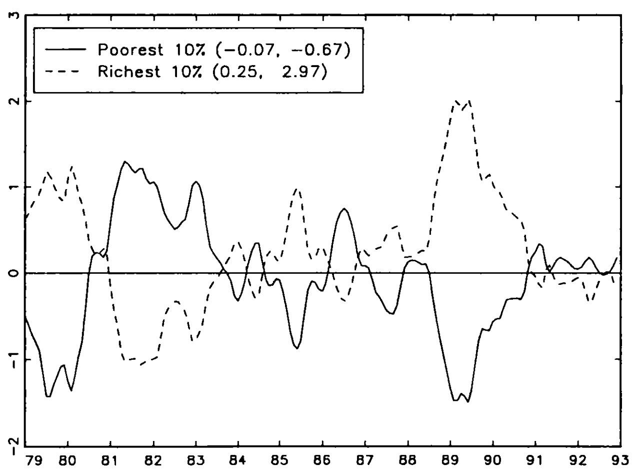
Figure 2.13: Inflation rates by tenure, RPI measure, per cent, 1979-92.



which enter the index from early 1990 had the reverse effect, cutting the rate of increase for home-owners relative to the average and allowing the cost of living for renters to catch up with the average as rents rose more sharply and interest rate cuts for home-buyers pulled the average down. By the end of the period, the average cost of living for households with mortgages rose 1.07 per cent more than the all-households average on this measure of shelter costs.

Figure 2.14 shows the difference in cost of living inflation for households in the top and bottom 10 per cent of the income distribution[15]. To a large extent, the differences are driven by differences in tenure types between the two groups. The increase in the cost of living for poorer households in early 1981 corresponds to the timing of the rent increase. Similarly, the fall in the mid- to late 1980s coincides with the increases in mortgage rates which are shown to impact on the richer households, most of whom are home-owners.

Figure 2.14: Difference in inflation rates, by income group, RPI measure, percentage points, 1979-92.



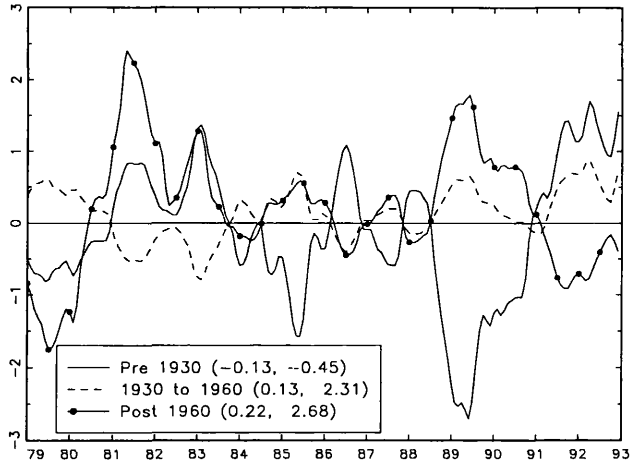
[15]Before-housing-costs measure: see Goodman and Webb (1994).

Compared with Figure 2.4, the inclusion of housing costs appears to amplify the cycles in the indices. Adding housing costs increases the average difference for poorer households from -0.01 to -0.07 percentage points and the final difference in growth levels from 0.32 per cent to 0.67 per cent less than the population mean. This is because increases in housing-costs inflation generally coincide with non-housing inflation. The 1981 rent increases, for example, coincided with a period of higher-than-average non- housing inflation for poorer households. Mortgage inflation at the end of the 1980s coincided with a period of higher-than-average inflation for the richer households. Only at the end of the sample period, in the 1990s, do the housing and non-housing effects appear to cancel each other out as rents rise once more relative to mortgages while non-housing inflation for the poorest 10 per cent fell.

Figure 2.15 shows the difference in inflation rates for three broad date-of-birth cohorts: households in which the head was born before 1930 (i.e. those in which the head-of-household was aged 50 or more at the start of the period and over 63 at the end), those in which the head was born after 1930 but before 1960, and those in which the head was born after 1960 (i.e. households in which the head was under 19 at the beginning and 32 at the end).

The path for the youngest cohort is similar to that for renters and poorer households until about 1983. They seemed to be particularly hard hit in early 1981 by the combined effects of the rent increase and other, non-housing inflation. During the mid-1980s, this cohort appears to take on some of the characteristics of richer home-owners, possibly as a result of the right to buy council houses and as part of the general shift towards owner-occupation.

Figure 2.15: Difference in inflation rates, by head of household's data-of-birth cohort, percentage points, 1979-92.



This turns out to be unfortunate since they then enter the period of high interest rates with more members who are mortgagors. The average difference from the all-households inflation rate is therefore quite high at 0.22 percentage points above average and consequently their cost of living level at the end has grown 2.68 per cent more than average. There therefore appears to be quite a strong cohort-specific effect in which an ill-timed move into owner-occupation increased the cost of living of younger households. In contrast to those born after 1960, the eldest households did relatively well, finishing the period with a cost of living with has grown 0.45 per cent slower than average.

2.5.2 The User Cost Approach

Figure 2.16 shows Törnqvist average inflation rates calculated exclusive and inclusive of housing costs, with shelter costs measured by the user cost method as well as the R.P.I. mortgage interest payments method. Because the user cost and mortgage interest payment indices start off similarly, the differences from Figure 2.12 in the time path of the all-items index up to the early 1980s are slight. From that point onwards, however, they are quite striking. As the expected capital gains on housing impact upon the shelter costs index during the mid-1980s, the user cost method gives an average all-items inflation index which goes negative in 1986. Increased interest rates and capital losses at the end of the 1980s combine to push the all-items user cost measure well above the R.P.I. measure.

Figure 2.16: Inflation rates, with and without housing, all households, user cost and RPI shelter costs measures, per cent, 1979-92.

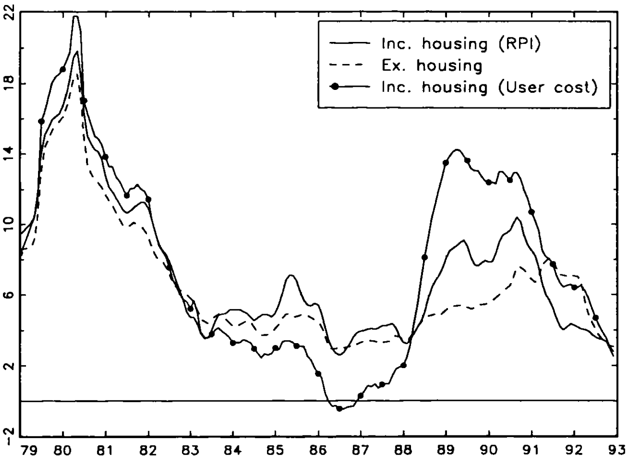
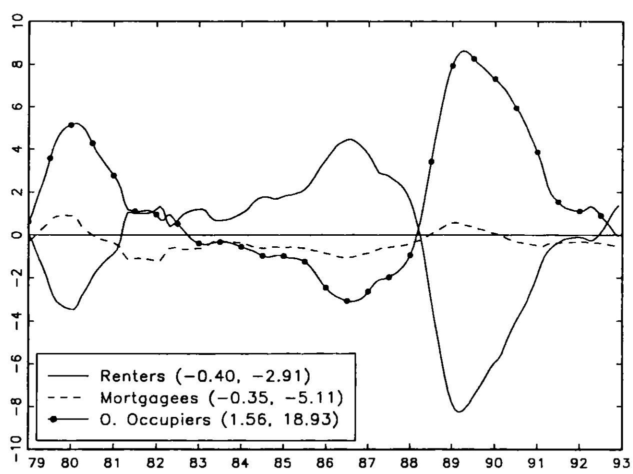


Figure 2.17 shows the effects of this pattern by tenure type. As expected, home-owners do relatively well during the housing boom, enjoying falls in their cost of living. Home-owners who own their houses outright in particular did very well in this period as their shelter costs reflect the capital gains without the mortgage costs. This, however, had the consequence that they were more exposed to the capital losses in the next few years. This gave owner-occupiers an inflation rate which was 1.56 percentage points higher

than average over the whole period, but by the end of 1992 their cost of living had grown nearly 19 per cent faster than average. This was due to their exposure to capital losses on their homes in the late 1980s.

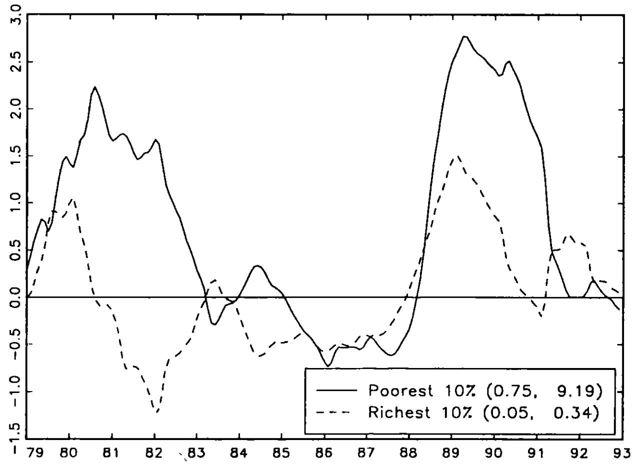
Figure 2.17: Differences in inflation rates, by tenure, user cost measure, percentage points, 1979-92.



The previous section picked up the relationship between the proportion of mortgage-payers and income bracket – households on higher incomes were more likely to be paying a mortgage, and the size of mortgage was likely to be greater than that of less-well-off households. The mortgage interest payments method, therefore, fails to pick up the large number of usually older households in the bottom income decile group which own their homes outright[16]. The user cost of housing does apply to these households because they experience capital gains and losses on the value of their homes. Figure 2.18 shows the difference for average inflation for the poorest and richest 10 per cent of

[16]In 1991-92, in the bottom income decile group (before-housing-costs measure), 24 per cent of households own their homes outright and 28 per cent have a mortgage (Department of Social Security, 1994).

Figure 2.18: Differences in inflation rates, by income group, percentage points, 1979-92.

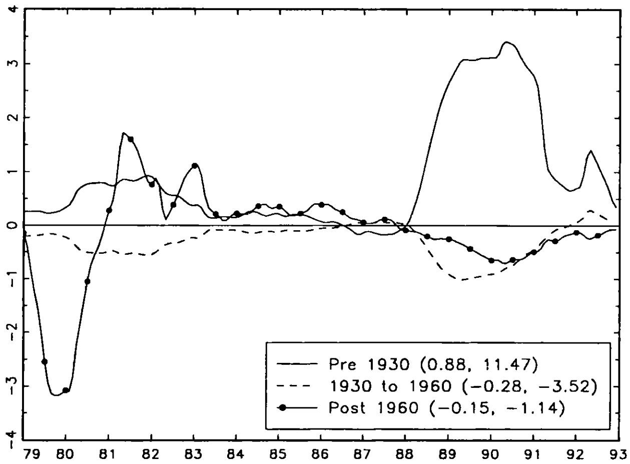


the population[17]. As expected, because of the number of owner-occupiers in the bottom decile group, this is quite different from the corresponding figure using the R.P.I. measure (Figure 2.14). Now, the inflation in the bottom decile group is 0.75 percentage points higher than average over the period, leaving the bottom decile group with a cost of living which has risen 9.19 per cent faster than average at the end of the period.

Figure 2.19 shows the cohort differences corresponding to Figure 2.15. The pattern here is again markedly different. The oldest households now do worst, with an average inflation rate 0.88 percentage points higher than average and a cost of living growth at the end of the period which is 11.47 per cent faster than average. This is clearly due to what happens after 1987.

[17]This uses the same definition of income as that used in Figure 2.4. It may be appropriate to extend the definition to include the capital gains and losses on housing. Figure 2.18, however, allows the comparison between the two measures using exactly the same population groups as in Figure 2.14.

Figure 2.19: Difference in inflation rates, by head of household's data-of-birth cohort, percentage points, 1979-92.



The reason for the large hump in inflation for the eldest cohort is probably the treatment of households that own their houses outright. These sorts of households were therefore exposed to the capital losses on their homes which the user cost measure includes and this completely alters the picture to one where the cohort-specific effect falls not on the young but on the old.

2.6 Summary and Conclusions

Several criticisms can be made of the approach adopted in this chapter. Many revolve around the general problems of measurement discussed in the introductory chapter. Firstly, differences in spending patterns could be a function of differences in prices which we do not observe in these data. Apart from the regional aspect, which is not examined, price differences could also be correlated with the household characteristics which are examined. For example, poorer households without private transport may be forced to buy goods at the corner shop rather than the edge-of-town superstore[18]. The prices they face may be higher than those paid by richer households. However, this only matters if the rates of change in these different sets of prices are different over time, or if households switch between the two sets of prices. This is related to the general problem of outlet bias discuss in Section 1.4.1

Secondly, the issues of quality change and the arrival of new goods (discussed in Section 1.4.2) have not been addressed. Quality improvements in goods and services over the period may mean that more utility is now derived from consumption of some goods than was formerly the case. This means that cost of living indices like those calculated here may overestimate cost increases because they do not adjust for quality improvements. The welfare effects of the introduction of new goods should also be included. The cost of living index relating two period in which one good did not exist in the first period should depend upon the virtual price of the new good for that period; i.e. the price consistent with the observed zero demand. This is typically a very high price which captures the welfare effects of the rationing of the good to zero consumption. When it is included in the index it tends to reduce inflation between the periods considered.

If these forms of bias apply equal across all household types then they may not matter much, simply affecting the rate of increase of all indices equally. But both quality bias and new goods bias may well be thought to apply more importantly to cost of living indices for richer households because they are likely to spend a greater proportion of their total expenditure on technological goods which are more affected by quality change, and because (assuming the demands for new goods are normal) they will also have a higher virtual price for new goods in the period prior to the one in which

[18] See Piachaud, D. & J. Webb (1996) for example.

they first exist. For both of these reasons the rate of increase in the cost of living for richer households calculated by normal methods may suffer greater upwards bias then it does for poorer households. These issues are discussed in Chapter 3 and Chapter 5

This chapter does not resolve the issue of the treatment of housing costs. The sensitivity of the results to different measures of shelter costs is illustrated but further work is required to develop truly sensible treatment of shelter costs with an appropriate weight. This would provide enough material for a long paper in its own right.

It is important to reiterate that these results, because they concentrate exclusively on the demand side, are entirely dependent upon the period studied. A different period would have given different results. The run of data from 1979 to 1992 does, however, nest two other papers (Bradshaw and Godfrey (1983) and Fry and Pashardes (1986)) and shows that their results, like those here, do not apply more widely than over the period from which they draw their data.

The object of this chapter has been to examine the extent and pattern of differences in the cost of living for subgroups of the population. The main result is that differences in the overall growth in cost of living at the end of the period studied are small. However, relative inflation rates for different households cycle over the period and there are several periods in which inflation rates differ widely between the top and bottom of the income distribution and between demographic groups.

The fall in the relative price of necessities and the corresponding increase in the price of luxuries over the period and the difference in expenditure patterns between rich and poor households have meant that the cost of living has increased faster for richer households than it has for poorer households.

The progressive nature of indirect taxes between 1979 and 1992 has been shown to have contributed to this effect. This means that the real income of poorer households is slightly higher at the moment, and the real income of richer households is slightly lower at the moment, than standard income statistics suggest, and this marginally narrows the increase in real income inequality. However, this does not imply that it is good to be poor. The differences are small and the welfare effects of low income massively outweigh the effects of a slightly lower-than-average increase in their cost of living[19].

Given that these differences between groups are small, the obvious question is whether they matter when up-rating benefits etc. Benefit up-rating is designed to compensate poor households for year-to-year increases in the cost of living. On average, cost of living increases in line with the average index over the period would have been broadly accurate (in fact, they have been overly generous by a very small amount). This should not be taken to imply, however, that there is no need for the government to use an index more representative of the cost of living of poorer households to up-rate benefits. This chapter has demonstrated that households in receipt of benefits have had both periods of higher-than-average and periods of lower-than-average increases in living costs in the order of around 2 per cent. These periods can last up to one or two years. Benefit up-rating on the basis of average increases has therefore overcompensated them for increases in their cost of living at some times and undercompensated them at other times. These period-to-period errors matter if there are liquidity constraints and households cannot, for example, borrow in order to smooth their consumption. There almost certainly are such constraints, and this means that using the 'wrong' index imposes costs on poorer households even if the overall increase is more or less

[19] See Stoker (1986).

right when viewed over a longer period.

Chapter 3

Revealed Preference Tests and Nonparametric Engel Curves

3.1 Introduction

The attraction of revealed preference 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 introduced to describe individual demands by Samuelson (1947, 1948, 1950) and Houthakker (1950) and revisited by Afriat (1973) and Diewert (1973), it has usually been applied to aggregate data. This has presented a number of problems which were outlined in the introductory chapter. 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), Bronars (1987) and Russell (1992)). Thus aggregate data often lack power to reject revealed preference conditions. Secondly, it has proven difficult to devise tests of the significance of rejections. This chapter develops and applies techniques which allow a nonparametric analysis of micro data to be conducted. Moving to a revealed preference analysis of the micro data in a framework of nonparametric regression allows the problems described to be overcome.

As well as power and significance considerations, there are also a number of other motivations for this chapter. First, parametric demand studies on micro data often reject Slutsky symmetry which is one of the implications of utility maximisation subject to a linear budget constraint. Amongst the many possible explanations for this rejection are either that we have the 'wrong' functional form or that there exists no well-behaved form of preferences which can rationalise the data. Nonparametric analysis allows this to be checked. Second, it has proven difficult to test for (global) negative semi-definiteness of the Slutsky matrix in parametric demand models. Using revealed preference nonparametric analysis allows one to simultaneously test for both symmetry and negative semi-definiteness. Third, if the integrability conditions are not rejected, one might often wish to go on and use demand estimates for policy analysis. Using parametric analysis there is always some uncertainty as to how much the welfare conclusions are driven by functional form. Employing nonparametric techniques allows the estimation of bounds on welfare effects and use of these bounds to judge the importance of the choice of functional form on welfare conclusions (this is the topic of Chapter 4). Fourth, the nonparametric analysis can aid in the development of new and parsimonious parametric demand systems. Finally, the nonparametric analysis can be extended to investigate revealed preference for conditional demands.

The immediate problem faced in using micro data is the enormous heterogeneity that is apparent in the data. Taking two households that are very similar in observable demographics, time, place and total expenditure it is usually found that demand patterns are different. This make the application of (revealed preference) nonparametric techniques to micro data problematic. Even taking a small number of households in different price regimes (periods) usually leads to a rejection of the nonparametric conditions (see Koo (1963), Mossin (1972) and Mattei (1994), for example, and the recent paper by Sippel (1997) on the use of experimental data). Rejection becomes certain with sample sizes of several thousand.

To overcome this heterogeneity problem this chapter exploits the idea that large numbers of individuals, in particular regions or time periods, face the same relative prices. It uses nonparametric regression analysis to construct a local average Engel curve for each common price regime. Averaging out the heterogeneity in this way provides a statistical structure within which to examine the consistency of data with revealed preference theory without imposing a global parametric structure to preferences. This provides an alternative to the Afriat inefficiency measure explored in Famulari (1995) and Mattei (1994). Using a long time series of repeated cross-sections raises the question of whether revealed preference theory can be rejected in a statistical sense for particular types of individuals or in particular subperiods of the data.

The plan of the chapter is as follows. Section 3.2 concerns individual data and tests of revealed preference. In section 3.2.1 we briefly review the way in which tests of the Generalised Axiom of Revealed Preference (GARP) are applied to data on prices and quantities demanded. Section 3.2.2 discusses the problem that data, particularly annual data, often lacks power to reject GARP because total expenditure growth swamps relative price movements with the result that budget planes seldom intersect. A way in which a sequence of total expenditures can be chosen in order to maximise the power of the GARP test with respect to a given preference ordering is suggested. Section 3.2.3 puts forward a method of implementing this procedure by using nonparametric Engel curves which, given common price regimes within periods, also define expansion paths. Apart from the benefit of allowing the power of tests of GARP to be improved, the use of nonparametric Engel curves also provide a stochastic structure within which one can construct pairwise tests of the significance of any violations of GARP. This is discussed in section 3.2.4. Section 3.2.5 considers the use of conditional demands and separability in tests of GARP when preferences for a particular good, or group of goods, or the quality of those goods may be changing over time. The technique outlined is a revealed preference analogue of the model put forward by Fisher and Shell (1971)[1] and may be used to correct the price of a good for quality variations (an important issue in the construction of cost of living indices). Section 3.3 discusses preference heterogeneity and examines the relation between the nonparametric Engel curves used to test GARP and the average demands of a set of heterogeneous households upon which they are based. Section 3.4 presents an empirical investigation and application of these ideas on a large repeated cross section dataset: the UK Family Expenditure Survey, 1974 to 1993. Section 3.4.1 briefly describes the data source used and the commodity grouping. Sections 3.4.2 and 3.4.3 set out the method employed for estimating Engel curves. This is a semi-parametric approach which allows complete flexibility in modelling the shape of Engel curves whilst controlling for some observed heterogeneity in our data and also for the endogeneity of the main right-hand-side variable, log total expenditure. Section 3.4.4 discusses the validity of the assumption of normality of demands which is necessary for the construction of our GARP test, and illustrates three typical Engel curves for three years of data. Section

[1]The central idea of Fisher and Shell's model was to introduce quality as an additional argument to the utility function in a way which allowed for easy interpretation in terms of prices or costs. The idea goes back to Prais and Houthakker (1955).

3.4.5 presents results. The results of the GARP tests are reported in 3.4.5.1. The effects of using conditional demands / quality adjustments are shown in 3.4.5.2. Section 3.5 concludes.

3.2 Individual Data and Revealed Preference

3.2.1 Revealed Preference and Observed Demands

Suppose we wished to test experimentally whether a particular agent had rational and stable preferences. In the context of demand, this means facing them with a series of prices and total expenditures and testing whether their demand responses satisfy the Slutsky conditions. Specifically, if we have T time periods and given an n-vector of prices $\mathbf{p}_t$ in each period t we could present the agent with a series of total expenditures x_t and test whether the resulting time series of n-vector demands $\mathbf{q}_t = \mathbf{q}(\mathbf{p}_t, x_t) = \mathbf{q}_t(x_t)$ satisfied revealed preference tests. To do this, construct a $(T \times T)$ matrix m in which, for each pairwise comparison, we define the (t, s) element as an indicator variable:

$$m^{ts} = 1[\mathbf{p}_t'\mathbf{q}_t(x_t) \geq \mathbf{p}_t'\mathbf{q}_s(x_s)] \text{ for all } t, s = 1, ..., T. \tag{3.1}$$

which is unity when the revealed preference comparison in parentheses is satisfied (see Varian (1982)). Thus $\mathbf{q}_t(x_t)$ is (directly) revealed at least as good as $\mathbf{q}_s(x_s)$, $(\mathbf{q}_t(x_t)\ R^0\ \mathbf{q}_s(x_s))$ if the latter vector of quantities is affordable at period t prices. Since transitive orderings can reveal an *indirect* revealed preference relationship between pairs, we also need to form the transitive closure. This is achieved by checking for all transitive links between cycles tk, $kl, .., js$ which also imply $m^{ts} = 1$ and is denoted in what follows by $\widetilde{m}^{ts} = 1$. Varian (1982) shows that this can be achieved relatively inexpensively using Warshall's algorithm.

For pairs ts that satisfy $\widetilde{m}^{ts} = 1$ (that is, $\mathbf{q}_t$ is (indirectly) revealed at least as good as $\mathbf{q}_s$: $\mathbf{q}_t \ R \ \mathbf{q}_s$) we are in a position to check for a possible violation with revealed preference theory. If for such a pair

$$\mathbf{p}_s'\mathbf{q}_t(x_t) < \mathbf{p}_s'\mathbf{q}_s(x_s) \equiv x_s, \tag{3.2}$$

or, if we can find a transitive cycle which implies this relation indirectly, then demands $\mathbf{q}_s(x_s)$ are revealed *strictly* preferred to demands $\mathbf{q}_t(x_t)$ ($\mathbf{q}_s(x_s) \ P^0 \ \mathbf{q}_t(x_t)$, or $\mathbf{q}_s(x_s) \ P \ \mathbf{q}_t(x_t)$) and we have a violation of GARP. In terms of the Afriat inequalities (Varian (1982, p 949)), $\mathbf{q}_t \ R \ \mathbf{q}_s$ implies that there exist numbers $U_s, U_t, \lambda_s > 0$ such that

$$U_t \leq U_s + \lambda_s \mathbf{p}_s'(\mathbf{q}_t - \mathbf{q}_s). \tag{3.3}$$

If (3.2) holds then $\mathbf{p}_s'(\mathbf{q}_t - \mathbf{q}_s) < 0$, and since $\lambda_s > 0$ it must be that $U_t < U_s$ which is a failure of $\mathbf{q}_t \ R \ \mathbf{q}_s$ and consequently a failure of GARP.

3.2.2 Choosing a Path for Comparison Points

The choice of the sequence of total expenditures x_t used in the comparisons above requires some discussion. There is a well known problem with applying GARP tests to data in practice to which Varian (1982) refers in his applied work. This is that income growth over time can swamp variations in relative prices (which are what we are interested in). This is because real income growth induces outward movements of the budget constraint and, combined with typically small period-to-period relative price movements, this means that budget lines may seldom cross (as illustrated in Figure 1.1 in Chapter 1). As a result, data often lack the power to reject GARP. Indeed, if we choose the x_t's so that budget lines never cross then we can never violate the GARP conditions. Clearly then, the power of the test will depend critically on the choice of $(x_1, x_2, ...x_T)$.

One solution is to choose a sequence of constant 'real' total expenditures. Thus given x_1 and a set of price indices $(P_1(\mathbf{p}_1), P_2(\mathbf{p}_2), ...P_T(\mathbf{p}_T))$ one could choose $x_t = x_1 P_t / P_1$ Although superficially attractive this begs the question of what price index to use. More importantly, even if the series of demands generated in this way did satisfy GARP, we cannot be sure that any other series of total expenditures starting from x_1 would also satisfy GARP.

To begin with, consider a simple two period (price regime), two good example. It is important to be sure that the probability of finding a rejection of GARP does not simply depend upon the total expenditure levels of the bundles which we are comparing. The solution is the following: from any starting bundle (say $\mathbf{q}_0\left(x_0\right)$), compute the expenditure level for the second period such that $\widetilde{x}_1 = \mathbf{p}_1'\mathbf{q}_0$. This gives the bundle $\widetilde{\mathbf{q}}_1\left(\widetilde{x}_1\right)$ such that $\mathbf{p}_1'\widetilde{\mathbf{q}}_1 = \mathbf{p}_1'\mathbf{q}_0$ (i.e. $\widetilde{\mathbf{q}}_1 \ R^0 \ \mathbf{q}_0$). If there is no rejection of GARP for these bundles (i.e. $\widetilde{\mathbf{q}}_1 \ R^0 \ \mathbf{q}_0$ and not $\mathbf{q}_0 \ P^0 \ \widetilde{\mathbf{q}}_1$) then no point on the expansion path for period 0 can cause a rejection of GARP compared to the starting bundle $\mathbf{q}_0$. This is defined to be the *Pairwise Most Powerful* test of GARP. The following proposition sets out the sense in which it is most powerful.

Proposition 3.1. *Suppose that for some pair of prices* $(\mathbf{p}_0, \mathbf{p}_1)$ *there is some pair of total outlays that follows the budget sequence* $(\hat{x}_0, \hat{x}_1)$ *and which reveals a certain utility ordering over bundles. Further suppose that the demands given by this path reject GARP. If all goods are normal then the demands for the PMP comparison which preserve this utility ordering also reject GARP.*

Proof.

Let $\widehat{\mathbf{q}}_0\left(\widehat{x}_0\right)$ and $\widehat{\mathbf{q}}_1\left(\widehat{x}_1\right)$ denote the demands corresponding to the budgets $(\hat{x}_0, \hat{x}_1)$. Suppose that $\widehat{\mathbf{q}}_1\left(\widehat{x}_1\right) \ R^0 \ \widehat{\mathbf{q}}_0\left(\widehat{x}_0\right)$ and $\widehat{\mathbf{q}}_0\left(\widehat{x}_0\right) \ P^0 \ \widehat{\mathbf{q}}_1\left(\widehat{x}_1\right)$ which is a violation of GARP. This implies

(i) $\widehat{x}_0 = \mathbf{p}_0'\widehat{\mathbf{q}}_0 > \mathbf{p}_0'\widehat{\mathbf{q}}_1$

(ii) $\widehat{x}_1 = \mathbf{p}_1'\widehat{\mathbf{q}}_1 \geq \mathbf{p}_1'\widehat{\mathbf{q}}_0$

Denote the PMP comparison starting from $\hat{x}_0$ by $\widetilde{x}_1$ which preserves the preference ordering in $(\hat{x}_0, \hat{x}_1)$. From the definition above we have $\widetilde{\mathbf{q}}_1(\widetilde{x}_1)$ such that $\mathbf{p}_1'\widetilde{\mathbf{q}}_1 = \mathbf{p}_1'\widehat{\mathbf{q}}_0$ and so

(iii) $\widetilde{x}_1 = \mathbf{p}_1'\widetilde{\mathbf{q}}_1 = \mathbf{p}_1'\widehat{\mathbf{q}}_0$.

Normality implies that since $\hat{x}_1 \geq \widetilde{x}_1$, then

(iv) $\mathbf{p}_0'\widehat{\mathbf{q}}_1 \geq \mathbf{p}_0'\widetilde{\mathbf{q}}_1$,

and hence, given (i) and (iv)

(iiv) $\widehat{x}_0 = \mathbf{p}_0'\widehat{\mathbf{q}}_0 > \mathbf{p}_1'\widetilde{\mathbf{q}}_1$

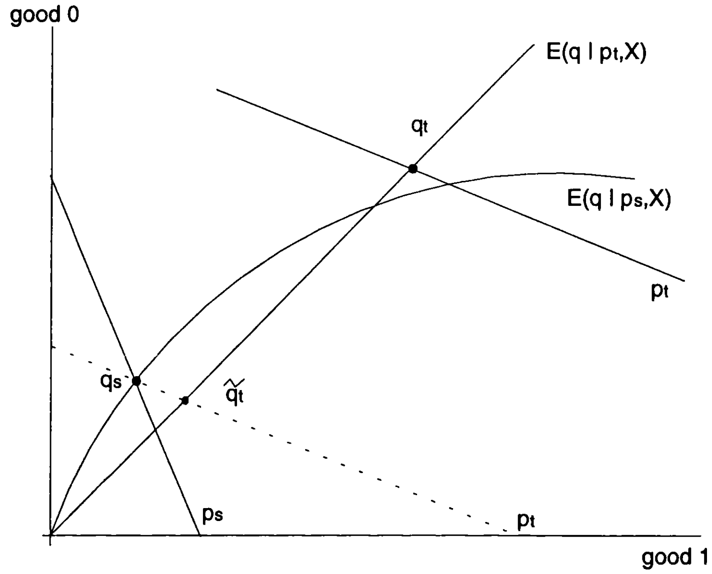
which is a violation of GARP

■

Figure 3.1 shows both the benefit of using the idea of the PMP comparison, and also its limitation. In Figure 3.1 the initial data which are observed are $\{\mathbf{p}_s, \mathbf{p}_t, \mathbf{q}_s, \mathbf{q}_t\}$ in which all of the information on preferences which is revealed is simply that $\mathbf{q}_t P^0 \mathbf{q}_s$. What the PMP procedure does is to compute the bundle $\widetilde{\mathbf{q}}_t$ such that $\widetilde{\mathbf{q}}_t R^0 \mathbf{q}_s$. By Proposition 3.1 we know that if this does not cause a rejection (specifically that $\widetilde{\mathbf{q}}_t R^0 \mathbf{q}_s$ and not $\mathbf{q}_s P^0 \widetilde{\mathbf{q}}_t$), then no point on the expansion path $E(\mathbf{q} \mid \mathbf{p}_t, X)$ can reject compared to the starting point $\mathbf{q}_s$. That this is true in the specific example of Figure 3.1 can be checked easily.

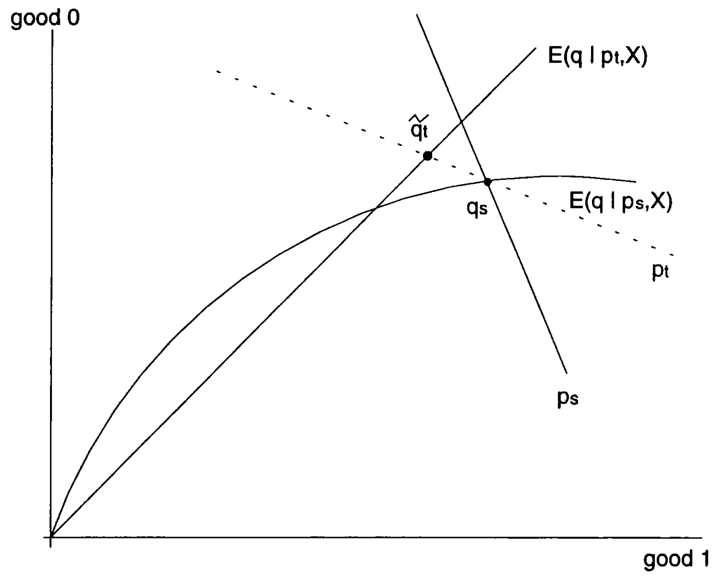
But the problem which remains is also obvious from the Figure. Given these prices, the expansion paths drawn cannot have been generated by well behaved preferences. And the PMP comparison does not identify this. The solution is to take a number of starting points in the expansion path $E(\mathbf{q} \mid \mathbf{p}_s, X)$ corresponding to a number of points in the total expenditure distribution. An alternative starting point is illustrated in Figure 3.2. Now the PMP procedure finds the rejection.

Figure 3.1: The PMP comparison - low initial total expenditure.



The PMP procedure, applied to a number of starting points in the based period's total expenditure distribution (say certain quantile points) gives a powerful method for finding pairwise rejections. Suppose though that we wish to test a dataset with more than two price vectors and demands. To do this consider a sequential version of the PMP procedure: the *Sequentially Most Powerful* (SMP) path. This is a simple algorithm for determining a sequence of x_t points through the data which maximises the chance of finding a rejection given a particular preference ordering of the data. The algorithm for choosing the SMP path following a sequence of relative prices from any starting point is a recursive scheme. Given first period total expenditure, in subsequent periods we choose $x_{t+1} = \mathbf{p}'_{t+1}\mathbf{q}_t(x_t)$. Thus second period total outlay is chosen so that the first period bundle is just affordable at the second period prices and so on. This maximises the possibility of finding a rejection

Figure 3.2: The PMP comparison - high initial total expenditure.



between period t and $t+1$ and between $t+1$ and $t+2$, $t+2$ and $t+3$ and so on, forming a sequence of bundles in a chain in which we maximise the possibility of finding a rejection for each link with respect to its two neighbours. The sense in which this is most powerful is given in the following proposition.

Proposition 3.2. *Suppose that for some set of prices* $(\mathbf{p}_1, ..., \mathbf{p}_T)$ *there is some total outlay path that follows the budget sequence* $(\hat{x}_1, \hat{x}_2, ...\hat{x}_T)$ *and which reveals a certain utility ordering over bundles. Further suppose that the demands given by this path reject GARP. If all goods are normal then the demands for the SMP path which preserve this utility ordering also reject GARP.*

Proof.

Without loss of generality we can take $T = 3$. Let the path of demands based on $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ reject GARP and let the preference ordering be, say, reverse chronological. That is, $\mathbf{q}_2(\hat{x}_2)\ R^0\ \mathbf{q}_1(\hat{x}_1)$, $\mathbf{q}_3(\hat{x}_3)\ R^0\ \mathbf{q}_2(\hat{x}_2)$. Suppose that $\mathbf{q}_1(\hat{x}_1)\ P^0\ \mathbf{q}_3(\hat{x}_3)$ gives the rejection. These imply:

(i) $\hat{x}_2 = \mathbf{p}_2'\mathbf{q}_2(\hat{x}_2) \geq \mathbf{p}_2'\mathbf{q}_1(\hat{x}_1)$,

(ii) $\hat{x}_3 = \mathbf{p}_3'\mathbf{q}_3(\hat{x}_3) \geq \mathbf{p}_3'\mathbf{q}_2(\hat{x}_2)$ and

(iii) $\hat{x}_1 = \mathbf{p}_1'\mathbf{q}(\hat{x}_1) > \mathbf{p}_1'\mathbf{q}_3(\hat{x}_3)$

Denote the MP path starting from $\hat{x}_1$ and which preserves the preference ordering in $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ by $(\hat{x}_1, \tilde{x}_2, \tilde{x}_3)$. By definition we have:

(iv) $\mathbf{p}_2'\mathbf{q}_1(\hat{x}_1) = \tilde{x}_2 = \mathbf{p}_2'\mathbf{q}_2(\tilde{x}_2)$ and

(v) $\mathbf{p}_3'\mathbf{q}_2(\tilde{x}_2) = \tilde{x}_3 = \mathbf{p}_3'\mathbf{q}_3(\tilde{x}_3)$.

This implies that $\mathbf{q}_2(\tilde{x}_2)\ R^0\ \mathbf{q}_1(\hat{x}_1)$ and $\mathbf{q}_3(\tilde{x}_3)\ R^0\ \mathbf{q}_2(\tilde{x}_2)$. From (i) and (iv) we have $\hat{x}_2 \geq \tilde{x}_2$ which implies $q_2^i(\hat{x}_2) \geq q_2^i(\tilde{x}_2)$ for all i (since all goods are normal). Combining this with (ii) and (v) we have:

$\mathbf{p}_3'\mathbf{q}_3(\hat{x}_3) \geq \mathbf{p}_3'\mathbf{q}_2(\hat{x}_2) \geq \mathbf{p}_3'\mathbf{q}_2(\tilde{x}_2) = \mathbf{p}_3'\mathbf{q}_3(\tilde{x}_3)$.

Normality implies that $\mathbf{q}_3(\hat{x}_3) \geq \mathbf{q}_3(\tilde{x}_3)$. Hence, from (iii) we have:

$\mathbf{p}_1'\mathbf{q}_1(\hat{x}_1) > \mathbf{p}_1'\mathbf{q}_3(\hat{x}_3) \geq \mathbf{p}_1'\mathbf{q}_3(\tilde{x}_3)$

which implies that $\mathbf{q}_1(\hat{x}_1)\ P^0\ \mathbf{q}_3(\tilde{x}_3)$ so that GARP is rejected for the MP path.

■

Thus if we test for violations of GARP along a given SMP path starting from a given total expenditure and we do not reject, then we can be confident that we would not reject for any other path which starts from the same total expenditure and maintains the preference ordering given by the SMP path. Conversely, if we have data with a particular preference ordering but in which budget lines seldom cross so that the path lacks power to test GARP, we can construct the SMP path sequence for that ordering and test GARP along it. This maximises the power of the test for the given ordering. There is no need for the chosen ordering of the SMP path to be chronological. In the statement of this proposition the normality assumption is necessary as without it we can construct counter-examples. In the data analysis below, the 22 commodities dealt with are all normal goods. This is illustrated in the data appendix to this chapter.

Figure 3.3: The SMP path

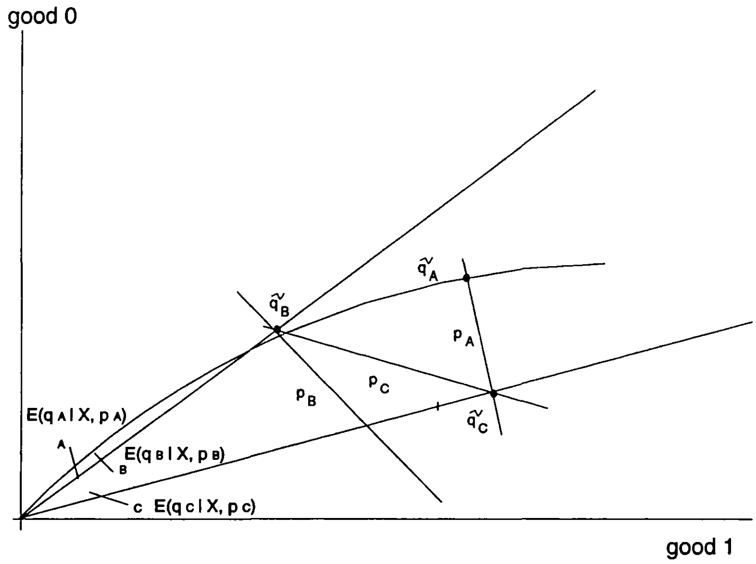
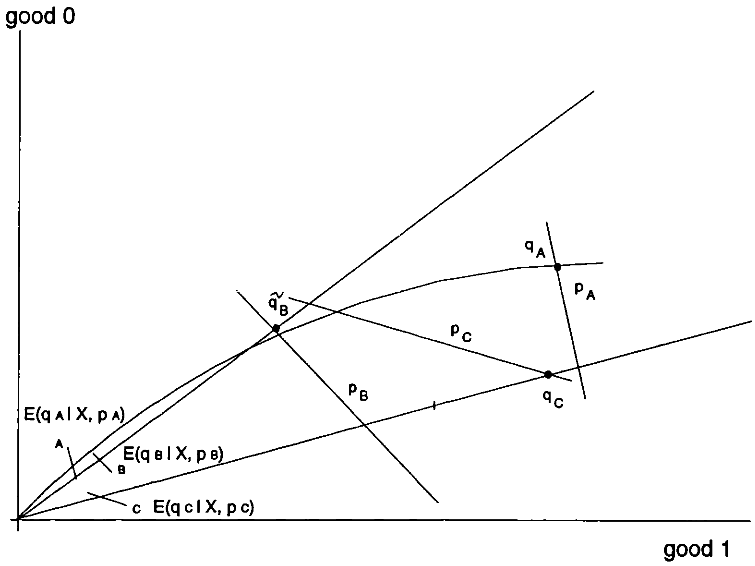


Figure 3.3 illustrates the way in with the SMP path works[2]. As with Figure 3.1 it illustrates both the benefit and the limitations of the technique. The SMP path which is illustrated, orders the data such that $\widetilde{\mathbf{q}}_A\ R^0\ \widetilde{\mathbf{q}}_C$, $\widetilde{\mathbf{q}}_C\ R^0\ \widetilde{\mathbf{q}}_B$ ($\Rightarrow \widetilde{\mathbf{q}}_A\ R\ \widetilde{\mathbf{q}}_B$). By Proposition 3.2, if this path passes (rejects) GARP then any other path which maintains the same utility ordering (i.e $(u(\widetilde{\mathbf{q}}_A) \geq u(\widetilde{\mathbf{q}}_C) \geq u(\widetilde{\mathbf{q}}_B)))$ will also pass (reject) GARP. Figure 3.4 shows that this is the case in this specific example. Here the preference ordering is $\mathbf{q}_A\ P^0\ \mathbf{q}_C$, $\mathbf{q}_C\ P^0\ \widetilde{\mathbf{q}}_B$ ($\Rightarrow \mathbf{q}_A\ P^0\ \widetilde{\mathbf{q}}_B$), i.e. it has the same utility ordering as

[2]The lines crossing the expansion paths near the origins indicate the corresponding relative prices.

the first path $(u(\widetilde{\mathbf{q}}_A) \geq u(\widetilde{\mathbf{q}}_C) \geq u(\widetilde{\mathbf{q}}_B))$, and these data pass GARP.

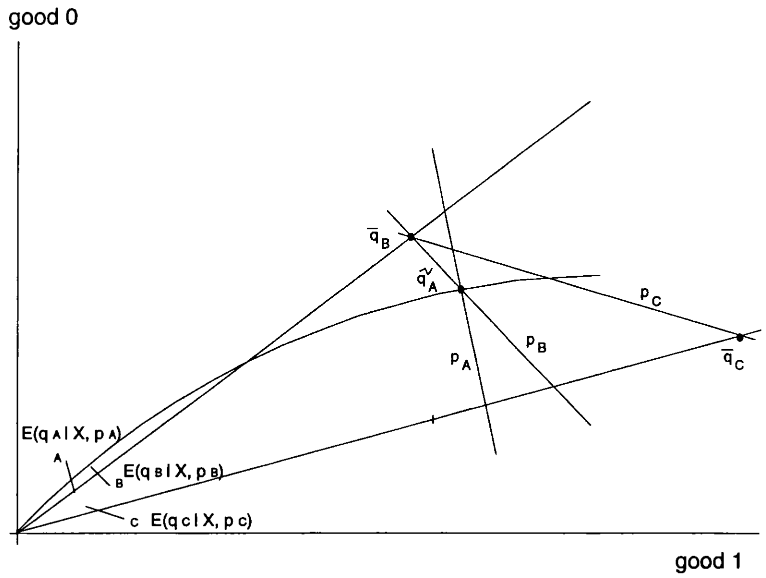
Figure 3.4: A non-SMP path with identical preference ordering.



The problem with the technique is also apparent in Figures 3.3 and 3.4: the expansion paths cross so cannot be consistent with the neo-classical model of consumer choice, and this is not picked up by testing on the SMP path. There are two reasons why the SMP path does not capture this rejection. First, as with the PMP path the result of the test may depend on the starting point. This is illustrated in Figure 3.5. Here the path starts at a different total expenditure level than in Figure 3.3 (in fact it starts at $\widetilde{\mathbf{q}}_A$ where the SMP path in Figure 3.3 ended) and a rejection is now revealed as the SMP path is constructed such that $\overline{\mathbf{q}}_C\ R^0\ \overline{\mathbf{q}}_B$, $\overline{\mathbf{q}}_B\ R^0\ \widetilde{\mathbf{q}}_A$ but $\widetilde{\mathbf{q}}_A\ P^0\ \overline{\mathbf{q}}_B$ gives the rejection. This problem can be partially solved by taking a variety of starting points as with the PMP comparison.

The second problem is that while every link in the SMP path is a PMP comparison, and therefore has the strong property that, compared to any

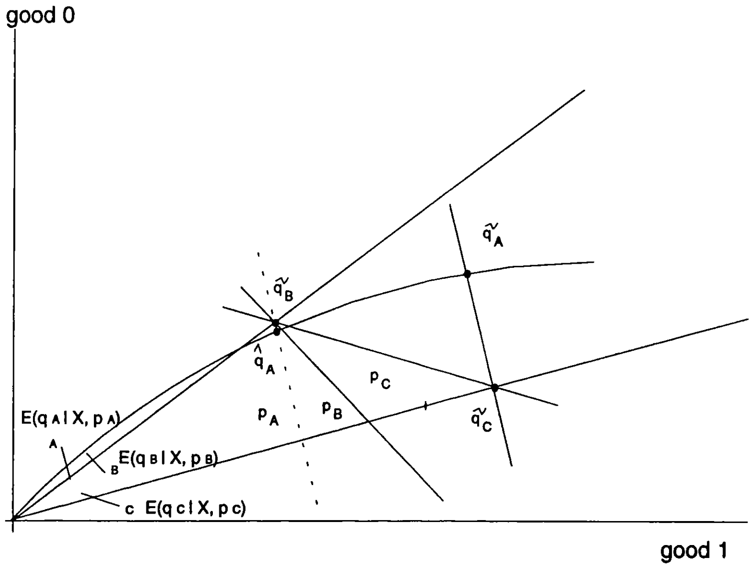
Figure 3.5: An SMP path with a rejection.



point on the SMP path, no point on the expansion path upon which the linking bundle is located can give a rejection if the PMP points pass, the SMP path does not have the equivalent property between points which are indirectly linked. In other words it is *not the case* that if non-adjacent bundles (i.e. points which are indirectly revealed preferred) pass GARP then no point on the expansion path on which the preferred bundle is located will reject with the not preferred point. This is illustrated in Figure 3.6

Figure 3.6 shows the SMP path which was used in Figure 3.3 $\{\widetilde{\mathbf{q}}_A, \widetilde{\mathbf{q}}_C, \widetilde{\mathbf{q}}_B\}$ (ordered best to worst). Each PMP link in the chain has the PMP property, so for example, no point on the expansion path $E(\mathbf{q}_A \mid \mathbf{p}_A, x)$ can reject with the point $\widetilde{\mathbf{q}}_C$ if the pair $(\widetilde{\mathbf{q}}_C, \widetilde{\mathbf{q}}_A)$ pass. But the non-adjacent bundles $\widetilde{\mathbf{q}}_A$ and $\widetilde{\mathbf{q}}_B$ do not share this property as the point $\widehat{\mathbf{q}}_A$ shows by revealing a rejection (the points $\{\widehat{\mathbf{q}}_A, \widetilde{\mathbf{q}}_C, \widetilde{\mathbf{q}}_B\}$ constitute another SMP path with a *different* preference ordering and which reveals the rejection).

Figure 3.6: An alternative SMP path



The SMP path is thus a weaker test between non-adjacent points than the PMP tests is between adjacent ones. However, as Proposition 3.2 states, the SMP path maximises the power of the test between non-adjacent points *conditional on the preference ordering.* In other words, if the SMP path is arranged $\{\widetilde{\mathbf{q}}_A, \widetilde{\mathbf{q}}_B, \widetilde{\mathbf{q}}_C\}$ (ordered best to worst), then while we know that $(\widetilde{\mathbf{q}}_A, \widetilde{\mathbf{q}}_B)$ and $(\widetilde{\mathbf{q}}_B, \widetilde{\mathbf{q}}_C)$ are PMP, the power of the test of the pair $(\widetilde{\mathbf{q}}_A, \widetilde{\mathbf{q}}_C)$ is not PMP but is maximised given that $\widetilde{\mathbf{q}}_B R^0 \widetilde{\mathbf{q}}_C$ and $\widetilde{\mathbf{q}}_A R^0 \widetilde{\mathbf{q}}_B$. That is, we could always make the test between $\widetilde{\mathbf{q}}_A$ and $\widetilde{\mathbf{q}}_C$ at least as powerful by making it a PMP comparison (i.e. by moving the period A bundle to a lower total expenditure level), but then we would lose the utility ordering of the SMP path (this would break the $(\widetilde{\mathbf{q}}_A, \widetilde{\mathbf{q}}_B)$ PMP pairing).

So far we have been assuming that we can take a single agent and present them with any path of total expenditures. In practice, of course, we cannot do this in anything but an experimental setting. Instead we have to use non-experimentally generated data on prices and quantities from a number of heterogeneous households observed only once.

To achieve the SMP path described in Proposition 3.2, we need to be able to move individual agents along their expansion paths. Movements along the expansion path are equivalent to movements along the Engel curve within a fixed price regime. So that if Engel curves are known then so are the expansion paths. To estimate the Engel curves we turn to nonparametric regression methods.

3.2.3 Nonparametric Regression

It is assumed that in period t, prices denoted $\mathbf{p}_t^j$ for each good $j = 1, ..., n$ are common to all individuals. For our purposes, it will be useful to think of t as time but it may alternatively reflect region or some other separation within which the same market price is set. Typically the number of different price regimes $t = 1, ...T$, will be small. Commodity demands and total expenditure, on the other hand, are indexed by both an individual index i and t, the dimension of i will be (very) large.

The advantage of micro demand data is that we can estimate Engel curves nonparametrically for each common price regime. At any point in time all individuals face the same relative prices and are characterised by differences in endowment or total budget x_{it}. For each individual i and good j there is an expenditure $p_t^j q_t^{ji}$ in period t. In the next sub-section we discuss allowing for heterogeneity in preferences; for now we simply define the nonparametric Engel curve for price regime $\mathbf{p}_t$ as the mean expenditure *conditional* on total outlay x i.e.

$$E(p_t^j q_t^j | x) = p_t^j g_t^j(x). \tag{3.4}$$

The price p_t^j is a constant in each period t so that, in any price regime $\mathbf{p}_t$, the conditional mean of each demand given total outlay x defines a set of cross-section demands:

$$E(q_t^j|x) = g_t^j(x) \text{ for } j = 1, ..., n. \tag{3.5}$$

The power of the nonparametric analysis comes from knowledge of the regression line $g_t^j(x)$ and its precision local to specific points of the x distribution.

From extensive earlier work on the Engel curve relationship in U.K. F.E.S. data (see Banks, Blundell and Lewbel (1994)), we know that budget shares that are linear in log total expenditure provide a good baseline specification for many commodities. For this reason we estimate the Engel curves using the nonparametric regression of budget shares on log total outlay. Defining budget shares as

$$w_t^j \equiv \frac{p_t^j q_t^j(x)}{x} \text{ for } j = 1, ..., n, \text{and } t = 1, ..., T, \tag{3.6}$$

the nonparametric regression estimates the conditional expectation

$$E(w_t^j|x) = m_t^j(\ln x) \text{ for } j = 1, ..., n, \text{ and } t = 1, ..., T. \tag{3.7}$$

In what follows we will refer to $m_t^j(\ln x)$ as the local average demand for good j in period t by individual i indexed by x.

3.2.4 Pointwise Inference for Pairwise Comparisons

At each stage in the above discussion we are comparing weighted sums of kernel regressions. The pairwise comparison in (3.2) can be written

$$\sum_{j=1}^{n} p_t^j g_t^j(x_t) > \sum_{j=1}^{n} p_t^j g_s^j(x_s) \text{ for } s \neq t. \tag{3.8}$$

Noting that adding-up implies

$$\sum_{j=1}^{n} p_t^j g_t^j(x_t) \equiv x_t \text{ for all } t$$

so condition (3.8) conveniently reduces to the comparison

$$\sum_{j=1}^{n-1} \alpha_{ts}^j g_s^j(x_s) < x_t - \delta_{ts} x_s, \tag{3.9}$$

where $\alpha_{ts}^j = p_t^j - \frac{p_t^n}{p_s^n} p_s^j$ and $\delta_{ts} = \frac{p_t^n}{p_s^n}$ are known constant weights in each price regime.

Since the nonparametric Engel curve has a pointwise asymptotic standard error we can evaluate the distribution of each $g_t^j(x_t)$ at a finite set of points x_t. For example, in what follows we consider certain quantile points on the SMP path. Pointwise standard errors for kernel regression are given in Härdle (1990).[3]

To evaluate (3.9) we need to find the distribution of the weighted sum of correlated kernel regression estimates. However, since the g^j kernel estimates are to be evaluated at the same point x using the same kernel smoother and the same bandwidth, the constants associated with the kernel function and the density $f_h(x)$ itself will be common to all variance and covariance terms. Pointwise standard errors and confidence bands for expression (3.9) are therefore tractable and are used extensively in the empirical application below.

[3]Briefly, for bandwidth choice h and sample size N the variance can be well approximated at point x for large samples by

$$var(g^j(x)) \simeq \frac{\sigma_j^2(x) c_K}{N h f_h(x)} \tag{3.10}$$

where c_K is a known constant and $f_h(x)$ is an (estimate) of the density of x

$$\sigma_j^2(x) = N^{-1} \sum_{i=1}^{N} \omega_h^i(x)(q^{ji} - g^j(x))^2$$

with weights from the kernel function

$$\omega_h^i(x) = K_h(x - x_t^i)/f_h(x).$$

3.2.5 Quality Change, Conditional Demands, Separability and GARP

It is common in empirical demand analysis to work with conditional demands (see for example Browning and Meghir (1991)). This is particularly convenient where some good, or group of goods, is considered to be rationed or subject to some unmeasured change in quality, preference or habit formation, and is also not separable from the group of goods under study. For example, demands for tobacco consumption are very likely to be subject to changes in preference and quality following government health announcements over the period of study. It is unlikely that the level and participation of tobacco consumption is therefore fully rationalisable by a set of stable preferences over this period. However, it is also likely that preferences over certain other goods of interest, such as beer, wine, spirits and entertainment are directly affected by tobacco consumption; that is, they do not form a subgroup which is separable from tobacco. Consequently, demands conditional on the level of tobacco consumption may be rationalised even though for the set of goods with tobacco included this would not be the case. Similarly, the set of goods excluding tobacco would also not be rationalised in the case where they were not separable from tobacco consumption.

Note that not only is the separability of tobacco intuitively implausible, given that there is an argument that preferences for tobacco may have changed, then there is good reason to expect that a dataset which includes tobacco will fail a test of GARP. If this is so then it means that separability is formally, as well as intuitively, rejected and we cannot simply omit tobacco from the set of goods considered. To see why this is so consider a dataset which is partitioned into two sub-sets of goods and prices, $((\mathbf{p}^k, \mathbf{q}^k),(\mathbf{p}^0, \mathbf{q}^0))$. A utility function $u(\cdot)$ is weakly separable if there exists a sub-utility function $w(\cdot)$ and a super-utility function $v\left(\mathbf{q}^k, w\right)$ which is strictly increasing in w such that

$$u\left(\mathbf{q}^k, \mathbf{q}^0\right) \equiv v\left(\mathbf{q}^k, w\left(\mathbf{q}^0\right)\right)$$

The criterion for separability is set out by Varian (1983)[4]. This is that if the data were generated by such a utility function, then the data $\left(\left(\mathbf{p}^k, \mathbf{p}^0\right), \left(\mathbf{q}^k, \mathbf{q}^0\right)\right)$ and the data $\left(\mathbf{q}^0, \mathbf{p}^0\right)$ must satisfy GARP. This is necessary. For sufficiency the data $\left(\mathbf{p}^0, \mathbf{q}^0\right)$ and $\left(\mathbf{p}^k, 1/\mu; \mathbf{q}^k, w\right)$ must satisfy GARP for some choice of (w, μ) which satisfy Afriat inequalities. In other words, the whole dataset has to pass GARP, the arguments of the sub-utility function have to pass GARP, and the whole dataset with the separable components replaced by their group 'price' ($1/\mu$ where μ is the marginal utility of income at $\mathbf{p}^{0\prime}\mathbf{q}^0$) and their group 'quantity' (w) must pass GARP. Given that it is likely that the necessary condition (the whole set of goods passing GARP) will be violated if one element is subject to preference or quality change, then the separability avenue is not open.

Consider instead the case of $n+1$ goods in which the 'conditioning' good q^0 is subject to some ration or quality change and the remaining 'goods of interest' $q^1, ..., q^n$ are thought to behave according to rational consumer theory. Note that if preferences over the goods of interest are assumed not to be separable from the conditioning good, and we do not observe the latter, then we can rationalise any set of prices and quantities for the goods of interest (see Varian (1986)). Thus a 'missing' good makes it impossible to test for GARP.

This idea of introducing a conditioning good is similar to the idea proposed by Prais and Houthakker (1955) and Fisher and Shell (1971) and generalised by Muellbauer (1975). They introduce a quality parameter directly

[4]Theorem 3, p. 105.

into the utility function with a multiplicative relation to the corresponding quantity. The simple choice model is as follows

$$\max U(\mathbf{q}_t)$$

$$\text{s.t. } \boldsymbol{\pi}_t'\mathbf{q}_t \leq x_t \quad t = 0, ..., T$$

where $\boldsymbol{\pi}_t = \mathbf{p}_t \cdot \boldsymbol{\mu}_t$ is the quality-adjusted price which can be decomposed into a quality-inclusive component $\mathbf{p}_t$ and a quality deflator $\boldsymbol{\mu}_t$. We assume that only one good (good 0) is subject to quality variation. Hence $\pi_t^k = p_t^k$ for $k \neq 0$, $\forall\ t$, and $p_0^0 = \pi_0^0$ (normalising μ_0^0 to 1), with $p_t^0 \gtrless \pi_t^0$ depending on whether quality is increasing or decreasing over time. If quality is increasing then $p_t^0 > \pi_t^0$ as $\mu_t^0 < 1$.

The first order condition is that

$$U'(\mathbf{q}_t) - \lambda_t (\mathbf{p}_t \cdot \boldsymbol{\mu}_t) = 0$$

where $\lambda_t > 0$ if the constraint binds. Substituting into the usual concavity conditions we have the Afriat inequality:

$$U_s \leq U_t + \lambda_t (\mathbf{p}_t \cdot \boldsymbol{\mu}_t)' (\mathbf{q}_s - \mathbf{q}_t) \tag{3.11}$$

or more specifically (denoting $\mathbf{q}_t \equiv (q_t^1, \ldots, q_t^n)$ and $\boldsymbol{\pi}_t \equiv (\pi_t^1, \ldots, \pi_t^n)$)

$$U_s \leq U_t + \lambda_t \boldsymbol{\pi}_t' (\mathbf{q}_s - \mathbf{q}_t) \qquad \text{if } t = 0$$

$$U_s \leq U_t + \lambda_t \boldsymbol{\pi}_t' (\mathbf{q}_s - \mathbf{q}_t) + \lambda_t \pi_t^0 (q_s^0 - q_t^0) \qquad \text{otherwise.}$$

We observe the quality-adjusted prices for all the other goods ($k \neq 0$), $\forall\ t$ (since there is assumed to be no quality variation). However, we do not observe the quality-adjusted prices for the 0th good (π_t^0), except in the 0th period. In order to calculate a price or cost-of-living index we need to calculate the quality-adjusted prices, $\pi_t^0 = (p_t^0 \mu_t)$.

The restrictions imposed by GARP in this case can be shown using a variant of Theorem 7 in Varian (1983) to imply a set of concavity conditions for the maximisation of some continuous, concave, monotonic and non-satiated utility function defined over $\mathbf{q} = (q^1, ..., q^K)'$ conditional on q^0.

Proposition 3.3. *The data* $(q_1^0, q_2^0, ...q_T^0, \mathbf{q}_1, \mathbf{q}_2, ...\mathbf{q}_T, p_1^0, p_2^0, ...p_T^0, \mathbf{p}_1, \mathbf{p}_2, ...\mathbf{p}_T)$ *can be rationalised iff there exist numbers* $U_s, U_t, \lambda_s > 0$ *and* μ_s *such that*

$$U_t \leq U_s + \lambda_s \mathbf{p}_s'(\mathbf{q}_t - \mathbf{q}_s) + \lambda_s p_s^0 \mu_s (q_t^0 - q_s^0) \tag{3.12}$$

Proof.

The proof is identical to that of Theorem 7 in Varian (1982).

■

Since the quality change model can be rewritten as the quality constant model with virtual prices the usual GARP restrictions apply to the data with the actual price of the quality-changed good replaced by the quality adjusted price. Thus allowing for a conditioning good is as though we can choose a price for this good that is different from the observed market price. If we can find μ_s's for each period that equal unity then we can rationalise the data on all $n + 1$ goods. But if GARP is rejected for the full set of goods, the addition of the extra free variables μ_s may make it possible to rationalise the conditional demands for the goods of interest. Formally, $\mu_s p_s^0$ is the virtual price for the conditioning good in period s. If agents like the conditioning good less over time then we would expect to find that $\mu_t > \mu_s$; that is, it is as though the *virtual* price of the conditioning good is rising over time. Adding more conditioning goods further relaxes the restrictions GARP places on the observed data.

The procedure for carrying out the conditioning is essentially a matter of finding the minimum price adjustment such that the data with the original price of the conditioning good replaced by its virtual price pass GARP. In general, for some rejection such as $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}_s$ and $\mathbf{p}_s'\mathbf{q}_s > \mathbf{p}_s'\mathbf{q}_t$ the minimum price adjustment to the price of the $0th$ good necessary such that $\mathbf{p}_s'\mathbf{q}_s \leq \mathbf{p}_s'\mathbf{q}_t$ is to set p_s^0 such that

$$p_s^0 = \frac{\sum_{k=1}^{K} p_s^k \left(q_t^k - q_s^k\right)}{\left(q_s^0 - q_t^0\right)} \tag{3.13}$$

If demand for the goods in question falls between s and t then $(q_s^0 - q_t^0) > 0$ and assuming that $\sum_{k=1}^{K} p_s^k \left(q_t^k - q_s^k\right) > 0$ then this is an upper bound on p_s^0 (denoted by $\overline{p}_s^0$); i.e., we need to set $p_s^0 \leq \overline{p}_s^0$ so that $\mathbf{p}_s'\mathbf{q}_s \leq \mathbf{p}_s'\mathbf{q}_t$.

Proposition 3.4. *If* $(q_t^0 - q_s^0) > 0$ *and* $\sum_{k=1}^{K} p_s^k \left(q_t^k - q_s^k\right) > 0$ *then set* $p_s^0 = \overline{p}_s^0$. *Any* $p_s^0 > \overline{p}_s^0$ *will violate GARP.*

Proof.

(1) Denote $\overline{\mathbf{p}}_s = \left(\overline{p}_s^0, p_s^1, ..., p_s^K\right)$

(2) $\overline{\mathbf{p}}_s$ is such that $\overline{\mathbf{p}}_s'\mathbf{q}_s = \overline{\mathbf{p}}_s'\mathbf{q}_t = x_s$

(3) Suppose $\underline{\mathbf{p}}_s > \overline{\mathbf{p}}_s$ where $\underline{\mathbf{p}}_s = \left(\underline{p}_s^0, p_s^1, ..., p_s^K\right)$.

(4) Then from (2) and (3) $\underline{\mathbf{p}}_s'\mathbf{q}_s > \overline{\mathbf{p}}_s'\mathbf{q}_s = \underline{\mathbf{p}}_s'\mathbf{q}_t \Rightarrow \mathbf{q}_s \; P \; \mathbf{q}_t$, but the SMP path induces that $\mathbf{q}_t R^0 \mathbf{q}_s$ which is a violation of GARP.

■

Similarly, if $(q_s^0 - q_t^0) < 0$ then this is a lower bound on p_s^0 (denoted by $\overline{p}_s^0$), and (analogously) any $p_s^0 < \overline{p}_s^0$ will violate GARP. The proof is analogous to that for Proposition 3.4.[5]

Now suppose that there are two rejections: $\mathbf{p}_r'\mathbf{q}_r > \mathbf{p}_r'\mathbf{q}_s \;\; \Rightarrow \;\; \mathbf{q}_r \; P^0 \; \mathbf{q}_s$ and $\mathbf{p}_r'\mathbf{q}_r > \mathbf{p}_r'\mathbf{q}_t \;\; \Rightarrow \;\; \mathbf{q}_r \; P^0 \; \mathbf{q}_t$, while the SMP path induces that $\mathbf{q}_t \; R^0 \, \mathbf{q}_s$, $\mathbf{q}_s \; R^0 \, \mathbf{q}_r$ and $\mathbf{q}_t \; R \; \mathbf{q}_r$. We need to find a single p_r^0 such that $\mathbf{p}_r'\mathbf{q}_r = \mathbf{p}_r'\mathbf{q}_s$ and

[5] Note that $\sum_{k=1}^{K} p_s^k \left(q_t^k - q_s^k\right)$ and $\left(q_t^0 - q_s^0\right)$ may have different signs. In this case the minimum necessary adjustment will give a negative price. If this is an upper limit then no positive price for this good can be found which can rationalise GARP.

$\mathbf{p}_r^{'}\mathbf{q}_r = \mathbf{p}_r^{'}\mathbf{q}_t$ (which is the minimum adjustment necessary). If $(q_s^0 - q_r^0) > 0$ and $(q_t^0 - q_r^0) > 0$ then we have two lower bounds of which the highest, $\max(\underline{p}_{rs}^0, \underline{p}_{rt}^0)$, encompasses the other and is the overall lower limit. Similarly if we have two upper limits then $\min(\overline{p}_{rs}^0, \overline{p}_{rt}^0)$ encompasses the other. But, if $(q_s^0 - q_r^0) > 0$ and $(q_t^0 - q_r^0) < 0$, say, then the first equation gives a lower limit for $p_r^0 \geq \underline{p}_r^0$, and the second gives an upper limit of $p_r^0 \leq \overline{p}_r^0$. If $\overline{p}_r^0 > \underline{p}_r^0$ then no value for p_r^0 in the interval will cause a violation of GARP. If $\overline{p}_r^0 < \underline{p}_r^0$ then there exists no value for p_r^0 which does not violate GARP.

Under some circumstances a single adjustment designed to address one particular rejection may cause rejections elsewhere. For example, suppose there is a single rejection from $\mathbf{p}_r^{'}\mathbf{q}_r > \mathbf{p}_r^{'}\mathbf{q}_t \quad \Rightarrow \quad \mathbf{q}_r\ P^0\ \mathbf{q}_t$ while the MP path induces that $\mathbf{q}_t\ R^0\mathbf{q}_s$, $\ \mathbf{q}_s\ R^0\ \mathbf{q}_r$ and $\mathbf{q}_t\ R\ \mathbf{q}_r$. We need to find a value for p_r^0 such that $\mathbf{p}_r^{'}\mathbf{q}_r = \mathbf{p}_r^{'}\mathbf{q}_t$. Suppose that $(q_t^0 - q_r^0) > 0$ for the rejecting pair (r, t) but that $(q_s^0 - q_r^0) < 0$ for the non-rejecting pair (r, s). In solving for p_r^0 such that $\mathbf{p}_r^{'}\mathbf{q}_r = \mathbf{p}_r^{'}\mathbf{q}_t$ we find a lower limit on $p_r^0 \geq \underline{p}_r^0$ and need to raise p_r^0 such that $p_r^0 \geq \underline{p}_r^0$. If the resulting increase in p_r^0 is large enough, then this can cause a rejection in the previously non-rejecting pair (r, s) by increasing $\mathbf{p}_r'\mathbf{q}_r$ more that $\mathbf{p}_r'\mathbf{q}_s$ (since $(q_r^0 > q_s^0)$). Since the solution to the (r, t) rejection is a lower bound, and the solution to the consequent (r, s) rejection is an upper bound no value for p_r^0 can resolve the conflict and pass GARP.[6]

To summarise this section: the problems with tests of GARP have been mainly to do with a lack of power to reject revealed preference conditions in aggregate or average demand data, and the lack of a stochastic framework.

[6]Note further that $(q_t^0 - q_r^0) < 0$ for the rejecting pair (r, t) but that $(q_s^0 - q_r^0) > 0$ for the non-rejecting pair (r, s), gives the opposite result if the decrease in p_a^0 which results from the upper limit is big enough. While this procedure can, therefore, cause rejections elsewhere in the data, by analogous arguments it can also fix rejections elsewhere. There is, as far as we know, no easy way to tell which will happen in advance.

This section has suggested that the use of nonparametric Engel curves can be used to address both of these problems. It has shown how to maximise the power of the test of GARP for a pair of price regimes and for a sequence of price regimes with a given preference ordering by moving demands along nonparametrically estimated expansion paths. The properties of these procedures have been discussed. The stochastic structure of nonparametric regressions was briefly reviewed and the way in which to operationalise a statistical test of GARP was described. In the event of a rejection of GARP a method of investigating adjustments to prices to account for quality change or changes in preferences was set out. The relationship between the (nonparametric) Engel curves used in the discussion above and the average demands of a set of heterogeneous agents is now considered.

3.3 Preference Heterogeneity

There are two alternative ways of interpreting the impact of heterogeneity on the average demands estimated from nonparametric Engel curve regression. We could assume individual demands are rational and then ask for conditions on preferences and/or heterogeneity that imply rationality for average demands. This is the approach of McElroy (1987), Brown and Walker (1991) and Lewbel (1996). Alternatively, we could make no rationality assumptions on individual demands and simply ask what conditions enable average demands to satisfy rationality properties. This is the approach of Hildenbrand (1994) and Grandmont (1992).

Let $\omega(x, \mathbf{p}, \varepsilon)$ be the budget share system of n equations for a household with heterogeneity vector $\boldsymbol{\varepsilon}$. This heterogeneity may be observable (for example, family composition or age) or it may be unobservable taste heterogeneity. The necessary condition for the average budget shares recovered by the nonparametric analysis discussed above to be equal to average budget shares is that:

$$\omega(x, \mathbf{p}, \boldsymbol{\varepsilon}) = \mathbf{f}(x, \mathbf{p}) + \Lambda(x, \mathbf{p})' \boldsymbol{\varepsilon} \text{ with } E(\boldsymbol{\varepsilon}|x, \mathbf{p}) = \mathbf{0} \tag{3.14}$$

where $\Lambda(.)$ is an $n \times m$ matrix. Given this combination of functional form restrictions and distributional assumptions, our nonparametric analysis recovers $\mathbf{g}^t(x) = \mathbf{f}(x, \mathbf{p}^t)$.

In the analysis below we apply the GARP tests to the mean function $\mathbf{f}(x, \mathbf{p})$. This function gives mean responses to changes in prices conditional on a given level of total expenditure. The reason that we are interested in testing for GARP using these mean responses is that without such a rationality condition holding it is difficult to see how we would ever conduct welfare analyses of price changes without some underlying utility function. Additionally, the utility function that is associated with an integrable set of demands $\mathbf{f}(x, \mathbf{p})$ is a prime candidate for use in equilibrium models that assume a representative agent.

Note that this aggregation structure is very different to those used in Gorman (1954) and Muellbauer (1976). In particular, we are not aggregating across different incomes. Additionally, we are not assuming that individual demands are integrable; that is, for given $\boldsymbol{\varepsilon}$ we may have that the Slutsky conditions fail for $\omega(x, \mathbf{p}, \varepsilon)$. In this respect, the structure here is close to that of Hildenbrand (1994) and Grandmont (1992).

In the heterogeneity structure given in (3.14) above we do not impose that individual demands satisfy the Slutsky conditions. If, however, we wish to impose integrability at the household level then there are restrictions on the $\Lambda(x, \mathbf{p})$ matrix and the distribution of the heterogeneity terms (see McElroy (1987), Brown and Walker (1989) and Lewbel (1996)).

To illustrate, suppose that each household's preferences are Piglog. This covers the class of Almost Ideal (see Deaton and Muellbauer (1980)) and Translog (see Jorgenson, Lau and Stoker (1982) demand systems. The budget share for good j can be expressed as

$$w_j^t = \alpha_j + \Gamma_j(\mathbf{p}^t) + \beta_j \left(\ln x - \boldsymbol{\alpha}' \ln \mathbf{p}^t - \Gamma_j(\mathbf{p}^t)' \ln \mathbf{p}^t\right) \tag{3.15}$$

where α_j and β_j are preference parameters, $\Gamma_j(\mathbf{p}^t)$ is a nonstochastic matrix of functions of prices and $\boldsymbol{\alpha} = (\alpha_1, ...\alpha_n)'$. Allowing α_j and β_j to have random components ε_j and η_j respectively results in a share model where the residual term is given by:

$$u_j = \varepsilon_j + \beta_j \boldsymbol{\varepsilon}' \ln \mathbf{p}^t + \eta_j \left(\ln x - (\boldsymbol{\alpha} + \boldsymbol{\varepsilon})' \ln \mathbf{p}^t\right)$$

If we assume $E(\boldsymbol{\varepsilon} | \ln x, \ln \mathbf{p}) = \mathbf{0}$, $E(\boldsymbol{\eta} | \ln x, \ln \mathbf{p}) = \mathbf{0}$ and $E(\boldsymbol{\varepsilon\eta} | \ln x, \ln \mathbf{p}) = \mathbf{0}$ we have the heterogeneity structure given in (3.14) above. In general it is not possible to generalise from the Piglog (see Lewbel (1996), for example) and we assume that the shape of Engel curve is locally Piglog which further justifies working with share regressions with $\ln x$ as an explanatory variable.

In general the error term in (3.14) will represent measurement and optimisation error as well as preference heterogeneity so it would seem natural to work with local average demands. Averaging locally to each x eliminates unobserved heterogeneity, measurement error and (zero mean) optimisation errors in demands but preserves any nonlinearities in the Engel curve relationship for each price regime.

3.4 An Empirical Investigation on Repeated Cross-Sections

3.4.1 Data

We use repeated cross-sections of household-level data from the UK Family Expenditure Survey (1974 to 1993). This is a somewhat aggregated version of the dataset used in Chapter 2 and described in section 2.2.3 of that chapter. In this case sub-sample of all the households with a car was drawn[7]. The first and last percentiles of the within-year total expenditure distribution in this sub-sample was then trimmed out. This leaves 86,733 households (between 3,655 and 4,913 in each year). Expenditures on non-durable goods by these households were aggregated into 22 commodity groups and chained Laspeyres price indices for these groups were calculated from the sub-indices of the UK Retail Price Index giving 20 annual price points for each group of goods.

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. More precise definitions and descriptive statistics are provided in the Tables 3.3 and 3.4 in the data appendix to this chapter.

3.4.2 Semi-parametric Estimation and Demographics

In the Engel curve analysis we add a set of demographic characteristics to the conditional mean specification. For each budget share we write a partially

[7]This was in order to allow us to include motoring and particularly petrol as commodity groups.

linear form:

$$w = m(\ln x) + z'\gamma + \varepsilon \tag{3.16}$$

in which $z'\gamma$ represents a linear index in terms of a finite vector of observable exogenous demographic variables z and unknown parameters γ which would differ across commodities. We assume $E(\varepsilon|z, \ln x) = 0$ and $Var(\varepsilon|z, \ln x) = \sigma^2(z, \ln x)$. Following Robinson (1988), a simple transformation of the model can be used to yield an estimator for γ.

Taking expectations of (3.16) conditional on $\ln x$, and subtracting from (3.16) yields

$$w - E(w|\ln x) = (z - E(z|\ln x))'\gamma + \varepsilon. \tag{3.17}$$

Replacing $E(w|\ln x)$ and $E(z|\ln x)$ by their nonparametric estimators $\widehat{m}_h^w(\ln x)$ and $\widehat{m}_h^z(\ln x)$, the ordinary least squares estimator for γ is $\sqrt{n}$ consistent and asymptotically normal.

The estimator for $m(\ln x)$ in equation (3.16) for band width h is then simply

$$\widehat{m}_h(\ln x) = \widehat{m}_h^w(\ln x) - \widehat{m}_h^z(\ln x)'\widehat{\gamma}. \tag{3.18}$$

Since $\widehat{\gamma}$ converges at $\sqrt{n}$ the asymptotic distribution results for $\widehat{m}_h^w(\ln x)$ remain unaffected by estimation of γ and follows from the distribution of $\widehat{m}_h^w(\ln x) - \widehat{m}_h^z(\ln x)'\gamma$.

3.4.3 Endogeneity and Semi-parametric Correction

Since x is total expenditure it is quite reasonable to suppose $\ln x$ may be endogenous to demands. Ignoring demographic variation for the moment, the budget share equation is:

$$w = m(\ln x) + \varepsilon \tag{3.19}$$

but now we have to allow for:

$$E(\varepsilon|\ln x) \neq 0. \tag{3.20}$$

In this case

$$E(w|\ln x) \neq m(\ln x) \tag{3.21}$$

and

$$\widehat{m}_h^w(\ln x) \overset{p}{\nrightarrow} m(\ln x)$$

so that the nonparametric estimator will not be consistent.

However, suppose there exists an instrumental variable y such that

$$\ln x = \pi y + v \text{ with } E(v|y) = 0. \tag{3.22}$$

In the Engel curve case an obvious variable to use for an instrument for log total expenditure is log disposable income which will be correlated with log total outlay, but not determined within the consumer allocation problem. This choice is used in the empirical results of this section. Moreover, assume the following linear conditional model holds

$$w = m(\ln x) + v\rho + u \tag{3.23}$$

with

$$E(u|\ln x) = 0. \tag{3.24}$$

In this case the semiparametric estimator described above can be used to mimic the augmented regression approach to instrumental variable regression. Note that

$$w - E(w|\ln x) = (v - E(v|\ln x))\rho + \varepsilon. \tag{3.25}$$

So that the estimator of $m(\ln x)$ is given by

$$\widehat{m}_h(\ln x) = \widehat{m}_h^w(\ln x) - \widehat{m}_h^v(\ln x)\widehat{\rho}. \tag{3.26}$$

In place of the unobservable error component v we use the first stage residuals

$$\widehat{v} = \ln x - y\widehat{\pi}$$

where $\widehat{\pi}$ is the least squares estimator of π. Since $\widehat{\pi}$ and $\widehat{\rho}$ converge at $\sqrt{n}$ the asymptotic distribution for $\widehat{m}_h(\ln x)$ follows the distribution of $\widehat{m}_h^y(\ln x) - \widehat{m}_h^v(\ln x)\rho$.

3.4.4 Estimated Engel Curves and Normality

The three figures below show the estimated Engel curves (budget share against log total nominal expenditure) for 3 of our 22 commodities, for 3 of our 20 periods (1975, 1980, 1985). These represent a typical necessity (bread), a luxury (entertainment) and beer which displays a roughly quadratic logarithmic Engel curve behaviour.

On each Engel curve we plot the points on the chronological SMP paths which correspond to the 1st, 10th, 25th, 50th, 75th, 90th and 99th percentile points in the base year (1974). Pointwise 95% confidence bands at these points are also drawn. Note that, as we would expect, the precision is much lower at the tails of the outlay distribution. The left to right drift of the Engel curves apparent in these figures illustrates the growth in nominal expenditure which took place between these periods.

Normality of demands was necessary for the proof of the properties of the SMP path which we intend to exploit in our test of GARP. Nonparametric regressions of quantity demanded for each commodity for 3 years of data against log total spending are presented in the data appendix for 1975, 1980 and 1985. There are no instances of goods behaving as inferior goods. The years illustrated are typical.

Figure 3.7: The Engel curve for Bread

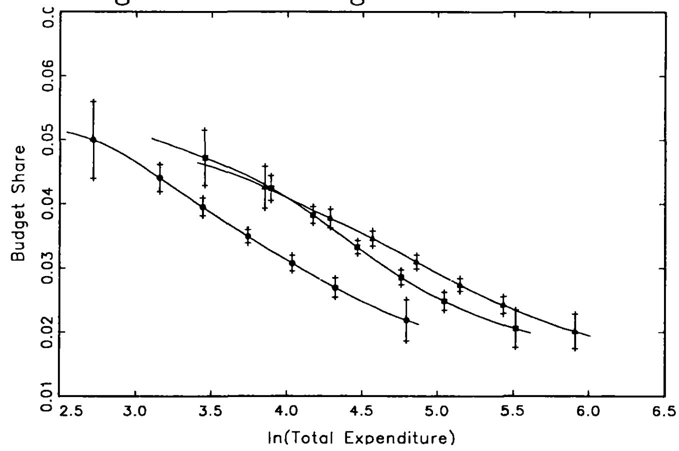
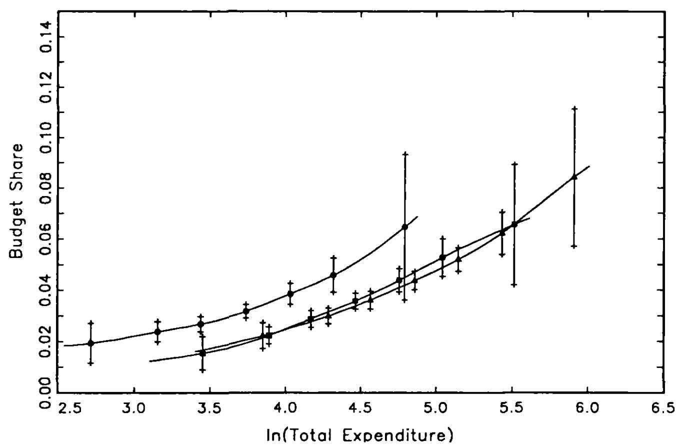


Figure 3.8: The Engel curve for Entertainment

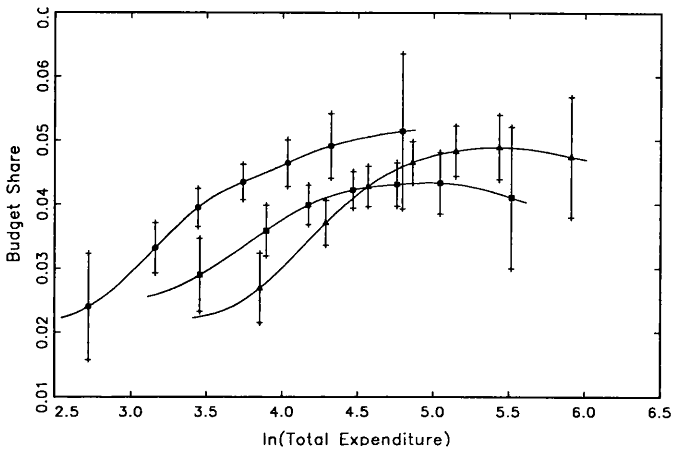


3.4.5 Results

Violations of GARP

We proceed by estimating non-parametric Engel curves for each commodity group within each time period and calculate $g_t^j(x_t)$ at various comparison

Figure 3.9: The Engel curve for Beer



points in the total expenditure distribution. We also control for the number of adults and children in the household as described in section 3.4.2 above, the parameters and standard errors of the reduced form are given in the appendix to this chapter. In the first year of our data we selected the comparison points to be at the 1st percentile, 1st decile, 1st quartile, median, 3rd quartile, 9th decile and 99th percentile points. The comparison points for the following years were chosen to maximise the power of the test on a chronological SMP path which orders the data according to $\mathbf{q}_{t+1}\ R^0\ \mathbf{q}_t\ R^0\ \mathbf{q}_{t-1}$ as described above. By Proposition 3.2 we know that if this path passes GARP then no path which preserves the same preference ordering will violate GARP. We also present the annual median and mean (non-SMP) paths for comparison. We define a $(T \times T)$ indicator matrix m and compute the transitive closure $\widetilde{m}$ in which we know that by construction of the SMP path every element in the lower triangle must be one since either $\mathbf{q}_t\ R\ \mathbf{q}_{t-i}$ or $\mathbf{q}_t\ P\ \mathbf{q}_{t-i}$. We then check for rejections in the corresponding direct and transitive comparisons i.e. if $\mathbf{q}_t\ R\ \mathbf{q}_{t-i}$ (or $\mathbf{q}_t\ P\ \mathbf{q}_{t-i}$) in the lower triangle then $\mathbf{q}_{t-i}\ P\ \mathbf{q}_t$

(or $\mathbf{q}_{t-i}\ R\ \mathbf{q}_t$) in the upper triangle indicates a rejection of GARP.

Table 3.1: Number of rejections of GARP, by size of test.

Comparison paths	**Size of test**				
	Raw	**0.300**	**0.200**	**0.100**	**0.050**
SMP path starting points:					
1st percentile point	5	1	1	1	0
1st decile point	0	0	0	0	0
1st quartile point	2	0	0	0	0
Median	1	1	1	1	1
3rd quartile point	1	0	0	0	0
9th decile point	10	10	0	0	0
99th percentile point	51	25	11	0	0
Median	0	0	0	0	0
Mean	0	0	0	0	0

Table 3.1 shows the number and pattern of rejections for the our full budget system of 22 goods. Each column refers to a different size of test at each point. As can be seen, GARP is rejected for a large number of points in the tail of the outlay distribution but these rejections are not very significant statistically. GARP is, however, rejected in the data at the 0.05% level for the median maximum power path. However the rejection only occurs for a single comparison point (1985 compared to 1986) and, except for this point, the large number of rejections in the raw data are considerably reduced by the use of pointwise confidence bands. It is interesting to observe that there are *no* rejections, even in the raw data, for the annual median or mean non-SMP paths. This is consistent with the observation which arises in tests of GARP on aggregate data that if the budget constraint is allowed to shift much either way between comparison points, as it does for median or mean

total expenditure, then there is little chance of being able to find demands that cannot be rationalised.

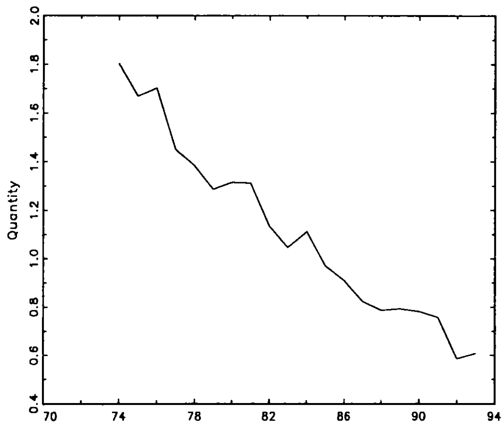
Conditional Demands

Given that the GARP test indicates that there is a significant rejection (for the median SMP path at least) of the idea that these data are rationalisable by a single stable set of well-behaved preferences, one obvious solution is to look around for a suitable conditioning good and apply the ideas outline in section 3.3 above. The GARP test itself gives us no clue as to the good or goods which are causing the rejection, but it does rule out the possibility of omitting a good on the groups of separability (as discussed above). We choose tobacco as our conditioning good since we argue that we have reasonable prior belief that preferences for tobacco may have changed over the period with the arrival of new information on the health effects of smoking. We apply the conditioning procedure to the Median SMP path to the path of demands illustrated below.

In this case the rejection is being caused by $\mathbf{p}'_{85}\mathbf{q}_{85} > \mathbf{p}'_{85}\mathbf{q}_{86}$ while the chronological SMP path requires $\mathbf{p}'_{86}\mathbf{q}_{86} = \mathbf{p}'_{86}\mathbf{q}_{85}$. The necessary minimum price adjustment is to reduce the tobacco price in 1985 from 5086 to 4474 (1974 = 1000)). The intuition is that the original price increase between 1985 and 1986 (5086 to 5463 (1974 = 1000)) is insufficient to rationalise the large fall in demand shown in Figure 3.10 between the same two years, even allowing for the confidence band around the comparison.

Since there is only one comparison which fails (1985-1986) only one price adjustment is necessary for these data to be rationalisable. Table 3.1 below report the results for the re-run of the GARP test with the actual price series for tobacco replaced by the virtual price. Of course this adjustment

Figure 3.10: The median SMP quantity path for tobacco.



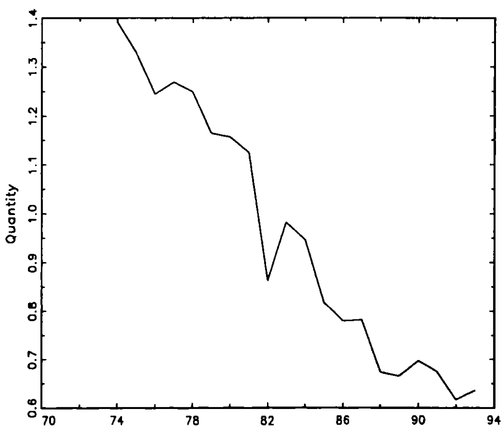
changes the relative prices not just between the years in question (85/86), but between 1985 and every other year. This does not appear to cause any rejections elsewhere as indicated by Table 3.2.

Table 3.2: Number of rejections of GARP, by size of test, virtual price of tobacco.

Comparison paths	**Size of test**				
	Raw	**0.300**	**0.200**	**0.100**	**0.050**
SMP path starting points:					
1st percentile point	4	0	0	0	0
1st decile point	0	0	0	0	0
1st quartile point	1	0	0	0	0
Median	0	0	0	0	0
3rd quartile point	0	0	0	0	0
9th decile point	10	10	0	0	0
99th percentile point	43	11	11	0	0
Median	0	0	0	0	0
Mean	0	0	0	0	0

Note that not only has the raw median SMP path rejection now gone, but also that the price adjustment seems to reduce the number of rejections on other SMP paths. Take, as an example of this, the 1st quartile MP path which originally showed two rejections. The 1st quartile SMP quantity path is shown in Figure 3.11.

Figure 3.11: The 1st quartile MP quantity path for tobacco.



The original rejections on this path were in the pairs of years 85/86 and 87/88. As in the case of the median SMP quantity path, there is a fall in demand between 1985 and 1986 and the virtual price solution indicates that the minimum necessary price adjustment is downwards (from 5086 to 4896 (1974=1000)) to rationalise this. The corresponding downwards adjustment for the median SMP path was greater and since in both cases $(q^0_{86} - q^0_{85})$ is negative the bound is an upper bound and the minimum median adjustment encompasses the minimum 1st quartile adjustment.

3.5 Summary and Conclusion

This chapter has applied nonparametric demand theory to the nonparametric statistical analysis of consumer demand. It exploited the idea that price taking individuals in the same market face the same relative prices, in order to smooth across the demands of individuals for each common price regime. This was shown to provide a method of improving the power of tests of revealed preference *and* to provide a stochastic structure within which to examine the consistency of individual data and revealed preference theory. Simple procedures which are able to maximise the power of a pairwise test (PMP) and a sequential test (SMP) conditional on a preference ordering were devised, and a method of working with conditional demands which may be convenient where some good, or group of goods, is considered to be rationed or subject to some unmeasured change in quality, preference or habit formation, and is also not separable from the group of goods under study was discussed. Conditions were also derived which can enable inference on the rationality of individual households to be made on the basis of local average demands. Using a long time series of repeated cross-sections from the 1974-1993 UK Family Expenditure Surveys it was possible to examine whether revealed preference theory is rejected, and to investigate the properties of demands conditional on tobacco expenditures.

Appendix 3.A: Data

Table 3.3: Commodity Group: Definitions.

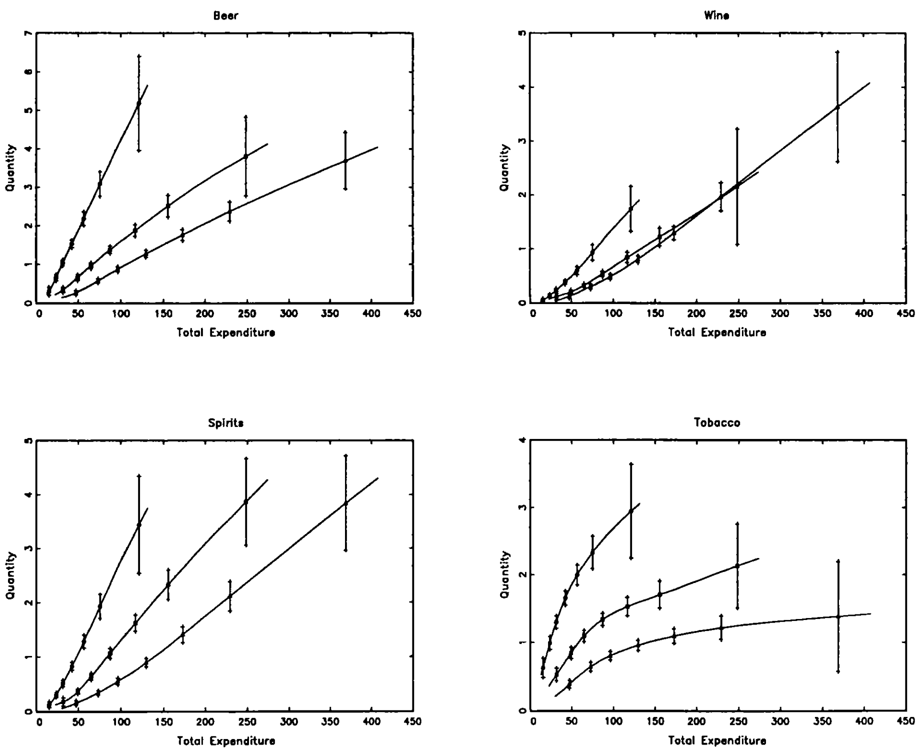
Commodity Group	
1) Beer	Beer, on and off licence sales.
2) Wine	Wine, on and off licence sales.
3) Spirits	Spirits, on and off licence sales.
4) Meat	All meat & fish
5) Dairy	All diary products, oils and fats.
6) Vegetables	Fresh, tinned and dried vegetables & fruit.
7) Bread	Bread, flour, rice & cereals.
8) Other foods	Tea, coffee, drinks, sugar, jams & sweets.
9) Food consumed outside the home	Restaurant & canteen meals.
10) Electricity	Account & slot meter payments.
11) Gas	Account & slot meter payments.
12) Adult clothing	Adult clothing
13) Children's clothing and footwear	Children's clothing & footwear
14) Household services	Post, phone, domestic services & fees.
15) Personal goods and services	Personal & chemist's goods & services.
16) Leisure goods	Records, CD's, toys, books & gardening.
17) Entertainment	Entertainment.
18) Leisure services	TV licences & rentals.
19) Fares	Rail, bus & other fares.
20) Tobacco	Cigarettes, pipe tobacco & cigars.
21) Motoring	Maintenance, tax & insurance.
22) Petrol	Petrol & oil

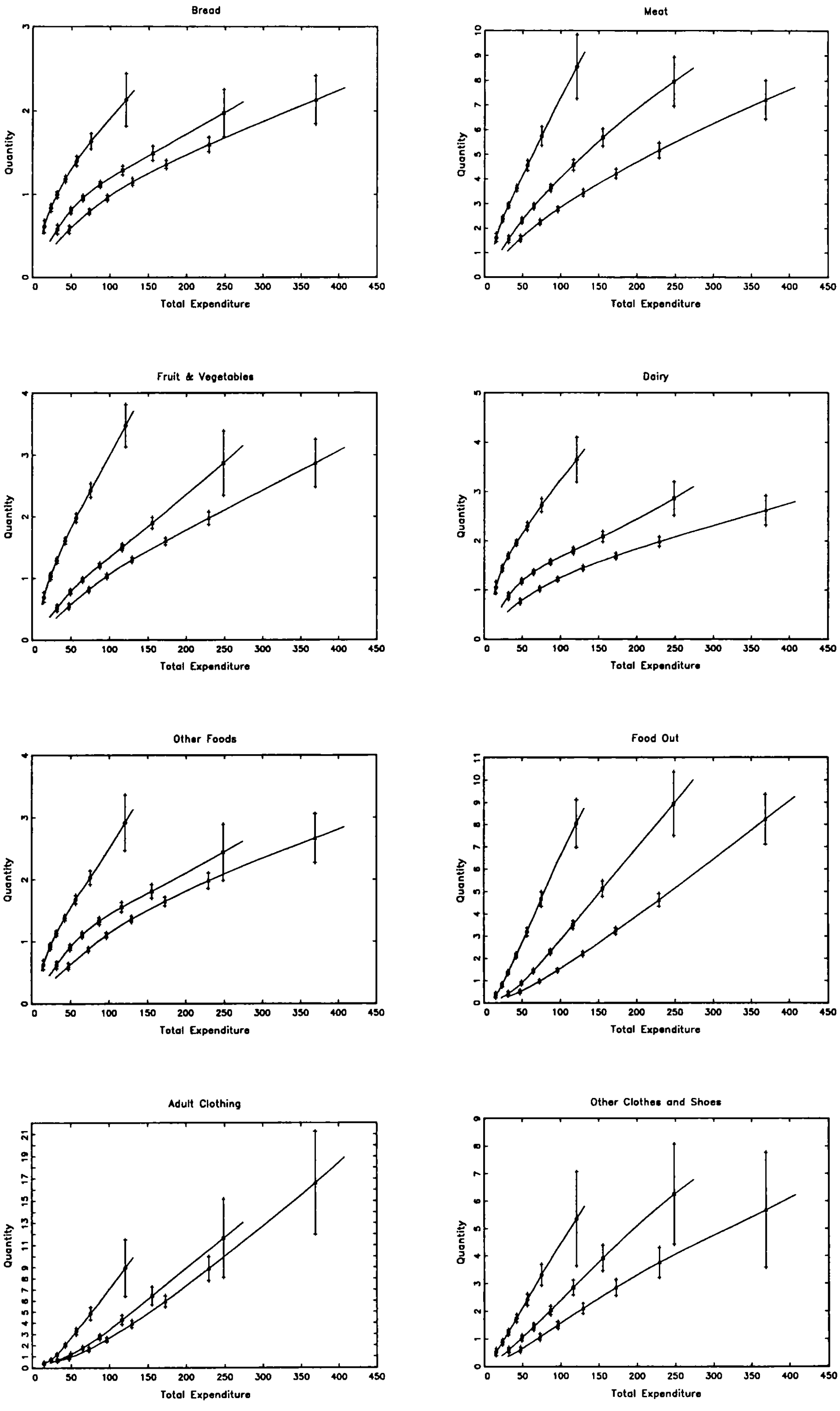
Table 3.4: Total Nominal Household Expenditure: Annual descriptive statistics.

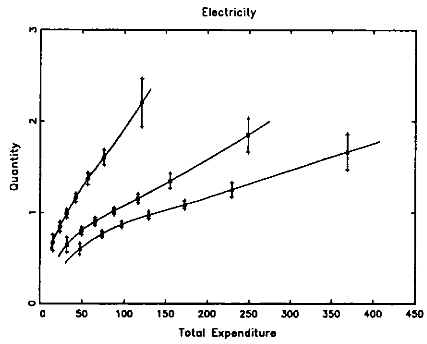
Year	No. of Obs	Mean	Std Dev.	10%	50%	90%
1974	3655	38.05	18.10	19.09	34.29	62.08
1975	4025	45.74	21.34	22.79	41.24	74.82
1976	3895	51.07	24.22	25.42	45.94	82.08
1977	4031	58.88	27.84	29.54	52.98	96.70
1978	3954	65.41	31.55	32.34	58.42	107.16
1979	3851	76.00	37.17	36.69	68.56	124.31
1980	4103	88.99	43.30	42.82	79.65	148.80
1981	4552	98.19	48.38	47.58	87.14	165.91
1982	4469	103.59	50.11	50.98	93.03	170.38
1983	4247	111.24	54.57	53.49	99.88	185.35
1984	4256	118.29	60.25	54.66	104.66	200.51
1985	4317	125.52	64.69	57.31	111.43	211.87
1986	4388	136.07	72.37	61.68	119.07	234.95
1987	4599	142.06	74.43	63.27	126.36	242.43
1988	4704	151.99	82.00	67.08	134.61	262.62
1989	4803	162.63	86.41	72.11	144.87	281.89
1990	4616	176.09	95.04	75.43	155.54	303.85
1991	4688	183.48	99.29	78.46	161.31	316.34
1992	4913	189.14	97.40	83.57	169.56	324.34
1993	4696	201.66	110.70	86.32	177.22	346.41

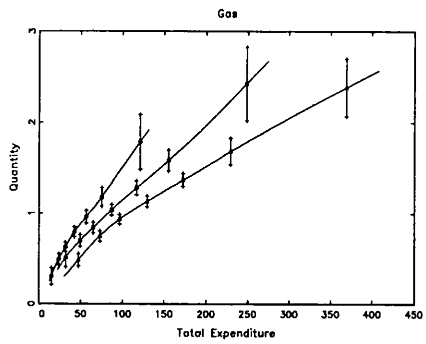
Appendix 3.B: Quantity against log Total Nominal Expenditure for each commodity group; 1975, 1980 and 1985.

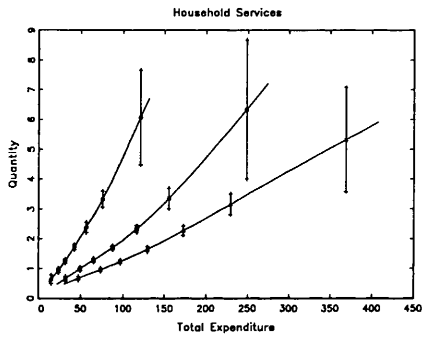
These figures show adaptive kernel estimates of the relationship between quantity demanded for each good and the log of total nominal household expenditure on all goods for three years of the data; 1975, 1980 and 1985. Each kernel regression has the percentile (1st, 10th, 25th, 50th, 75, 90th, 99th) chronological SMP path points and corresponding pointwise 95% confidence intervals marked on the curve.

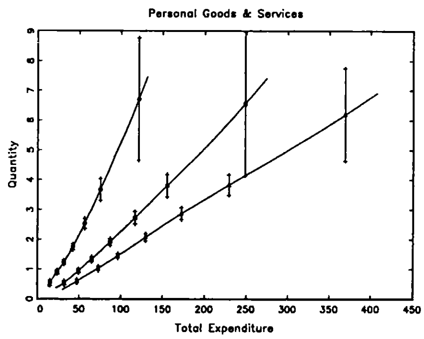


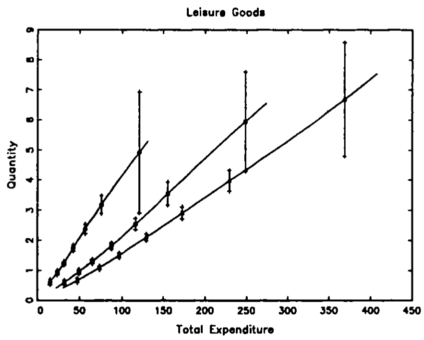

Bread
Meat
Fruit & Vegetables
Dairy
Other Foods
Food Out
Adult Clothing
Other Clothes and Shoes
Quantity
Total Expenditure

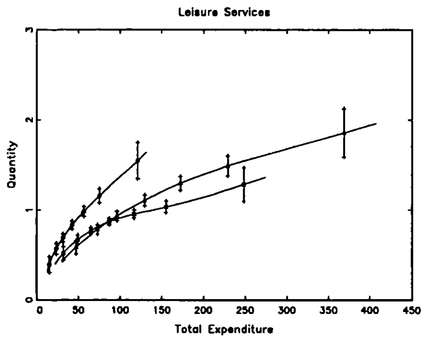

Electricity
Quantity
Total Expenditure

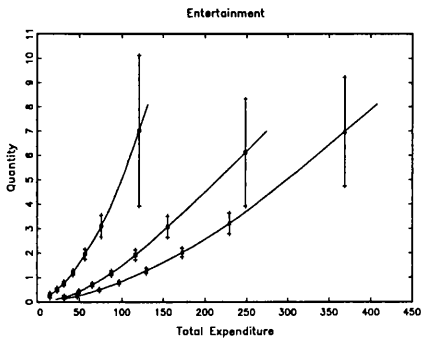

Gas
Quantity
Total Expenditure

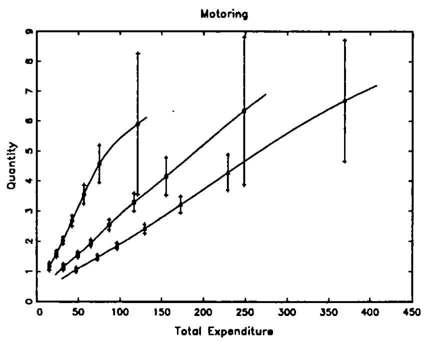

Household Services
Quantity
Total Expenditure

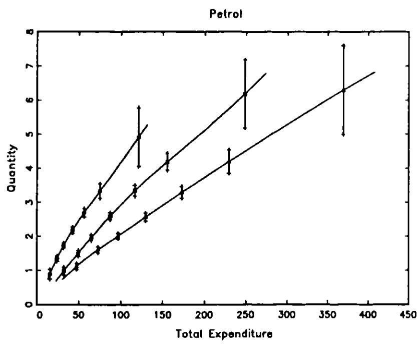

Personal Goods & Services
Quantity
Total Expenditure

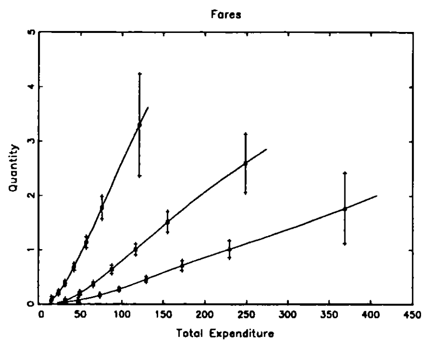

Leisure Goods
Quantity
Total Expenditure


Leisure Services
Quantity
Total Expenditure


Entertainment
Quantity
Total Expenditure


Motoring
Quantity
Total Expenditure


Petrol
Quantity
Total Expenditure


Fares
Quantity
Total Expenditure

Appendix 3.C: Reduced form Equations

Table 3.5 is the first stage equation which shows the least squares regression of log total expenditure on log income and demographic variables for household i in period t.

$$\ln x_{it} = \alpha_t + \beta_t \ln y_{it} + \boldsymbol{\delta}_t' \mathbf{z}_{it} + v_{it}$$

this is equation (3.22) in the text above.

Tables 3.6, 3.7, 3.8, 3.9, 3.10, 3.11, 3.12 and 3.13 refer to the least squares endogeneity-corrected regression

$$\begin{aligned} w_{jit} - E(w_{jit}|\ln x_{it}) = \quad & \pi_{jt}(v_{it} - E(v_{it}|\ln x_{it})) + \\ & \boldsymbol{\rho}_{jt}'(\mathbf{z}_{it} - E(\mathbf{z}|_{it} \ln x_{it})) + \varepsilon_{jit} \end{aligned}$$

for household i, in period t, for good j, this refers to equation (3.17) in the text above. The order of the commodity groups is the same as that in table 3.3 above (Eq (1) is beer, Eq(22) is petrol). The three parameters in each equation refer to (in order) π_{jt}, and $\boldsymbol{\rho}_{jt}$ which is (2×1) (the demographic vector $\mathbf{z}_{it}$ is (2×1) for household i, with the number of children as the first element and the number of adults minus 2 as the second). Parameters which are significant at 95% are identified by **, those significant at 90% are identified by *.

Table 3.5: First stage equation.

Year	Constant	log Income	No. Children	No. Adults-2
1974	1.7831**	0.4301**	0.0521**	0.1726**
1975	1.9069**	0.4244**	0.0558**	0.1845**
1976	2.1322**	0.3863**	0.0679**	0.1756**
1977	2.0277**	0.4347**	0.0549**	0.1758**
1978	2.0962**	0.4272**	0.0570**	0.1791**
1979	2.0159**	0.4631**	0.0572**	0.1606**
1980	2.4614**	0.3820**	0.0670**	0.2012**
1981	2.4873**	0.3891**	0.0707**	0.1965**
1982	2.4251**	0.4122**	0.0623**	0.1800**
1983	2.5719**	0.3937 **	0.0569**	0.1747**
1984	2.4475**	0.4212**	0.0696**	0.1983**
1985	2.4612**	0.4233**	0.0731**	0.1788**
1986	2.7977**	0.3656**	0.0794**	0.2196**
1987	2.6777**	0.3948**	0.0652**	0.1991**
1988	2.7592**	0.3836**	0.0833**	0.2053**
1989	2.8236**	0.3808**	0.0753**	0.2131**
1990	2.8705**	0.3798**	0.0851**	0.2044**
1991	3.0462**	0.3512**	0.0894**	0.2425**
1992	3.0574**	0.3566**	0.0764**	0.2289**
1993	3.0082**	0.3754**	0.0681**	0.2277**

Table 3.6: Endogeneity-corrected Regression, goods 1 to 5, 1974 to 1983.

	Eq(1)	Eq(2)	Eq(3)	Eq(4)	Eq(5)
1974	0.0109**	-0.0077**	-0.0056*	0.0093**	0.0150**
	-0.0016**	-0.0024**	-0.0037**	0.0087**	0.0031**
	0.0128**	-0.0058**	-0.0047**	0.0136**	0.0193**
1975	0.0013	-0.0120**	-0.0036	0.0032*	0.0187**
	-0.0022**	-0.0030**	-0.0038**	0.0084**	0.0025**
	0.0106**	-0.0064**	-0.0017	0.0122**	0.0210**
1976	0.0044	-0.0110**	-0.0067*	0.0045**	0.0236**
	-0.0039**	-0.0034**	-0.0055**	0.0074**	0.0022**
	0.0191	-0.0082**	-0.0077**	0.0111**	0.0196**
1977	0.0021	-0.0089**	-0.0093**	0.0066**	0.0298**
	-0.0032**	-0.0025**	-0.0049**	0.0075**	0.0030**
	0.0101**	-0.0067**	-0.0059**	0.0110**	0.0228**
1978	0.0013	-0.0072**	-0.0117**	0.0057**	0.0189**
	-0.0019**	-0.0026**	-0.0066**	0.0086**	0.0037**
	0.0112**	-0.0050**	-0.0078**	0.0125**	0.0179**
1979	-0.0009	-0.0105**	-0.0113**	0.0063**	0.0208**
	-0.0030**	-0.0031**	-0.0061**	0.0074**	0.0029**
	0.0089**	-0.0075**	-0.0073**	0.0107**	0.0177**
1980	0.0064*	-0.0103**	-0.0076**	0.0064**	0.0060
	-0.0030**	-0.0031**	-0.0047**	0.0082**	0.0022**
	0.0084**	-0.0078**	-0.0051**	0.0120**	0.0147**
1981	0.0012	-0.0064**	-0.0036	0.0030*	0.0137**
	-0.0044**	-0.0034**	-0.0054**	0.0078**	0.0031**
	0.0101**	-0.0070**	-0.0040**	0.0109**	0.0168**
1982	-0.0031	-0.0124**	-0.0035	0.0087**	0.0250**
	-0.0046**	-0.0038**	-0.0048**	0.0071**	0.0027**
	0.0074**	-0.0090**	-0.0037**	0.0110**	0.0220**
1983	0.0044	-0.0124**	-0.0028	0.0054**	0.0208**
	-0.0060**	-0.0037**	-0.0046**	0.0058**	0.0034**
	0.0106**	-0.0090**	-0.0038**	0.0099**	0.0176**

Table 3.7: Endogeneity-corrected Regression, goods 6 to 11, 1974 to 1983.

	Eq(6)	Eq(7)	Eq(8)	Eq(9)	Eq(10)	Eq(11)
1974	0.0047*	0.0024	0.0036	-0.0398**	-0.0009	-0.0101**
	0.0098**	0.0055**	0.0066**	-0.0029**	0.0019**	0.0009**
	0.0132**	0.0086**	0.0096**	-0.0180**	0.0011	-0.0032**
1975	0.0039*	-0.0030	0.0039	-0.0405**	-0.0046*	-0.0089**
	0.0103**	0.0061**	0.0095**	-0.0034**	0.0023**	0.0003
	0.0137**	0.0081**	0.0127**	-0.0183**	0.0002	-0.0032**
1976	0.0090**	0.0042	0.0091**	-0.0378**	-0.0000	-0.0090**
	0.0106**	0.0059**	0.0084**	-0.0034**	0.0023**	0.0015**
	0.0146**	0.0106**	0.0114**	-0.0183**	0.0024**	-0.0033**
1977	0.0137**	0.0049*	0.0113**	-0.0449**	0.0029	-0.0111**
	0.0115**	0.0059**	0.0072**	-0.0014**	0.0031**	0.0013**
	0.0152**	0.0085**	0.0105**	-0.0149**	0.0019*	-0.0025**
1978	0.0086**	-0.0026	0.0088**	-0.0441**	0.0015	-0.0163**
	0.0110**	0.0050**	0.0081**	-0.0018**	0.0024**	0.0013**
	0.0157**	0.0058**	0.0117**	-0.0159**	0.0024**	-0.0052**
1979	0.0045**	-0.0008	0.0148**	-0.0456**	-0.0058**	-0.0134**
	0.0099**	0.0048**	0.0073**	-0.0030**	0.0030**	0.0010**
	0.0133**	0.0061**	0.0097**	-0.0140**	0.0022**	-0.0042**
1980	0.0061**	-0.0077**	0.0121**	-0.0485**	0.0006	-0.0124**
	0.0113**	0.0052**	0.0082**	-0.0017**	0.0028**	0.0012**
	0.0137**	0.0042**	0.0094**	-0.0140**	0.0025**	-0.0044**
1981	0.0058**	-0.0069**	0.0105**	-0.0346**	-0.0064**	-0.0090**
	0.0112**	0.0049**	0.0088**	-0.0019**	0.0028**	0.0008*
	0.0140**	0.0052**	0.0108**	-0.0113 **	0.0012	-0.0027**
1982	0.0067**	-0.0040*	0.0100**	-0.0483**	-0.0008	-0.0188**
	0.0100**	0.0055 **	0.0073**	-0.0025**	0.0032**	0.0013**
	0.0146**	0.0062**	0.0100**	-0.0174**	0.0028**	-0.0032**
1983	0.0005	-0.0073**	0.0052**	-0.0415**	0.0016	-0.0181**
	0.0064**	0.0045**	0.0082**	-0.0057	0.0037**	0.0022**
	0.0120**	0.0041**	0.0062**	-0.0102**	0.0033**	-0.0047**

Table 3.8: Endogeneity-corrected Regression, goods 1 to 5, 1984 to 1993.

	Eq(1)	Eq(2)	Eq(3)	Eq(4)	Eq(5)
1984	0.0078**	-0.0151**	-0.0021	0.0063**	0.0162**
	-0.0035**	-0.0034**	-0.0049**	0.0076**	0.0025**
	0.0096**	-0.0090**	-0.0030**	0.0107**	0.0185**
1985	0.0051*	-0.0164**	-0.0080**	0.0021	0.0132**
	-0.0035**	-0.0045**	-0.0044**	0.0072**	0.0035**
	0.0138**	-0.0102**	-0.0039**	0.0100**	0.0184**
1986	-0.0038	-0.0142**	0.0007	0.0060**	0.0129**
	-0.0052**	-0.0042**	-0.0044**	0.0073**	0.0035**
	0.0071**	-0.0097**	-0.0021*	0.0101**	0.0175**
1987	0.0074*	-0.0145**	-0.0021	0.0061**	0.0189**
	-0.0045**	-0.0043**	-0.0056**	0.0080**	0.0038**
	0.0110**	-0.0103**	-0.0022**	0.0106**	0.0204**
1988	0.0188**	-0.0097**	-0.0016	0.0049**	0.0191**
	-0.0036**	-0.0038**	-0.0047**	0.0075**	0.0049**
	0.0125**	-0.0079**	-0.0024**	0.0108**	0.0203**
1989	0.0080**	-0.0099**	0.0021	0.0055**	0.0119**
	-0.0034**	-0.0034**	-0.0041**	0.0070**	0.0033**
	0.0096**	-0.0079**	-0.0019**	0.0104**	0.0171**
1990	0.0077*	-0.0141**	-0.0010	0.0059**	0.0164**
	-0.0046**	-0.0037**	-0.0040**	0.0072**	0.0047**
	0.0061**	-0.0091**	-0.0026**	0.0111**	0.0201**
1991	0.0048	-0.0132**	-0.0014	0.0011	0.0185**
	-0.0051**	-0.0041**	-0.0045**	0.0066**	0.0032**
	0.0085**	-0.0087**	-0.0017*	0.0096**	0.0205**
1992	0.0016	-0.0112**	0.0003	0.0030**	0.0159**
	-0.0053**	-0.0036**	-0.0047**	0.0070	0.0035**
	0.0082**	-0.0088**	-0.0029**	0.0099**	0.0199**
1993	0.0046	-0.0082**	-0.0028	0.0021	0.0143**
	-0.0060**	-0.0037**	-0.0046**	0.0058**	0.0034**
	0.0103**	-0.0069**	-0.0035**	0.0095**	0.0186**

Table 3.9: Endogeneity-corrected Regression, goods 6 to 11, 1984 to 1993.

	Eq(6)	Eq(7)	Eq(8)	Eq(9)	Eq(10)	Eq(11)
1984	0.0056**	-0.0041*	0.0042	-0.0368**	0.0035	-0.0149**
	0.0094**	0.0054**	0.0075**	-0.0030**	0.0033**	0.0007
	0.0126**	0.0052**	0.0091**	-0.0121**	0.0039**	-0.0043**
1985	0.0040**	-0.0066**	0.0037	-0.0319**	0.0045*	-0.0110**
	0.0087**	0.0062**	0.0066**	-0.0037**	0.0034**	0.0015**
	0.0122**	0.0037**	0.0082**	-0.0110**	0.0047**	-0.0030**
1986	0.0048**	-0.0078**	0.0050*	-0.0435**	0.0065**	-0.0184**
	0.0078**	0.0051**	0.0075**	-0.0051**	0.0036**	0.0001
	0.0122**	0.0043**	0.0067**	-0.0163**	0.0047**	-0.0041**
1987	0.0035**	-0.0056**	0.0061**	-0.0451**	0.0070**	-0.0158**
	0.0086**	0.0063**	0.0076**	-0.0061**	0.0037**	0.0013**
	0.0114**	0.0059**	0.0067**	-0.0187**	0.0043**	-0.0034**
1988	0.0049**	-0.0068**	0.0082**	-0.0416**	0.0024	-0.0093**
	0.0078**	0.0060**	0.0082**	-0.0081**	0.0032**	0.0015**
	0.0113**	0.0044**	0.0095**	-0.0157**	0.0050**	-0.0033**
1989	0.0029*	-0.0067**	0.0037*	-0.0385**	-0.0023	-0.0102**
	0.0074**	0.0052**	0.0074**	-0.0060**	0.0031**	0.0019**
	0.0109**	0.0049**	0.0083**	-0.0145**	0.0035**	-0.0017*
1990	0.0078**	-0.0012	0.0010	-0.0502**	0.0051**	-0.0079**
	0.0075**	0.0061**	0.0075**	-0.0067**	0.0032**	0.0008*
	0.0127**	0.0081**	0.0079**	-0.0164**	0.0048**	-0.0016*
1991	0.0071**	-0.0030	0.0024	-0.0403**	0.0050**	-0.0081**
	0.0080**	0.0053**	0.0072**	-0.0057**	0.0032**	0.0016**
	0.0127**	0.0066**	0.0083**	-0.0149**	0.0056**	-0.0023**
1992	0.0020	-0.0034*	0.0032	-0.0414**	0.0012	-0.0002
	0.0070**	0.0053**	0.0086**	-0.0065**	0.0036**	0.0022**
	0.0113**	0.0065**	0.0090**	-0.0184**	0.0037**	0.0018*
1993	0.0013	-0.0036*	-0.0012	-0.0408**	0.0072**	-0.0065**
	0.0064**	0.0045**	0.0082**	-0.0057**	0.0037**	0.0022**
	0.0103**	0.0060**	0.0091**	-0.0141**	0.0053**	-0.0005

Table 3.10: Endogeneity-corrected Regression, goods 12 to 16, 1974 to 1983.

	Eq(12)	Eq(13)	Eq(14)	Eq(15)	Eq(16)
1974	0.0270**	0.0215**	-0.0390**	-0.0073	-0.0066
	-0.0072**	0.0104**	-0.0055**	-0.0048**	-0.0017**
	-0.0021	-0.0013	-0.0236**	-0.0061**	-0.0082**
1975	0.0151**	0.0210**	-0.0316**	-0.0054	-0.0042
	-0.0092**	0.0107**	-0.0059**	-0.0054**	-0.0011
	-0.0062**	-0.0030	-0.0234**	-0.0048**	-0.0059**
1976	0.0203**	0.0204**	-0.0261**	-0.0028	-0.0105**
	-0.0061**	0.0120**	-0.0045**	-0.0049**	-0.0020**
	-0.0052**	-0.0010	-0.01738**	-0.0040**	-0.0067**
1977	-0.0049	0.0158**	-0.0269**	-0.0070	-0.0070
	-0.0087**	0.0100**	-0.0036**	-0.0055**	-0.0012
	-0.0144**	-0.0006	-0.0196**	-0.0055**	-0.0072**
1978	0.0245**	0.0274**	-0.0328**	-0.0118**	-0.0110**
	-0.0082**	0.0116**	-0.0050**	-0.0060**	-0.0007
	-0.0055**	-0.0001	-0.0196**	-0.0079**	-0.0053**
1979	0.0074	0.0292**	-0.0277**	0.0003	-0.0117**
	-0.0096**	0.0128**	-0.0029**	-0.0049**	-0.0025**
	-0.0113**	-0.0014	-0.0179**	-0.0038**	-0.0076**
1980	0.0238**	0.0251**	-0.0383**	-0.0006	-0.0041
	-0.0096**	0.0115**	-0.0062**	-0.0050**	-0.0003
	-0.0025	-0.0029	-0.0220**	-0.0065**	-0.0035*
1981	0.0192**	0.0202**	-0.0324**	-0.0116**	-0.0178**
	-0.0081**	0.0100**	-0.0042**	-0.0058**	-0.0021**
	-0.0035*	-0.0031	-0.0193**	-0.0076**	-0.0096**
1982	0.0171**	0.0219**	-0.0309**	-0.0070	-0.0025
	-0.0067**	0.0124**	-0.0042**	-0.0047**	-0.0007
	-0.0026	-0.0022	-0.0178**	-0.0051**	-0.0089**
1983	0.0298**	0.0164**	-0.0335**	-0.0102**	-0.0058
	-0.0079**	0.0124**	-0.0051**	-0.0039**	-0.0002
	-0.0011	-0.0014	-0.0218**	-0.0059**	-0.0041**

Table 3.11: Endogeneity-corrected Regression, goods 17 to 22, 1974 to 1983.

	Eq(17)	Eq(18)	Eq(19)	Eq(20)	Eq(21)	Eq(22)
1974	-0.0477**	-0.0033	-0.0144**	0.0144**	0.0597**	0.0140**
	-0.0036**	-0.0003	-0.0013**	0.0004	-0.0076**	-0.0047**
	-0.0188**	-0.0004	0.0007	0.0135**	0.0003	-0.0005
1975	-0.0291**	-0.0024	-0.0125**	0.0349**	0.0424**	0.0133**
	-0.0033**	0.0004	-0.0020**	0.0014*	-0.0074**	-0.0050**
	-0.0220**	0.0001	0.0024**	0.0214**	-0.0072**	-0.0002
1976	-0.0277**	-0.0003	-0.0170**	0.0220**	0.0275**	0.0040
	-0.0046**	0.0005	-0.0028**	0.0003	-0.0060**	-0.0038**
	-0.0207**	0.0006	0.0006	0.0158**	-0.0034	-0.001
1977	-0.0498**	-0.0019	-0.0145**	0.0329**	0.0474**	0.0188**
	-0.0039**	0.0004	-0.0024**	0.0019**	-0.0097**	-0.0049**
	-0.0244**	0.0007	0.0017	0.0202**	-0.0053*	0.0042**
1978	-0.0363**	-0.0038**	-0.0101**	0.0318**	0.0559**	0.0033
	-0.0055**	-0.0002	-0.0021**	0.0009	-0.0087**	-0.0033**
	-0.0255**	0.0001	0.0023	0.0182**	-0.0019	0.0021
1979	-0.0365**	-0.0049**	-0.0123**	0.0259**	0.0596**	0.0126**
	-0.0028**	0.0006**	-0.0016**	0.0019**	-0.0067**	-0.0053**
	-0.0173**	-0.0006	0.0027**	0.0143**	0.0061**	0.0013
1980	-0.0371**	0.0018	-0.0141**	0.0296**	0.0588**	0.0039
	-0.0031**	0.0003	-0.0032**	0.0027**	-0.0077**	-0.0061**
	-0.0220**	0.0012**	0.0010	0.0180**	0.0033	0.0023
1981	-0.0406**	0.0001	-0.0117**	0.0328**	0.0511**	0.0232**
	-0.0037**	0.0004*	-0.0025**	0.0021**	-0.0064**	-0.0041**
	-0.0221**	0.0009*	0.0021**	0.0179**	-0.0022	0.0027
1982	-0.0439**	-0.0001	-0.0153**	0.0354**	0.0476**	0.0180**
	-0.0056**	0.0003	-0.0033**	0.0022**	-0.0067**	-0.0044**
	-0.0223**	0.0006	0.0004	0.0182**	-0.0021	0.0011
1983	-0.0327**	-0.0010	-0.0161**	0.0348**	0.0402**	0.0221**
	-0.0034**	0.0002	-0.0031**	0.0013	-0.0107**	-0.0041**
	-0.0200**	0.0004	-0.0016	0.0191**	-0.0011	0.0013

Table 3.12: Endogeneity-corrected Regression, goods 12 to 16, 1984 to 1993.

	Eq(12)	Eq(13)	Eq(14)	Eq(15)	Eq(16)
1984	0.0203**	0.0240**	-0.0347**	0.0015	-0.0241**
	-0.0081**	0.0125**	-0.0053**	-0.0048**	-0.0022**
	-0.0050**	-0.0000	-0.0233**	-0.0042**	-0.0113**
1985	0.0214**	0.0204**	-0.0318**	0.0018	-0.0109**
	-0.0078**	0.0156**	-0.0043**	-0.0044**	-0.0020**
	-0.0052**	-0.0006	-0.0204**	-0.0031*	-0.0080**
1986	0.0065	0.0225**	-0.0193**	-0.0030	-0.0111**
	-0.0112**	0.0119**	-0.0026**	-0.0045**	-0.0025**
	-0.0092**	0.0013	-0.0165**	-0.0028*	-0.0083**
1987	0.0057	0.0176**	-0.0297**	0.0065	-0.0091**
	-0.0083**	0.0147**	-0.0047**	-0.0044**	-0.0013
	-0.0033	0.0022	-0.0231**	-0.0032*	-0.0083**
1988	0.0157**	0.0179**	-0.0333**	-0.0031	-0.0053
	-0.0097**	0.0142**	-0.0055**	-0.0049**	-0.0011
	-0.0007	0.0002	-0.0249**	-0.0127**	-0.0054**
1989	0.0137**	0.0143**	-0.0203**	-0.0023	-0.0070*
	-0.0103**	0.0110**	-0.0027**	-0.0052**	-0.0030**
	-0.0051**	0.0001	-0.0215**	-0.0111**	-0.0079**
1990	0.0191**	0.0177**	-0.0257**	-0.0062	-0.0130**
	-0.0092**	0.0115**	-0.0038**	-0.0051**	-0.0032**
	-0.0022	0.0023	-0.0262**	-0.0137**	-0.0095**
1991	0.0091	0.0184**	-0.0230**	0.0060	-0.0066*
	-0.0098	0.0131**	-0.0014**	-0.0016	-0.0025**
	-0.0083**	0.0019	-0.0233**	-0.0092**	-0.0069**
1992	0.0203**	0.0134**	-0.0250**	0.0029	-0.0080*
	-0.0097**	0.0135 **	-0.0035**	-0.0056**	-0.0028**
	-0.0007	-0.0010	-0.0254**	-0.0034*	-0.0103**
1993	0.0052	0.0145**	-0.0127**	-0.0080*	-0.0096**
	-0.0073**	0.0121**	-0.0040**	-0.0047**	-0.0045**
	-0.0091**	0.0005	-0.0262**	-0.0051**	-0.0095**

Table 3.13: Endogeneity-corrected Regression, goods 17 to 22, 1984 to 1993.

	Eq(17)	Eq(18)	Eq(19)	Eq(20)	Eq(21)	Eq(22)
1984	-0.0318**	0.0001	-0.0112**	0.0320**	0.0408**	0.0125**
	-0.0038**	0.0002	-0.0021**	0.0019**	-0.0066**	-0.0034**
	-0.0204**	0.0013**	-0.0009	0.0206**	-0.0005	0.0028
1985	-0.0374**	0.0002	-0.0087**	0.0366**	0.0275**	0.0221**
	-0.0042**	0.0007**	-0.0024**	0.0018**	-0.0087**	-0.0053**
	-0.0215**	0.0014**	-0.0007	0.0200**	-0.0024	-0.0024
1986	-0.0256**	-0.0030**	-0.0131**	0.0273**	0.0438**	0.0269**
	-0.0017	0.0002	-0.0022**	0.0040**	-0.0061**	-0.0015
	-0.0221**	0.0012**	-0.0013	0.0165**	0.0059**	0.0048**
1987	-0.0257**	0.0015	-0.0152**	0.0362**	0.0229**	0.0234**
	-0.0022**	0.0001	-0.0029**	0.0030**	-0.0094**	-0.0034**
	-0.0213**	0.0011**	-0.0013	0.0190**	-0.0015	0.0040**
1988	-0.0281**	-0.0009	-0.0132**	0.0291**	0.0306**	0.0011
	-0.0023**	0.0005**	-0.0030**	0.0039**	-0.0076**	-0.0034**
	-0.0225**	0.0012**	-0.0004	0.0177**	0.0009	0.0022
1989	-0.0234**	-0.0016*	-0.0063**	0.0244**	0.0227**	0.0191**
	-0.0007	-0.0000	-0.0024**	0.0023**	-0.0066**	-0.0008
	-0.0203**	0.0010**	0.0010	0.0156**	0.0032	0.0066**
1990	-0.0343**	0.0017	-0.0115**	0.0314**	0.0410**	0.0104**
	-0.0019*	0.0002	-0.0018**	0.0029**	-0.0075**	-0.0000
	-0.0218**	0.0014**	0.0001	0.0162**	0.0051**	0.0070**
1991	-0.0275**	0.0012	-0.0127**	0.0302**	0.0221**	0.0098**
	-0.0033**	0.0002	-0.0027**	0.0036**	-0.0083**	-0.0029**
	-0.0237**	0.0019**	-0.0004	0.0206**	-0.0040	0.0070**
1992	-0.0364**	0.0003	-0.0147**	0.0388**	0.0307**	0.0068*
	-0.0044**	0.0006**	-0.0020**	0.0060**	-0.0079**	-0.0014
	-0.0287**	0.0020**	-0.0007	0.0225**	-0.0003	0.0058**
1993	-0.0234**	-0.0014	-0.0096**	0.0354**	0.0304**	0.0127**
	-0.0012	0.0007**	-0.0018**	0.0040**	-0.0076**	-0.0000
	-0.0222**	0.0013**	-0.0004	0.0205**	0.0000	0.0061**

Chapter 4

Revealed Preference Methods for Bounding True Cost of Living Indices

4.1 Introduction

This chapter is about measuring a true cost of living index, or at least approximating it as closely as possible *without* making any assumptions regarding the form of the preferences underlying the index. The existence of the index is tested using similar techniques to those developed in Chapter 3. The reason why approximations to true indices have been sought in the long literature on the subject is that, while a true cost of living index is a reasonably straightforward theoretical concept, the informational requirements necessary for true indices to be estimated are onerous and fully parametrised models may often reject the restrictions which theory places upon them. This is because, as discussed in the introductory chapter, the calculation of a true index based upon a demand system carries the dual implicit maintained hypotheses that, (a) the data are rationalisable by a stable and well-behaved set of preferences, and that, (b) the functional form of those preferences is known (and is the one estimated). Either or both of these assumption may be untrue and we have no way of knowing which because the hypotheses are imposed jointly.

As a result of the difficulties in estimating true cost of living indices, simple approximations to true indices like the Paasche and Laspeyres price indices are often used since these require only information on observed equilibrium price and quantity outcomes, and, unless one wishes to interpret them as exact versions of a true index, then they do not require any assumption on the functional form of underlying preferences.

This chapter suggests a revealed preference method which, without the need for functional and parametric assumptions, allows *two-sided* bounds to be place on a given true index recovered from average demands. These bounds, are shown to be as least as tight as the classical nonparametric bounds discussed in the literature and are also the tightest bounds obtainable given only the (testable) assumption that a given finite dataset of prices and average demands was generated by a well-behaved utility function. The algorithm which computes the bounds turns out to be a powerful method of performing this test. This improved bound is compared to other, popular approximations and is used to see how close these indices seem to approximate the truth. The plan of this chapter is as follows. Section 4.2 discusses the principal theorems in the literature on nonparametric bounds to true cost of living indices. Section 4.3 discusses how revealed preference information can be used to supplement simple indices. A method based on the Generalised Axiom of Revealed Preference which improves these classical bounds further without recourse to functional assumptions in then presented and discussed. Section 4.4 discusses the conditions under which local average demands provide a valid basis upon which to compute average welfare bounds. This is an extension of the discussion on heterogeneity in Chapter 3. Section 4.5 briefly describes the U.K. micro data used in the empirical application. Section 4.6 reports results on the subperiods within which bounds can be recovered and compares a number of popular, exact and superlative indices with classical revealed preference bounds and the improved GARP-based bounds on the true index. It also provides further evidence on differences in cost of living indices by total expenditure level (i.e. non-homotheticity which was discussed in Chapter 2), and presents estimates of the bounds on the substitution bias incurred by the use of the simple Laspeyres approximation. Section 4.7 concludes.

4.2 True Indices and Approximations

In this section the main theorems of nonparametric bounds on true cost of living indices are presented and discussed[1]. Many of these theorems have been proved by several authors over time and under sometimes somewhat different assumptions/regularity conditions. I have not reproduced the proofs here. Diewert's (1978) survey brings all of these results together *and* provides proofs which are either directly borrowed from or are slightly modified versions of the originals in which he has relaxed some of the author's conditions. The literature on bounds to true indices has grown up because of the strong wish to avoid parametric estimation. The idea is to use only the (usually) observable price and (typically average) demand data to approximate a true index $(P_T(\mathbf{p}_0, \mathbf{p}_1, u(\mathbf{q})))$. This section starts with the simplest cases and then traces the improvements to the bounds which additional data can bring about.

To begin, consider the case in which there is only one good and two

[1] To recap the notation: $\mathbf{x} \geq 0$ means that each element of the vector $\mathbf{x}$ is non-negative. $\mathbf{x} \gg 0$ means that each element of the vector $\mathbf{x}$ is positive. $\mathbf{x} > 0$ means $\mathbf{x} \geq 0$ but $\mathbf{x} \neq 0$.

periods $(1, 0)$. Then the true index becomes $P_T(\mathbf{p}_0, \mathbf{p}_1, u(\mathbf{q})) = p_1^1/p_0^1$ for all $\mathbf{q} > 0$. Note that if price in the two periods are proportional then P_T is equal to the common factor of proportionality for any $\mathbf{q} > 0$, i.e. for $\lambda > 0$ and $\mathbf{p}_0, \mathbf{p}_1 \gg 0$, we have $P_T(\mathbf{p}_0, \lambda\mathbf{p}_0, u(\mathbf{q})) = \lambda$ and $P_T(\mathbf{p}_0, \mathbf{p}_1, u(\mathbf{q})) = 1/P_T(\mathbf{p}_1, \mathbf{p}_0, u(\mathbf{q}))$. If prices are not proportional then P_T will depend upon $u(\mathbf{q})$ unless $u(\mathbf{q})$ is homothetic (see Proposition 2.1 in Chapter 2).

Now suppose that there are more than two goods but quantities are not known – only prices in each period are observed.

Theorem 4.1. (Due to Lerner (1935), Joseph (1935-36), Samuelson (1947), Pollak (1971)). *If $u(\cdot)$ is continuous, then for every $\mathbf{p}_0, \mathbf{p}_1 \gg 0$, $\overline{\mathbf{q}} > 0$ and $u(\mathbf{q}) > u(0)$, $\min_i\{p_1^i/p_0^i : i = 0, 1, .., K\} \leq P_T(\mathbf{p}_0, \mathbf{p}_1, u(\mathbf{q})) \leq \max_i\{p_1^i/p_0^i : i = 0, 1, .., K\}$.*

That is, the true index must lie between the smallest and the largest price ratio of all the goods which enter the utility function $u(q)$.

Now suppose that there is a little more information than is available in the above case. Suppose that quantities as well as prices are observed in each period $(\mathbf{p}_0, \mathbf{q}_0)$ and $(\mathbf{p}_1, \mathbf{q}_1)$. Further assume that the consumer is a neo-classical utility maximiser/cost minimiser. There are now two obvious candidates for the quantity vector which enters $u(\mathbf{q})$; either $\mathbf{q}_0$ or $\mathbf{q}_1$. Taking these reference points ($u(\mathbf{q}_0)$ and $u(\mathbf{q}_1)$) in chronological order we have the following famous bounds:

Theorem 4.2. (Due to Konüs (1924)). *Suppose that $u(\cdot)$ is continuous and also assume cost-minimising behaviour. Then $P_T(\mathbf{p}_0, \mathbf{p}_1, u(\mathbf{q}_0)) \leq \mathbf{p}_1'\mathbf{q}_0/\mathbf{p}_0'\mathbf{q}_0 \equiv P_L$, where P_L is the Laspeyres price index.*

Corollary 4.2.1.(Due to Pollak (1971)). $\min_i\{p_1^i/p_0^i : i = 0, 1, .., K\} \leq$

$P_T(\mathbf{p}_0, \mathbf{p}_1, u(\mathbf{q}_0)) \leq \mathbf{p}_1'\mathbf{q}_0/\mathbf{p}_0'\mathbf{q}_0 \equiv P_L.$

Theorem 4.3. (Also due to Konüs (1924)). *Suppose that $u(\cdot)$ is continuous and also assume cost-minimising behaviour. Then $P_T(\mathbf{p}_0, \mathbf{p}_1, u(\mathbf{q}_1)) \geq \mathbf{p}_1'\mathbf{q}_1/\mathbf{p}_0'\mathbf{q}_1 \equiv P_P$, where P_P is the Paasche price index.*

Corollary 4.3.1.(Also due to Pollak (1971)). $\max_i \{p_1^i/p_0^i : i = 0, 1, .., K\} \geq P_T(\mathbf{p}_0, \mathbf{p}_1, u(\mathbf{q}_1)) \geq \mathbf{p}_1'\mathbf{q}_1/\mathbf{p}_0'\mathbf{q}_1 \equiv P_P.$

The intuition behind the Konüs single-sided bounds in theorems 4.2 and 4.3 is straightforward and was discussed in Chapter 1[2]. As shown in equation (1.5) in Chapter 1, only in the case where preferences exhibit no substitution effects are the Paasche and Laspeyres indices identical to their corresponding true indices. In other words, the only case in which the Paasche and Laspeyres indices are precisely equal to their true counterparts is one in which preferences are Leontief, when the cost function takes the particular form

$$c(\mathbf{p}, u) = \sum_i \alpha_i(u) p_i,$$

since, by Shephard's lemma

$$\frac{\partial c(\mathbf{p}, u)}{\partial p_i} = q_i = \alpha_i(u).$$

Corollaries 4.2.1 and 4.3.1 which provide two-sided bounds by combining the results of theorems 4.1 and 4.2 and 4.3 respectively, generally give tighter bounds than their corresponding price-only bounds (theorem 4.1). This is

[2]The idea is that, in the case of the Laspeyres index, $\mathbf{q}_0$ may be the least cost way of achieving the reference welfare level $u_0(\mathbf{q}_0)$ with prices $\mathbf{p}_0$, but this is not necessarily the case once prices have changed to $\mathbf{p}_1$. As a result $\mathbf{p}_1'\mathbf{q}_0 \geq c(\mathbf{p}_1, u_0)$ (this result was also presented in equation (1.7) in Chapter 1). The argument for the Paasche index is analogous.

because a weighted average of price changes/ratios must be bounded from above by the largest single price change.

Note that the Konüs inequalities refer to two *different* true indices; one in which the reference utility level is the period 0 level (u_0), and the other which takes u_1, i.e. the Paasche and Laspeyres provide one-sided bounds on two correspondingly different true indices. As it stands, then, the claim that they provide a two-side bound on a true single index is false. This, and the claim that the Laspeyres is always greater than or equal to the Paasche[3] requires the further (parametric) assumption of homotheticity this is set out in theorem 4.4.

Theorem 4.4. (Due to Frisch (1936)). *If $u(\cdot)$ is homothetic then for $q \gg 0$, $P_L \geq P_T(\mathbf{p}_0, \mathbf{p}_1, u(\mathbf{q})) \geq P_P$.*

As discussed in Chapter 2, homotheticity results in a unique index. If there is only one index, and if the Paasche approximates it from below and the Laspeyres from above (i.e. the Konüs bounds in theorems 4.2 and 4.3), then the result in theorem 4.4 follows immediately. It also follows that the Paasche index cannot exceed the Laspeyres, and that the true index lies somewhere in the interval. Further, *every* point in the interval can be a constant-utility or true index and, unless we have more information than is present in $(\mathbf{p}_0, \mathbf{q}_0)$, $(\mathbf{p}_1, \mathbf{q}_1)$ and the strong (and as Chapter 2 showed) empirically rejected assumption that preferences are homothetic, then no point in the interval has any claim to be truer that any other.

The homotheticity assumption can, however, be relaxed and the Paasche and Laspeyres indices can provide two-sided nonparametric bounds on a true index with a utility level (indifference surface) between u_0 and u_1.

[3]Shown in Afriat (1977).

Theorem 4.5. (Due to Konüs (1924)). *Suppose that $u(\cdot)$ is well behaved and also assume cost-minimising behaviour. Then $\exists$ some λ^* such that $0 \geq \lambda^* \geq 1$ and $P_T(\mathbf{p}_0, \mathbf{p}_1, u(\lambda^* \mathbf{q}_1 + (1 - \lambda^*) \mathbf{q}_0))$ lies between P_P and P_L, either so that $P_L \leq P_T \leq P_P$, or $P_P \leq P_T \leq P_L$.*

Since λ^* can be chosen so that $u(\lambda^* \mathbf{q}_1 + (1 - \lambda^*) \mathbf{q}_0)$ is a reference utility level which is somewhere between $u(\mathbf{q}_0)$ and $u(\mathbf{q}_1)$,the Paasche and Laspeyres indices bound a true index for a reference welfare level lying between the base and end period welfare levels. There is no requirement here for the Laspeyres to exceed the Paasche.

These theorems and corollaries provide *all* of the approximations and nonparametric bounds which can be placed on the true index under various circumstances and with access to various amounts of data. As Diewert points out, these are the tightest bounds available in the sense that they are unimprovable "unless we are willing to make specific assumptions about the functional for the aggregator [utility] function"[4]. The rest of this chapter presents a GARP-based method for improving these classical bounds, under the (testable) assumption that a given finite dataset is rationalisable by a well-behaved utility function, *without* specifying the form of the utility function.

4.3 Bounds from Revealed Preference

4.3.1 Classical Bounds

Afriat (1977) described the way in which revealed preference information can be used to provide classical bounds on the welfare effects of a price change. In many cases the revealed preference approach simply amounts to a restate-

[4]Diewert (1978), p. 170.

ment of the theorems given above in a different framework. This is shown in an example below. The general idea is that the axioms of revealed preference can be used to gain information on the curvature of indifference surfaces in commodity space and that this can be used to improve classical two sided bounds on the welfare effects of price changes. If the position of an indifference curve can be approximated, then the minimum cost of achieving that welfare at some different set of prices level can also be approximated by reading off the expenditure level of a budget surface (with the new prices) placed tangentially to the indifference curve bounds. The better the approximation to the indifference curve, the better the approximation to the cost of living index. For an example consider Figures 4.1 and 4.2.

Figure 4.1 shows a single price-quantity observation $(\mathbf{p}_0, \mathbf{q}_0)$ with expenditure $x_0 = \mathbf{p}_0'\mathbf{q}_0$. Assume that this observation was generated by a rational consumer. The bounds on the possible position of the indifference curve through that point are given by the shaded areas and these bounds are wide. The area revealed preferred to $\mathbf{q}_0$ (denoted $RP(\mathbf{q}_0)$) results simply from the monotonicity of utility. The area to which $\mathbf{q}_0$ is revealed preferred is the shaded area below the x_0 budget line (denoted $RW(\mathbf{q}_0)$). The indifference curve cannot pass through the boundary of either set but it can lie anywhere in between (or even along) these extremes which represent indifference curves which are either Leontief or straight lines — since GARP is a generalisation of the Strong Axiom of Revealed Preference (SARP) which allows for flat areas on indifference surfaces, while SARP does not. The resulting bounds on the welfare effects of the new price regime $\mathbf{p}_t$ are illustrated by the dashed lines. These bounds are wide and show the upper and lower bounds on the compensating expenditure level for the new budget constraint. The revealed preference bounds correspond exactly to the nonparametric bounds derived by Pollak (1971) (and described in the corollaries to theorems 4.2 and 4.3 above) who shows that the base period referenced true cost of living index is bounded from above by the Laspeyres price index and from below by the minimum relative price change between the periods considered.

Figure 4.1: Classical bounds, one observation.

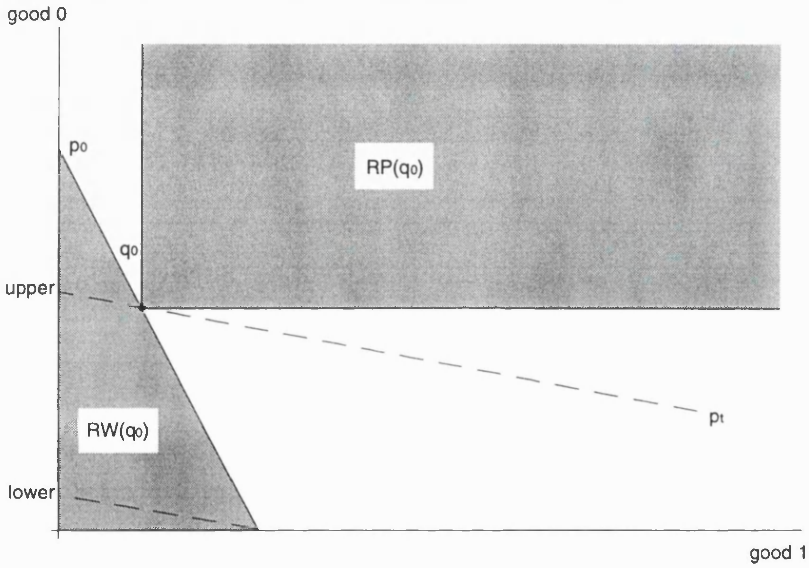
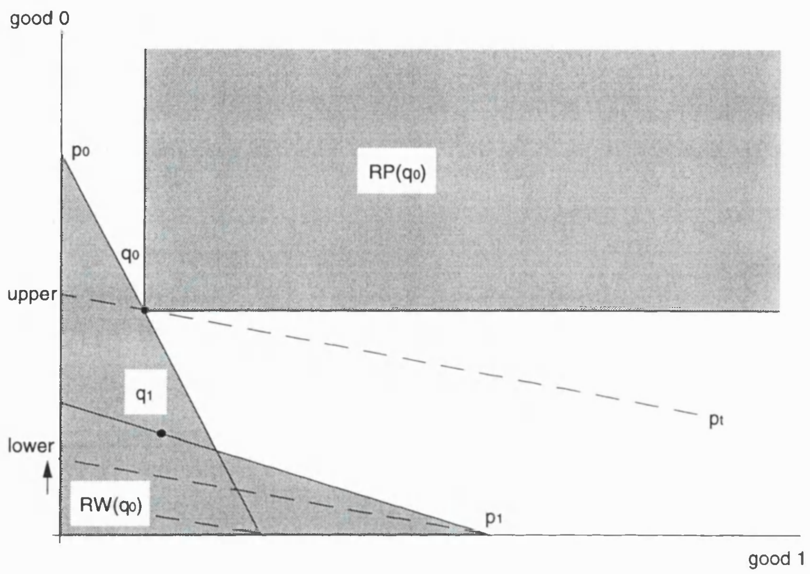


Figure 4.2 introduces a second price-quantity observation $(\mathbf{p}_1, \mathbf{q}_1)$ and the $RP(\mathbf{q}_0)$ and $RW(\mathbf{q}_0)$ sets are redrawn utilising this new information. In this case the $RP(\mathbf{q}_0)$ set is unchanged and the welfare effects remain bounded from above by Leontief preferences, but new information has been gained on the lower bound on the indifference curve giving evidence of some degree of curvature. The new lower bound uses both the x_0 and the $x_1 = \mathbf{p}_1' \mathbf{q}_1$ budget lines. And, since $\mathbf{q}_0 P^0 \mathbf{q}_1$, both $\mathbf{q}_1$ and points to which it is revealed preferred must lie below the indifference curve. In other word, while the indifference curve can run either along or above the $\mathbf{p}_0'\mathbf{q}_0$ budget line from the good 0 axis to the point where it meets the $\mathbf{p}_1'\mathbf{q}_1$ line, it must lie *above* the $\mathbf{p}_1'\mathbf{q}_1$ budget

Figure 4.2: Classical bound, two observations.



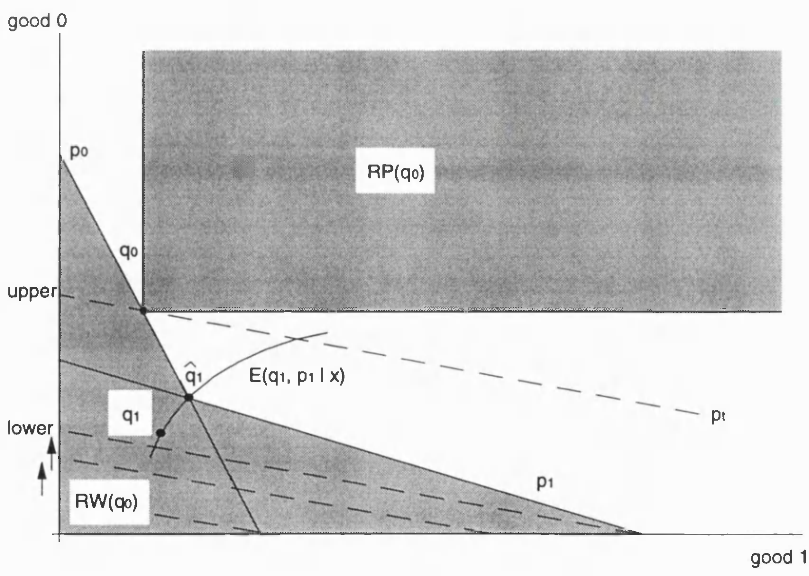
line from that point to the good 1 axis. There is no way, conditional on the present information, of telling how far above this line it can lie (except that it cannot cross the boundary of the $RP(q_0)$ set). Nevertheless, the extra information allows the lower bound to be tighter than was the case with the single $(\mathbf{p}_0, \mathbf{q}_0)$ observation. This motivates the procedure discussed in the next section which is designed to improve the bound further.

A second motivation is as follows. It is obvious that this sort of improvement to the welfare bounds from revealed preference is only possible when budget surfaces cross as only then do they convey much information on the curvature of indifference surfaces. This may, for the reasons discussed in Chapter 3, be a rare occurrence, particularly with aggregate or average demand data; indeed in Varian's (1982) applied work on GARP bounds on cost of living indices, improvements were only possible to the classical bounds for two years out of thirty two studied. In practice this has limited the usefulness of GARP-based bounds.

4.3.2 Improving the Bounds

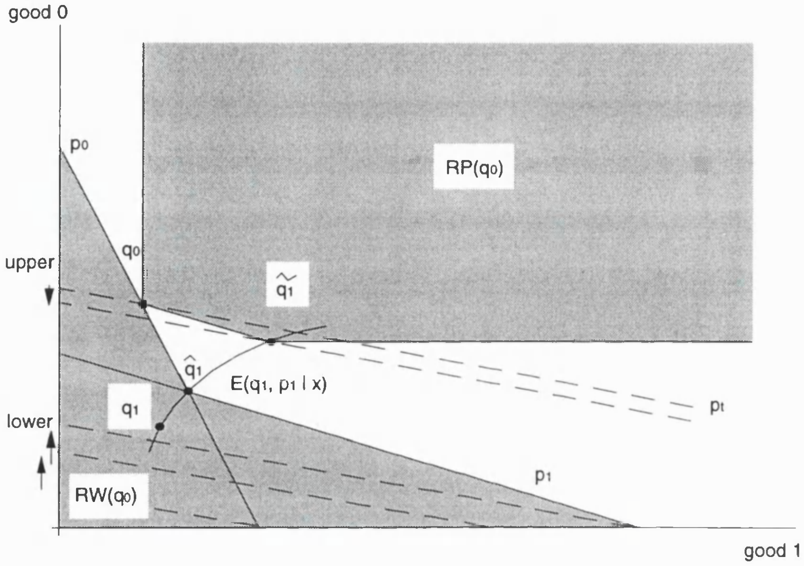
These problems of GARP-based bounds do not arise if it is possible to move budget surfaces around. And so, as in Chapter 3, knowledge of (nonparametric) expansion paths can be used to improve the informational content of revealed preference restrictions. Movements along expansion paths allow the maximum information on the curvature of the indifference curve through a given point to be utilised. This is because, by varying total expenditure, the budget surfaces can be placed as desired. To illustrate the idea, consider Figure 4.3.

Figure 4.3: GARP-improved lower bound, two observations.



The expansion path for demands with prices $\mathbf{p}_1$ is now added to the information available in Figure 4.2. This allows the budget line for prices $\mathbf{p}_1$ to be moved out to a higher total expenditure level (since $\mathbf{p}_1' \, \widehat{\mathbf{q}}_1 > \mathbf{p}_1' \, \mathbf{q}_1$).

Figure 4.4: GARP- improved upper and lower bounds, two observations.



Setting total spending and hence placing the budget line so that it lies on the budget surface which the base bundle $\mathbf{q}_0$ is also on, tightens the previously available bound on the indifference curve since now $\mathbf{p}_0'\mathbf{q}_0 = \mathbf{p}_0'\widehat{\mathbf{q}}_1$ and hence it remains the case that $\mathbf{q}_0 R^0 \widehat{\mathbf{q}}_1$. If the expenditure level were set any small amount (Δ) higher than $\mathbf{p}_1'\widehat{\mathbf{q}}_1$ then there would no longer be any additional revealed preference restrictions on the path of the indifference curve. This is because $\mathbf{p}_0'(\widehat{\mathbf{q}}_1 + \Delta) > \mathbf{p}_0'\widehat{\mathbf{q}}_1$ implies $\mathbf{p}_0'\mathbf{q}_0 = \mathbf{p}_0'\widehat{\mathbf{q}}_1 < \mathbf{p}_0'(\widehat{\mathbf{q}}_1 + \Delta)$ i.e. it is no longer the case that $\mathbf{q}_0 R^0 (\widehat{\mathbf{q}}_1 + \Delta)$, and if Δ is small enough, then neither is it the case that $\mathbf{p}_1'(\widehat{\mathbf{q}}_1 + \Delta) \geq \mathbf{p}_1'\mathbf{q}_0$). If there is no revealed preference ranking of bundles, then there are no restrictions

A major additional benefit, and the third motivation for using this technique, is that each budget surface can be used *twice*; once to improve the lower bound on the indifference curve and once to improve the upper bound. This is illustrated in Figure 4.4. Here, the budget line using the $\mathbf{p}_1$ price vector is placed at an expenditure level such that $\mathbf{p}_1'\widetilde{\mathbf{q}}_1 = \mathbf{p}_1'\mathbf{q}_0$ which implies $\widetilde{\mathbf{q}}_1 R^0 \mathbf{q}_0$. By similar arguments as before, the indifference curve cannot pass above $\widetilde{\mathbf{q}}_1$ or the plane connecting $\mathbf{q}_0$ and $\widetilde{\mathbf{q}}_1$.

In practice the issue of placing the budget lines in this way so as to maximise the information on the indifference surface is not so straightforward as the budget surfaces are unlikely to be only two or three dimensional and it is therefore hard to see how to place the budget lines. The following algorithms provide upper and lower bounds on an indifference curve through a given point in commodity space with $T+1$ price regimes $\{0, 1, ..., T\}$:

$RP\,(\mathbf{q}_0)$ **Bound Algorithm**

Output is the set RP of boundary points of which $\mathbf{q}_0$ is a member and which has $T+1$ elements where $\mathbf{p}_i'\mathbf{q}_i \leq \mathbf{p}_i'\mathbf{q}_j \ \forall \ \mathbf{q}_i, \mathbf{q}_j \in RP$ and either $\mathbf{q}_i \ R^0 \ \mathbf{q}_0$ or $\mathbf{q}_i \ R \ \mathbf{q}_0$ for all $\mathbf{q}_i \in RP$.

1) Set $W = \{\mathbf{q}_0\}\,,\ \tau = \{0, 1, ..., T\}\,,\ E = \varnothing$

2) Set $R = \{q_t\,(\min\{x \mid \mathbf{p}_t'\mathbf{q}_t\,(x) = \mathbf{p}_t'\mathbf{q}_w\})\,, \mathbf{q}_w \in W, t \in \tau\}$

3) Set $E = \{\mathbf{q}_i \in R : \mathbf{p}_i'\mathbf{q}_i > \mathbf{p}_i'\mathbf{q}_j \text{ for some } \mathbf{q}_j \in R\}$

4) Set $W = R \backslash E$

5) If $E = \varnothing$ set $RP = W$ and stop. Otherwise go to (2).

$RW\,(\mathbf{q}_0)$ **Bound Algorithm**

Output is the set RW of boundary points of which $\mathbf{q}_0$ is a member and which has $T+1$ elements where $\mathbf{p}_i'\mathbf{q}_i \leq \mathbf{p}_i'\mathbf{q}_j \ \forall \ \mathbf{q}_i, \mathbf{q}_j \in RW$ and either $\mathbf{q}_0 \ R^0 \ \mathbf{q}_i$ or $\mathbf{q}_0 \ R \ \mathbf{q}_i$ for all $\mathbf{q}_i \in RW$.

1) Set $B = \{\mathbf{q}_0\}\,,\ \tau = \{0, 1, ..., T\}\,,\ E = \varnothing$

2) Set $R = \{q_t\,(\max\{x \mid \mathbf{p}_b'\mathbf{q}_b\,(x) = \mathbf{p}_b'\mathbf{q}_t\})\,, \mathbf{q}_b \in B, t \in \tau\}$

3) Set $E = \left\{\mathbf{q}_i \in R : \mathbf{p}_j'\mathbf{q}_j > \mathbf{p}_j'\mathbf{q}_i \text{ for some } \mathbf{q}_j \in R\right\}$

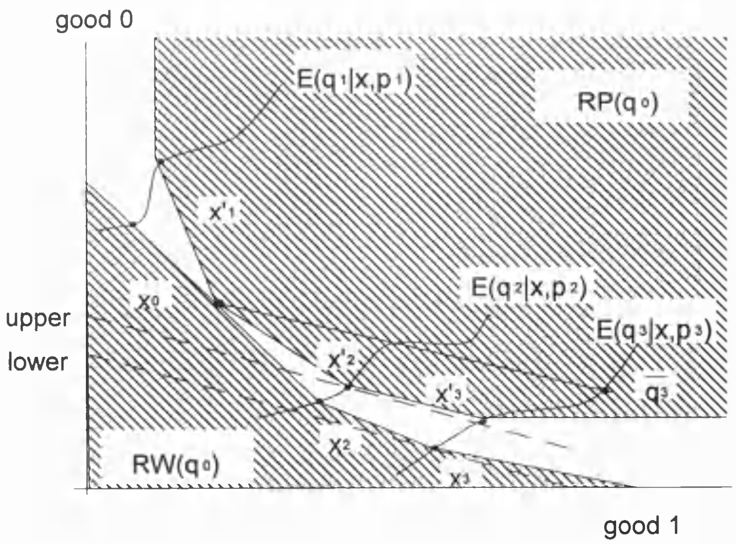
4) Set $B = R \backslash E$

5) If $E = \varnothing$ set $RW = B$ and stop. Otherwise go to (2).

Consider the $RP\left(\mathbf{q}_0\right)$ algorithm. The idea is to find a set of points which are either directly or transitively revealed preferred to the base bundle $(\mathbf{q}_0)$ but in which none are strictly preferred to it (either directly or transitively). At the first execution of step (2) the bundles $(\mathbf{q}_t)$ such that $\mathbf{q}_t \; R^0 \; \mathbf{q}_0$ $(\mathbf{p}_t'\mathbf{q}_t = \mathbf{p}_t'\mathbf{q}_0)$ are computed and these are placed in the set R. The next step then identifies which of the points in R are revealed strictly preferred to others and these are placed in the intermediate set E. Such points must also be revealed strictly preferred to the base bundle and must therefore be above the indifference curve through $\mathbf{q}_0$. The remaining points (which are directly but not strictly preferred to $\mathbf{q}_0$) are placed in the set W. At the next execution of (2) compute the set of points which are directly but not revealed strictly preferred to every point in W (this will include points already in W) are computed and R is updated. R now consists of bundles which are both directly and transitively revealed preferred to the base bundle and for each member of W we compute a bundle for each of the $T+1$ periods. For example, if there are two members of W, two $\mathbf{q}_k$ bundles are found, each one is constructed such that $\mathbf{p}_k'\mathbf{q}_k = \mathbf{p}_k'\mathbf{q}_w$ for each $\mathbf{q}_w \in W$. Again, step (3) and (4) remove those points which are either directly or indirectly strictly preferred to $\mathbf{q}_0$ (for example, in the above case one of the two $\mathbf{q}_k$ bundles must be strictly preferred to the other if they are not the same). The remainder are placed in W which now consists of bundles which are directly (but not strictly) preferred to $\mathbf{q}_0$ and bundles which are transitively preferred (again not strictly so) to the base bundle.

At each stage, the algorithm computes a set of bundles which are (not strictly) revealed preferred to each member of the previous set W (which are also preferred but not strictly preferred to the base bundle). The two sets are then compared and any points which are above the indifference curve (transitively or directly strictly preferred to $\mathbf{q}_0$) are discarded thus improving the bound on each iteration. Each iteration is an attempt to find a set of points which are worse than the previous set, but which are still better than the base bundle. The algorithm converges when it cannot find any such points (that is, when W is not improvable). At this point E becomes an empty set. Note that the set W is being continually up-dated and bundles may be removed if points added in later iterations reveal that they have a transitive strict preference relation to the base bundle. The algorithm for the boundary of the revealed worse $(RW(\mathbf{q}_0))$ set is entirely analogous.

Figure 4.5: GARP-improved bounds.



A simple two dimensional illustration of the algorithms to compute the improved bounds is shown in figure 4.5. In this example budgets are set such that $\mathbf{p}_1'\mathbf{q}_1(x_1') = \mathbf{p}_1'\mathbf{q}_0$, $\mathbf{p}_2'\mathbf{q}_2(x_2') = \mathbf{p}_2'\mathbf{q}_0$, are directly preferred to $\mathbf{q}_0$ and are added to W in the first iteration. The point $\overline{\mathbf{q}_3}$ would also be identified in the first iteration since $\mathbf{p}_3'\overline{\mathbf{q}_3} = \mathbf{p}_3'\mathbf{q}_0$ but $\mathbf{p}_3'\overline{\mathbf{q}_3} > \mathbf{p}_3'\mathbf{q}_1$ and $\mathbf{p}_3'\overline{\mathbf{q}_3} > \mathbf{p}_3'\mathbf{q}_2$ imply $\overline{\mathbf{q}_3}\ P\ \mathbf{q}_0$ so $\overline{\mathbf{q}_3}$ must be above the indifference curve and so can be improved. The next iteration computes points revealed preferred to each of the (now three) members of $W = \{\mathbf{q}_0, \mathbf{q}_1(x_1'), \mathbf{q}_2(x_2')\}$. The points revealed preferred to $\mathbf{q}_0$ will just be $\{\mathbf{q}_0, \mathbf{q}_1(x_1'), \mathbf{q}_2(x_1'), \overline{\mathbf{q}_3}\}$ again.

There will be another four points (one on each expansion path including $E(\mathbf{q}_0|x, \mathbf{p}_0)$ which is not shown) which are revealed preferred to each of the other members of W $(\mathbf{q}_1(x_1'), \mathbf{q}_2(x_1'))$. These are placed in R, replacing the previous set and any strictly preferred bundles are removed to E. This leaves the $\mathbf{q}_3(x_3')$ bundle at an expenditure level (x_3') such that $\mathbf{p}_3'\mathbf{q}_3(x_3') = \mathbf{p}_3'\mathbf{q}_2(x_2')$. Now $\mathbf{q}_3(x_3')\ R^0\ \mathbf{q}_2(x_2')$ and $\mathbf{q}_2(x_2')\ R^0\ \mathbf{q}_0$ give $\mathbf{q}_3(x_3')\ R\ \mathbf{q}_0$ and the algorithm ends with the upper bound illustrated as the next iteration will find no improvements, R will be identical to W and E will be the empty set. The budget lines using each price vector as the final total expenditure levels are denoted $\{x_1', x_2', x_3'\}$.

The illustration used nicely behaved expansion paths which were rationalisable with GARP. However, such data may not be on offer, and in fact if the data in the vicinity[5] of the reference bundle $\mathbf{q}_0$ violate GARP then the algorithms will not converge. This is shown below.

Proposition 4.1 *If the data local to the reference bundle* $\mathbf{q}_0$ *reject GARP, then the algorithm for the boundary to the set* $RP(\mathbf{q}_0)$ *will not converge.*

Proof.

Without any loss in generality take the simplest case. In which there are

[5] The reason that the argument is restricted to commodity space local to the reference bundle is that expansion paths may cross and un-cross as they move through higher levels of total expenditure. Thus GARP may be rejected for, say high income households, but pass for low income households. The convergence of the algorithms requires that GARP is not violated in the region around the reference bundle.

two periods. Denote the reference period bundle as $\mathbf{q}_0$, the other as $\mathbf{q}_1$.

1) By Step (1) $W = \mathbf{q}_0$, $t = \{0, 1\}, E = \varnothing$.

2) By Step (2) set x_1 such that $\mathbf{p}_1'\mathbf{q}_1(x_1) = \mathbf{p}_1'\mathbf{q}_0$. Set $R = \{\mathbf{q}_0, \mathbf{q}_1(x_1)\}$

3) Suppose that these data reject GARP. Since by construction $\mathbf{q}_1 R^0 \mathbf{q}_0$ this means that $\mathbf{q}_0 P^0 \mathbf{q}_1$.

4) By Step (3) $E = \{\mathbf{q}_0\}$ since $\mathbf{q}_0 P^0 \mathbf{q}_1$ through violation of GARP and by Step (4) $W = \{\mathbf{q}_1\}$.

5) Since $E \neq \varnothing$, return to Step (2) and set x_0' such that $\mathbf{p}_0'\mathbf{q}_0(x_0') = \mathbf{p}_0'\mathbf{q}_1$. Set $R = \{\mathbf{q}_0(x_0'), \mathbf{q}_1(x_1)\}$

6) Now by construction $\mathbf{q}_0(x_0') R^0 \mathbf{q}_1$ but since $\mathbf{q}_0 P^0 \mathbf{q}_0(x_0')$ then (2) implies that $\mathbf{q}_1(x_1) P^0 \mathbf{q}_0(x_0')$ which is a violation of GARP.

8) Hence by Step (3) $E = \{\mathbf{q}_1(x_1)\}$ through violation of GARP, and by Step (4) $W = \{\mathbf{q}_0(x_0')\}$.

9) Since $E \neq \varnothing$. Return to Step (2). etc.

■

Proposition 4.2 *If the data local to the reference bundle* $\mathbf{q}_0$ *reject GARP, then the algorithm for the boundary to the set* $RW(\mathbf{q}_0)$ *will not converge.*

Proof.

The proof is analogous with that for Proposition 4.1.

■

What this means is that these algorithms provide a test of GARP in the region around the reference bundle. Indeed, at each step bundles are found such that $\mathbf{p}_t'\mathbf{q}_t = \mathbf{p}_t'\mathbf{q}_0$. Using the argument for the PMP and SMP paths described in Chapter 3, this maximises the possibility of finding the rejection $\mathbf{p}_0'\mathbf{q}_0 > \mathbf{p}_0'\mathbf{q}_t$, although the 'sequence' here is a set of pairwise comparisons either directly or transitively to $\mathbf{q}_0$. Propositions 4.1. and 4.2 tell us that if the algorithms fail to converge then there are no coherent indifference curves to bound in that region of the data because the data reject GARP.

To sum up, revealed preference restrictions can be used to bound indifference curves through commodity space (as described in Varian (1982)), but the information gleaned by this method may not be terribly useful. This section has described a method which, using the idea of nonparametric expansion paths which was developed in Chapter 3, is designed to improve revealed preference bounds on indifference curves and hence to improve bounds on the true cost of living referenced on a curve passing through some bundle of goods. These algorithms will only converge if the data pass GARP locally. Equivalently, they will only converge if there are coherent regions of indifference to bound. The procedure then, is to run the algorithms to bound an indifference curve through a chosen point (say average demands in the first year of the data). If the algorithms fail to converge (this is clear once we find a point, on the boundary of the $RP(\mathbf{q}_0)$, say $\mathbf{q}_r$ such that $\mathbf{q}_r R \mathbf{q}_0$ and $\mathbf{q}_0 P^0 \mathbf{q}_r$), then identify the rejection in question and remove it and subsequent years from the data and rerun. This will give the longest available sub-periods of convergence/non-rejection of GARP.

4.4 Local Average Demands and Local Average Welfare

This section turns to the relationship between the (nonparametric) Engel curves used to move the budget lines and described in Chapter 3, the average demands of a set of heterogeneous agents and the average welfare of those agents. Chapter 3 discussed the issue of heterogeneity in section 3.3. The heterogeneity structure assumed here is exactly the same but now the further question arises. If the procedure outlined above uses local average demands as a basis for welfare measurement, under what circumstances do such welfare measures make sense ?

To recap slightly, as in Chapter 3 let $\omega(x, \mathbf{p}, \varepsilon)$ be the budget share system of n equations for a household with heterogeneity vector $\boldsymbol{\varepsilon}$. This heterogeneity may be observable (for example, family composition or age) or it may be unobservable taste heterogeneity. The necessary condition for the average budget shares recovered by the nonparametric analysis discussed above to be equal to average budget shares is that:

$$\omega(x, \mathbf{p}, \boldsymbol{\varepsilon}) = \mathbf{f}(x, \mathbf{p}) + \Lambda(x, \mathbf{p})' \boldsymbol{\varepsilon} \text{ with } E(\boldsymbol{\varepsilon}|x, \mathbf{p}) = \mathbf{0} \tag{4.1}$$

where $\Lambda(.)$ is an $n \times m$ matrix. Given this combination of functional form restrictions and distributional assumptions, the nonparametric analysis recovers $\mathbf{g}^t(x) = \mathbf{f}(x, \mathbf{p}^t)$. In the analysis in Chapter 3 and here the procedures and estimates refer to the mean function $\mathbf{f}(x, \mathbf{p})$. This function gives mean responses to changes in prices conditional on a given level of total expenditure. Recall that this heterogeneity structure does not impose that individual demands satisfy the Slutsky conditions. But additional restrictions (such as preferences being locally Piglog) can do this.

In general the error term will represent measurement and optimisation error as well as preference heterogeneity so it was argued that it would seem natural to work with local average demands. Averaging locally to each x is designed to eliminate unobserved heterogeneity, measurement error and (zero mean) optimisation errors in demands but preserves any nonlinearities in the Engel curve relationship for each price regime.

Suppose that now one wishes to estimate bounds on welfare based on average demands. Is it reasonable to compute welfare bounds using local average demands? Under what circumstances can we speak of the local average welfare gain to a set of households indexed by x and $\mathbf{p}_t$?

For utility measures based on local average demands to make sense we need further conditions. The Piglog assumption is important but needs to be strengthened. In particular, heterogeneity has to be restricted to the intercept of the share equation and β_j cannot vary with unobserved heterogeneity. From this we can compute the local average welfare cost of a price change by

$$E(\Delta \ln c(\mathbf{p}^t; u)|x, \mathbf{p})) = \alpha_0 + \overline{\boldsymbol{\alpha}}'\Delta \ln \mathbf{p}^t + \Delta \left(\boldsymbol{\Gamma}(\mathbf{p})' \ln \mathbf{p}^t\right) + \Delta \mathbf{b}(\mathbf{p}^t)\overline{u}.$$

for individual households indexed by average utility $\overline{u}$. By assuming the implicit cardinalisation of utility, average utility $\overline{u}$ is given by the indirect utility

$$\overline{u}(\mathbf{p}^t; y) = \frac{\ln y - \alpha_0 - \overline{\boldsymbol{\alpha}}' \ln \mathbf{p}^t - \boldsymbol{\Gamma}(\mathbf{p})' \ln \mathbf{p}^t}{\mathbf{b}(\mathbf{p}^t)}.$$

Of course we may not care whether or not individual households are rational or not. Then we only care about average demands and we can take them to serve as the basis for estimating the effects of price changes on the welfare of a representative consumer.

4.5 Results

The data to be used are the same data which were used in Chapter 3. That was 20 years of cross section data from a car-owning subsample of the Family Expenditure Survey. Expenditures on non-durable goods by these households were aggregated into 22 commodity groups and chained Laspeyres price indices for these groups were calculated from the sub-indices of the UK Retail Price Index giving 20 annual price points for each group of goods.

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.[6]

4.5.1 Non-Rejecting Sub-Periods

As discussed above the algorithms presented in section 4.3.2 are designed to bound well-behaved indifference curves will only converge if there are a (locally) coherent set of indifference curves to bound. Nevertheless if there are rejections of GARP, the algorithms can be run for non-rejecting sub-periods and the results of this are presented in Table 4.1.

Table 4.1: Continuous periods of convergence.

<table>
<tr><th rowspan="2"></th><th colspan="20">Years</th></tr>
<tr><th>74</th><th>75</th><th>76</th><th>77</th><th>78</th><th>79</th><th>80</th><th>81</th><th>82</th><th>83</th><th>84</th><th>85</th><th>86</th><th>87</th><th>88</th><th>89</th><th>90</th><th>91</th><th>92</th><th>93</th></tr>
<tr><td>1st</td><td colspan="10"></td><td colspan="2"></td><td colspan="8"></td></tr>
<tr><td>10th</td><td colspan="13"></td><td colspan="7"></td></tr>
<tr><td>25th</td><td colspan="12"></td><td colspan="2"></td><td colspan="6"></td></tr>
<tr><td>50th</td><td colspan="12"></td><td colspan="8"></td></tr>
<tr><td>75th</td><td colspan="13"></td><td colspan="4"></td><td colspan="3"></td></tr>
<tr><td>90th</td><td colspan="12"></td><td colspan="4"></td><td colspan="1"></td><td colspan="3"></td></tr>
<tr><td>99th</td><td colspan="8"></td><td colspan="5"></td><td colspan="1"></td><td colspan="2"></td><td colspan="4"></td></tr>
</table>

The table shows the largest continuous sub-period in which the algorithm is able to bound the indifference curve. For example, taking the median demand in 1974 as the starting point (and hence $u(\mathbf{q}_{74})$ as the reference welfare level), the procedure is able to bound a curve using the expansion paths and price data for 1974 to 1985 inclusive, and using any of the periods within the interval. If 1986 in added to the set of admissible periods the

[6] A more detailed description can be found in Appendix 3.A.

algorithm fails to converge[7]. The algorithm is then started again using the 1986 bundle as the new starting point. In all, for the median the entire period breaks down into two sub-periods within which an indifference curve can be bounded in each. Similarly the two decile paths break into three sub-periods, while the 99th percentile breaks down into five coherent continuous sub-periods.

4.5.2 Cost of Living bounds and Approximations

From the longest consistent sub-period (1974 to 1985) GARP-based bounds are derived on the true cost of living index $c(\mathbf{p}_t, u_{74}) / c(\mathbf{p}_{74}, u_{74})$ based at the welfare level given by median total expenditure in the base period. Table 4.2 summarises the results for the GARP based bounds, the classical bounds from revealed preference restrictions, the Paasche and Laspeyres first order approximations, Fisher's Ideal Index and the Törnqvist index.

Table 4.2: GARP bounds and approximations to the 1974-based true index.

Year	Paasche	Laspeyres	Fisher's	Törnqvist	Classical	GARP
1974	1000	1000	1000	1000	1000	1000
1975	1213	1229	1221	1219	[1025,1229]	[1212,1226]
1976	1493	1526	1509	1517	[1182,1526]	[1510,1526]
1977	1726	1777	1751	1759	[1239,1777]	[1752,1770]
1978	1891	1948	1920	1932	[1385,1948]	[1929,1948]
1979	2046	2109	2077	2088	[1461,2109]	[2084,2108]
1980	2417	2506	2461	2481	[1734,2506]	[2472,2498]
1981	2731	2838	2784	2806	[1770,2838]	[2800,2833]
1982	3031	3181	3105	3128	[1821,3181]	[3119,3159]
1983	3219	3378	3297	3315	[1828,3378]	[3313,3363]
1984	3371	3556	3462	3476	[1790,3556]	[3471,3524]
1985	3530	3734	3630	3640	[1836,3734]	[3634,3692]

[7]This is not surprising given that in Chapter 3 1985 and 1986 were found to cause a rejection on the median SMP path.

Figure 4.6 shows the classical bounds. The upper and lower classical bounds are illustrated by the solid lines, the improved GARP-based bound is shown by the dashed lines. The upper bound corresponds to the Laspeyres bound indicating zero substitution. The lower bound indicates extreme substitution towards the cheapest of the 22 goods available in each period. This shows that, as in Varian (1982), revealed preference restrictions based on the average demand data gives little additional information on the curvature of the indifference curve through commodity space.

Figure 4.6: Classical and GARP bounds on the 1974-based true index.

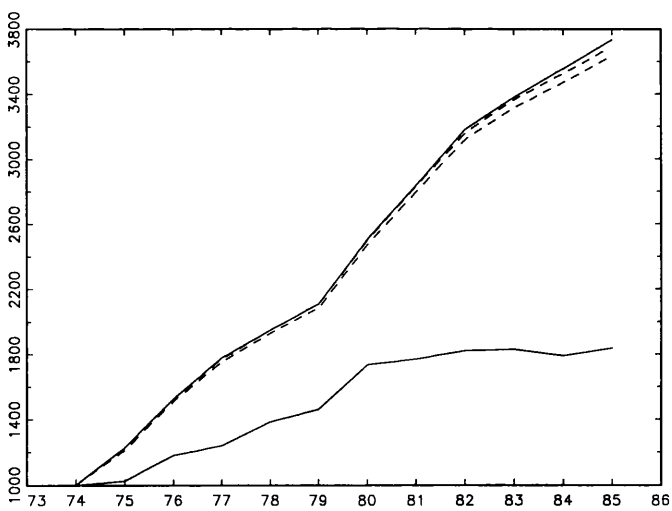


Figure 4.7 shows the Paasche and Laspeyres indices and the improved bounds derived from Engel curves using the algorithms described above. The upper solid line is the Laspeyres index, the lower solid line is the Paasche. The dashed lines indicate the upper and lower GARP-derived bounds on the true 1974 referenced cost of living index. As expected the Laspeyres

Figure 4.7: Paasche, Laspeyres and Fisher's indices and GARP bounds.

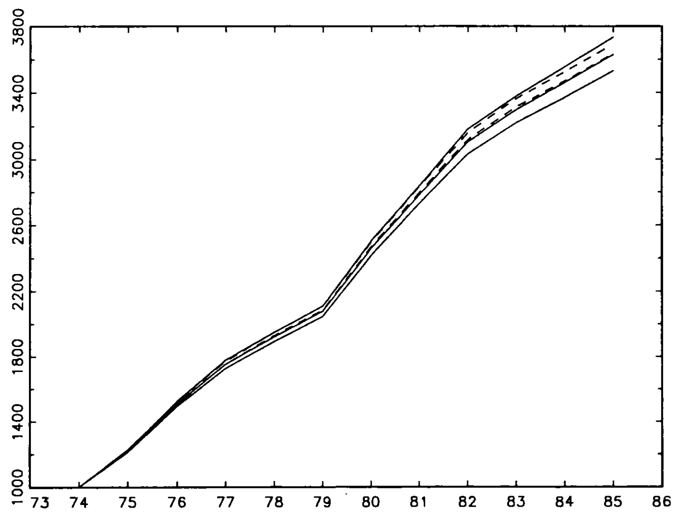
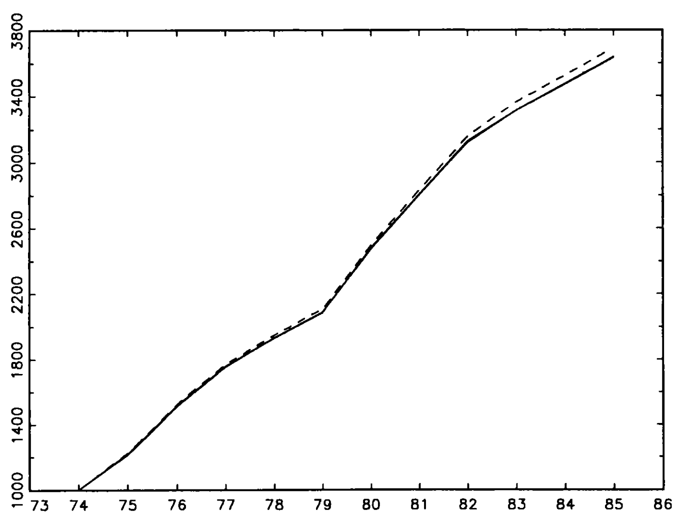


Figure 4.8: The Törnqvist index and GARP bounds.



approximates the true base referenced cost of living index from above. The Paasche bounds the t period index (which is not shown) from below. Fisher's Ideal index is the solid line between the Paasche and Laspeyres.

Figure 4.8 compares the improved GARP-derived bounds with the Törnqvist superlative index. As discussed in Chapter 2, the Törnqvist index is generally preferred to the Paasche or Laspeyres since, while they provide first order approximations to true indices, the Törnqvist (which is based on translog preferences) can provide a second order (local) approximation to any cost function[8] hence allowing for substitution and given that it is based upon translog preferences it is in sympathy with the estimation of Engel curves which are locally Piglog. Empirically, the Törnqvist seems to perform well here since it remains within the GARP-derived bounds on the true u_{74} based index.

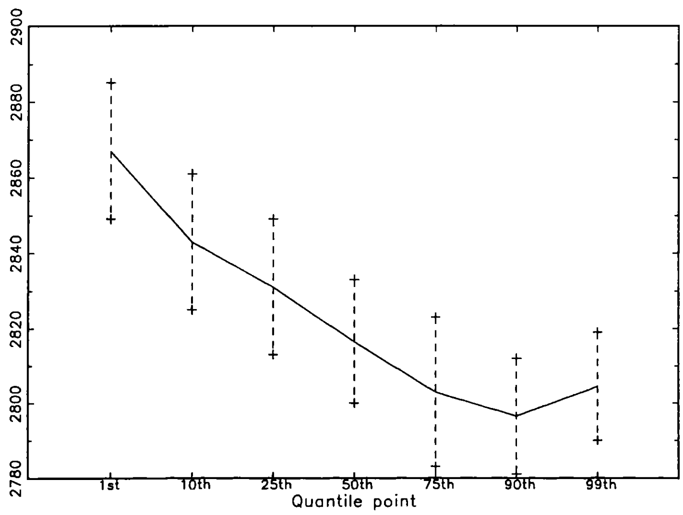
4.5.3 Substitution Bias and Non-Homotheticity

Recall that Malmquist (1953) and Afriat (1981) show that a unique and unambiguous price index can only exist if preferences are homothetic. If they are not, then different income groups will have different cost of living indices. In this section evidence of non-homotheticity is put forward by illustrating differences in the bounds on the true indices calculated at different points in the total expenditure distribution. Bounds on the true cost of living were computed for the other quantile points of the total expenditure distribution listed in Table 4.2. Figure 4.9 illustrates the way in which the mid-point of the bounds on the 1974 referenced true cost of living indices for each comparison path in 1981, and the bounds themselves, vary with total expenditure. Table 4.2 showed that this was the last year in which all indices could be computed

[8]See Christensen, Jorgenson and Lau (1975).

without some sort of chaining procedure.

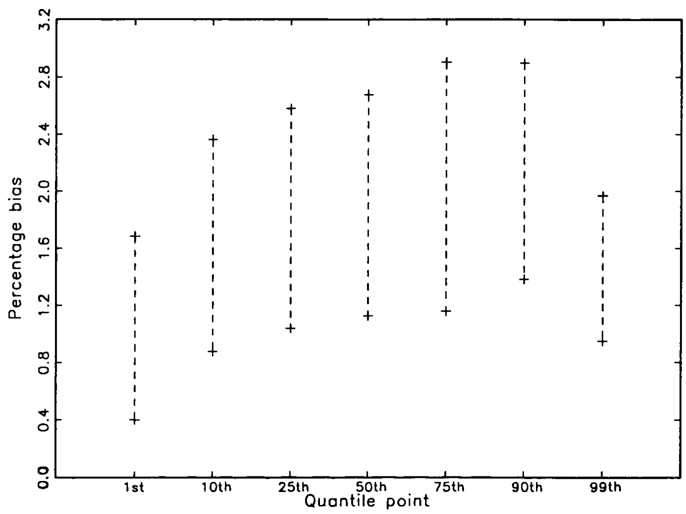
Figure 4.9: Bounds in 1981, by total expenditure quantile, (1974=1000)



The pattern which emerges is of greater increases in the cost of living of poorer (low spending) households over the period 1974 to 1981. This is in keeping with previous studies (for example Bradshaw and Godfrey (1983), Fry and Pashardes (1986) and the discussion in Chapter 2) and is unsurprising considering some of the events which took place in the 1970's including membership of the Common Market (in particular the common agricultural policy) and the oil crisis which put up the prices on necessities relatively fast compared to luxuries. The width of the bounds varies slightly by quantile.

Figure 4.10 shows bounds on the percentage substitution bias of the Laspeyres index for each quantile point. For example, the substitution bias for demands at the 1st percentile indicates the Laspeyres index with its fixed base-period weights overstates the true increase in the cost of living by between about 0.4% and 1.85%. There is less evidence of substitution bias in

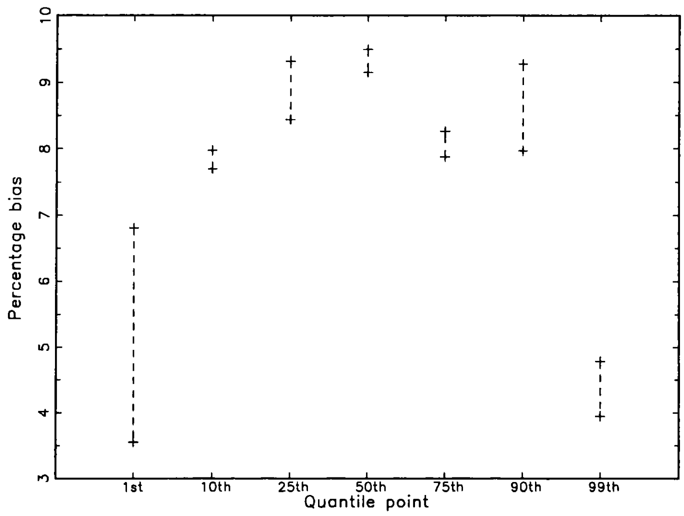
Figure 4.10: Percentage substitution bias, by total expenditure quantile (1981)



the tails of the total expenditure distribution perhaps indicating that high-spending households are too rich to care much about relative price movements and that low-spending households are already consuming mostly necessities in rough the necessary proportions. Results from these groups should be treated with some caution, however, as they are based on estimates of the expansion paths in the tails of the spending distribution where the confidence bounds are widest. Bias at average demand is higher with the mid-point of the bounds at around 1.9% by the end of the period.

Figure 4.11 shows the percent effects of substitution bias on measures of inflation. The percentage errors shown are the maximum's for each quantile point for the period. These mostly occurred around 1979/1980 when prices were rising fastest. For example, in 1980 with inflation at around, say, 20%, the rate calculated from the Laspeyres would have been out by 5% compared to the mid-point of the bounds on the true index. As with the bias in the

Figure 4.11: Maximum percentage substitution bias in inflation rates, by total expenditure quantile



levels, the bias in the differences seems to be concentrated in the middle of the spending distribution.

4.6 Summary and Conclusion

This chapter suggests a revealed preference method which, without the need for functional and parametric assumptions, allows two-sided bounds to be place on a given true index recovered from average demands. These bounds, are shown to be as least as tight as the classical nonparametric bound and are also the tightest possible bound obtainable given only the (testable) assumption that a given finite dataset of prices and average demands was generated by a well-behaved utility function. The algorithm which computes the bounds was shown to provide a powerful method of performing this test. These ideas were applied to UK micro data from 1974 to 1992 and it was found that coherent indifference curves were only recoverable up to the mid

1980's for most points in the total expenditure distribution.

This improved bound for average demands was compared to other, popular nonparametric approximations and was used to see how closely these indices seem to approximate the truth. In particular the bounds on the true index were used to assess the extent of the substitution bias which is incurred by using simple approximations like the Laspeyres compared to second order approximations. The Törnqvist index was shown to preform well in this respect. This procedure was also used to provide evidence on differences in cost of living indices by total expenditure level of the household (i.e. non-homotheticity), and estimates of the substitution bias incurred by the use of the simple Laspeyres approximation were shown to vary with total expenditure. In general, substitution bias of the Laspeyres formulation seemed to increase with total spending indicating relatively little substitution responses to relative price changes amongst the poorest (low spending) households.

Chapter 5

New Goods Bias in Cost of Living Indices: A Revealed Preference Approach

5.1 Introduction

Judging the effects of changing economic conditions over time or across states of nature is often a matter of trying to isolate the change of interest while holding everything else constant. True cost of living indices, for example, try to identify how the cost of achieving some constant level of economic welfare has changed as prices have changed. Many papers (including some of the previous chapters) have been written on the accuracy with which true indices can be approximated, but whatever the approach adopted for calculating the index itself[1], an additional complication, common to all indices, occurs when the number and quality of goods, as well as the prices of goods, changes between periods. The problems which these issues create were discussed in the introductory chapter and the question of accounting for quality change was analysed further in Chapter 3. This chapter concerns the problem of

[1]For example, the price index approach (Hicks (1940) and Rothbarth (1941)), the parametric approach to cost-of-living indices (Braithwait (1980)) or the non-parametric approach (Varian (1982)).

new goods bias.

The complication caused by new goods arises because, when the number of goods changes across time periods, the full price vector will not be observed in all periods. For example, in order to compare two periods when a new good is introduced in the second period (using either a true cost of living index or an approximation which includes the new good in the reference bundle[2]) we need to calculate what the (virtual) price of the new good was in the first period. This is usually taken to be the price which would just have driven demand for the good to zero in that period.

The most common approach to calculating the virtual price of a new good is parametric estimation and extrapolation of demand curves. This requires the imposition of a particular functional form for preferences, upon which the results will be heavily dependent. This is potentially worrying; indeed as discussed in section 1.4 in Chapter 1, Moulton (1996) argues that this problem is so severe that empirical studies have done little to identify a plausible range of bias arising from new goods.

This chapter presents an alternative revealed preference method for calculating the virtual price for a good for the period immediately prior to the one in which it first exists. This method does not require the estimation of a parametric demand system, and is consistent with the maximisation of a well-behaved utility function which is stable over time, with no further restrictions on the exact form of preferences necessary. Although this chapter focuses on the issue of including new goods in price and cost-of-living indices, the method can equally accommodate obsolete goods.

The plan of the chapter is as follows. Section 5.2 presents a framework for the valuation of new goods. Section 5.3 outlines some of the problems

[2] A Laspeyres price index, for example, would not include the new good.

inherent in a parametric approach to computing virtual prices. In section 5.4 an alternative revealed preference method for calculating a lower bound on the virtual price using observed choice outcomes is described. We then describe a way of improving this bound by using non-parametric estimates of expansion paths. Section 5.5 illustrates these ideas using data on the UK National Lottery, which was introduced in November 1994. The virtual price of the Lottery one year prior to its introduction is calculated, and the effect of including the new good in some measures of annual non-housing inflation rates over the period is examined. Section 5.6 concludes.

5.2 The Problem of New Goods

Suppose that there are $T+1$ periods, $t = 0, ..., T$, and $K+1$ goods, $k = 0, ..., K$, where the 0th good does not exist in period $t = 0$. All other goods are available in every period. Let $\mathbf{p}_t$ be the $K+1$ vector of prices in period t, and u_R be the reference utility level. The consumer's problem is to

$$\begin{array}{rl} & \max\ U(\mathbf{p}_t) \\ s.t. & \mathbf{p}_t'\mathbf{q}_t \leq x_t \end{array}$$

where $\mathbf{q}_t$ is the $K+1$ vector of goods bought in period t and x_t is the total budget in that period. The solution to this problem allows us to derive the cost function $c(\mathbf{p}_t, U)$.

Suppose that we wish to compute the true cost-of-living index:

$$P_T = \frac{c(\mathbf{p}_1, u_R)}{c(\mathbf{p}_0, u_R)}$$

In order to make the comparison between the two periods, we require (of course) that the form of the utility function is stable over time; since utility functions are ordinal we cannot compare the level u_R across two different utility functions and say that it has the same meaning. Therefore, we need to assume that the introduction of a new good does not change the utility function, i.e. that goods enter the utility function even before they exist — possibly even before the consumer can imagine them.

This assumption might be rationalisable using a Gorman-Lancaster[3] characteristics argument. In this case, consumers have preferences over basic characteristics and preferences over goods are derived from the characteristics they embody. If a new good can be described as a combination of existing characteristics, then consumers can rank consumption bundles which include the new good even though it does not currently exist. Of course, this requires the (strong) assumption that new goods do not add to the set of characteristics, and in that sense new goods are not really new at all.

A stable utility function also implies that goods do not become obsolete because of changes in preferences. Goods can become obsolete because they become more expensive to produce, or are superseded by higher quality and/or cheaper substitutes, and this can happen without preferences changing. However, it seems apparent that, in reality, another reason for obsolescence could be changing social conventions, fashions, arrival of new information about the good and so on, i.e. changes in preferences. A good example of this is the decline in the popularity of hats. They have not been superseded by a close substitute, nor is there any reason to suppose they are relatively more expensive to produce than in the past. What might be the case is that polite society no longer regards it as inappropriate to go bareheaded out of doors, i.e. preferences have changed in clothing space. Because of the need to assume a stable utility function, cost-of-living indices (based on preferences defined over goods)[4] cannot accommodate this possibility.

[3]Gorman (1956), Lancaster (1966).

[4]As mentioned in Chapter 1, an alternative formulation, due to Sen, is to define utility as deriving from capabilities rather than goods. Thus utility depends on the ability to

Unless we wish to interpret an approximation to a true index as a true index itself, then, because there is no corresponding utility structure there is no equivalent problem for the price index

$$P = \frac{\mathbf{p}_1'\overline{\mathbf{q}}}{\mathbf{p}_0'\overline{\mathbf{q}}}$$

However, if $\overline{\mathbf{q}}$ contains the new good then we still need to calculate its corresponding period 0 price. The lack of a utility structure to the problem means that it is not clear how to do this. This may be an instance in which the utility-based approach contributes a "structure which is an essential part of the matter"[5], as it allows the question to be addressed and restrictions to be placed on the possible values of the period 0 price of the new good. This is described next.

Supposing that the utility function does not change, in order to calculate $c(\mathbf{p}_0, u_R)$ or $\mathbf{p}_0'\overline{\mathbf{q}}$ (when $\overline{\mathbf{q}}$ includes the new good) we need all $K + 1$ prices in the $\mathbf{p}_0$ vector. The problem is, that as good 0 does not exist in period $t = 0$, we do not observe p_0^0. In order to proceed further, we need to devise a method of calculating a value for the missing price p_0^0.

John Hicks (1940) discussed the question of how to value new goods and, more generally, the issue of how to deal with rationed goods when constructing index numbers. He showed that the price of a rationed good in an index number should not be the actual price, but the price which would make the rationed level of consumption consistent with an unrationed choice. New goods are a special case of rationing, for which the ration level in the period prior to the one in which they first exist is zero. Thus the virtual prices for

be respectable or the ability to be well nourished, rather than on hats owned or food consumed. A capability-based view might be able to accomodate our example without the need for preferences to have changed. Utility depends on the capability to be respectable, and the cost of achieving respectability has fallen because of changing social conventions.

[5]Afriat (1977), also cited in the introductory chapter.

new goods would be those which "just make the demand for these commodities (from the whole community) equal to zero"[6].This approach captures the benefit of the introduction of a new good by imagining that its price has reached its period t value from a level in period 0 which was marginally above the maximum value of the good to consumers. Put another way, we are trying to recover the consumer surplus generated by the introduction of a new good at a given price. Price falls in levels above the reservation price will be irrelevant to consumers - they do not benefit if the price of a good falls to a level which is still too high for them to purchase the good. Thus the reservation price is the correct price to choose as the virtual price of the new good, since the benefit of a price fall from a level exceeding the reservation price is no greater than if the price had fallen to the same level from the reservation price itself[7].

More formally, consider a simple model of rationing in which our $T+1$ observed demand bundles are the outcome of the following choice

$$\max U(\mathbf{q}_t)$$

$$\text{s.t.} \quad \begin{array}{l} \mathbf{p}_t'\mathbf{q}_t \leq x_t \quad t = 0, ..., T \\ R_t\mathbf{q}_t \leq \mathbf{\Omega}_t \end{array}$$

The R_t is a $(K+1) \times (K+1)$ matrix of zeros and ones on the diagonal which pick out rationed goods and zeros elsewhere, and $\mathbf{\Omega}_t$ is a $(K+1)$ vector indicating ration levels. In what follows we suppose that good 0 is rationed

[6] Hicks (1940), p.114.

[7] Hicks uses the virtual price which drives aggregate demand to zero since his population has an implicit representative consumer. When the population consists of heterogeneous individuals, we could expect them to have different virtual prices for the new good. In this case, when calculating individual costs-of-living indices on which to base the aggregate index, we should use the individual's own reservation price, and not the one which would drive aggregate demand to zero (ie the highest virtual price). Again, the individual's virtual price captures the benefit to him of having the new good introduced, and falls in price above this level are irrelevant. This point is further discussed in Section 5.3.

in period $t = 0$ such that $q_0^0 = 0$. The element of R_0 corresponding to q_0^0 is one, with the other elements on the diagonal equal to zero $k \neq 0$ (the un-rationed goods) and R_t is a matrix of zeros for $t \neq 0$ (the un-rationed periods). Correspondingly $\boldsymbol{\Omega}_t$ is a vector of zeros for all periods[8]. The corner solution $q_0^0 = 0$ is the outcome of the rationing constraint. The first order condition is that

$$U'(\mathbf{q}_t) - \lambda_t \left(\mathbf{p}_t + R_t \cdot \frac{\mu_t}{\lambda_t} \right) = 0$$

where $\lambda_t > 0$, and $\mu_t > 0$ if the constraints bind. More specifically

$$\begin{aligned} &U'(\mathbf{q}_0) - \lambda_0 \left(\mathbf{p}_0 + R_0 \cdot \tfrac{\mu_0}{\lambda_0} \right) = 0 && \text{if } t = 0 \\ &U'(\mathbf{q}_t) - \lambda_t \mathbf{p}_t = 0 && \text{otherwise.} \end{aligned}$$

For $t = 0$ the vector $\left(\mathbf{p}_0 + R_0 \cdot \frac{\mu_0}{\lambda_0} \right)$ is the vector of actual prices $\left(p_0^k \right)$ for goods $k \neq 0$ and the virtual price $\left(p_0^0 + \frac{\mu_0}{\lambda_0} \right)$ for good 0.

The $\mathbf{q}_t$ vector which is the outcome of such a rationed choice is the same as the solution vector to the unrationed problem

$$\max U(\mathbf{q}_t)$$

$$\text{s.t.} \quad \boldsymbol{\pi}_t' \mathbf{q}_t \leq x_t \quad t = 0, ..., T$$

where $\pi_0^0 = \left(p_0^0 + \frac{\mu_0}{\lambda_0} \right)$ and $\pi_0^k = p_0^k$ for $k \neq 0$, and $\pi_t^k = p_t^k$ for $k \neq 0$ and $t \neq 0$. In order to calculate a price or cost-of-living index we need to calculate the virtual price, π_0^0.

[8] $\Omega_0^0 = 0$ indicates the ration for q_0^0, the other elements of Ω_t could, in fact, take any value as the corresponding elements of r_t are zero.

5.3 A Parametric Approach

The usual parametric approach to estimating virtual prices proceeds by assuming a particular functional form for demand which is consistent with maximisation of a particular form for the utility function. A system of demand equations, $\mathbf{q}_t(\boldsymbol{\pi}_t, x_t)$ is estimated using data from periods in which all goods are available (i.e. in which $\boldsymbol{\pi}_t = \mathbf{p}_t$), and this is used to predict π_0^0, the lowest price which would result in zero demand for good 0 (for a given x_0) in period 0. A recent example of this technique is Hausman (1997).

There are a number of possible problems associated with this approach. In particular, the estimate of the virtual price will be heavily dependent on the maintained hypothesis concerning functional form. Furthermore, determining the best functional form is difficult when non-nested models are being compared. In addition, parametric methods are reliant upon (possibly suspect) out-of-sample predictions of the demand curve to solve $q_0^0(\boldsymbol{\pi}_0, x_0) = 0$ for π_0^0. This is because of the necessary extrapolation of the demand curve across regions over which relative price variations have never been observed in the data. Finally, parametric models are computationally time-consuming. For these reasons, we propose using a revealed preference technique, which is described below.

5.4 A Revealed Preference Approach

The attraction of revealed preference conditions is that they apply to any well behaved utility function and, beyond this, no additional restrictions on the precise form of preferences underlying consumer demands are required. This property was set out in Afriat's Theorem (described in the Introductory chapter) which shows that, if consumers' observed choices, given the prices they face, satisfy the Generalised Axiom of Revealed Preference (GARP) (defined below), then these choices could have been generated by the maximisation of some non-satiated utility function.

The starting point is the model of choice under rationing described in section 5.2. Unfortunately, revealed preference restrictions *cannot* be applied to data which is the outcome of choice under rationing[9]. But this does not pose a problem, because, as noted, the outcome of a rationed choice is equivalent to the outcome of the unrationed problem with virtual prices. Because of this, we can apply revealed preference conditions to data in periods subsequent to the introduction of the new good, to learn about preferences without the need for functional form assumptions. We can place an upper bound on the indifference curve passing through the period 0 bundle and use this to place a lower bound on the virtual price π_0^0. Afriat's Theorem tells us that this bound will be such that the data $(\boldsymbol{\pi}, \mathbf{q})$ (which include π_0^0 and $q_0^0 = 0$) are consistent with the maximisation of a well behaved utility function, under the maintained assumption that the utility function does not change with the introduction of the new good.

More precisely, Afriat's Theorem shows that the maintained hypothesis that consumers have a stable utility function is equivalent to an assumption that the entire data set $(\boldsymbol{\pi}, \mathbf{q})$ satisfies GARP. If the subset of data $(\boldsymbol{\pi}_t, \mathbf{q}_t)$ for $t \neq 0$ satisfy GARP, then we can use it to place restrictions on the set of possible values which π_0^0 can take by requiring that π_0^0 is such that the *full* data set, including $q_0^0 = 0$, satisfy GARP. If the subset of data, $(\boldsymbol{\pi}_t, \mathbf{q}_t)$ for $t \neq 0$, did not satisfy GARP (which is easily testable using the techniques set out in Chapter 3), then, of course, there could not exist a π_0^0 which would rationalise $(\boldsymbol{\pi}, \mathbf{q})$ and the utility-base approach is called into question.

[9] Varian (1983), Theorem 7, page 108.

Taking the case in which there is only one rationed good, we observe $\boldsymbol{\pi}_t$ (equal to the actual prices $\mathbf{p}_t$) for all goods from period 1 onwards, and for all goods in the 0th period except good 0 ($\pi_0^k = p_0^k$ for $k \neq 0$). We can then use the data on periods $t \neq 0$ (if they satisfy GARP) to find the lower limit on π_0^0 consistent with non-violation of GARP for $(\boldsymbol{\pi}, \mathbf{q})$.

Denote the set of consumption bundles which are revealed preferred to $\mathbf{q}_0$ by $RP(\mathbf{q}_0)$. With $K+1 > 1$ and $T+1 > 1$, and $(\boldsymbol{\pi}_t, \mathbf{q}_t)$ for $t \neq 0$ satisfying GARP, then for non-violation of GARP for the entire data set $(\boldsymbol{\pi}, \mathbf{q})$, we cannot have $\mathbf{q}_0 P \mathbf{q}_s$ for $\mathbf{q}_s \in RP(\mathbf{q}_0)$, which implies

$$\boldsymbol{\pi}_0' \mathbf{q}_s \geq \boldsymbol{\pi}_0' \mathbf{q}_0$$

$$\Rightarrow \quad \pi_0^0(q_s^0 - \ q_0^0) \geq \sum_{k=1}^{K} \pi_0^k (q_0^k - q_s^k)$$

$$\Rightarrow \quad \pi_0^0 \geq \frac{\sum_{k=1}^{K} \pi_0^k (q_0^k - q_s^k)}{q_s^0}, \qquad \text{if } q_s^0 > 0$$

Thus, each $\mathbf{q}_s \in RP(\mathbf{q}_0)$ gives a lower bound on π_0^0 - call this set $\pi_0^0(\mathbf{q}_s)$. The highest value in this set encompasses all the other lower bounds and is the lower limit on π_0^0 given the data. This is proved below. Denote $\max\{\pi_0^0(\mathbf{q}_s)\}$ by $\overline{\pi}_0^0$.

Proposition 5.1. *Any $\pi_0^0 < \overline{\pi}_0^0$ violates GARP for $(\boldsymbol{\pi}, \mathbf{q})$.*

Proof.

(1) Denote $\overline{\boldsymbol{\pi}}_0 = \left(\overline{\pi}_0^0, \pi_0^1, ..., \pi_0^K\right)$

(2) $\overline{\pi}_0^0$ is such that $\overline{\boldsymbol{\pi}}_0' \overline{\mathbf{q}}_s = \overline{\boldsymbol{\pi}}_0' \mathbf{q}_0 = x_0$ where $\overline{\mathbf{q}}_s \in RP(\mathbf{q}_0)$

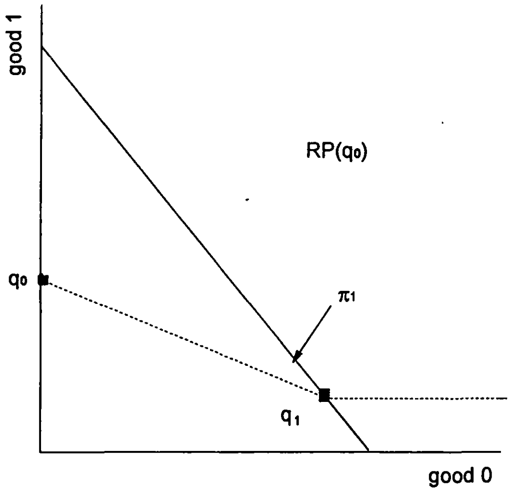
(3) Suppose $\underline{\pi}_0^0 < \overline{\pi}_0^0$, where $\underline{\boldsymbol{\pi}}_0 = \left(\underline{\pi}_0^0, \pi_0^1, ..., \pi_0^K\right)$

(4) Then from (2) and (3) $\underline{\boldsymbol{\pi}}_0' \overline{\mathbf{q}}_s < \overline{\boldsymbol{\pi}}_0' \overline{\mathbf{q}}_s = \overline{\boldsymbol{\pi}}_0' \mathbf{q}_0 = \underline{\boldsymbol{\pi}}_0' \mathbf{q}_0$ (since $q_0^0 = 0$) $\Rightarrow \mathbf{q}_0 P^0 \overline{\mathbf{q}}_s$ which is a violation of GARP.

$\blacksquare$

A two good, two period case is illustrated in Figure 5.1 below. The budget line in period 1 is given by $\boldsymbol{\pi}_1$, the period 1 bundle by $\mathbf{q}_1$ and the corner solution in period 0 by $\mathbf{q}_0$. Clearly, $\mathbf{q}_1 P^0 \mathbf{q}_0$. As a result, any period 0 price ($\boldsymbol{\pi}_0$) shallower than the line connecting $\mathbf{q}_0$ and $\mathbf{q}_1$ would violate GARP for the data set $(\boldsymbol{\pi}_0, \boldsymbol{\pi}_1; \mathbf{q}_0, \mathbf{q}_1)$. So π_0^0/π_0^1 must be greater than or equal to the gradient of the $\mathbf{q}_0$ to $\mathbf{q}_1$ line. The area above the dashed line is the set $RP(\mathbf{q}_0)$. The boundary of this set represents the best upper bound on the indifference curve passing through $\mathbf{q}_0$ available from the raw data.

Figure 5.1: Bounding the virtual price, a two good, two period example.



We note at this point that, since we are interested in changes in relative prices over time, the choice of base year at which to index prices is irrelevant. Suppose the series $\boldsymbol{\pi}_t$, $t = 0, ..., T$, derives from data $\boldsymbol{\theta}_t$, $t = 0, ..., T$, using period s as the base, i.e. $\pi_t^k = \frac{\theta_t^k}{\theta_s^k}$, $\forall\ k, t$. Thus, $\frac{\pi_t^i}{\pi_t^j} = \frac{\theta_t^i}{\theta_t^j} \cdot \frac{\theta_s^j}{\theta_s^i}$, $\forall\ i, j, t$. Suppose, instead, that period r was chosen as the base period, giving a different price series $\phi_t^k = \frac{\theta_t^k}{\theta_r^k}$, $\forall\ k, t$, and so $\frac{\phi_t^i}{\phi_t^j} = \frac{\theta_t^i}{\theta_t^j} \cdot \frac{\theta_r^j}{\theta_r^i}$, $\forall\ i, j, t$. It should be evident that the change in relative prices over time is the same for both indexed price series, i.e. comparing periods t and u gives

$$\frac{\left[\frac{\pi_t^i}{\pi_t^j}\right]}{\left[\frac{\pi_u^i}{\pi_u^j}\right]} = \frac{\left[\frac{\theta_t^i}{\theta_t^j} \cdot \frac{\theta_s^j}{\theta_s^i}\right]}{\left[\frac{\theta_u^i}{\theta_u^j} \cdot \frac{\theta_s^j}{\theta_s^i}\right]} = \frac{\left[\frac{\theta_t^i}{\theta_t^j}\right]}{\left[\frac{\theta_u^i}{\theta_u^j}\right]} = \frac{\left[\frac{\theta_t^i}{\theta_t^j} \cdot \frac{\theta_r^j}{\theta_r^i}\right]}{\left[\frac{\theta_u^i}{\theta_u^j} \cdot \frac{\theta_r^j}{\theta_r^i}\right]} = \frac{\left[\frac{\phi_t^i}{\phi_t^j}\right]}{\left[\frac{\phi_u^i}{\phi_u^j}\right]} \quad \forall\, t, u$$

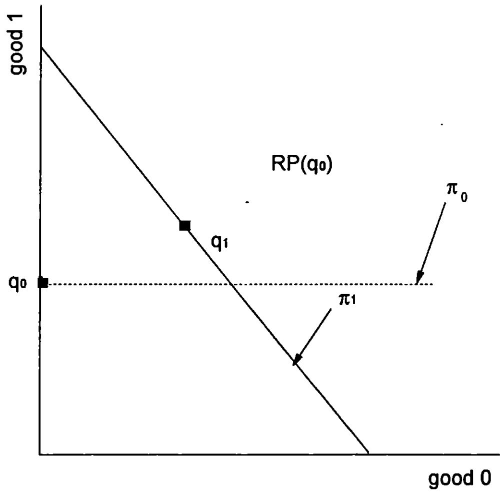
The problem with using the bundles observed in actual data is that[10], because movements of the budget line between periods are generally large and relative price changes are typically small, budget lines seldom cross. As a result as we saw in Chapters 3 and 4, data may lack power either to reject, *or* to usefully invoke GARP. This means that, when applying revealed preference restrictions to observed bundles, it is possible that the lower bounds we can recover are not particularly enlightening. To see this consider the two good, two period case illustrated in Figure 5.2 below.

We observe two bundles; the corner solution $\mathbf{q}_0$, and $\mathbf{q}_1$ which contains more of both commodities than $\mathbf{q}_0$. Since $\mathbf{q}_1$ lies in the interior of the $RP(\mathbf{q}_0)$ set by monotonicity, the data contain no additional information on the shape of the indifference curve through $\mathbf{q}_0$. The dashed budget line π_0 indicates the lowest price for π_0^0 consistent with the observed pattern of demand and GARP. In this case $\overline{\pi}_0^0 = 0$ and hence *any* non-negative price can rationalise the data.

The second problem with this approach is that, unlike parametric models, we cannot use data for periods when $\mathbf{q}_t \notin RP(\mathbf{q}_0)$. This is because these periods do not provide any revealed preference restrictions at all on π_0^0.

[10]As pointed out by, amongst others, Varian (1982) and Blundell *et al* (1997).

Figure 5.2: A negative virtual price, a two good, two period example.



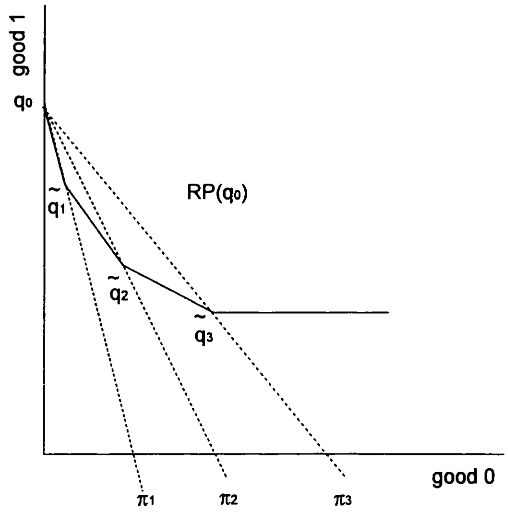
5.4.1 Improving the Bounds

Chapter 3 outlined a procedure designed to improve the power of tests of GARP. This made use of income expansion paths (Engel curves) estimated from micro data. These paths allow budget constraints to be moved in and out and are used to find a sequence of price-demand combinations designed to maximise the chances of finding rejections. This was necessary to overcome the problem posed for GARP tests by income growth over time. An adaptation of this technique was described in Chapter 4 where a method of improving the bounds to an indifference curve was described. This chapter also makes use of this technique to improve upon the bound on the virtual price which can be obtained from the raw data.

We can use the Engel curve/expansion path for $\mathbf{q}_t$ to find $\widetilde{\mathbf{q}}_t = E\left(\mathbf{q}_t \mid \widetilde{x}_t\right)$, where $\widetilde{x}_t = \boldsymbol{\pi}_t'\widetilde{\mathbf{q}}_t = \boldsymbol{\pi}_t'\mathbf{q}_0$, i.e. $\widetilde{\mathbf{q}}_t$ is the bundle which would be chosen at period t prices with a budget which makes $\mathbf{q}_0$ just affordable.

By shifting the budget constraint inwards in this manner, we improve

Figure 5.3: Bounding the virtual price, a two good, four period example.



the upper bound on the indifference curve passing through $\mathbf{q}_0$. In addition, we can now use information from all periods in which the full price vector is observed rather than just the subset of these periods which are revealed preferred to $\mathbf{q}_0$. That is, we can move budget lines *out* as well as in. We apply this procedure to *all* periods in which the full vector of prices is observed thereby defining a K-dimensional convex set representing the boundary of the $RP(\mathbf{q}_0)$ set (of which all $\widetilde{\mathbf{q}}_t$ are members). This is illustrated for $K+1 = 2, T+1 = 4$ in Figure 5.3.

As we know that $\widetilde{\mathbf{q}}_t R^0 \mathbf{q}_0$ (since by construction, $\boldsymbol{\pi}_t' \widetilde{\mathbf{q}}_t = \boldsymbol{\pi}_t' \mathbf{q}_0$, and so $\widetilde{\mathbf{q}}_t$ was chosen when $\mathbf{q}_0$ was affordable), we can use the set $\widetilde{\mathbf{q}}_t \in RP(\mathbf{q}_0)$ where $t = 1, ..., T$, to compute an improved lower bound on π_0^0 by the same argument as before. That is, for non-violation of GARP, $\widetilde{\mathbf{q}}_t R \mathbf{q}_0$ implies not $\mathbf{q}_0 P \widetilde{\mathbf{q}}_t$, and so

$$\boldsymbol{\pi}_0'\widetilde{\mathbf{q}}_t \geq \boldsymbol{\pi}_0'\mathbf{q}_0$$

$$\Rightarrow \quad \pi_0^0(\widetilde{q}_t^0 - q_0^0) \geq \sum_{k=1}^{K} \pi_0^k(q_0^k - \widetilde{q}_t^k)$$

$$\Rightarrow \quad \pi_0^0 \geq \frac{\sum_{k=1}^{K} \pi_0^k(q_0^k - \widetilde{q}_t^k)}{\widetilde{q}_t^0}, \qquad \text{as } q_0^0 = 0, \text{ and if } \widetilde{q}_t^0 > 0$$

Thus, each $\widetilde{\mathbf{q}}_t \in RP(\mathbf{q}_0)$ gives a lower bound on π_0^0 - call this set $\pi_0^0(\widetilde{\mathbf{q}}_t)$. As with the $\pi_0^0(\mathbf{q}_s)$ set, this will contain a highest value $(\max\{\pi_0^0(\widetilde{\mathbf{q}}_t)\}. \equiv \overline{\pi}_0^0)$, and it is this value which should be taken as the lower limit on π_0^0.

Proposition 5.2. *Any* $\pi_0^0 < \overline{\pi}_0^0$ *violates GARP for* $(\boldsymbol{\pi}, \widetilde{\mathbf{q}})$.

Proof.

The proof is analogous with that for Proposition 5.1.

■

The lower bound obtained by this method of using expansion paths is always an improvement over the bound obtained from raw data (unless $\{\widetilde{\mathbf{q}}_t\} = \{\mathbf{q}_s\}$).

Proposition 5.3. $\max\{\pi_0^0(\widetilde{\mathbf{q}}_t)\} \geq \max\{\pi_0^0(\mathbf{q}_s)\}$.

Proof.

(1) $\boldsymbol{\pi}_s'\mathbf{q}_s \geq \boldsymbol{\pi}_s'\mathbf{q}_0 = \boldsymbol{\pi}_s'\widetilde{\mathbf{q}}_s \Rightarrow \mathbf{q}_s R \widetilde{\mathbf{q}}_s \ \forall \ \mathbf{q}_s \in RP(\mathbf{q}_0)$

(2) The bound $\pi_0^0(\widetilde{\mathbf{q}}_s)$ comes from setting $\boldsymbol{\pi}_0'\widetilde{\mathbf{q}}_s = \boldsymbol{\pi}_0'\mathbf{q}_0 = x_0$

(3) Denote $\widetilde{\boldsymbol{\pi}}_0 = \left(\pi_0^0(\widetilde{\mathbf{q}}_s), \pi_0^1, ..., \pi_0^K\right)$

(4) The bound $\pi_0^0(\mathbf{q}_s)$ comes from setting $\boldsymbol{\pi}_0'\mathbf{q}_s = \boldsymbol{\pi}_0'\mathbf{q}_0 = x_0$

(5) Suppose that $\pi_0^0(\widetilde{\mathbf{q}}_s) < \pi_0^0(\mathbf{q}_s)$

(6) Since $\widetilde{\pi}_0^k = \pi_0^k \ \forall \ k \neq 0$ steps (2), (4) and (5) imply that $\widetilde{\boldsymbol{\pi}}_0'\mathbf{q}_s < x_0 = \widetilde{\boldsymbol{\pi}}_0'\widetilde{\mathbf{q}}_s \Rightarrow \widetilde{\mathbf{q}}_s P \mathbf{q}_s$, but this is a violation of GARP, from (1)

(7) $\pi_0^0(\widetilde{\mathbf{q}}_s) \geq \pi_0^0(\mathbf{q}_s)\ \forall\, s \Rightarrow \max\{\pi_0^0(\widetilde{\mathbf{q}}_s)\} \geq \max\{\pi_0^0(\mathbf{q}_s)\}$

(8) Since $\{\widetilde{\mathbf{q}}_s\} \subset \{\widetilde{\mathbf{q}}_t\}$, $\max\{\pi_0^0(\widetilde{\mathbf{q}}_t)\} \geq \max\{\pi_0^0(\widetilde{\mathbf{q}}_s)\} \Rightarrow \max\{\pi_0^0(\widetilde{\mathbf{q}}_t)\} \geq \max\{\pi_0^0(\mathbf{q}_s)\}$

■

The improved $RP(\mathbf{q}_0)$ set that comes from using expansion paths to calculate $\widetilde{\mathbf{q}}_t$ such that $\boldsymbol{\pi}_t'\,\widetilde{\mathbf{q}}_t = \boldsymbol{\pi}_t'\,\mathbf{q}_0\ \forall\, t \neq 0$ may not give the tightest upper bound on the indifference curve through $\mathbf{q}_0$ that we can obtain. This can be seen by considering the following. Amongst the $RP(\mathbf{q}_0)$ set, we may be able to find one or more members $\widetilde{\mathbf{q}}_i$ for which there exists some $\widetilde{\mathbf{q}}_j \in RP(\mathbf{q}_0)$, $j \neq i$, such that $\widetilde{\mathbf{q}}_i P^0 \widetilde{\mathbf{q}}_j$, i.e. $\boldsymbol{\pi}_i'\widetilde{\mathbf{q}}_i > \boldsymbol{\pi}_i'\widetilde{\mathbf{q}}_j$. The algorithms described in Chapter 4 were designed to find exactly this type of point. In this case, we can use expansion paths to find a $\widehat{\mathbf{q}}_i$ for each $\widetilde{\mathbf{q}}_i$ such that $\boldsymbol{\pi}_i'\widehat{\mathbf{q}}_i = \boldsymbol{\pi}_i'\widetilde{\mathbf{q}}_j$, i.e. $\widehat{\mathbf{q}}_i R^0 \widetilde{\mathbf{q}}_j$. Since $\widetilde{\mathbf{q}}_j R^0 \mathbf{q}_0$ and $\widehat{\mathbf{q}}_i R^0 \widetilde{\mathbf{q}}_j$ this implies that $\widehat{\mathbf{q}}_i R \mathbf{q}_0$. In addition $\boldsymbol{\pi}_i'\widetilde{\mathbf{q}}_i > \boldsymbol{\pi}_i'\widetilde{\mathbf{q}}_j = \boldsymbol{\pi}_i'\widehat{\mathbf{q}}_i$ tells us that $\widetilde{\mathbf{q}}_i P^0 \widehat{\mathbf{q}}_i$. Hence $\widetilde{\mathbf{q}}_i P^0 \widehat{\mathbf{q}}_i R \mathbf{q}_0$, which implies that $\widehat{\mathbf{q}}_i$ tightens the boundary on the indifference curve passing through $\mathbf{q}_0$ as compared to $\widetilde{\mathbf{q}}_i$. As in Chapter 4, it may be possible to iterate this procedure several times as each improvement may introduce new $\mathbf{q}_i P^0 \mathbf{q}_j$ relationships, where $\mathbf{q}_i$ and $\mathbf{q}_j$ are members of the current best $RP(\mathbf{q}_0)$ set. It might seem that this would allow us to further improve the bound on π_0^0. However, this proves not to be the case as the following proposition shows.

Proposition 5.4. *None of these further boundary improvements on the original improved $RP(\mathbf{q}_0)$ set will enable us to tighten the lower bound on π_0^0.*

Proof.

(1) Take $\widetilde{\mathbf{q}}_i$, $\widehat{\mathbf{q}}_i \in RP(\mathbf{q}_0)$ where $\widetilde{\mathbf{q}}_i P^0 \widehat{\mathbf{q}}_i$

(2) Then $\exists$ a $\mathbf{q}_j \in RP(\mathbf{q}_0)$ s.t. $\boldsymbol{\pi}_i'\widetilde{\mathbf{q}}_i > \boldsymbol{\pi}_i'\mathbf{q}_j = \boldsymbol{\pi}_i'\widehat{\mathbf{q}}_i$, i.e. $\widetilde{\mathbf{q}}_i P^0 \widehat{\mathbf{q}}_i R^0 \mathbf{q}_j R \mathbf{q}_0$

(3) Denote the bound on π_0^0 from setting $\boldsymbol{\pi}_0'\mathbf{q}_j = \boldsymbol{\pi}_0'\mathbf{q}_0 = x_0$ by $\pi_0^0(\mathbf{q}_j)$

(4) Let $\boldsymbol{\pi}_{0j}$ be the price vector for period 0 when π_0^0 is set to $\pi_0^0(\mathbf{q}_j)$

(5) Denote the bound on π_0^0 from setting $\boldsymbol{\pi}_0'\widehat{\mathbf{q}}_i = \boldsymbol{\pi}_0'\mathbf{q}_0 = x_0$ by $\pi_0^0(\widehat{\mathbf{q}}_i)$

(6) Let $\boldsymbol{\pi}_{0i}$ be the price vector for period 0 when π_0^0 is set to $\pi_0^0(\mathbf{q}_i)$

(7) Suppose that $\pi_0^0(\mathbf{q}_j) < \pi_0^0(\widehat{\mathbf{q}}_i)$

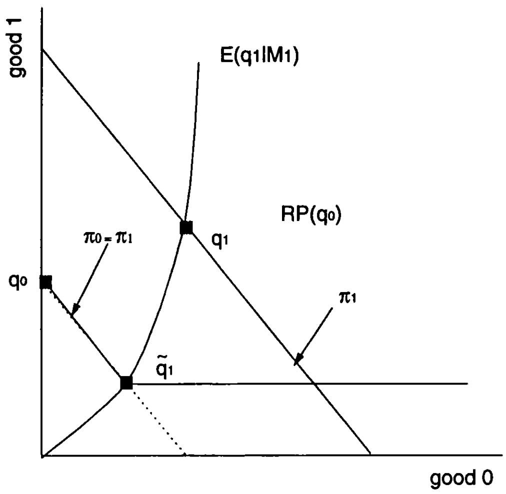
(8) Since $\pi_{0i}^k = \pi_{0j}^k \ \forall \ k \neq 0$ steps (3), (5) and (7) imply that $\boldsymbol{\pi}_{0j}'\widehat{\mathbf{q}}_i < \boldsymbol{\pi}_{0i}'\widehat{\mathbf{q}}_i = x_0 = \boldsymbol{\pi}_{0j}'\mathbf{q}_0 \Rightarrow \mathbf{q}_0 P \widehat{\mathbf{q}}_i$, which is a violation of GARP

(9) Hence $\pi_0^0(\mathbf{q}_j) \geq \pi_0^0(\widehat{\mathbf{q}}_i)$, so improving the boundary point from $\widetilde{\mathbf{q}}_i$ to $\widehat{\mathbf{q}}_i$ cannot give a higher lower bound on π_0^0 than can already be obtained from $\mathbf{q}_j$

■

The use of the expansion path is illustrated for the two good, two period case below. The curve $E(\mathbf{q}_1 \mid x_1)$ is the expansion path through the bundle chosen in period 1 ($\mathbf{q}_1$).

Figure 5.4: The improved bound, a two good, two period example.



The dashed line shows the budget constraint which makes $\mathbf{q}_0$ just affordable at period 1's prices, and the bundle which would be chosen under these circumstances, $\widetilde{\mathbf{q}}_1$, is given by the intersection of the expansion path with this budget constraint. As $\widetilde{\mathbf{q}}_1 R \mathbf{q}_0$, the line through $\widetilde{\mathbf{q}}_1$ and $\mathbf{q}_0$ gives the lowest value for π_0^0 consistent with GARP for the data set $(\boldsymbol{\pi}_0, \boldsymbol{\pi}_1; \mathbf{q}_0, \widetilde{\mathbf{q}}_1)$.

As is evident from the illustration, in the two good case, the lower bound obtained for π_0^0 is simply that the price of good 0 relative to good 1 must be greater than or equal to the period 1 relative price, i.e. the lowest price we have subsequently observed it being purchased at. This is only the case for examples with only two goods.

Proposition 5.5. *For $K+1=2$ and $T+1 \geq 2$, the lower bound on π_0^0 is $\frac{\pi_0^0}{\pi_0^1} = \frac{\pi_r^0}{\pi_r^1}$, where $\frac{\pi_r^0}{\pi_r^1}$ is the highest (by Proposition 5.1) observed relative price.*

Proof.

(1) $\widetilde{\mathbf{q}}_r R \mathbf{q}_0$ and GARP $\Rightarrow \pi_0^0 \widetilde{q}_r^0 + \pi_0^1 \widetilde{q}_r^1 \geq \pi_0^1 q_0^1$

(2) $\Rightarrow \frac{\pi_0^0}{\pi_0^1} \geq \frac{q_0^1 - \widetilde{q}_r^1}{q_0^1}$

(3) $\widetilde{x}_r = \boldsymbol{\pi}_r' \mathbf{q}_0 \Rightarrow \pi_r^0 \widetilde{q}_r^0 + \pi_r^1 \widetilde{q}_r^1 = \pi_r^1 q_0^1$

(4) $\Rightarrow \frac{\pi_r^0}{\pi_r^1} = \frac{q_0^1 - \widetilde{q}_r^1}{q_0^1}$

(5) From (2) and (4) $\frac{\pi_0^0}{\pi_0^1} \geq \frac{\pi_r^0}{\pi_r^1}$

■

Therefore, with only two goods, the lower bound for π_0^0 is the highest relative price at which we have since observed it being bought — which is not particularly insightful. However, for $K+1 > 2$, $\frac{\pi_0^0}{\pi_0^K} \geq \frac{\pi_t^0}{\pi_t^K}$ iff $\frac{\pi_0^k}{\pi_0^K} = \frac{\pi_t^k}{\pi_t^K}$ $\forall\, k \neq 0$, (taking the Kth good as the numeraire), that is, the lower bound on the period 0 relative price for good 0 is equal to the highest subsequent observed relative price for good 0 *only* if there is no relative price movement in the other $k = 1, .., K$ goods between period 0 and period t.

Figure 5.5: Bounding the virtual price,a three good, two period example.

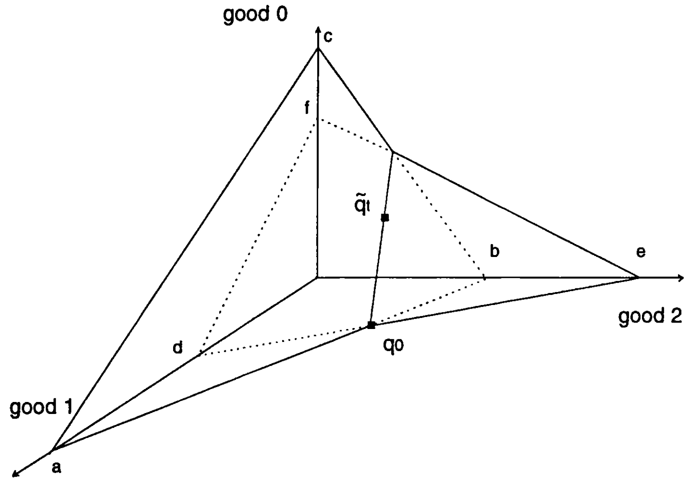


Figure 5.5 illustrates a three good, two period case. Suppose ab is the period 0 budget line over goods 1 and 2, which is known since π_0^1 and π_0^2 are observed. The simplex def is the budget set which makes $\mathbf{q}_0$ affordable at period t prices. The point $\widetilde{\mathbf{q}}_t$ must be on the boundary of the period 0 budget set, and so the line ab and the point $\widetilde{\mathbf{q}}_t$ define the period 0 budget set, given by the plane abc. This shows that the only circumstance under which the limit we obtain for the price of good 0 relative to goods 1 and 2 is identical to the relative prices in period t, is one in which relative prices for goods 1 and 2 do not change between periods, so that ab=de.

Note that, unlike the two good case, the limit on the virtual price we obtain for good 0 can be below the period t relative price. This is easily seen if de is taken as the period 0 budget constraint, and abc is the period t budget set (adjusted to make $\mathbf{q}_0$ just affordable at period t prices), in which case the gradients of ef and df define the price limit on good 0 relative to goods 1 and 2 respectively. In this case, the price level of good 0 relative to good 1 will fall between period 0 and period t, and its price relative to good 2 rises. It is also easy to construct an example where the price of good 0 relative to both good 1 and good 2 falls between period 0 and period t. This can be seen by imagining the case where $\widetilde{\mathbf{q}}_t$ has more of good 2 and less of goods 1 and 0 than currently illustrated. This shows that it is possible that the virtual relative price of good 0 can be lower than observed prices in later periods which have positive demands for good 0, and that this does not reject GARP.

5.5 An Empirical Application: Valuing the National Lottery

5.5.1 Data

The particular example of a new good which we have chosen to examine is the UK National Lottery. There are three reasons for this. First, spending on the Lottery appeared as a separately identified expenditure category in the our data source, the UK Family Expenditure survey (F.E.S.), immediately upon commencement in November 1994. This is comparatively rare since spending data on most new goods are usually allocated to residual expenditure categories, and only given separate codings once they have proved themselves sufficiently important. This was the case, for example, with both video recorders and personal computers until 1984 and 1985 respectively. The National Lottery, however, was recognised as interesting enough at the time of its launch (November 1994) for it to warrant separate recording immediately. This makes the effects of its introduction much cleaner in the data. Secondly, unlike many new goods, particularly technological goods, in the time period covered by our data set the Lottery has not been subject to much change in quality since its introduction[11] - its characteristics have remained largely unaltered[12]. Since the distinction between issues of newness and those of quality variation are not always clear (since a substantially improved good may thought of as new, in some sense), this also makes the example a cleaner one. Finally, the conceptual basis for pricing a new good relies on a characteristics model in which the new good is a combination of existing characteristics. The National Lottery might fit this description quite well, since it can be thought of as a bet or an investment with risk/return attributes which were already available as a combination of pre-existing gambles and assets.

The most up-to-date survey covers the financial year 1994/95. We define monthly price regimes using sub-price indices from the UK Retail Price Index. This gives us data for five months during which the full set of goods, including the Lottery, were available (November 94 to March 95). However, the Lottery was only introduced part way through November, so, as we do not have a full month's observations, we drop November from the sample. We take December 93 as our 0th period, and calculate the virtual price of the National Lottery just under one year before its introduction.

We allocate household spending to 24 commodity groups (bread, meat, dairy, vegetables and fruit, other foods, food out, beer, wine and spirits, tobacco, electricity, gas, other fuels, household goods, household services, men's clothing, women's clothing, children's clothing, other clothing and footwear, personal goods and services, motoring, fares, leisure goods, leisure

[11]There are now Wednesday draws which may have affected the demand for the initial Saturday only draws, however our data ends before these were introduced.

[12]The expected value of a ticket may vary from week to week (particularly on a rollover week). However, since our data are monthly we expect to average out most of such variation. In particular Farrel and Walker (1997) show that in roll-over weeks increased demand for the tickets keeps the expected value of a ticket largely constant.

services, National Lottery), and calculate price indices for these groups using within-period weights for the 70 or so sub-categories included in the non-housing RPI. We also trim the top and bottom percentiles of the within-period total (non-housing) expenditure distribution. This leaves us with 4,578 observations.

5.5.2 Results

To estimate the virtual price π_0^0 by applying revealed preference conditions to the dataset $(\boldsymbol{\pi}_t, \mathbf{q}_t)$ where $t \neq 0$, we are exploiting the assumption that the data were generated by a stable, well-behaved utility function. The validity of this assumption would be compromised immediately if the data $(\boldsymbol{\pi}_t, \mathbf{q}_t)$ where $t \neq 0$, did not itself satisfy GARP. However, even if GARP is a good approximation to the behaviour of real households, optimisation errors by households and measurement errors in data mean that large micro-level datasets are likely to reject GARP (see Famulari (1995) and the discussion in Chapter 3 for example). This is particularly the case when we move budget lines so that they cross more frequently than otherwise, in order to improve the bound on the virtual price. As with Chapter 3 and 4, the estimated expansion paths provide the stochastic structure necessary to assess the seriousness of any rejections

If rejection of GARP for the $t \neq 0$ periods is insignificant for some acceptable size of test, we can proceed with the ideas outlined in section 5.4. While it is obvious that, given violations in the $t \neq 0$ periods, there cannot exist values for the virtual price π_0^0 which can rationalise the whole dataset *exactly*, the virtual price we calculate will not introduce any *further* violations (by Proposition 5.1) and so will be consistent with the idea that the entire dataset does not statistically reject GARP.

To begin with, we check for violations of GARP in the four periods in which we observe all demands by estimating consumption bundles at four different points in the total expenditure distribution; the within-period median and mean demands for the raw data, and two sets of adjusted data, using estimated expansion paths to set $x_t = \pi_t' \mathrm{q}_0$, which take the median and mean demands respectively in period 0 (i.e. December 93) as their expenditure base. The results are shown in Tables 5.1.

Table 5.1: Violations of GARP by size of test.

Comparison paths	Size of Test[13]				
	Raw Count	0.30	0.20	0.10	0.05
Median	0	0	0	0	0
Mean	0	0	0	0	0
Improved bound					
Median	3	1	1	0	0
Mean	0	0	0	0	0

There are no rejections of GARP for the within-period raw mean and median demands. Usually this is not surprising since, as pointed out earlier, average growth in expenditure is often large enough to swamp relative price movements particularly with annual data. However, with these monthly data, rejections may be thought more likely, even at average demands, since growth in total expenditure over the period may be small, and one reason why GARP can be rejected is seasonality in demands; particularly between short time periods like adjacent months as we have here. There are rejections of GARP for the improved $RP(\mathrm{q}_0)$ set calculated using 0th period median demand, though not for that using the mean demand. Violations are certainly not surprising since these points are located on budget lines which cross frequently (see Figure 5.5). However, these violations are not significant at a 90% confidence level. We therefore conclude that, at least for average or median demand, these data do not significantly reject GARP.

Table 5.2 reports the virtual prices consistent with each consumption bundle on the improved and non-improved boundaries of the $RP(\mathbf{q}_0)$ sets. First consider the within-period median and mean paths. We find, for the median path, that all four periods are revealed preferred to December 93, whereas this is only true for one period (December 94) for the mean. We compute bounds on the virtual price of the new good based on each of these bundles and report them in descending order of size (setting March 95 = 1000) and give the number of violations of GARP for the whole dataset, including $t = 0$, based on each virtual price. In accordance with Proposition 5.1, the maximum virtual price in each set of points ($\pi_0^0 = -490$, $\pi_0^0 = -25302$ respectively) gives no additional rejections, while, when there is more than one revealed preferred period as for the median, the lower prices do introduce GARP violations. However, the bound that either of these sets gives us on the virtual price is less than useful. Even the highest virtual prices recovered are negative, implying they are obtained from points which are revealed preferred to demand in December 93 simply by monotonicity (as illustrated in Figure 5.2). In effect, all these observations can tell us is that any non-negative virtual price for the lottery would be consistent with GARP and zero demand in December 93. This illustrates the problems which arise for testing or invoking GARP when growth in total spending overwhelms relative price movements.

Table 5.2: Virtual Prices and violations of GARP by size of test.

Comparison Path	Virtual Price (March 95 = 1000)	Size of Test Raw Count	0.30	0.20	0.10	0.05
Median	-490	0	0	0	0	0
	-1247	1	1	1	1	1
	-3015	3	3	3	3	3
	-31551	6	6	6	6	6
Mean	-25302	0	0	0	0	0
Improved bound						
Median	1206	3	1	1	0	0
	1104	10	1	1	0	0
	1096	10	1	1	0	0
	924	10	1	1	0	0
Mean	1356	0	0	0	0	0
	1307	3	0	0	0	0
	1247	3	0	0	0	0
	998	6	0	0	0	0

The results from the improved revealed preferred sets are given in the next two blocks. With these we are able to set the budget lines so that all periods are revealed preferred to period 0 (December 93) by setting x_t such that $\pi_t' \mathbf{g}_t (x_t) = \pi_t' \mathbf{q}_0 (x_0)$. Thus we maximise the informational content of the data by pulling the budget lines for already revealed preferred periods *in,* and pushing those for the previously un-usable not revealed preferred periods (a problem for the mean, though not the median, path) *out.* All virtual prices computed for these bundles are now positive and the highest from each set (1205 and 1356 for the median and mean based paths respectively) are such that they do not introduce further violations of GARP (consistent with Proposition 5.2). Note that they are both greater than the solutions for the corresponding unimproved paths (consistent with Proposition 5.3).

We use our results to examine the effect of including the new good in a range of price indices. In constructing an aggregate price index, we choose to take a weighted average of individual price indices, rather than calculating an index based on average demands and prices; the two are unlikely to be equal, from the familiar result that the average of a function does not necessarily equal the function of the averages.

It could be expected that the virtual price of the new good will differ across heterogeneous individuals, and, if so, the price which would drive *aggregate* demand for the new good to zero would be the highest reservation price we can find amongst all individuals. Only then will each individual demand none of the good. However, in calculating the individual price indices which form the basis of the aggregate index, it should not be the virtual price which drives aggregate demand to zero which is used, but the individual's own virtual price. That is, the shadow price for each individual should be used, rather than the single market price which would just result in zero aggregate demand. This, again, captures the idea that an individual receives no benefit from falls in price above his own reservation price, particularly in thinking of price indices as under- or over-estimates of true cost-of-living indices. For every other price needed to construct the indices we use the observed market price. This involves the assumption that individual shadow prices are identical and equal to market prices in all cases apart from the virtual price of the new good. In other words, that an individual's reservation price is never below the market price for any other good in any period, or for the new good from the 1st period onwards.

In our specification of Engel curves, we assume that the only source of heterogeneity is in the size of households' budgets, and that all individuals face identical prices for all goods in a given period (excepting the virtual price for the new good in the 0th period). Given these assumptions, the virtual price could be expected to vary with total budgets in a predictable way for certain types of new good. For example, if the new good is a normal good, then we should observe a virtual price which increases with income. Normality would mean that, starting from a given virtual price which drives a particular individual's demand for the new good to zero, if the income endowment is increased, then demand for the good will become positive. To reduce demand to zero again, when the prices of all other goods remain the same, the price of the new good needs to increase, since normality also implies a negative own price effect on demand. Thus, virtual price will rise with total budgets if the good of interest is normal.

Ideally, we would like to construct the aggregate price indices as an average of all individual indices, but, since the FES does not follow individuals over time, this will not usually be possible. Take, for example, the Paasche index, $P_i = \frac{\pi_1' q_{1i}}{\pi_{0i}' q_{1i}}$, where i is an index of individuals in period 1, $i = 1, ..., I$, and $\boldsymbol{\pi}_{0i} = (\pi_{0i}^0, \pi_0^1, ..., \pi_0^K)$. This index uses period 1 demand as its base, thus the problem which arises is that we cannot calculate π_{0i} for these particular I period 1 individuals since we do not observe them in period 0. We can, though, follow observationally similar individuals across time. Given our assumption that individuals differ only in total expenditure, this means treating individuals from the same part of the expenditure distribution over time as being similar. Thus individual i in period 1 can be assigned the virtual price calculated for an individual in period 0 occupying the same position in the spending distribution. In reality, of course, the individuals in each month's sample will not be drawn from exactly the same points in the distribution, and so we must work with average demands to a certain extent, as constructing similar individuals will entail averaging across the month's data within a given range of the distribution. We choose budget deciles as our level of disaggregation, and calculate virtual prices for each decile based on within-decile mean demand, which we then use to calculate a price index for the decile[14]. Whilst the data we have available leads us to choose this approach, these within-decile averages, naturally, suffer from the criticism that they are based on average demand and price, and are therefore not necessarily equal to the true averages, and that the virtual price for the decile is calculated from average demand, and, in the same manner, may not equal the average of individual reservation prices across the decile, could they be calculated.

Table 5.3: Virtual Price by budget decile.

Comparison Path	Virtual Price (March 95=1000)	$t \neq 0$ data pass GARP at
1st decile mean	1173	raw
2nd decile mean	1131	70%
3rd decile mean	1132	70%
4th decile mean	1142	90%
5th decile mean	1150	80%
6th decile mean	1276	70%
7th decile mean	1419	raw
8th decile mean	1477	raw
9th decile mean	1481	90%
10th decile mean	1692	90%

Table 5.3 reports the virtual price recovered for the within-decile mean expenditures, and the rounded confidence level at which the data excluding

[14]Note that, as the Laspeyres index uses q_0 as its base, it would be possible to construct an index for each individual. However, for comparability purposes, we calculate all the aggregate price indices based on budget deciles.

the 0th period passes GARP. The results show that, broadly speaking, a positive relationship between total budget and the virtual price of the lottery, consistent with it being a normal and ordinary good, is observed in our data. Again, in all cases, any violations of GARP are contained in the data excluding period 0, and the virtual price calculated does not introduce new violations of GARP. But as we must be dubious of attaching any meaning to a virtual price derived from data characterised by significant violations of GARP, we need to check whether this occurs, and the table reports that no violations are significant beyond a 90% confidence level.

These prices appear rather low — only about 70% higher than the actual price charged at the introduction of the Lottery even for the top expenditure decile. The first thing to bear in mind is that these numbers are a *lower bound* on the possible value of the virtual price (the lowest consistent with utility maximisation). There are a number of other possible reasons as well. Firstly, the changes in price weighted demands between the base period and later periods $\left(\pi_0^k\left(q_0^k - q_s^k\right)\right)$ may be small. This could be because we are only looking at variations in demand over one year, in which time we might not expect big demand changes; or possibly because there are demand changes taking place outside our group of goods, perhaps in housing expenditure or, more likely, in savings decisions. Secondly, we may be smoothing preferences across two heterogeneous groups; one of which will buy lottery tickets while the other is made up of abstainers who would never buy a ticket whatever the price. Thus the virtual price we calculate may be an average between a price of zero for one group (the abstainers) and a value somewhat higher than our estimate for the other group.

Nevertheless, this virtual price solution is the lowest consistent with the maintained hypothesis that these FES data were generated by the maximisation of a uniform, stable and well-behaved utility function, without invoking an additional assumption on the specific form for that function. Further, given that we only have five price points at which we observe demands for 24 goods, even if we wanted to estimate a parametric demand model, it is highly unlikely that it would be possible to obtain reliable estimates of price coefficients on these data.

We calculate three commonly used prices indices (excluding housing costs); the Laspeyres, Paasche and Törnqvist indices. In constructing the aggregate index, we weight the individual indices by the proportion of that individual's (or, rather, decile's) spending out of aggregate spending in the base period used in the index (the starting period for the Laspeyres, the end period for the Paasche, and both periods for the Törnqvist). For example, denoting the aggregate Paasche index by P, the individual index by P_i, for $i = 1, ..., I$, and the weight by W_i^p, then

$$P = \sum_{i=1}^{I} P_i W_i^p$$

where

$$P_i = \frac{\boldsymbol{\pi}_1' \mathbf{q}_{1i}}{\boldsymbol{\pi}_{0i}' \mathbf{q}_{1i}} \quad \text{and} \quad W_i^p = \frac{\boldsymbol{\pi}_1' \mathbf{q}_{1i}}{\sum\limits_{i=1}^{I} \boldsymbol{\pi}_1' \mathbf{q}_{1i}}$$

Period 1 prices are the same for everyone, hence $\boldsymbol{\pi}_1$ is not indexed by i, and $\pi_{0i} = (\pi_{0i}^0, \pi_0^1, ..., \pi_0^K)$, the vector of period 0 prices containing individual $i's$ own virtual price for the new good. Table 5.4 shows how the inclusion of the Lottery affects inflation rates compared to its exclusion, over the period December 93 to December 94.

Table 5.4: Annual inflation rates (%), with and without the new good.

Nov. 93 to 94	Laspeyres	Tornqvist	Paasche
Exc. Lottery	1.87	1.86	1.84
Inc. Lottery	1.87	1.67	1.58

Naturally, the Laspeyres measure of inflation is unaffected by the presence of the new good since it calculates the relative cost of buying the starting period's bundle over time, which, of course, does not contain the new good. The Paasche index tracks the cost of purchasing the end period demand, which does include the new good. As is often the case, it gives lower values for inflation than the Laspeyres index. The Paasche index will be affected by whether the new good is included or not. In this example, inclusion of the Lottery has the effect of reducing the measure of inflation by 0.26 percentage points, about a 14% reduction, compared to when the index ignores the new good. The fall in value follows because the virtual price of the Lottery is higher than its observed price one year later; this falling price will tend to reduce the measure of inflation. The Törnqvist index lies between the Laspeyres and Paasche indices, and may be a closer approximation to a true cost-of-living index under a more plausible model of consumer behaviour than those implicit in the other two[15]. The effect on the Törnqvist index of including the new good is more muted than for the Paasche, and reduces the inflation rate by 0.19 percentage points, or by around 10% of the lottery-exclusive measure.

[15] See Diewert (1976).

5.6 Summary and Conclusion

This chapter presents a revealed preference method of calculating the lower bound on the virtual price of a new good for the period immediately prior to the one in which it first exists. This bound is chosen such that the data are consistent with the Generalised Axiom of Revealed Preference and, therefore, it is also consistent with the maximisation of a well-behaved utility function. As a result this bound encompasses all parametric solutions which arise from the estimation of integrable demand systems. In this chapter a method for improving the bounds recoverable, wherever appropriate data exist, by using non-parametric expansion paths is presented. This allows us to increase the amount of information on the position of the indifference curve through the consumption bundle chosen when the new good did not exist and hence improve the bound on the virtual price. It is argued that this approach has two principal merits compared to parametric estimation. First, it does not require a maintained assumption regarding the form of the utility function. Second, it is computationally simple. These ideas are applied to UK Family Expenditure Survey data on the National Lottery and its virtual price, one year before its introduction is computed at various points across the expenditure distribution. These virtual prices are used to examine the effect of including the new good in annual non-housing measures of inflation over the year December 93 to December 94. We find the value obtained can be up to 0.26 percentage points lower than when the Lottery is excluded from the calculation, depending on the price index used.

Bibliography

[1] Allen, R. G. D. (1958), "Movements in Retail Prices since 1953", *Economica*, **25**, 14-25.

[2] Afriat, S.N. (1965),"The Construction of Utility Functions from Expenditure Data", *International Economic Review*, **8**, 67-77.

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

[4] Afriat, S.N. (1976), *The Combinatorial Theory of Demand*, London: Input-Output Publishing Co.

[5] Afriat, S.N. (1977), *The Price Index*, Cambridge: Cambridge University Press.

[6] Afriat, S.N. (1981), "On the Constructability of Consistent Price Indices between Several Periods Simultaneously", in *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*, ed. by Angus Deaton, Cambridge: Cambridge University Press.

[7] Atkinson, A.B., J. Gomulka, and N. H. Stern (1990),"Spending on Alcohol: Evidence from the Family Expenditure Survey 1970-1983", *Economic Journal*, **100**, 808-827.

[8] Banks, J. Blundell, R.W. and Lewbel, A. (1994), "Quadratic Engel Curves, Indirect Tax Reform and Welfare Measurement", UCL Discussion Paper, 94-04, forthcoming *Review of Economics and Statistics.*

[9] Banks, J., Blundell, R. and Lewbel, A. (1996). "Tax Reform and Welfare Measurement: Do We Need Demand System Estimation?", *Economic Journal*, **106**, 1227-1241.

[10] Berndt, E.R., Griliches, Z. and N.J. Rappaport, (1995), "Econometric Estimates of Price Indices for Personal Computers in the 1990's", *Journal of Econometrics*, **68**, 243-268.

[11] Bierens, H.J. and Pott-Buter H.A. (1990) "Specification of Household Engel Curves by Nonparametric Regression", *Econometric Reviews*, **9**, 123-184.

[12] Blundell, R. W., (1988), "Consumer Behaviour: Theory and Empirical Evidence", *Economic Journal*, **98**, 16-65.

[13] Blundell, R., Pashardes, P. and Weber, G. (1994), "What Do We Learn About Consumer Demand Patterns from Micro-data?", *American Economic Review*, **83**, 570-83.

[14] Bradshaw, J. and Godfrey, C. (1983), "Inflation and the Poor", *New Society*, **65**, 247-8.

[15] Braithwait, S.D. (1980), "Substitution bias in the Laspeyres Price index: An Analysis using Estimated Cost of Living Indices", *American Economic Review*, **70**, 64-77.

[16] Brittain, J. A. (1960), "Some Neglected Features of Britain's Income Levelling", *American Economic Review*, **50**, Papers and Proceedings, 335-44.

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

[18] Brown, D.J. and Matskin, R.L. (1995),"Estimation of nonparametric functions in simultaneous models with an application to consumer demand", mimeo, Northwestern University.

[19] Brown, B.W. and Walker, M.B., (1989), "The Random Utility Hypothesis and Inference in Demand Systems" *Econometrica*, **57**, 815-829.

[20] Browning, M.J. and Meghir, C. (1991), "The Effects of Male and Female Labour Supply on Commodity Demands", *Econometrica*, **59**, 925-951.

[21] Christensen, A. G., Jorgenson, D. W. and Lau, L. J. (1975), 'Transcendental Logarithmic Utility Functions', *American Economic Review*, **65**, 367-83.

[22] Crawford, I., Laisney, F and I. Preston (1996), "Estimation of Household Demand Systems with Unit Values", Institute for Fiscal Studies, Working Paper, 96/16.

[23] Cross, R. (1982), "The Duheme-Quine Thesis, Lakatos and the Appraisal of Theories in Macroeconomics", *Economic Journal*, **92**, 320-340.

[24] Cunningham, A.W.F. (1996), "Measurement Bias in Price Indices: an Application to the U.K.'s R.P.I.", Bank of England, Working Paper, No. 47.

[25] Deaton, A.S. (1981) Editors preface in *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*, ed. by Angus Deaton, Cambridge: Cambridge University Press.

[26] Deaton, A.S. (1988), "Estimation of Own and Cross-Price Elasticities from Household Survey Data", *Journal of Econometrics*, **36**, 7-23.

[27] Deaton, A.S. (1988), "Quality, Quantity and the Spacial Variation of Price", *American Economic Review*, **78**, 418-443.

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

[29] Deaton A. S. and J. Muellbauer (1980), *Economics and Consumer Behaviour*, Cambridge, Cambridge University Press.

[30] Department of Social Security (1994), *Households Below Average Income: A Statistical Analysis 1979-1991/92*, London: HMSO.

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

[32] Diewert, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics*, **4**, 115-145.

[33] Diewert, W.E (1978), "Duality Approaches to Microeconomic Theory", IMSS Technical Report, No. 281, Stanford University, cited in Diewert (1981).

[34] Diewert, W.E. (1981), "The Economic Theory of Index Numbers", in *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*, ed. by Angus Deaton, Cambridge: Cambridge University Press.

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

[36] Dougherty, A. and van Order, R. (1982), "Inflation, Housing Costs and the Consumer Price Index", *American Economic Review*, **72**, 154-64.

[37] Engel, E. (1895), "Die Lebenskosten Belgischer Arbeiter-Familien früher und jetzt", *International Statistical Institute Bulletin*, **9**, 171-80.

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

[39] Fan, J. (1992),"Design Adaptive Nonparametric Regression", *Journal of the American Statistical Association*, **87**, 998-1004.

[40] Farrell, L and I. Walker, (1997), "It could be you ! But What's it Worth ? The Welfare Gain from the National Lottery", Institute for Fiscal Studies, Working Paper, W97/4.

[41] Fisher, F.M. and K. Shell (1971, "Taste and Quality Change in the Pure Theory of the True Cost of Living Index" in Z. Griliches (ed) *Price Indexes and Quality Change: Studies in New Methods of Measurement*, Harvard University Press, Cambridge, M.A.

[42] Frisch, R. (1936), "Annual Survey of General Economic Theory: The Problem of Index Numbers", *Econometrica*, 4, 1-39.

[43] Fry, V. and Pashardes, P. (1985), "Constructing True Cost of Living Indices from Engel Curves", Institute for Fiscal Studies, Working Paper no. 65.

[44] Fry, V. and Pashardes, P. (1986), *The RPI and the Cost of Living*, Report Series no. 22, London: Institute for Fiscal Studies.

[45] Goodman, A. and Webb, S. (1994), *For Richer, For Poorer: The Changing Distribution of Income in the United Kingdom, 1961-91*, Commentary no. 42, London: Institute for Fiscal Studies.

[46] Gordon, R.J. (1990), *The Measurement of Durable Goods Prices*, Chicargo: University of Chicago Press.

[47] Gorman, W. M. (1953), "Community Preference Fields", *Econometrica*, **21**, 63-80

[48] Gorman, W.M. (1956), "A possible procedure for analysing quality differentials in the egg market", *London School of Economics*, mimeo, reprinted in *Review of Economic Studies*, **47**, 843-856 (1980).

[49] Gorman, W. M. (1981), "Some Engel Curves", in *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*, ed. by Angus Deaton, Cambridge: Cambridge University Press.

[50] Grandmont, J.M. (1992), "Transformation of the Commodity Space, Behavioral Heterogeneity, and the Aggregation Problem", *Journal of Economic Theory*, **57**, 1-35.

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

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

[53] Härdle W. and Linton, O. (1995), "Nonparametric Regression analysis" in Engle and McFadden, *Handbook of Econometrics* Vol IV, North-Holland.

[54] Hausman, J. (1997), "Valuation of New Goods under Perfect and Imperfect competition", in Bresnahan, T. and R.J. Gordon, eds., *The Economics of New Goods*, Chicago: University of Chicago Press.

[55] Hausman, J. A., W. K. Newey, H. Ichimura, and J. L. Powell (1991), "Identification and Estimation of Polynomial Errors in Variables Models", *Journal of Econometrics*, **50**, 273-296.

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

[57] Heckman, J.J. (1974), "Effect of Child-Care Programs on Women's Work Effort", *Journal of Political Economy.*

[58] Hicks, J. (1940), "The Valuation of the Social Income", *Economica*, 105-124.

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

[60] Holly, A. (1982), "A Remark on Hausman's Specification Test", *Econometrica*, **50**, 749-59.

[61] Houthakker, H.S. (1950), "Revealed Preference and the Utility Function", *Economica*, New Series **17**, 159-174.

[62] Howe,H., R.A,Pollak, and T.J.Wales (1979) "Theory and Time Series Estimation of the Quadratic Expenditure System", *Econometrica*, **47**, 1231-1248.

[63] Jorgenson,D.W., L.J.Lau, and T.M.Stoker (1982),"The Transcendental Logarithmic Model of Aggregate Consumer Behavior", in *Advances in Econometrics*, Vol. 1, ed. by R. Basmann and G. Rhodes, Greenwich, Connecticut: JAI Press.

[64] Joseph, M.F.W. (1935-36) "Mr. Lerners' Supplementary Limits for Index Numbers", *Review of Economic Studies*, **3**, 155-157.

[65] Konüs, A. A. (1924), "The Problem of the True Index of the Cost of Living", reprinted in *Econometrica*, **7**, 10-29.

[66] Koo, A.Y.C. (1963),"An Empirical Test of Revealed Preference Theory", *Econometrica,* **31**, 646-664.

[67] Lancaster, K.J. (1966), "A New Approach to Consumer Theory", *Journal of Political Economy*, **74**, 132-157.

[68] Lerner, A.P. (1935-36), "A Note on the Theory of Price Index Numbers", *Review of Economic Studies*, **3**, 50-56.

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

[70] Lewbel, A. (1991), "The Rank of Demand Systems: Theory and Nonparametric Estimation", *Econometrica*, **59**, 711-730.

[71] Lewbel, A (1996), " Demand systems with and without errors: reconciling econometric, random utility and GARP models", mimeo, Brandeis University, September 1996.

[72] Malmquist, S. (1953), "Index Numbers and Indifference Surfaces", *Trabajos de Estatistica*, **4**, 209-242. cited in Diewert (1981).

[73] Manser, M.E. and McDonald, R.J. (1988), "An Analysis of Substitution Bias in Measuring Inflation, 1959-1985", *Econometrica,* **56**, 909-930.

[74] Marshall, A (1890), *The Principles of Economics*, London, 8th edition 1946.

[75] Mattie, A. (1994), "La Coherence des Choix des Cosommateurs Suisses", *Revue Suisse d'Economie Politique et de Statistique*, **130**, 3-20.

[76] McElroy, M.B. (1987), "Additive General Error Models for Production, Cost and Derived Demand or Share Equations" *Journal of Political Economy*, **95**, 737-757.

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

[78] Moulton, B.R. (1996), "Bias in the Consumer Price Index: What is the Evidence", *Journal of Economic Perspectives*, **10**, 159-177.

[79] Muellbauer, J. (1974), "Prices and Inequality: the United Kingdom Experience", *Economic Journal*, March, 32-55.

[80] Muellbauer, J. (1975), "The Cost of Living and Taste and Quality Change", *Journal of Economic Theory*, **10**, 269-283.

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

[82] Muellbauer, J. (1977), "The Cost of Living", in *Social Security Research*, London: HMSO.

[83] Piachaud, D. (1978), "Prices and the Distribution of Incomes", in Lord Diamond (Chairman), Royal Commission on the Distribution of Income and Wealth, Selected Evidence for Report 6, London: HMSO.

[84] Piachaud, D. and Webb, J. (1996) *The Price of Food: Missing out on Mass Consumption* LSE/STICERD., London.

[85] Pollak,R.A. (1971), "The Theory of the Cost of Living Index", Research Discussion Paper No. 11, Bureau of Labor Statistics, Washington, D.C.

[86] Popper, K. R. (1963), *Conjectures and Refutations*, Routledge.

[87] Prais, S.J. (1959), "Whose cost of living ?", *Review of Economics Studies*, **26**, 126-134.

[88] Prais, S.J. and Houthakker H.S. (1955), *The Analysis of Family Budgets*, Cambridge: Cambridge University Press.

[89] Reinsdorf, M. (1993), "The Effects of Outlet Price Differentials in the U.S. Consumer Price Index", in Foss, M.F., Manser, M.E. and A.H. Young, eds., *Price Measurements and their Uses*, Chicago: University of Chicago Press.

[90] Robinson, P.M. (1988), "Root n-consistent Semiparametric Regression", *Econometrica*, **56**, 931-954.

[91] Robinson, W. and Skinner, T. (1989), *Reforming the RPI: A Better Treatment of Housing Costs?*, Commentary no. 15, London: Institute for Fiscal Studies.

[92] Rothbarth, E. (1941), "The Measurement of Changes in Real Income under Conditions of Rationing", *Review of Economic Studies*, 100-107.

[93] Russell, R.R. (1992),"Remarks on the Power of Nonparametric tests of Consumer-Theory Hypotheses", in L. Phlips and L.D.Taylor (eds), *Aggregation, Consumption and Trade*, Dordrecht: Kluwer Academic Publishers, 121-135.

[94] Ryan, R.A. and Wales, T.J., (1996), "A Simple Method for Imposing Local Curvature in some Flexible Consumer Demand Systems", University of British Columbia, Discussion Paper, 96-25.

[95] Samuelson, P.A. (1947), *Foundations of Economic Analysis*, Cambridge, M.A., Havard University Press.

[96] Samuelson, P.A. (1948), "Consumption Theory in terms of Revealed Preference", *Economica*, **15**, 243-253.

[97] Samuelson, P.A. (1950), "The Problem of Integrability in Utility Theory", *Economica*, **17**, 355-385.

[98] Samuelson, P.A., and S. Swamy, (1974) "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis", *American Economic Review*, **64**, 566-593.

[99] Sen, A. (1977), "Starvation and exchange entitlements: A General approach and its application to the Great Bengal Famine", *Cambridge Journal of Economics*, **1**, 33-60, cited in Deaton and Muellbauer (1980).

[100] Sen, A. (1985), "The standard of living: lecture I - concepts and critiques", in G. Hawthorn (ed.), *The Standard of Living*, Cambridge: Cambridge University Press.

[101] Sen, A.K. (1985), *Commodities and Capabilities*, Amsterdam: North-Holland.

[102] Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.

[103] Sippel, R. (1997), "An Experiment on the Pure Theory of Consumer's Behaviour", *Economic Journal*, September.

[104] Stoker, T. (1986), 'The Distributional Welfare Effects of Rising Prices in the United States', *American Economic Review*, **76**, 335-49.

[105] Stone, J.R.N. (1956), *Quantity and Price Indexes in National Accounts*, Paris: Organisation for European Economic Co-operation.

[106] Tanner, S. (1996), "How Much do Consumers Spend ? - Comparing the FES and National Accounts", mimeo, Institute for Fiscal Studies.

[107] Tipping, D. G. (1970), "Price Changes and Income Distribution", *Applied Statistics*, vol. 19, pp. 1-17.

[108] Törnqvist, L. (1936), "The Bank of Finland's consumption price index", Bank of Finland Monthly Bulletin, **10**, 1-8.

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

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

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

[112] Vartia, Y. (1983), "Efficient Methods of Measuring Welfare Change and Compensated Income in terms of Ordinary Demand Functions", *Econometrica*, **51**, 79-98.

[113] Warshall, S., (1962), "A Theorem on Boolean Matrices", *Journal of the American Association of Computing Machinery*, **9**, 11-12.

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