TESTING THE RPI DATA FOR CONSISTENCY WITH THE THEORY OF THE COST-OF-LIVING INDEX

Ian Crawford
Isabella Image

THE INSTITUTE FOR FISCAL STUDIES
WP02/19

Testing the RPI data for consistency with the theory of the cost-of-living index

Ian Crawford* and Isabella Image**

* Institute for Fiscal Studies & University College London
** Office for National Statistics

April 2002

Abstract

This paper tests the published section level price and weight data used in the compilation of the UK Retail Prices Index for consistency with the theory of the cost-of-living index. We use a nonparametric test of theoretical consistency and bootstrap statistical methods to estimate the probability of consistency.

Acknowledgements: We are grateful to Adrian Ball, Laura Blow, Anna Brueton, David Fenwick and Ian Preston for helpful comments and advice. Financial support for Ian Crawford from the Leverhulme Trust via the Centre for Microdata Methods and Practice and the ESRC via the IFS Centre for the Microeconomic Analysis of Public Policy, and for Isabella Image from the ONS sanctioned by Adrian Ball and David Fenwick is gratefully acknowledged. Material from the FES made available by the ONS through the ESRC Data Archive has been used by permission of the controller of HMSO. Neither the ONS nor the ESRC Data Archive bear responsibility for the analysis or the interpretation of the data reported here. The views expressed here are those of the authors and not of the Office for National Statistics nor the IFS, which has no corporate view. The authors are responsible for all errors.

Summary

If the UK were to adopt a cost-of-living index formulation for its measure of inflation instead of the cost-of-goods index it presently uses, then one option would be simply to take the published price and weight data used to calculate the UK Retail Prices Index at present, and plug them into a cost-of-living index formulation. In this paper we aim to answer the question: are the published data on weights and prices which are used to compile the UK's measure of consumer price inflation consistent with the economic theory of the cost-of-living index applied to some representative UK consumer? We use a nonparametric test of theoretical consistency and bootstrap statistical methods to estimate the probability of consistency. This probability turns out to be quite high. This means that, in principle, the data are consistent with the theory and *could* be plugged into a cost-of-living index formula. We then discuss whether they *should*.

1 Introduction

One of the most important issues facing National Statistical Institutes is the conceptual basis that should be used for their published inflation rates. This is important for two reasons. Firstly, an underlying framework may be very useful when addressing practical issues such as the calculation of the entire index and, at a lower level, quality adjustment, introduction of new goods or sampling[1]. Secondly, the users of the index will have different needs; for example inflation as a macroeconomic indicator will not necessarily need to be calculated on the same basis as inflation used for price adjustment and taxation.

There are two main alternative conceptual frameworks: cost-of-goods indices (COGI's) and cost-of-living indices (COLI's). COGI's compare the relative costs of buying a fixed set of commodities under two different sets of prices, COLI's compare the minimum costs of purchasing a fixed level of economic welfare under two different sets of prices. Countries such as the UK publish a COGI; the UK's measure of consumer price inflation is the Retail Prices Index (RPI) which is defined as "an average measure of change in the prices of goods and services for the purpose of consumption by the vast majority of households in the UK"[2]. Eurostat regulations require all member countries to produce an inflation rate according to this conceptual basis[3]. The COLI approach is often argued (mostly by economists) to be superior to the COGI but it rests on the assumption that consumers behave as if they are making rational and consistent decisions about their purchases according to their preferences over commodities, and that these preferences conform to certain regularity conditions which allow them to be represented by a technical device known as a utility function. This assumption may not hold. As Sydney Afriat pointed out,

> "One feature which seems never to be noticed, but should nevertheless be recognized as outstanding, is that the basic hypothesis can be inadmissable on the data... There has been neglect of [this fact] for price indices generally." (Afriat (1977) p. 103)

If consumers are not rational and consistent in the fairly narrow sense which the basic economic theory requires, then the COLI framework is not suitable, at least not without modification[4]. Afriat, who did much to develop the methods which we will use to test data for consistency with the theory was rather sceptical about the role of economic theory in the measurement of consumer price indices:

> "Utility functions give service in theoretical discussions where they contribute structure which is an essential part of the matter. But the data used in practice cannot bear that structure."(Afriat (1977) p. 3)

In this paper we test whether the UK data can, or cannot, "bear that structure". If the UK were to adopt a COLI formulation for its measure of inflation

[1]See Triplett (2001) and Schultze and Mackie (eds.) (2002) for two recent discussions of these and other issues relating to the conceptual basis for price measurement.

[2]Baxter (ed) (1998) p.3.

[3]Report from the Commission to the Council on Harmonization of Consumer Price Indices in the European Union (COM(2000)742 final, 21.11.2000), p.19.

[4]See Triplett (2001) for a discussion of this issue.

then one option would be simply to take the published price and weight data used to calculate the UK RPI at present, and plug them into a COLI formulation. We adopt a simple nonparametric test for the consistency of these data with the economic theory behind the COLI and we apply this test to the published price and weight data used to calculate the UK RPI. We aim to answer the question: are the published data on weights and prices which are used to compile the UK's measure of consumer price inflation consistent with the theory of the COLI applied to some representative UK consumer?

The plan of the paper is as follows. In section 2 we review the theory of the cost-of-living index. Section 3 describes the test we will be using. Section 4 describes the data to which we apply the test. Section 5 presents the results and a discussion. Section 6 concludes.

2 Cost-of-living indices

The notion of the COLI is based upon the idea is that there may be a number of combinations of commodities which would all make the consumer equally well off (in economic terms), and that the consumer will choose the one (given the prices they face) which minimises the cost of reaching any given level of economic welfare. If they did not do this then their behaviour would be wasteful in the sense that they could improve their economic welfare simply and costlessly by changing their consumption patterns. Let $\mathbf{p}$ and $\mathbf{q}$ denote K-vectors of market prices and corresponding purchase quantities. The consumer is assumed to behave *as if*[5] their preferences are representable by a utility function $U(\mathbf{q})$ and they are solving the following problem.

$$C(\mathbf{p}, u) \equiv \min_{\mathbf{q} \geq \mathbf{0}} \{\mathbf{p}'\mathbf{q} : U(\mathbf{q}) \geq u\} \tag{1}$$

where the solution is given by the cost or expenditure function $C(\mathbf{p}, u)$ which describes the minimum cost of reaching an arbitrary utility level given an arbitrary set of prices. The key element of this model of consumer behaviour is the utility function. A utility function is an analytically convenient method used in classical economic theory to summarise a *rational* consumer's preferences over goods and services[6]. It is a continuous function $U : \mathbb{R}_+^K \rightarrow \mathbb{R}$ such that the commodity vector $\mathbf{q}$ is considered by the consumer to be at least as good as commodity vector $\mathbf{q}'$ if and only if $U(\mathbf{q}) \geq U(\mathbf{q}')$[7]. The solution to the consumer's cost minimisation problem (1) is completely determined by the

[5]It is important to note that the standard model of consumers behaviour is not designed for its literal, descriptive properties but its analytical properties. It is not necessary to believe that consumers actual *do* the following, just that their behaviour is consistent with it. However, if you want to give a "welfare" interpretation to the model then it is necessary to believe that consumers optimise something which matters to them, and that that something is economic well-being.

[6]The term "rational" is used here in a technical sense idiosyncratic to Economics. If preferences are rational in this sense then it means that they satisfy certain regularity conditions/assumptions: completeness, reflexivity and transitivity. Conversely if a consumers preferences conform to these regularity conditions (plus monotonicity and continuity) then they can be represented by a utility function. See Varian (1978) pp. 111-113.

[7]A utility function is a handy technical device nothing more. It not necessary to give it any any psychological interpretation. It is simply the thing which gets maximised, whatever that is. The particular utility number it produces from a given bundle of goods is not in itself

constraints they face (i.e. the prices of goods commodities and their available budget) and their utility function.

The 'true' cost of living index, also known as the Konüs index[8], is the minimum cost of achieving the standard of living indexed by the utility u under (say) period 1 prices $\mathbf{p}_1$, relative to the minimum cost of achieving that same standard of living under (say) period 0 prices $\mathbf{p}_0$. In other words this is expressed as the ratio of cost functions

$$P_K = \frac{C(\mathbf{p}_1, u)}{C(\mathbf{p}_0, u)}$$

One of the key differences between the COGI and the COLI, therefore, is that the COLI is built upon behavioural assumptions about consumers. Imagine that we have T observations on price and quantity data for a list of commodities $\{\mathbf{p}_t, \mathbf{q}_t\}$ for $t = 1, ..., T$. If the theory is right then at each observation the quantity data will have been generated by a consumer behaving *as if* they were solving a problem like (1). In the next section we describe how this hypothesis can be tested without making any assumptions as to the existence or functional form of the utility functions involved nor about their associated cost functions; the only necessary data are the prices and quantities involved (or alternatively the corresponding price and quantity indices). If the data show that the data generating consumer does not have a consistent utility function, then this would suggest that the COLI is not a suitable conceptual basis for a CPI. Conversely, if the data show that a utility function may be constructed then this would suggest that the theoretical basis for the COLI is supported by actual behaviour.

3 Testing the theory

Suppose we have T observations on K-vectors of price and quantities $\{\mathbf{p}_t, \mathbf{q}_t\}$ for $t = 1, ..., T$. Varian (1982) shows how to test these data for consistency with the cost-minimisation model. Following Varian (1983) we set out the following definitions

Definition 1 *A utility function $U(\mathbf{q})$ rationalises the dataset $(\mathbf{p}_t, \mathbf{q}_t)$ $t = 1, ..., T$ if $U(\mathbf{q}_t) \geq U(\mathbf{q})$ for all $\mathbf{q}$ such that $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$, for $t = 1, ..., T$.*

Definition 2 *An observation $\mathbf{q}_t$ is directly revealed preferred to a bundle $\mathbf{q}$, written $\mathbf{q}_t R^0 \mathbf{q}$, if $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$. An observation $\mathbf{q}_t$ is revealed preferred to a bundle $\mathbf{q}$, written $\mathbf{q}_t R \mathbf{q}$, if there is some sequence of observations $(\mathbf{q}_j, \mathbf{q}_k, ..., \mathbf{q}_l)$ such that $\mathbf{q}_t R^0 \mathbf{q}_j$, $\mathbf{q}_j R^0 \mathbf{q}_k$, ..., $\mathbf{q}_l R^0 \mathbf{q}$. In this case we say R is the transitive closure of the relation R^0. An observation $\mathbf{q}_t$ is directly revealed strictly preferred to a bundle $\mathbf{q}$, written $\mathbf{q}_t P^0 \mathbf{q}$, if $\mathbf{p}_t'\mathbf{q}_t > \mathbf{p}_t'\mathbf{q}$.*

Definition 3 *Data 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 $\mathbf{q}_s P^0 \mathbf{q}_t$.*

significant. The only relevant feature is its ordinal character. Thus if $U(\mathbf{q})$ represents some preferences and $f : \mathbb{R} \rightarrow \mathbb{R}$ is strictly increasing then $f(U(\mathbf{q}))$ represents exactly the same preferences as $U(\mathbf{q})$ since $f(U(\mathbf{q})) \geq f(U(\mathbf{q}'))$ if and only if $U(\mathbf{q}) \geq U(\mathbf{q}')$.

[8] Konüs (1924).

It turns out that the GARP condition is an observable consequence of cost minimisation, and that it is also a sufficient condition for it. If a dataset is consistent with GARP then there exists a utility function which can rationalise it and it is also consistent with the cost-minimisation model of behaviour. This is summarised by the following theorem due to Afriat (1967), Diewert (1973) and Varian (1982).

Afriat's Theorem[9]**:** *The following statements are equivalent:*
(1) There exists a non-satiated utility function which rationalises the dataset $(\mathbf{p}_t, \mathbf{q}_t)$ $t = 1, ...T..$
(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.
(4) There exists a non-satiated, continuous concave, monotonic utility function which rationalises the data.

In the rest of this paper we ask whether the RPI section level data are rationalisable by a stable utility function. That is, do they look like they were generated by a representative UK consumer solving the cost minimisation problem in equation (1)? If they do then the theoretical basis of the COLI is supported and a COLI representation of inflation using these data is admissable. If they are not then there exists no utility function (or cost function) consistent with these data and hence that there exists no data-consistent COLI representation of inflation[10]. In order to carry out the test we will apply Afriat's Theorem to these data: to prove the existence of a data-consistent utility function it is sufficient to show that statement (3) holds, namely that these data satisfy the Generalised Axiom of Revealed Preference. We then apply tests of GARP to these data. This is computationally very straightforward. Following Varian (1982) we construct a $T \times T$ matrix m whose entry in the i'th row and j'th column is given by:

$$M_{ij} = \begin{cases} 1 & \text{if } \mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j \text{, that is } \mathbf{q}_i R^0 \mathbf{q}_j \\ 0 & \text{otherwise} \end{cases}$$

M is constructed directly from the data and summarises the R^0 relations in the dataset. We then use Warshall's algorithm[11] to compute the transitive closure of M giving the matrix Mt with $i - j$ element

$$Mt_{ij} = \begin{cases} 1 & \text{if } \mathbf{q}_i R \mathbf{q}_j \\ 0 & \text{otherwise} \end{cases}$$

[9]This Theorem with a different, but equivalent, version of condition (3) was proved by Afriat (1967). Diewert (1974) also provided a proof which omits condition (3). Varian (1982) provides a proof which exchanges Afriat's version of (3) for GARP.

[10]Note that the test establishes necessary and sufficient conditions for the existence of a cost-of-living representation of inflation for those commodities within the domain of the RPI. It does not consider other factors such as publicly provided goods. The test might therefore also be thought of a necessary condition for the existence of a sub cost-of-living index defined over the domain of the RPI and embedded within a more general set of factors. See Appendix B for a discussion of the testable implications for this type of structure.

[11]Warshall (1962).

We then look for violations of GARP using Mt by checking through this matrix for each $Mt_{ij} = 1$ and checking the condition $\mathbf{p}_j'\mathbf{q}_j > \mathbf{p}_j'\mathbf{q}_i$ (that is, $\mathbf{q}_j P^0 \mathbf{q}_i$). If we find any such instances there is a violation of GARP.

4 The Retail Prices Index data

We use the published section level price indices and weights for the Retail Prices Index and aim to test (nonparametrically) whether these data are consistent with their having been generated by a rational economic agent[12]. The published section indices are essentially sub-price indices for groups of similar commodities (bread, cereals, biscuits and cakes,....,UK holidays). There are, at present, 77 of these published each month and they represent the lowest degree of aggregation at which these sort of data are published in the UK. Over the period we are considering here (January 1975 to December 2000) there have been a number of changes in classifications of these sections and therefore, in order to get consistent definitions over time we have had to aggregate the section data a little in some cases[13]. This gives us 69 sections/commodity groups in all. The complete list can be found in Appendix A. It should be noted that the composition of the sections is not continuous but rather changes over time in order to keep the basket of goods in RPI up-to-date. Changes in the composition of the sections (perhaps due to changes in preferences) and consequent quality change ought to be *prima facie* causes of violations of GARP.

Our starting point is the published monthly price indices for each of these commodity groups and corresponding annual weight data. We then need to get these data into a form which is amenable to testing GARP, that is a conformable set of corresponding price and quantity vectors. In what follows, p_t^i and q_t^i will represent these section level indices rather than the actual prices or physical quantities themselves. Note that this does not affect our test. We now discuss how we derive the required price and quantity information from these data.

The published section weight for section i in year t is defined as[14]

$$w_t^i = p_{Jan,t}^i Q_{b,t}^i \tag{2}$$

where $p_{Jan,t}^i$ is the section price index for in January of year t, and $Q_{b,t}^i$ is the "quantity of the i'th item purchased in the base period"[15] corresponding to year t (i.e. July of year $t-2$ to June of year $t-1$). This quantity is not literally a quantity (dozens of eggs, loaves of bread etc.) but rather a normalised quantity index. The normalisation is such that the published weight is expressed in parts-per-thousand, i.e. $\sum_i w_t^i = 1000$. Therefore if we denote the non-normalised quantity index for the purchases made in the base period

[12]Our data source is CSO (1991) plus a number of issues of *The Economy, Consumer Price Indices* MM23 (formerly *Consumer Price Indices, Business Monitor* MM23) which is a National Statistics publication published by HM Stationery Office for ONS.

[13]For example, there was a single combined section index and weight published for poultry and other meat until 1987 when they were split. To get consistent data we have had to recombine the separate sections after 1987. We do this using the RPI methodology for aggregating section indices. See Baxter (ed) (1998) pp. 68-70.

[14]Baxter (ed) (1998), p. 12.

[15]Baxter (ed) (1998), p. 12.

by $q^i_{b,t}$ then the published weight can be written as

$$w^i_t = \frac{p^i_{Jan,t} q^i_{b,t}}{\sum_j p^j_{Jan,t} q^j_{b,t}} \times 1000 \tag{3}$$

where the normalisation of the quantity indices is given by

$$Q^i_{b,t} = \frac{q^i_{b,t}}{\sum_j p^j_{Jan,t} q^j_{b,t}} \times 1000$$

As a result all the normalised quantities are equal to the non-normalised demands scaled by a common factor within each year, and demands in different base periods are scaled differently such that $\mathbf{p}'_{Jan,t}\mathbf{Q}_{b,t} = \mathbf{p}'_{Jan,s}\mathbf{Q}_{b,s} = 1000$ for all t, s. This normalisation essentially scales the non-normalised demands along rays going through the origin in q-vector space and so may affect the GARP tests. As a result even though the normalised quantities are proportional to the non-normalised quantities we are interested in recovering the non-normalised quantities. These are given by

$$q^i_{b,t} = \frac{Q^i_{b,t} \sum_j p^j_{Jan,t} q^j_{b,t}}{1000} = \frac{w^i_{b,t}}{p^i_{Jan,t}} \left[\sum_j x^j_{b,Jan,t} \right] \frac{1}{1000} \tag{4}$$

and to calculate them we need to know $\sum_j x^j_{b,Jan,t}$, total expenditure in the base period expressed in January year t prices. We calculate this from Family Expenditure Survey (FES) data corresponding to each base period.

The FES is one of the major sources of information for weights in the RPI. To get an estimate of total expenditure in the base period at January year t prices we use the same methodology which is employed when the FES expenditure data is used to calculate sectional weights[16]. Denote the FES based estimate of average nominal weekly spending on section i during the base period by $x^i_{b,t}$. This is then uprated to January of year t prices by dividing by the price in the January of the previous year (i.e. the mid-point of the base period) and multiplying by the price in the January of year t :

$$x^i_{b,Jan,t} = x^i_{b,t} \frac{p^i_{Jan,t}}{p^i_{Jan,t-1}}$$

where $x^i_{b,Jan,t}$ is now base period spending on section i expressed in January of year t prices[17] i.e.$x^i_{b,Jan,t} = p^i_{Jan,t} q^i_{b,t}$. Summing over sections gives the cost of

[16]Baxter (ed.) (1998) p. 39. Note that some section weights are calculated from three years of FES data, and other are calculated from alternative sources when the FES is thought to underestimate expenditure on the relevant commodities. Other weights such as council tax, mortgage interest payments and depreciation are calculated using models (Baxter (ed) (1998)). For these reasons the quantities which correspond to the published weights are not the same as those which we would recover directly from the FES section expenditures. As a result, because we are interested in testing the published data, we do not use the FES to estimate individual section weights but only total expenditure. See the discussion below.

[17]Note that this step in the RPI methodology is an approximation since it implicity assumes that $x^i_{b,t} = p^i_{jan,t-1} q^i_{b,t}$, that is, that the mid-point of the base period serves as a good approximation to the average prices over the base period. A better approximation may be to

purchasing the base period set of demands at January year t prices[18]. From (4) this gives us the following way of recovering the non-normalised section quantity indices

$$q_{b,t}^{i} = \frac{w_{b,t}^{i}}{p_{Jan,t}^{i}} \left[\sum_{j} x_{b,t}^{j} \frac{p_{Jan,t}^{j}}{p_{Jan,t-1}^{j}} \right] \frac{1}{1000}$$

where the term in square brackets is an FES based estimate of average base period weekly spending in January of year t prices. This gives us 24 observations on the quantity demand vector[19]. For certain commodities like spending on alcohol and tobacco the FES is known to under-record household expenditure (and consequently RPI section weights for these commodities are usually derived from more reliable alternative sources). As a result we would expect an estimate of total expenditure based on the FES to be somewhat lower than the true figure. Within a base period this has no relevant effect on our estimates of non-normalised quantities, it just scales them all down by the same proportional amount. It will have an effect on our GARP tests, however, if the extent of the underestimate varies much between years. Luckily this appears not to be the case as the relationship between FES-based measures of total spending and National Accounts measures is stable over this period[20]. Therefore, whilst we could gross up the FES based estimate, it is unnecessary.

We now turn to the calculation of the corresponding prices. The price data used in a test of GARP must be contemporaneous with the quantity data. That is to say the price data used to test GARP must refer to the prices which the data-generating agent faced when they chose each of the observed set of quantity vectors. If the demands in question are the average demands for some period, then the relevant set of prices are the average prices which obtained in the corresponding period. We therefore construct the price vectors by averaging the monthly section prices over the base periods corresponding to the quantity/weight data.

---

uprate each month ($m \in b$) in the base period individually

$$x_{b,Jan,t}^{i} = \frac{1}{12} \sum_{m \in b,t} \frac{p_{Jan,t}^{i}}{p_{m \in b,t}^{i}} x_{m \in b,t}^{i}$$

where $x_{m \in b,t}^{i}$ is average expenditure in month m of the base period and $p_{m \in b,t}^{i}$ is the correp-sonding price index for that month.

[18]Note that the FES based section weights are then finally defined in terms of these up-rated section expenditures by normalising them into (rounded) parts-per-thousand

$$w_{t}^{i} = \frac{x_{b,Jan,t}^{i}}{\sum_{j} x_{b,Jan,t}^{j}} \times 1000 = \frac{p_{Jan,t}^{i} q_{b,t}^{i}}{\sum_{j} p_{Jan,t}^{j} q_{b,t}^{j}} \times 1000 \tag{5}$$

which is identical to the weight defined above in (3).

[19]We have organised the FES into commodity groups corresponding to our RPI section weight data from 1975 to 1999. The first set of weights we have which overlap with these FES data are those for 1977 which refer to the July 1975 to June 1976 base period, the last are the weights for 2000 which correspond to the base period July 1998 to June 1999.

[20]See Banks and Johnson (eds) (1998).

5 Results

To recap our data and approach; we have weight and price data for 69 consistently defined section level (and occasionally slightly above section level) commodity groups for 24 periods. We calculate the implied (non-normalised) quantities using estimates of total expenditure from the FES. Although measures of the sampling variation associated with the section weights are not published we would still like to get some idea of the robustness of our results. Therefore, treating the RPI section weights as known constants, in our procedure sampling variation is present in the FES-based estimates of total spending in each RPI base period. We use bootstrap resampling methods applied to the FES household level data within each base period (10,000 resamples of N from N with replacement), to estimate the probability of our data passing GARP. The results are in Table 1 below

Table 1: Estimated probabilities.

$P(GARP)$	$P(not\ GARP)$	Standard error
0.8704	0.1296	0.00336

We find that the probability of accepting GARP on the population from which this sample was drawn is high; 87%. Of course bootstrap procedures do not help with the choice of critical values but we take this to indicate that the basic result which we find, which is that if we carry out a straight test of GARP on these data there are no violations, is robust to sampling variation in the estimate of total spending in the FES. The RPI price and weight data, at this level of aggregation, are consistent with the theory of the cost-of-living index: there exists a stable utility function which can rationalise these data, and these data are consistent with their having been generated by a representative UK consumer solving the cost-minimisation problem described in the theory. We now turn to a general discussion of these results and their implications

Firstly even though this is the first time we are aware of this sort of exercise being carried out on a country's CPI data should we be surprised at the results? The central issue is whether this type of test has much power to reject GARP. There has been relatively little work on this[21] but what there has been (which has essentially taken the form of Monte Carlo studies rather than analytic results) indicates that the power to reject, against a range of alternatives ranging from random behaviour to "close" to rational behaviour, depends on: *i)* K (the number of commodity groups) relative to T (the number of observations), and *ii)* the balance between relative price changes and variation in total expenditure levels between observations. In our data we disaggregate as much as possible ($K = 69$ is the maximum number of commodity groups we can define consistently) and T is fixed at 24 by the availability of the necessary data. Clearly GARP has only been tested at the section level. The sections are specifically constructed to be representative of different parts of a consumer's budget but it is possible that violations of GARP exist at the disaggregated level, but it

[21] See Bronars (1987) and Blundell *et al* (2002) for details on this issue.

is not possible to test this with the published data[22]. We could aggregate our data further (perhaps to the group level: Food, Catering,..., Leisure Services) and test at this level too. We may be more likely to find a violation of GARP. See Appendix B for a discussion of the extra structure on the consumer's utility function which would be implied by this type of exercise. The second issue is beyond our control. Consumers' average income increases over time, and usually at a rate faster than prices. As income goes up, so does expenditure and the consumer will be able to afford better bundles of goods over time. Ultimately this could mean that consumer choices observed later on in the data were not previously affordable and neither was a similar level of expenditure attainable. As a result it would be difficult to identify whether consumers were behaving inconsistently with regard to their preferences even if they were. Whether or not this is going on in these data is an open question.

What about the implications of this test? The data appear to be able to bear the weight of the theory at this level of aggregation. So what? Our test establishes that the UK *could* use these data to adopt a COLI definition of its measure of the consumer price index, but *should* it?

It is certainly the case that a COGI can mismeasure the change in the cost of living (the classic example of this is to do with the tendency for consumers to substitute away from more expensive goods as prices change whilst no account is taken of this commodity substitution under COGI frameworks) but with frequent rebasing and chaining as carried out in the UK this problem may not be so serious. Proponents of COLI point out that its main benefit is that it provides a useful conceptual framework when addressing many practical issues that will be faced by National Statistical Institutes in the actual calculation of their index such as quality change and the incorporation of new goods[23]. In these and many other areas the theory behind the COLI approach does "contribute structure which is an essential part of the matter", to use Afriat's phrase[24].

However, in practice it is impossible to construct a plausible index that is fully consistent with cost of living theory and to publish the index in a timely manner. Firstly, the practical production of a COLI would probably use either a Fisher or Törnqvist index number formula and these both use current period weights which are only available after a time-lag which would not be appropriate for the RPI's monthly schedule of publication. Only a Laspeyres-type index can be produced in a timely manner and this is one of the main reasons why the UK uses such an index. Secondly, while the effects of substitution under a COLI can be allowed for, this is only within the domain covered by the CPI (namely market goods and services purchased by the average consumer); the effects of other factors on consumers' welfare, such as crime or pollution, cannot be easily be included in the calculation even of a COLI. Further, opponents of the COLI notion point out that there are other uses of a CPI beyond price adjustment and that in these contexts a COLI may not be the correct conceptual notion to adopt. Its uses in the calculation of the GDP expenditure and as a measure of the change in the purchasing power of the currency are not so closely related to the behaviour of individual consumers; for these, a COGI may be more suitable.

[22]It would be possible to check that data passes GARP at the individual item level with the necessary (unpublished) data.

[23]See, for example, Gorman (1956).

[24]See the discussion in Triplett (2001) for further examples.

6 Conclusions

In this paper we have tested whether the assumptions behind the theory of the COLI are consistent with the data used to compile the RPI. The results indicate that they are and hence that the COLI approach, applied to these data, *could* be adopted as a basis for inflation measurement in the UK. Our results do not say that it *should*. Far from it; they merely show that there is no reason to reject the COLI framework on purely empirical grounds. The debate as to whether the cost of living framework meets the needs of the users of a country's CPI, and as to whether it is a practical means of calculating that CPI, will continue. For the meantime, the practical advantages, to do with timely publication, of maintaining the COGI approach to the publication of the monthly index probably outweigh the disadvantages. As Afriat says:

> "Practice can stand without theory; put in another way, it constitutes its own theory. All that is accomplished by expressing anything in practice in terms of utility theory is to show it does not conflict with that theory. Practice is consistent with a poor specialization of that theory, so poor that nothing is added by bringing it in." (Afriat (1977) p. 3).

References

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

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

[3] Banks, J. and P. Johnson (eds.), (1998), *How Reliable is the Family Expenditure Survey? Trends in Income and Expenditures over Time*, London: Institute for Fiscal Studies

[4] Baxter, M. (ed), (1998), *The Retail Prices Index, Technical Manual*, London: HMSO

[5] Blackorby, C., D. Primont, R. Russell, (1979), *Duality, Separability and Functional Structure: Theory and Applications*, Amsterdam: North Holland.

[6] Blundell, R.W., M. Browning and I.A. Crawford (2002), "Nonparametric Engel curves and revealed preference", forthcoming *Econometrica*

[7] Boskin *et al* (1996), "Toward a more accurate measure of the cost of living", *Final Report to the Senate Finance Commission from the Advisory Commission to Study the Consumer Price Index.* Washington DC.

[8] Bronars, S.G. (1987),"The Power of Nonparametric Tests of Preference Maximisation", *Econometrica*, **55**, 693-698.

[9] CSO, (1991), *Retail Prices 1914-1990*, London, HMSO

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

[11] 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).

[12] Konüs, A.A. (1924), "The problem of the true index of the cost of living", reprinted in *Econometrica*, **7**, pp 10-29.

[13] Schultze, C.L. and C. Mackie (eds), (2002) *At What Price: Conceptalising and Measuring Cost-of-Living and Price Indexes*, Washington DC: National Academy Press

[14] Triplett, J. "Should the cost-of-living index provide the conceptual framework for a consumer price index?", *Economic Journal*, vol 111, no. 472, pp. 311-334.

[15] Varian, H. (1978), *Microeconomic Analysis*, 2nd Edition (1984), New York: Norton

[16] Varian, H. (1982), "The non-parametric approach to demand analysis", *Econometrica*, **50**, 945-973.

[17] Varian, H. (1983), "Non-parametric tests of consumer behaviour", *Review of Economic Studies*, **L**, 99-110.

[18] Warshall, S., (1962), "A Theorem on Boolean Matrices", *Journal of the American Association of Computing Machinery*, **9**, 11-12.

Appendices

A Commodity group definitions

1: Bread, 2: Cereals, 3: Biscuits & cakes, 4: Beef, 5: Lamb, 6: Pork, 7: Bacon, 8: Poultry & other meat (until 1987), Poultry combined with Other meat (thereafter), 9: Fish, 10: Butter, 11: Oils & fats, 12: Cheese, 13: Eggs, 14: Milk, fresh, 15: Milk Products, 16: Tea, 17: Coffee & other hot drinks, 18: Soft drinks, 19: Sugar & preserves, 20: Sweets & chocolates, 21: Potatoes, 22: Other vegetables, 23: Fruit, 24: Other foods (excl. petfood), 25: Canteen meals, 26: Other meals & snacks (until 1980), Restaurant meals combined with Takeaway meals & snacks (thereafter), 27: Beer, 28: Wines and spirits, 29: Cigarettes, 30: Other tobacco, 31: Rent, Dwelling insurance & ground rent, 32: Mortgage interest payments, 33: Rates and water charges, Rates (& community charge from 1989 onwards), Water charges, 34: Repair/maintenance charges, 35: Do-it-yourself charges, 36: Coal & solid fuels, 3:7 Electricity, 38: Gas, 39: Oil & other fuel, 40: Furniture, 41: Furnishings, 42: Electrical appliances, 43: Other household equipment, 44: Household consumables, 45: Pet care, 46: Postage, 47: Telephones charges etc., 48: Domestic services, 49: Fees & subscriptions, 50: Men's outerwear, 51: Women's outerwear, 52: Children's outerwear, 53: Other clothing, 54: Footwear, 55: Chemist's goods, 56: Personal services, 57: Purchase of motor vehicles, 58: Maintenance of motor vehicles, 59: Petrol & oil, 60: Vehicle tax & insurance, 61: Rail fares, 62: Bus & coach fares, 63: Other travel costs, 64: Audio-visual equipment, 65: Personal articles, Records, toys, photo & sports goods & personal articles, Records & tapes (inc CD's from 1995 onwards), Toys, photo & sports goods, 66: Books & newspapers, 67: Gardening products, 68: TV licences & rentals, 69: Entertainment & recreation, Foreign holidays, UK holidays.

B Weak functional separability

We discuss the criteria necessary to test for weak separability and the existence of rational subgroups of commodities[25]. Weak separability is a restriction on the functional structure of the utility function which has a number of desirable properties. Chief amongst these, for our purposes, is the ability to derive sub-cost-of-living indices for sub-groups of goods. Let us partition the set of goods $\mathbf{q}$ into J subgroups $\mathbf{q}^1$, $\mathbf{q}^2$, ..., $\mathbf{q}^J$ and let us define the macro-utility function for the whole dataset $U(\mathbf{q})$ as

$$U(\mathbf{q}) = U(V^1(\mathbf{q}^1), V^2(\mathbf{q}^2), ..., V^J(\mathbf{q}^J))$$

where the $V^i(\mathbf{q}^i)$ are the sub-utility functions for each subgroup of commodities. Think of the subgroups as, for example, the RPI group level data (Food, Catering,..., Leisure Services), and the sub-utility functions as describing the consumers' preferences for commodities within these groups independently of the consumption level of other commodities in other goods. We therefore may derive the sub-cost functions and the Konüs sub-indices to be

$$\frac{C^1(\mathbf{p}_1^1, V^1)}{C^1(\mathbf{p}_0^1, V^1)}, \frac{C^2(\mathbf{p}_1^2, V^2)}{C^2(\mathbf{p}_0^2, V^2)}, ..., \frac{C^J(\mathbf{p}_1^J, V^J)}{C^J(\mathbf{p}_0^J, V^J)}$$

where sub-cost functions are defined by

$$C^i(\mathbf{p}^i, V^i) = \min_{\mathbf{q}^i} \mathbf{p}^{i\prime}\mathbf{q}^i \text{ such that } V^i(\mathbf{q}^i) \geq \overline{v}^i$$

In exactly the same way that demonstration of GARP is equivalent to the statement that the utility function exists, we may see that demonstration of a stronger version of GARP is equivalent to the statement that the utility function is weakly separable. This proof for this theorem is given in Varian (1983):

Theorem: *The following statements are equivalent:*

(1) There exists a weakly separable, concave, monotonic continuous non-satiated utility function that rationalises the data;

[25] For a more detailed discussion see Blackorby, Primont and Russel (1979) and Varian (1983).

(2) There exist numbers $V_s^i, V_t^i, \mu_t^i \geq 0$*, for time periods* $s, t = 0, ..., T$ *such that*

$$\begin{array}{rcl} V_s^1 & \leq & V_t^1 + \mu_t^1 \mathbf{p}_t^{1\prime} \left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right) \\ V_s^2 & \leq & V_t^2 + \mu_t^2 \mathbf{p}_t^{2\prime} \left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right) \\ & & \cdots \\ V_s^J & \leq & V_t^J + \mu_t^J \mathbf{p}_t^{J\prime} \left(\mathbf{q}_s^J - \mathbf{q}_t^J\right) \end{array}$$

and further there exist $U_s, U_t, \lambda_t > 0$ *such that*

$$U_s \leq U_t + \lambda_t \left[\mathbf{M}_t\right]' \left[\mathbf{V}_s - \mathbf{V}_t\right]$$

where $\mathbf{M}_t$ *is the vector whose i'th element is* $\left(\frac{1}{\mu_t^i}\right)$ *and* $\mathbf{V}_s, \mathbf{V}_t$ *are the J-vectors whose i'th elements are* V_s^i *or* V_t^i*.*
(3) The data for each group $(\mathbf{p}^i, \mathbf{q}^i)$ *and for the whole dataset* $(\mathbf{M}_t, \mathbf{V}_t)$ *satisfy GARP for some choice of* (V_t^i, μ_t^i) *that satisfy the Afriat inequalities.*

The V_s^i are known as the Afriat numbers and the inequalities they satisfy in condition (2) are known as the Afriat inequalities. They may be interpreted as equivalent to quantities, and the reciprocals of their associated μ_t^i may be interpreted as prices. The dataset $(\mathbf{M}_t, \mathbf{V}_t)$ in statement (3) thus corresponds to the whole dataset of prices and quantities $(\mathbf{p}, \mathbf{q})$. This Theorem corresponds exactly to Afriat's Theorem, but places the same restrictions on each of the subutility functions as on the macro-utility function for the whole dataset.

How, then, is this established in practice? In order to establish the weak separability of statement (1) we need only show that statement (3) holds (and conversely to disprove weak separability we need only to disprove statement (3)). The vectors for the whole dataset $(\mathbf{M}_t, \mathbf{V}_t)$ may be calculated by linear programming on the inequalities for the J individual groups[26]. These are then tested for GARP together with the price and quantities sets for the individual subgroups $(\mathbf{p}^i, \mathbf{q}^i)$.

The Theorem is presented in a form that allows all goods to be separated into J commodity groups, but it is also possible to check weak separability for individual groups; thus for example it is possible to check if the food group alone is separable from all other goods by finding only the relevant Afriat number in statement (2), and then checking GARP only for the full dataset and the one food subgroup $(\mathbf{p}^1, \mathbf{q}^1)$ in statement (3).

If the Afriat numbers can be found and GARP is satisfied, then we have shown that the dataset is consistent with weak separability. However, the converse does not hold. If there is an apparent violation of GARP in the $(\mathbf{M}_t, \mathbf{V}_t)$ dataset, then this may simply imply that we have chosen the wrong Afriat numbers. The inequalities in statement (2) only give bounds and there may be many numbers which lie within them, other than those returned by the algorithm. If GARP appears to be violated, then it may be that we just need to find the appropriate V_s^i values and it is then necessary to search for alternatives. This can be computationally very expensive. A test of GARP on the RPI group level data can be thought of as a separability test where the group level weights and price indices correspond to the data $(\mathbf{M}_t, \mathbf{V}_t)$. Violations of GARP may mean, not that the data are not weakly separable, but that the published weights and prices at this level are not the right Afriat numbers.

[26] Varian (1982) provides a suitable algorithm.