A NONPARAMETRIC BOUND ON SUBSTITUTION BIAS IN THE UK RETAIL PRICES INDEX

Laura Blow
Ian Crawford

THE INSTITUTE FOR FISCAL STUDIES
Working Paper Series No. W99/15

A nonparametric bound on substitution bias in the UK Retail Prices Index

Laura Blow[*] and Ian Crawford[y]

May 28, 1999

Abstract

This paper uses revealed preference restrictions and nonparametric statistical methods to bound true cost-of-living indices. These are compared to the popular price indices including the type used to calculated the UK RPI. This is used to assess the method of calculating the RPI for substitution bias.

Key Words: Cost-of-living indices, substitution bias, revealed preference.

JEL Classi…cation: C43, D11.

Acknowledgements: We are very grateful to Richard Blundell and Martin Browning for many helpful comments and for allowing us to draw upon jointly authored work. We are also grateful to J. Peter Neary, Zoe Smith and seminar particpants at University College Dublin, the Institute for Fiscal Studies and Southampton University. Financial support from the ESRC (award number: R000222725) and the Bank of England 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 Bank of England, the ESRC, ONS nor the ESRC Data Archive bear responsibility for the analysis or the interpretation of the data reported here. The authors are responsible for all errors.

[*] Institute for Fiscal Studies

[y] Institute for Fiscal Studies and University College London.

Address for correspondence: Institute for Fiscal Studies, 7 Ridgmount St., London, WC1E 7AE. laura_blow@ifs.org.uk or ian_crawford@ifs.org.uk

Summary

- This paper uses revealed preference restrictions and nonparametric statistical methods to bound the extent of substitution bias in the UK Retail Prices Index (RPI). In doing this we apply a method suggested by Blundell Browning and Crawford to bound true cost-of-living indices to the true index most closely approximated by the RPI formula.

- We show that, unlike a true Laspeyres price index, the direction of bias in the RPI formula is unknown a priori.

- We show that over the period considered (1976 to 1997) the RPI has overstated the true increase in the cost-of-living by as much as 3.5%

- We show that the RPI formula has generally but not always overstated the annual (January to January) rate of in‡ation over the period. The maximum absolute error was around 0.35 percentage points (in 1976) or 8.5% (in1993).

1. Introduction

Substitution bias is the systematic difference between a fixed-weight price index and a true economic cost-of-living index. A fixed-weight price index measures the proportional change in the cost of buying some fixed bundle of goods as prices change. A true economic cost-of-living index measures the proportional change in the minimum cost of maintaining some fixed level of economic welfare as prices change. Bias arises because commodity weights which are fixed for, say, one year at a time (as is the case with the RPI) cannot account for the possibility that spending patterns might adjust in response to relative price changes as consumers substitute during that period away from relatively expensive goods towards relatively cheaper ones (i.e. that consumers are cost-minimisers). The economic definition of a true cost-of-living index, and hence the notion of substitution bias, is therefore founded on the (in principle) testable assumption of optimising economic behaviour on the part of households. The aim of this paper is to estimate the extent of substitution bias in the UK Retail Prices Index (RPI) by comparing it to the true cost-of-living index which most closely corresponds to it.

It is important to remember that the RPI is not a true cost-of-living index, nor does it represent an attempt by the Office for National Statistics (ONS) to calculate or approximate one[1]. Why, then should we seek to compare the RPI with a true index? Afriat (1977) for one has strongly questioned whether, in most practical situations, the economic theory of the cost-of-living index and its associated paraphernalia of utility functions and cost functions really contributes very much of any potential use to statistical agencies like the ONS. He argues that

"Utility functions give service in theoretical discussions where they

[1] Baxter, M. (ed) (1997), p. 85.

contribute structure which is an essential part of the matter. But the data used in practice cannot support that structure. Practice can stand without theory"[2].

One highly practical alternative approach to the index number problem (the axiomatic approach) is not concerned with notions of cost-minimising economic behaviour at all; it is simply concerned with constructing an aggregate measure of price changes which possesses certain reasonable/desirable empirical properties[3].

Stone (1956) however, suggested three reasons why the utility-based economic approach to index numbers is useful in defining and constructing indices, and as a comparator for perhaps more pragmatic measures like the RPI.

> "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."[4]

Practicality is therefore a prime concern of much of the literature on economic index numbers; as Deaton (1981) puts it, index number theory "combines side-by-side some of the most difficult and abstruse theory with the most immediate practical issues of everyday measurement". We argue that, as a practical matter, very many of the uses to which the RPI is put require that it is interpreted either as a true cost-of-living index , or as a reasonable approximation to one (for

[2] Afriat (1977), p. 3.
[3] Diewert and Nakamura (1993) provide a survery of approaches to index number calculation.
[4] Stone, (1956), pp. 18-19.

example, up-rating of social security programmes, the measurement of real wages etc.). Consequently, its relationship to a true cost-of-living index is a matter both of interest and importance.

The plan of this paper is as follows. In section 2 we first review the formal definition of a true cost-of-living index. We then discuss the construction of the RPI and how it relates to a true index. In section 3 we briefly outline the main approaches to constructing true cost-of-living indices. The method which is adopted in calculating the true index is important. Some approaches may give results which are partly driven by their underlying assumptions regarding the form of the relationship between economic welfare and consumption patterns. We use a method of calculating bounds on a true index which is not subject to this criticism as it does not require the form of this relationship be specified. We propose to use the revealed preference method which is described in Varian (1982) and elsewhere. The cost of this approach is that it cannot provide a precise measure of the true index and we must therefore make do instead with bounds. The general idea is to place bounds on the set of consumption bundles which would yield the same level of welfare as the base bundle. Blundell, Browning and Crawford (1998) show that these bounds can be significantly improved by means of nonparametric expansion paths estimated from household-level data. Further, this procedure allows bounds to be developed for nonhomothetic preferences. We apply this approach in section 4 and report results which compare chained fixed-weight indices of the type used in the construction of the RPI with bounded true cost-of-living indices. Section 5 concludes.

2. Cost-of-living indices and the RPI

The economic approach to the measurement of changes in the cost-of-living is due to Konüs (1924). This defines a true index as the minimum cost of achieving some reference welfare level u_R when prices are p_t, relative to the minimum cost of achieving the same welfare with prices p_s. In notation this is written as

$$\frac{c(p_t; u_R)}{c(p_s; u_R)} \tag{2.1}$$

where $c(p_t; u_R)$ is the consumer's cost or expenditure function evaluated at the period t price vector p_t and the reference welfare level u_R. The cost function[5] is central to the whole area of cost-of-living indices. It is defined by

$$c(p_t; u_R) \equiv \min_q \left\{ p_t' q : u(q) \geq u_R \right\}. \tag{2.2}$$

The reference welfare level can be chosen more or less arbitrarily, however u_s and u_t (where, for example, $u_t \equiv u(q_t)$) are popular and obvious choices since, by the assumption of cost minimising behaviour on the part of consumers, the cost functions evaluated at $(p_t; u_t)$ and $(p_s; u_s)$ are directly observable: $c(p_t; u_t) \equiv p_t' q_t$ and $c(p_s; u_s) \equiv p_s' q_s$. Note that as the cost function depends upon utility, comparing two indices with two different reference welfare levels (u_{R1} and u_{R2}) generally gives

$$\frac{c(p_t; u_{R1})}{c(p_s; u_{R1})} \neq \frac{c(p_t; u_{R2})}{c(p_s; u_{R2})}, \tag{2.3}$$

except under certain special circumstances in which these ratios are independent of utility — homothetic preferences (see, for example Deaton and Muellbauer (1980)).

The classic result on the substitution bias inherent in a fixed-weight price index was set out in Konüs's famous inequalities[6] concerning the Paasche and Laspeyres indices. For example, in the case of the Laspeyres index which measures the changing cost of buying the fixed-basket of good q_0 at the contemporaneous

[5] It is a continuous real value function which is homogeneous of degree one in prices, is increasing in u and nondecreasing in p, and increasing in at least one price, and is concave with respect to prices.

[6] Konüs (1924).

price vector p_0 with the cost as prices change to p_1

$$\frac{p_1'q_0}{p_0'q_0} \geq \frac{c(p_1; u_0)}{c(p_0; u_0)} \tag{2.4}$$

where $u_0 \equiv u(q_0)$. That is, the Laspeyres index (which holds the initial basket fixed) is always greater than or equal to the corresponding base-referenced true cost-of-living index (which holds the welfare associated with that basket fixed).

The intuition behind this result is straightforward. Assuming cost minimising behaviour, the fixed bundle q_0 was the cheapest way of reaching u_0 under the original set of prices p_0. But this is not necessarily the case once prices have changed to p_1. All that is required to prove this is the reflexivity of preferences, as this ensures that q_0 is at least as good as itself. One way in which to purchase the welfare $u_0(q_0)$ is to purchase q_0 itself, at a cost of $p_1'q_0$, and hence the minimum cost of purchasing u_0 at p_1 cannot exceed $p_1'q_0$. The result that the Laspeyres index overstates the true change in the cost-of-living follows immediately. We now turn to the construction of the RPI.

The January to January RPI is a weighted average of the relative prices of various goods across the two months. The weights chosen correspond to the mean proportion of total spending accounted for by spending on the good in question and are calculated using data principally from the Family Expenditure Survey (FES) but also from other sources like the National Food Survey.

Consider the monthly RPI for the year t based in January. If the weights were based on the spending patterns in January then the RPI would simply be a Laspeyres price index. In fact, the weights are calculated using the shares of goods in consumer expenditure from July in the year $t-2$ to June in year $t-1$, expressed in January of year t prices. For example, the weights used in the 1997 RPI are calculated using spending data from July 1995 to June 1996 (we denote the base period for the RPI in year t as bt) expressed in January

97 prices. Therefore the RPI in month m in year t based in January of year t (denoted $RPI_{m;t=Jan;t}$, where $m = Jan; Feb; \ldots; Dec; Jan$ as the RPI is extended into January of the following year in order to chain years together — see below) is given by

$$RPI_{m;t=Jan;t} = \frac{p'_{m;t}\overline{q}_{bt}}{p'_{Jan;t}\overline{q}_{bt}}$$

where $\overline{q}_{bt}$ is the average demand vector in the base period for the year t index: July in year $t - 2$ to June in year $t - 1$. Thus, while the RPI which compares January in one year to the January in the following year does indeed compare the cost of buying a certain bundle of goods across the two times, this bundle of goods was bought, on average, 12 months before the start of the index. The RPI is not, therefore, a Laspeyres index and the Konüs result on the direction of bias does not apply.

The within-year RPI could be interpreted as approximating the cost of achieving the level of utility enjoyed from consumption in the base period as prices change, i.e.

$$RPI_{m;t=Jan;t} = \frac{p'_{m;t}\overline{q}_{bt}}{p'_{Jan;t}\overline{q}_{bt}} \approx \frac{c(p_{m;t}; u(\overline{q}_{bt}))}{c(p_{Jan;t}; u(\overline{q}_{bt}))}$$

where the RPI holds the basket of goods $\overline{q}_{bt}$ fixed, and the corresponding true cost-of-living index holds the economic welfare $u(\overline{q}_{bt})$ fixed.

Whereas we know that a Laspeyres index will always overstate its corresponding true cost-of-living index the comparison of the RPI to its corresponding true index is ambiguous since both

$$p'_{m;t}\overline{q}_{bt} \geq c(p_{m;t}; u(\overline{q}_{bt})) \text{ and } p'_{Jan;t}\overline{q}_{bt} \geq c(p_{Jan;t}; u(\overline{q}_{bt}))$$

by an argument analogous to that behind the Konüs result above.

Thus, in the case of the RPI there is no immediate presumption of upward bias. This is because both the numerator and the denominator overstate the true cost of achieving $u(\overline{q}_{bt})$ as neither account for the possibility of substitution in the consumption bundle since the data for $\overline{q}_b$ was collected.

To take a somewhat extreme example, if relative prices were constant except for an increase in January from which they then return to their original (period b) levels, we would have

$$p^0_{m;t}\overline{q}_{bt} = c(p_{m;t}; u(\overline{q}_{bt})) \text{ and } p^0_{Jan;t}\overline{q}_{bt} \geq c(p_{Jan;t}; u(\overline{q}_{bt}))$$

and therefore

$$\frac{p^0_{m;t}\overline{q}_{bt}}{p^0_{Jan;t}\overline{q}_{bt}} \leq \frac{c(p_{m;t}; u(\overline{q}_{bt}))}{c(p_{Jan;t}; u(\overline{q}_{bt}))}$$

We now turn to comparisons between years. Between years, the RPI is a chained index[7], so, for example, the index for May 1990 taking January in 1988 as the base is

$$RPI_{May;90=Jan;88} = RPI_{May;90=Jan;90} \times \frac{RPI_{Jan;90=Jan;89}}{100} \times \frac{RPI_{Jan;89=Jan;88}}{100} \tag{2.5}$$

(recall that the basket is held fixed for 13 month to allow for this). That is, it is the product of within-year indices, each with a different reference bundle. Because the bundle of goods changes across the multiplied expressions, it is difficult to interpret this as an approximation to any true fixed welfare base cost-of-living index. The corresponding idea is to chain together true, annually rebased, cost-of-living indices,

$$\frac{c(p_{May;90}; u(\overline{q}_{b90}))}{c(p_{Jan;90}; u(\overline{q}_{b90}))} \times \frac{c(p_{Jan;90}; u(\overline{q}_{b89}))}{c(p_{Jan;89}; u(\overline{q}_{b89}))} \times \frac{c(p_{Jan;89}; u(\overline{q}_{b88}))}{c(p_{Jan;88}; u(\overline{q}_{b88}))}. \tag{2.6}$$

However, this does not have any particular economically meaningful interpretation unless preferences are homothetic[8] in which case the changing welfare base[9]

[7] The term is due to Fisher (1911).

[8] The cost function can be written in the form $c(p; u) = a(u)b(p)$.

[9] $\overline{q}_{b90}$ refers to the mean demands from July 88 to June 89 which are used as the base for the 1990 index. $\overline{q}_{b89}$ refers to the mean demands from July 87 to June 88 etc.

cancels out of each individual index giving

$$\frac{c(p_{May;90})}{c(p_{Jan;88})} \tag{2.7}$$

If we were willing to assume homotheticity, the resulting index would be independent of the base and so would be the true for any welfare level. If we are unwilling to assume homotheticity then it is not obvious exactly what such an index is a comparison of.

Chaining has two main merits, both of them practical. Firstly it allows for the introduction of new goods and the removal of obsolete ones from the price and weight data. Secondly, it reduces the differences between series derived from different index number formulations. Diewert (1978) argues for chaining indices on the basis that discrepancies between different index numbers (the Paasche and Laspeyres, for example) will be smaller between periods which are closer together rather than further apart. One problem with chaining other than the lack of an easy economic interpretation is that chained indices like the RPI fail Walsh's multiperiod identity test, namely that

$$\frac{p'_{Jan;89}\overline{q}_{b88}}{p'_{Jan;88}\overline{q}_{b88}} \times \frac{p'_{Jan;90}\overline{q}_{b89}}{p'_{Jan;89}\overline{q}_{b89}} \times \frac{p'_{Jan;91}\overline{q}_{b90}}{p'_{Jan;90}\overline{q}_{b90}} \times \frac{p'_{Jan;88}\overline{q}_{b91}}{p'_{Jan;91}\overline{q}_{b91}} \neq 1$$

Chained true indices with a moving base like the one corresponding to the RPI will also fail a similar test. We now turn to the way in which the true index might be calculated.

3. Calculating true cost-of-living indices

There are three main methods in the literature on calculating cost-of-living indices[10]. These are (in descending order of the strength of the assumptions necessary for them to be valid):

[10] For detailed discussions see Diewert and Nakamura (1993).

1. Direct estimation of the cost function by means of a demand system[11].

2. The calculation of a superlative price index[12].

3. The calculation of nonparametric bounds to the true index[13].

Direct estimation of the cost function requires that the cost function exists, that the functional form of the cost function is known, and that the parameters of the cost function can be estimated. This is computationally expensive and the maintained hypothesis on functional form may be hard to test.

The use of a superlative[14] index requires that the cost function exists and that the form of cost function is known or that it can be approximated closely by the function chosen for the index. It does not require the parameters themselves to be estimated. The problem with superlative indices, as far as an assessment of substitution bias is concerned, is that they are typically not based on the same welfare level as the true index which corresponds most closely to the RPI. The Törnqvist index for example, which is a non-homothetic superlative index, is based on the geometric mean of the welfare levels of the two periods being compared. As a result, comparisons between the RPI and a Törnqvist index would con‡ate substitution bias with di¤erences which were to do with their di¤erent reference welfare levels.

The use of nonparametric bounds does not require either any assumptions on the form of the cost function, or that the function be estimated. Its validity rests only the existence of a stable cost function over the period under consideration

[11] Examples of this approach are comparatively rare because of its computationally expensive nature. One example, however, is Braithwait (1980).

[12] For example, Diewert (1976).

[13] For example, Lerner (1935-36), Pollak (1971), Afriat (1977), Varian (1982) and Manser and McDonald (1988)

[14] A superlative index is one which is based upon a cost function which can provide a second order approximation to any well behaved cost function (see Deiwert (1976) and (1981)). The term, on this de…nition, is due to Diewert.

and this can be tested nonparametrically in terms of the Generalised Axiom of Revealed Preference[15],[16]. The benefit of revealed preference restrictions is that they do not presuppose the existence of utility functions/cost functions which might rationalise the data, so they can be used to test a dataset for consistency with a utility function/cost function with well behaved properties but an unspecified functional form.

For these reasons — the nondependence of the resulting index on functional form, and the ability to test and reject the economic approach to true indices itself — we propose to utilise a nonparametric/revealed preference method. This basic idea has been used in the past, for example by Varian (1982) and by Manser and McDonald (1988). However, the problem which these authors discovered was that these nonparametric bounds can be imprecise — the possible values for the true index can lie anywhere within a wide range. They tightened the bounds by making an assumption of homotheticity. We discuss the general approach below and set out an alternative method for tightening the bounds using nonparametric Engel curves proposed by Blundell, Browning and Crawford (1998). They show that this method can dramatically improve the bounds recoverable from revealed preference restrictions. Further, the approach also allows the bounds to be calculated for non-homothetic preferences — Banks, Blundell and Lewbel (1996) show that Engel curves are nonlinear for a range of commodities in the FES.

[15] Note: We adopt the following notation.

Definition 1. q_t is directly revealed preferred to q, written $q_t R^0 q$, if $p_t' q_t \geq p_t' q$.

Definition 2. q_t is directly revealed strictly preferred to q, written $q_t P^0 q$, if $p_t' q_t > p_t' q$

Definition 3. q_t is revealed preferred to q, written $q_t R q$, if $p_t' q_t \geq p_t' q_s$, $p_s' q_s \geq p_s' q_r, \ldots, p_m' q_m \geq p_m' q$, for some sequence of observations $(q_t, q_s, \ldots, q_m)$. In this case, we say that the relation R is the transitive closure of the relation R^0.

Definition 4. q_t is revealed strictly preferred to q, written $q_t P q$, if there exist observations q_s and q_m such that $q_t R q_s$; $q_s P^0 q_m$; $q_m R q$.

[16] Data can be said to satisfy the Generalised Axiom of Revealed Preference (GARP) if $q_t R q_s \Rightarrow p_s' q_s \leq p_s' q_t$. Equivalently, the data satisfy GARP if $q_t R q_s$ implies not $q_s P^0 q_t$.

3.1. Revealed Preference Bounds

Afriat (1977) described the way in which revealed preference information can be used to provide classical bounds on the welfare e¤ects of a price change. We begin with the simplest example. Figure 3.1 shows a single price-quantity observation $(p_0; q_0)$ with expenditure $x_0 = p_0' q_0$. Assume that this observation was generated by a rational consumer. The bounds on the possible position of the indi¤erence curve through that point are given by the shaded areas and these bounds are wide. The area revealed preferred to q_0 (denoted RP (q_0)) results simply from the monotonicity of utility. The area to which q_0 is revealed preferred is the shaded area below the x_0 budget line (denoted RW (q_0)). The indi¤erence curve cannot pass through the boundary of either set but it can lie anywhere in between (or even along) these extremes. The resulting bounds on the welfare e¤ects of the new price regime 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.

Figure 3.1: Revealed preference bounds, one observation

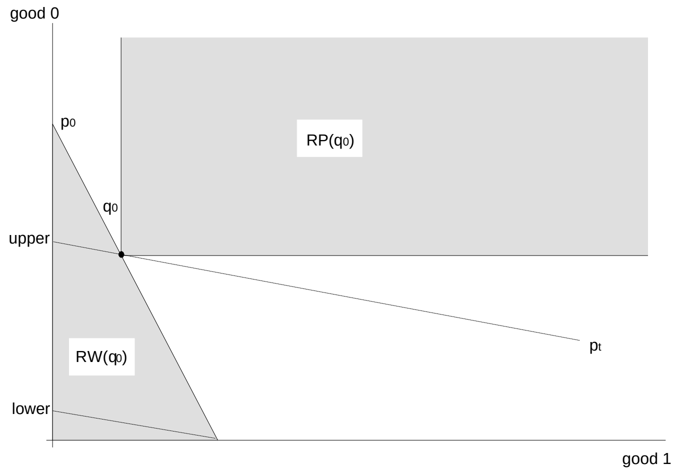
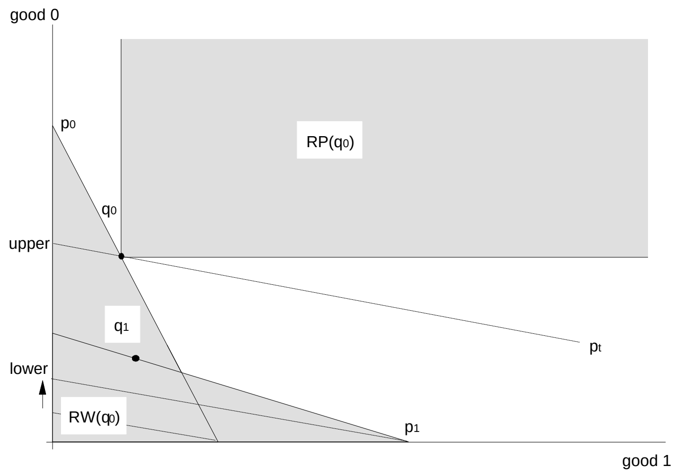


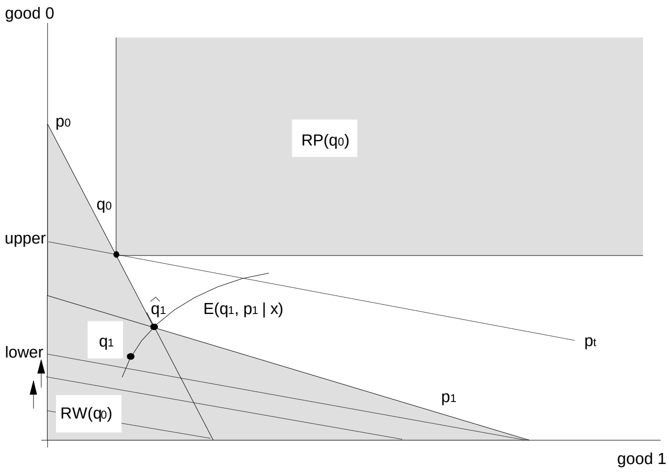
Figure 3.2 introduces a second price-quantity observation $(p_1; q_1)$ and the RP (q_0) and RW (q_0) sets are redrawn using this new information. In this example the RP (q_0) set is unchanged and the welfare e¤ects remain bounded from above by Leontief preferences, but new information has been gained on the lower bound on the indi¤erence curve giving evidence of some degree of curvature. The new lower bound uses both the x_0 and the $x_1 = p_1^{0} q_1$ budget lines. And, since $q_0 P^0 q_1$, both q_1 and points to which it is revealed preferred must lie below the indi¤erence curve. In other words, while the indi¤erence curve can run either along or above the $p_0^{0} q_0$ budget line from the good 0 axis to the point where it meets the $p_1^{0} q_1$ line, it must lie above the $p_1^{0} q_1$ budget line from that point to the good 1 axis. The extra information allows the lower bound to be tighter than was the case with the single observation.

Figure 3.2: Revealed preference bounds, two observations.



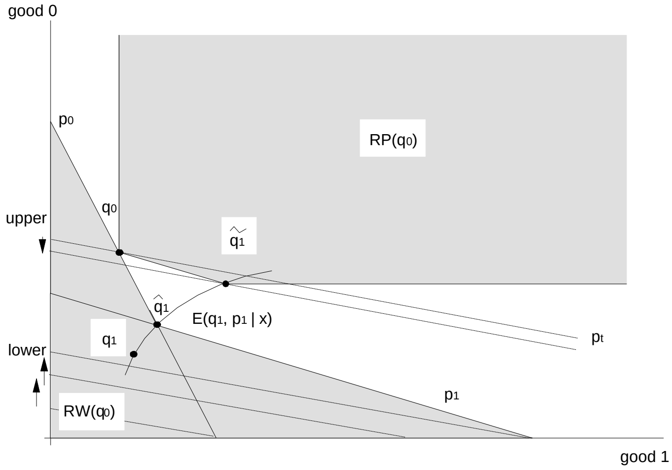
Blundell, Browning and Crawford (1998) propose the use of nonparametrically estimated expansion paths to improve the bounds which can be derived. Expansion paths allow the position of a bundle of goods to be varied by changing total expenditure whilst holding prices constant. In other words, by varying total expenditure, the budget surfaces can be placed as desired and greater precision can be generated. Expansion paths can be estimated using non-parametric regression techniques from the micro-data in the FES in which demands are observed to vary cross-sectionally with total expenditure within fixed price regimes. To illustrate their idea, consider Figure 3.3.

Figure 3.3: Improved lower bounds, two observations.



The expansion path for demands with prices p_1 is now added to the information available in Figure 3.2. This allows the budget line for prices p_1 to be moved out to a higher total expenditure level (since $p_1' \hat{q}_1 > p_1' q_1$). Setting total spending equal to $p_0' q_0$, and hence placing the budget line so that it lies on the budget surface which the base bundle q_0 is also on, tightens the previously available bound on the indifference curve since now $p_0' q_0 = p_0' \hat{q}_1$ and hence it remains the case that $q_0 R^0 \hat{q}_1$. A major additional benefit, and an additional 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 3.4. Here, the budget line using the p_1 price vector is placed at an expenditure level such that $p_1' \hat{q}_1 = p_1' q_0$ which implies $\tilde{q}_1 R^0 q_0$. By similar arguments as before, the indi¤erence curve cannot pass above $\tilde{q}_1$ or the plane connecting q_0 and $\tilde{q}_1$.

Figure 3.4: Improved upper and lower bounds, two observations.



Blundell, Browning and Crawford (1998) provide an algorithm for carrying out this bounding procedure using many periods/price regimes. They show that it also provides a test of GARP. For completeness we reproduce their algorithms and provide a proof that this procedure constitutes a test of GARP as appendices.

4. Empirical Application

4.1. Data

The indices calculated in this paper use information on price movements from the section indices of the retail price index for the period 1976 to 1997, and correspondingly grouped household expenditure data from the FES from July 1974 to June 1996. We are unable to look at substitution bias with sections as the price data required to do this are not publicly available. The original aim of the FES was to provide the basis of an average basket of goods to be used in the calculation of the U.K. RPI. The FES is an annual random cross section survey of around 7,000 households (this represents a response rate of around 70%). The FES 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 FES 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 FES 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 FES 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[17].

These problems may not be terribly serious, but there are other potential problems in the FES which might be more substantive. In particular, there may be problems of under or over-reporting of expenditures either through genuine forgetfulness, or active concealment, or a combination of forgetfulness and 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 FES

[17] Tanner (1996).

data are supplemented with data from other sources (clearances from bonded warehouses, for example) for use in national accounting and the RPI. 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 FES 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 FES means that large infrequent purchases (of durables for example) may be under estimated. Data on durables from the FES are bolstered by data from other sources in the computation of the RPI and the weights are computed as a moving average.

The degree to which data are aggregated across goods and services can affect the results in two ways. Firstly, a greater degree of substitution responses might be expected in disaggregated data (see Manser and McDonald (1988)). Secondly, un-warranted grouping of goods may cause rejections of GARP which relate to the separability structure imposed by the grouping (see Varian 1983). For this reason we use data which are disaggregated as far as possible. This amounts to 62 groups of goods and services. This is dictated by the level of disaggregation available in the published price data and the need to construct consistent groups of goods over time. For example, separate price indices are currently published for "poultry" and "other meat". However, these goods were grouped together in 1976 and so we have to maintain this grouping throughout. Details of the groups are given in the Appendix C. Finally, we only consider non-housing expenditure.

4.2. Econometric Considerations

We estimate the expansion paths we require by nonparametric smoothing across the cross section of households within each month/price regime. That is, within each month, prices are assumed constant across households and we use the cross-section variation in total expenditure to identify the expansion path. To be more explicit denote log expenditure for the i'th household by $\ln x_i$ and budget share for the i'th household by w_{ij} for the jth good. For each commodity j and each household i, we assume a Piglog structure

$$w_{ij} = f_j(\ln x_i) + \varepsilon_{ij}. \tag{4.1}$$

Since $\ln x$ is endogenous then $E(\varepsilon_{ij}|\ln x_i) \neq 0$ or $E(w_{ij}|\ln x_i) \neq f_j(\ln x_i)$ and the nonparametric estimator will not be consistent for the function of interest. To adjust for endogeneity in $\ln x$ we use the augmented regression technique in a semiparametric estimation framework due to Robinson (1988). We use log income ($\ln y$) as an instrumental variable such that

$$\ln x_i = \pi \ln y_i + v_i \tag{4.2}$$

with $E(v|\ln y) = 0$, and we assume that the following linear model holds

$$w_{ij} = f_j(\ln x_i) + v_i\rho_j + \varepsilon_{ij}. \tag{4.3}$$

We assume

$$E(\varepsilon_{ij}|v, \ln x) = 0 \text{ and } Var(\varepsilon_{ij}|v, \ln x) = \sigma_j^2(v, \ln x). \tag{4.4}$$

Following Robinson (1988), a simple transformation of the model can be used to give an estimator for the parameter ρ_j. Taking expectations of (4.3) conditional on $\ln x_i$, and subtracting from (4.3) yields

$$w_{ij} - E(w_{ij}|\ln x) = (v_i - E(v_i|\ln x))\rho_j + \varepsilon_{ij}. \tag{4.5}$$

Replacing $E(w_{ij}|\ln x)$ and $E(v_i|\ln x)$ by their nonparametric estimators, the parameter ρ_j can be estimated by ordinary least squares and is $\sqrt{n}$ consistent and asymptotically normal. The estimator for $f_j^h(\ln x)$ with bandwidth h is then

$$\hat{f}_j^h(\ln x) = E^h(w_{ij} - v_i \hat{\rho}_j | \ln x) \tag{4.6}$$

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

$$\hat{v} = \ln x - \hat{\pi} \ln y \tag{4.7}$$

where $\hat{\pi}$ is the least squares estimator of π. Since $\hat{\pi}$ and $\hat{\rho}$ converge at $\sqrt{n}$ the asymptotic distribution for $\hat{f}_h(\ln x)$ follows the distribution of $E^h(w_{ij} - v_i \rho_j | \ln x)$.

In our empirical application we use a Nadaraya-Watson kernel regression estimator of the jth share equation with bandwidth h,

$$\hat{f}_j^h(\ln x) = \frac{N^{-1} \sum_i K_h(\ln x - \ln x_l) w_{ij}}{N^{-1} \sum_i K_h(\ln x - \ln x_l)} \tag{4.8}$$

with sample size N, where $K_h(\cdot) = h^{-1} K(\cdot / h)$ is chosen to be a Gaussian kernel weight function $K(\cdot)$, and $\ln x_l$ is the l'th point in the $\ln x$ distribution at which we evaluate the kernel. Using the same bandwidth to estimate each $f_j^h(\ln x)$ guarantees adding up across equations.

To compute demand bundles at some given total expenditure level $(\ln \tilde{x})$ from these semiparametric Engel curves, we utilise our common price regime assumption (dropping the bandwidth)

$$E(q_j | \ln \tilde{x}; \pi_j) = \hat{f}_j(\ln \tilde{x}) \left(\frac{\tilde{x}}{\pi_j} \right).$$

4.3. Results

In this section we compare the series RPI* with various chained and un-chained true index bounds. RPI* does not equal the RPI because, as discussed above, of the differences in data sources and because the RPI weights households by shares out of the sum over all households of total expenditure. Our measure RPI* is calculated on FES data using equation 2.5 and we do not differentially weight households. Comparing our indices with the published non-housing RPI would con‡ate substitution bias with these other sources of di¤erence. Di¤erences between RPI* and the true indices to which we compare it are, therefore, purely a matter of substitution bias. We compare RPI* to the published non-housing RPI series in Appendix D.

Figure 4.1: Bounded true cost-of-living indices, 1976 to 1997, based 7/74-6/75 to 7/84-6/85

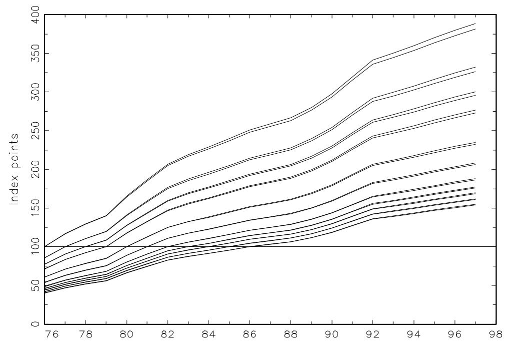
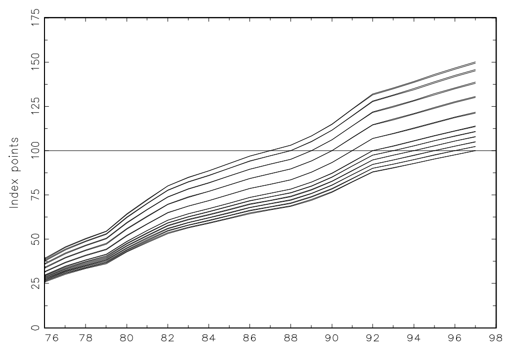


Figure 4.2: Bounded true cost-of-living indices, 1976 to 1997, based 7/85-6/86 to 7/95-6/96



Since our aim is to compare the RPI (equation 2.5) with the corresponding chained true index (equation 2.6), we begin by computing bounds on true cost-of-living indices for the period 1976 to 1997. In each index we move the base period by one year. Each period coincides with a link RPI in the chained RPI. For example, the bounded index based in 1976 uses the period July 1974 to June 1975 as the reference period and bounds

$$\frac{c\left(p_{(Jan;t)}, u\left(\bar{q}_{(July;75 - June;76)}\right)\right)}{c\left(p_{(Jan;77)}, u\left(\bar{q}_{(July;75 - June;76)}\right)\right)} \quad t = 76, 77, \ldots, 97.$$

The bounded index based in 1997 uses the period July 1995 to June 1996 as the reference period etc. These fixed-base bounded true indices form the basis of the true chained index defined in equation 2.6. Figures 4.1, and 4.2 show the GARP bounds on the annually rebased true indices for the period. Figure 4.1 shows bounds on the indices based during the periods July 1974 to June 1975, up until the index based in July 1984 to June 1985. Figure 4.2 shows the remaining indices.

4.3.1. The RPI overstates the increase in the cost-of-living

The index calculated by chaining together these true indices is the cumulative product of the annual changes from the base period of each index to the following period (as described in equation 2.6). This chained true index is shown in Table 4.1 The table also shows RPI*.

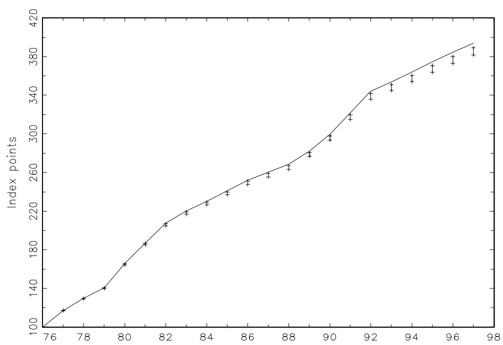
Figure 4.3 illustrates the bounds on the chained true index and RPI* over the period. The bounds at each point are shown by the crosses joined by a line. Figures 4.4 and 4.5 illustrate the differences between RPI* and the chained true index: figure 4.4 shows the absolute difference in index points, figure 4.5 shows these differences as percentages on the level of the index[18]. For example, figure 4.4 shows that by 1997 the absolute difference between the chained true index

[18] Note that in drawing these and subsequent figures we have not rounded the data.

Table 4.1: Cost-of-living indices, chained true index and RPI*, 1976-1997.

	Indices	
Year	GARP	RPI*
1976	100.0	100.0
1977	[116.9, 117.1]	117.2
1978	[128.9, 129.5]	129.6
1979	[139.8, 140.5]	140.7
1980	[164.1, 165.4]	165.6
1981	[184.9, 186.7]	187.1
1982	[204.7, 206.9]	207.5
1983	[216.8, 219.3]	220.1
1984	[226.4, 229.3]	230.2
1985	[236.9, 240.0]	241.1
1986	[247.7, 251.0]	252.2
1987	[255.4, 258.9]	260.4
1988	[263.2, 266.8]	268.4
1989	[276.5, 280.5]	282.2
1990	[293.6, 297.8]	299.6
1991	[314.8, 319.7]	321.8
1992	[336.1, 341.7]	344.3
1993	[344.9, 351.0]	354.0
1994	[354.1, 360.6]	364.1
1995	[363.7, 370.7]	374.6
1996	[372.9, 380.2]	384.5
1997	[381.6, 389.1]	393.6

Figure 4.3: RPI* and bounds on the chained true index, 1976 to 1997



RPI* is between 4.5 and 12 index points. Figure 4.5 shows that this is a di¤erence of between about 3.2% and 1.2% of the level of the index. This shows that RPI* has overestimated the increase in the cost-of-living (de…ned by equation 2.6) by between 1.2% and 3.2% by the end of the period.

Figure 4.4: Di¤erences between RPI* and the chained true index, 1976 to 1997, index points

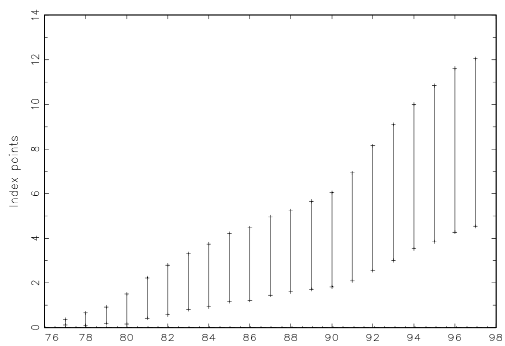
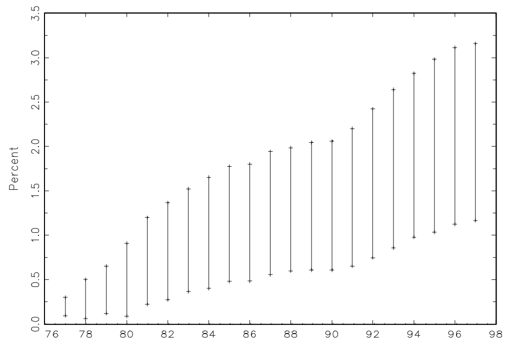


Figure 4.5: Percentage di¤erences between RPI* and the chained true index, 1976 to 1997,

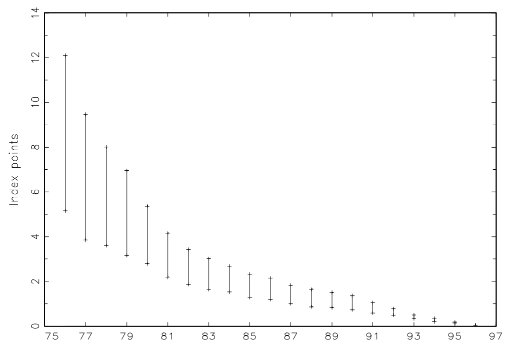


So far we have compared RPI* to the most closely corresponding true index, which happens to be annually re-based true indices chained together. We argued above that, unless we decide to assume that preferences are homothetic, it is hard to think of an economic question to which a chained index provides a sensible answer. Suppose we want to use RPI* to answer an economically meaningful question like how much has the cost of living increased between year t and 1997? Figures 4.6 and 4.7 show the differences between RPI* and the true indices comparing each year with 1997. The bounds for 1981, for example, show the difference between the cost-of-living increase from 1981 to 1997, measured by RPI* and the true cost-of-living index which compares these periods based on the average demand in for the 1981 base period (July 78 to June 79). That is

$$RPI_{Jan,98/Jan,81} - \frac{c(p_{jan,97}, u(\overline{q}_{b81}))}{c(p_{jan,81}, u(\overline{q}_{b81}))}.$$

This shows that RPI* overstates the true increase by between 2.2 and 4.1 index points, or as figure 4.7 indicates, between 1% and 2%.

Figure 4.6: Differences between true indices and RPI*, 1976 to 1997, index points

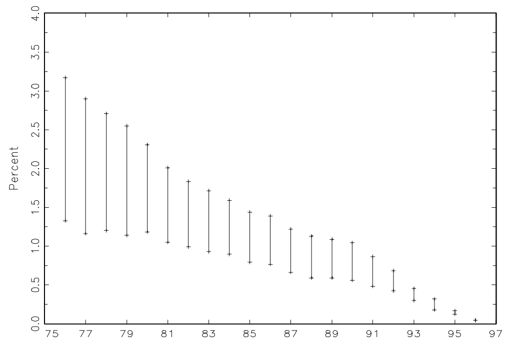


Suppose we wanted to use the RPI to answer the question: how much lower was the cost-of-living in year t than in 1997? Figure 4.8 and figure 4.9 show the differences between RPI* and the true index based in 1997, that is for, say 1981, the absolute difference in percentage points

$$RPI_{Jan,81/Jan97} - \frac{c(p_{jan,81}, u(\overline{q}_{b97}))}{c(p_{jan,97}, u(\overline{q}_{b97}))}$$

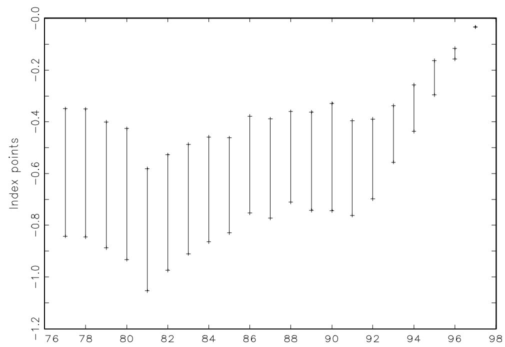
is the number in figure 4.8, whilst figure 4.9 reports these differences as percentages of the level. Just as the RPI overstates the increase in the cost of living since

Figure 4.7: Percentage di¤erences between true indices and RPI*, 1976 to 1997



1981, based in 1981; it understates the lower cost of living in 1981 based in 1997.

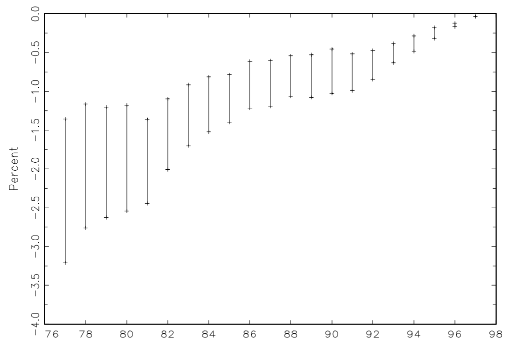
Figure 4.8: Di¤erences between the true index based in 1997 and RPI*, 1976 to 1997, index points



4.3.2. The RPI generally, but not always, overstates the true annual rate of in‡ation

Table 4.2 reports the in‡ation rate for RPI* and the chained true index. We follow the ONS's convention for calculating in‡ation (i.e. based upon the index which is already rounded to one decimal place and reported to one decimal place). In the …gures below we illustrate the underlying (unrounded data).

Figure 4.9: Percentage di¤erences between the true index based in 1997 and RPI*, 1976 to 1997



The di¤erences in in‡ation rates between RPI* and the chained true index is shown in …gures 4.10 and 4.11. Figure 4.10 shows that, for example, the absolute di¤erence in the in‡ation rate between RPI* and the true chained index for 1977 was between 0.1 and 0.35 percentage points. Figure 4.11 shows that since the true rate of in‡ation in 1977 was between 16.9% and 17.1%, this amounted to a 0.6% to 2.1% error in measuring the true rate. In general, in‡ation measured by RPI* has overstated in‡ation as measured by the chained true index. In percentage terms the greatest overstatement was in 1993 when the error was between 8.4% and 4.1%, whilst the percentage point di¤erence was between 0.22 and 0.11, and the true rate of in‡ation was low (between 2.6% and 2.7%). However, in three years, the rate of in‡ation measured by RPI* is within the bounds on the chained true index (1978, 1980 and 1990).

5. Conclusions

This paper has used revealed preference restrictions, improved by use of nonparametric statistical methods, to bound true cost-of-living indices which are then

Table 4.2: In‡ation rates, chained true index and RPI*, 1976-1997.

Year to	In‡ation Rates GARP	RPI*
1977	[16.9, 17.1]	17.2
1978	[10.3, 10.6]	10.6
1979	[8.4, 8.5]	8.5
1980	[17.4, 17.7]	17.7
1981	[12.7, 12.8]	13.0
1982	[10.7, 10.8]	10.9
1983	[5.9 ,6.0]	6.1
1984	[4.4, 4.5]	4.6
1985	[4.6, 4.7]	4.7
1986	[4.6, 4.6]	4.6
1987	[3.1, 3.2]	3.2
1988	[3.1, 3.1]	3.1
1989	[5.1, 5.1]	5.1
1990	[6.2, 6.2]	6.2
1991	[7.2, 7.3]	7.4
1992	[6.8, 6.9]	7.0
1993	[2.6, 2.7]	2.8
1994	[2.7, 2.7]	2.9
1995	[2.7, 2.8]	2.9
1996	[2.5, 2.6]	2.7
1997	[2.3, 2.3]	2.4

Figure 4.10: Di¤erences in in‡ation rates, RPI* and the chained true index, 1977 to 1997, percentage points

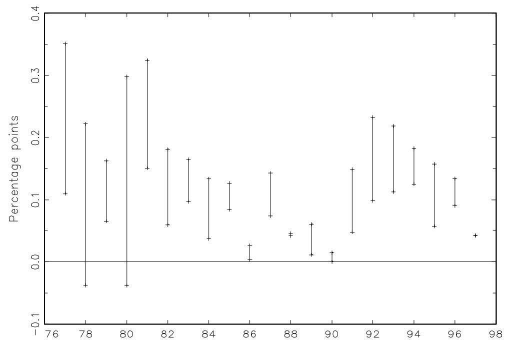
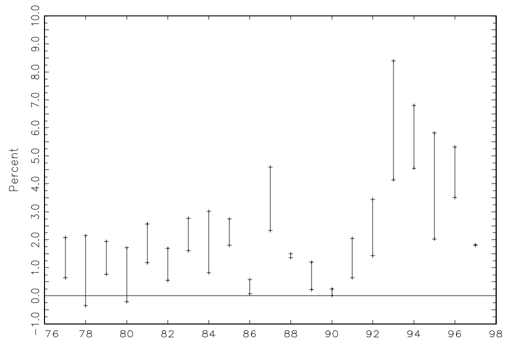


Figure 4.11: Percentage di¤erences in in‡ation rates, RPI* and the chained true index, 1977 to 1997



compared to an index calculated by the methodology used in preparation of the RPI. We do this at the section index level which is the most disaggregated level possible given the availability of price data. The …rst thing to note is that there is no theoretical presumption of an upward bias in the RPI formula. However, the empirical evidence shows that the RPI does tend to overstate year-to-year increases in the cost-of-living (only in three years was the annual rate of in‡ation as measured by the RPI method within the bounds on the true index). This means that over the period these errors are compounded. Over the period 1976 to 1997 the RPI overstated the true increase in the cost-of-living by up to 3.2%.

Appendices

A. The bounding algorithms

$RP(q_0)$ Bound Algorithm

Input is a set of $T+1$ expansion paths $F(q|p_t; x)$ for $t = 0, 1, \ldots, T$ and a base bundle q_0.

Output is the set RP of boundary points of which q_0 is a member and which has $T+1$ elements where $p_i' q_i \leq p_i' q_j \; \forall \; q_i, q_j \in RP$ and $q_i \, R \, q_0$ for all $q_i \in RP$.

1) Set $W = \{q_0\}$; $E = \emptyset$

2) For each t we define $q_t = F(q|p_t; x)$ where $\hat{x}_t = \arg\min \{x \mid p_t' q_t = p_t' q_w\}$ for all $q_w \in W$. Call this set R.

3) Set $E = \{q_i \in R : p_i' q_i > p_i' q_j \text{ for } q_j \in R\}$

4) Set $W = R \backslash E$

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

$RW(q_0)$ Bound Algorithm

Input is a set of $T+1$ expansion paths $F(q|p_t; x)$ for $t = 0, 1, \ldots, T$ and a base bundle q_0. This base bundle must be on an expansion path and have an observed corresponding price vector.

Output is the set RW of boundary points of which q_0 is a member and which has $T+1$ elements where $p_i' q_i \leq p_i' q_j \; \forall \; q_i, q_j \in RW$ and either $q_0 \, R^0 \, q_i$ or $q_0 \, R \, q_i$ for all $q_i \in RW$.

1) Set $B = \{q_0\}$; $E = \emptyset$

2) For each t we define $q_t = F(q|p_t; x)$ where $\hat{x}_t = \arg\max \{x \mid p_b' q_b = p_b' q_t\}$ for all $q_w \in W$. Call this set R.

3) Set $E = \left\{ q_i \in R : p_j' q_j > p_j' q_i \text{ for } q_j \in R \right\}$

4) Set $B = R \backslash E$

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

B. Convergence as a test of GARP

Proposition 1. If the data in the set R violate GARP, the algorithm for the boundary to the set $RP(q_0)$ does not converge.

Proof.
1) Suppose that the algorithm has complete n iterations, denote the sets $R; W$ and E at this point by $R_n; W_n$ and E_n:
2) Suppose that the data in the set R_n violate GARP: hence $\exists\, q_i; q_j \in R_n$ such that $q_i P^0 q_j$ and $q_j R q_i$.
3) By Step (3) of the algorithm $q_j \in E_n$ where $E_n \subset R_n$:
4) Hence $E_n \neq \emptyset$ and the algorithm has not converged and proceeds to evaluate $R_{n+1}; W_{n+1}$ and E_{n+1} where $W_{n+1} \neq W_n$.
■

Proposition 2. If the data in the set B violate GARP, then the algorithm for the boundary to the set $RW(q_0)$ does not converge.

Proof.
The proof is analogous with that for Proposition 1. ■

C. Commodity groups

The following is a list of the commodity groups used in the empirical work. The de…nitions are, for the most part, the same as for the RPI, (CSO, 1991, p.90-91). Where it di¤ers we have noted this

Bread, Cereals, Biscuits & cakes, Beef, Lamb, Pork, Bacon, Poultry and other meats (as RPI up to 1987, the sum of the separated groups thereafter), Fish, Butter, Oils and fats, Cheese, Eggs, Milk (sum of RPI groups "milk, fresh" and "milk products"), Tea, Co¤ee, Soft Drinks, Sugar and Preserves, Sweets and chocolates, Potatoes, Other vegetables, Fruit, Other Foods, Restaurants and other meals out (RPI group "Other meals and snacks" until 1979, the sum of RPI groups "Restaurant meals" and "Take-away meals and snacks" thereafter), Canteen Meals, Beer, Wines and Spirits, Cigarettes, Other tobacco, Solid fuels, Electricity, Gas, Oil and other fuels, Furniture and Furnishings (sum of the RPI groups of same names), Electrical Appliances, Other household equipment, Household consumables, Petcare, Postal charges, Telephone charges, Domestic services, Fees and subscriptions, Men's wear, Women's wear, Children's wear, Other clothing, Footwear, Personal articles and services (sum of RPI groups of the same name), Chemists goods, Purchase of motor vehicles, Maintenance of motor vehicles, Petrol and oils, Vehicle taxes and insurance, Rail fares, Bus fares, Other fares, Audio-visual equipment, records and tapes and toys & photographic and sports goods (sum of RPI groups of the same names post 1986), Books and newspapers, Gardening products, TV Licenses, Entertainment and recreation.

D. The non-housing RPI and RPI*

Table D.1: RPI* and the non-housing RPI, 1976-1997.

	Indices	
Year	Non-housing RPI	Non-housing RPI*
1976	100.0	100.0
1977	116.9	117.2
1978	128.9	129.6
1979	139.9	140.7
1980	164.5	165.6
1981	184.3	187.1
1982	203.3	207.5
1983	215.2	220.1
1984	224.5	230.2
1985	234.3	241.1
1986	244.8	252.2
1987	252.4	260.4
1988	260.4	268.4
1989	273.8	282.2
1990	289.2	299.6
1991	309.7	321.8
1992	332.1	344.3
1993	340.7	354.0
1994	351.6	364.1
1995	360.6	374.6
1996	370.5	384.5
1997	380.3	393.6

References

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

[2] 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.

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

[4] 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.

[5] Blundell, R.W., Browning, M.J. and I. A. Crawford, "Nonparametric Engel curves and revealed preference", University College London Discussion Paper, 98-08.

[6] Deaton, A.S. (1981) Editor's 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.

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

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

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

[10] 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.

[11] Diewert, W.E. and Nakamura, A. O., (ed.s) (1993), Essays in Index Number Theory, Vol. 1, North-Holland: Amsterdam.

[12] Fisher, I. (1911), The Purchasing Power of Money, London: Macmillan

[13] Konüs, A. A. (1924), "The Problem of the True Index of the Cost of Living", reprinted in Econometrica, 7, 10-29.

[14] Lerner, A.P. (1935-36), "A Note on the Theory of Price Index Numbers", Review of Economic Studies, 3, 50-56.

[15] Manser, M.E. and McDonald, R.J. (1988),"An Analysis of Substitution Bias in Measuring In‡ation, 1959-1985", Econometrica, 56, 909-930.

[16] Pollak,R.A. (1971), "The Theory of the Cost of Living Index", Research Discussion Paper No. 11, Bureau of Labor Statistics, Washington, D.C.

[17] Robinson, P.M. (1988), "Root n-consistent Semiparametric Regression", Econometrica, 56, 931-954.

[18] Stone, J.R.N. (1956), Quantity and Price Indexes in National Accounts, Paris: Organisation for European Economic Co-operation.

[19] Tanner, S. (1996), "How Much do Consumers Spend ? - Comparing the FES and National Accounts", mimeo, Institute for Fiscal Studies.

[20] Törnqvist, L. (1936), "The Bank of Finland's consumption price index", Bank of Finland Monthly Bulletin, 10, 1-8.

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

[22] Varian, H. (1983),"Nonparametric Tests of Consumer Behaviour", Review of Economic Studies, 99-110.