

Changes in Relative Consumer Prices and the Substitution Bias of the Laspeyres Price Index: Ireland, 1985-2001

R. A. SOMERVILLE*
Trinity College Dublin

Abstract: This paper shows that Irish relative consumer prices have changed significantly, 1985-2001, at the ten commodity-group level. A “true” cost-of-living index is derived from Madden’s (1993) parameter estimates for an *Almost Ideal Demand System*. Despite relative price changes, the substitution bias of a computed Laspeyres index is small, and the official Consumer Price Index tracks the computed index closely. Superlative indices are also constructed, but are not satisfactory cost-of-living indices in this context. Cost-of-living indices are computed for different income groups, and the impact of inflation in recent years is found to be negatively correlated with income.

I INTRODUCTION

“Accurately measuring prices and their rate of change, inflation, is central to almost every economic issue” (Boskin *et al.*, 1998, p. 3).

In Ireland and elsewhere, the index of consumer prices (hereinafter, the CPI) is widely used as an input to economic decisions. In this context it is often used implicitly as a cost-of-living index, although since the work of Konüs (1924, 1939), it has been known that fixed-weight price indices are biased approximations to cost-of-living indices because they fail to take account of the substitution effects of changes in relative prices. Further sources of bias arise

* The author is most grateful to Dr David Madden of University College Dublin for releasing unpublished data from his research for Madden (1993), and for generously giving permission for these data to be used in the preparation of this paper. An early application of the methodology of Section 4.3 to these data was undertaken, under the author’s supervision, by Mr Anthony McGuinness. Finally, the author would like to thank an anonymous reviewer for helpful comments and suggestions.

from quality-change and from lags in the entry or exit respectively of new or obsolete goods.

Section II introduces the problem: major uses of the CPI are discussed, and the difference between cost-of-living indices and statistical price indices is explored, along with the implications of treating the CPI as a cost-of-living index. Later sections report empirical investigations of Irish data. Section III contains an analysis of changes in relative consumer prices, 1985-2001. Section IV contains a discussion of alternative methodological approaches to measuring substitution bias. Section V sets out the empirical findings on bias, including an exploration of the differential impact of price-changes across the income range. Finally, conclusions are drawn in Section VI.

The core of the paper is Section V, which utilises Madden's (1993) work on consumer demand for 1958-1988. Some overlap with Madden's sample-period is desirable, so our sample starts at 1985, with base-year¹ set at 1987, because the 1987 Household Budget Survey (Central Statistics Office (CSO), 1989) provides base expenditure data and, implicitly, a base utility level. The terminal date is December 2001, that being the last month of the old CPI series before the recent re-weighting, which included significant methodological changes (CSO, 2003). Also, it was the last month before the arrival of Euro notes and coins, which was an exceptional event with possible repercussions for relative prices. Most of the analysis uses Madden's ten-group classification of consumer goods and services, as follows:²

- | | | |
|---|------------------|-----------------|
| 1. Food; | 2. Alcohol; | 3. Tobacco; |
| 4. Clothing & footwear; | 5. Fuel & power; | 6. Petrol; |
| 7. Transport [excluding
petrol] & equipment; | 8. Durables; | 9. Other goods; |
| | 10. Services. | |

Two sources of price data are used here. To maintain the connection with the CPI, the primary focus is on CPI price-series for the ten main commodity groups defined by Madden. The data, quarterly until November 1996 and monthly from January 1997, were obtained online from the CSO's website (CSO, online), supplemented for earlier years by the *Statistical Bulletin*.

¹ Strictly, the starting date of a price index I, at which I=100, is called the reference period. The base period is the date at which fixed expenditure weights are estimated. In this paper these usually coincide, in which case the term "base" covers both.

² Madden's categories differ from the ten major sub-groups in the CPI (up to 2001) in two ways: he treats *petrol* separately from other *transport*, and does not identify *housing* separately. For this paper, a *petrol* index is computed from CPI data, and removed from CPI *transport*. CPI *housing* and *services* are merged, except that the *repairs and decorations* element of *housing* is merged with *household durables*.

Alternatively, national accounts data for aggregate annual expenditure in eighteen consumption categories are published in CSO (online) and in hard copy in Tables 13 and 14 of *National Income and Expenditure* (hereinafter NIE), and the ratios of expenditure at current and constant prices provide price indices for each category. The NIE and CPI price-series differ, because of differences in weighting, and because the former are Paasche indices while the latter are Laspeyres (see Appendix A1.). We use seventeen of the NIE series, which for some purposes we aggregate into ten. "Housing" is excluded because this paper is concerned only with market activities: it is necessary therefore to exclude imputed rent from owner occupation, a major element of the housing series that is not identified separately before 1995.

II COST-OF-LIVING INDICES, PRICE INDICES, AND SUBSTITUTION BIAS

2.1 *Uses of the CPI*

In most countries, including Ireland, the CPI is a major economic indicator. It provides the most widely-used measure of inflation, and is central to monetary policy. Much recent research on the CPI comes from the USA, notably the report of the Boskin Commission (Boskin *et al.*, 1996). According to the Commission, "Slightly under one-third of total [US] Federal outlays...are indexed to changes in consumer prices. Several features of the individual income tax...are indexed; the individual income tax accounts for a little under half of federal revenues" (Boskin *et al.*, 1996, p. 13).

In Ireland, government spending and taxes are not formally indexed, but measured inflation influences them informally. For example, in December 2002 the Minister for Finance stated in his budget speech that "The rate of payment for old age pensioners will, by 2003, have increased by 59 per cent over the rate payable in 1997. This is well ahead of inflation..." (Government of Ireland, 2002, p. A. 12). In the 1997 budget, the then minister announced that "These increases [in weekly welfare payments] are about 4 per cent or twice the rate of inflation" (Government of Ireland, 1997, p. 19).

Another major use of the CPI is for deflating nominal magnitudes, *ex post*. For example, it is one of the deflators used to produce estimates of real GNP, and it is used by participants in wage-bargaining to assess outcomes. Inflation expectations are important across a range of economic decisions, and measured inflation may be an input to their formation, depending on the extent to which expectations are formed rationally. For example, in a study of EU countries, including Ireland, Madsen (1996, p. 1337) reports that "[inflation] expectations are somewhat adaptive and extrapolative".

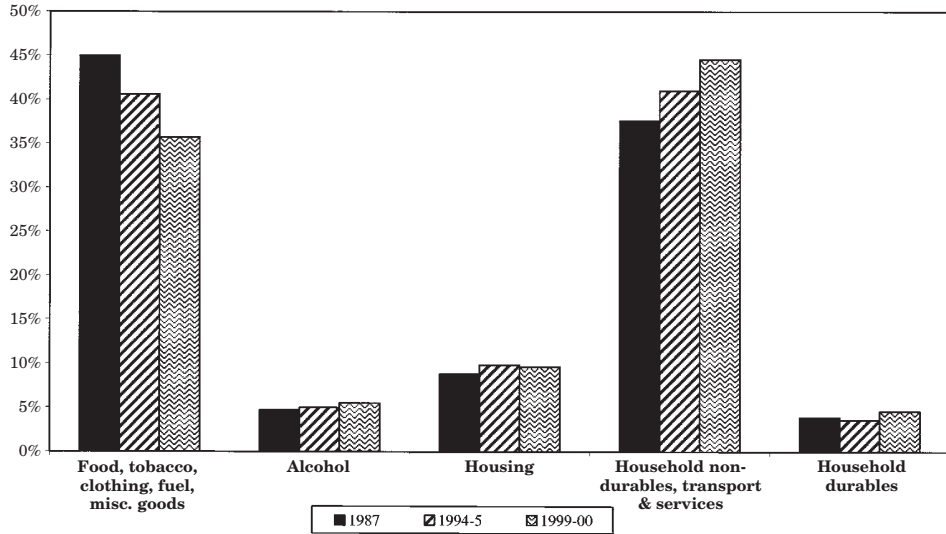
2.2 *The Cost-of-Living Index and Statistical Price Indices*

The “cost of living” is the income required to achieve a given living standard, or more formally, *the least expenditure required to obtain a given level of utility*. Optimisation is inherent in this concept, and any measure of the cost of living should in principle reflect all possible changes in the opportunity set, because optimising behaviour involves adapting optimally to such changes. The most obvious case in point is a change in the average price level, while other possibilities include changes in relative prices, or in the quality and characteristics of existing goods, or in the set of goods when new goods enter or others leave.

The uses of the CPI cited in Section 2.1 exemplify its common implicit interpretation as an index of the cost of living. However, statistical price indices are typically designed to track the cost of a fixed vector of goods, which means that consumer price indices are not generally “cost-of-living indices”. For example, the Commission of the EU maintains that “The HICP [i.e. Harmonised Index of Consumer Prices] is not a cost of living index” (Eurostat, 2001, p. 19). Similarly in Ireland, “The CPI measures price change. It is specifically designed *not* to take into account changes made by households to their pattern of expenditure...in response to changes in prices...The CPI is a price index, not a cost of living index” (CSO, 1997a, p. 257, and 2003, p. 37). In fact, in common with most official price-series, the Irish CPI is based on the Laspeyres methodology (see Appendix A1.): that is, it is designed to measure the cost of buying a fixed vector of consumption goods and services over time. It tracks the cost of the base level of utility, but not the *least* cost, because its construction does not allow for all possibilities of adapting optimally to changes in opportunities. Consequently, it overstates the least cost of achieving base-period utility.

From 1968 to 1996 the CPI in Ireland was revised on a seven-year cycle (which has since been reduced to five), and the reference date of each series came approximately two years after the base period, i.e. the date of the Household Budget Survey that provided the expenditure weights. Thus by November 2001, the CPI (referred to November 1996=100) was based on a Survey taken seven or more years earlier (1993-4). Successive Surveys show that average expenditure patterns have been changing significantly. For example, between 1987 and 1999-2000 the average expenditure shares of food, clothing, fuel, miscellaneous goods and tobacco have all fallen consistently, from a total of 45 per cent to just under 36 per cent, while the shares of household non-durables, transport and services have all risen, from just under 38 per cent in total to nearly 45 per cent (see Figure 1).

Figure 1: *Household Budget Survey, Trends in Composition of Spending, 1987-1999/2000*



Data Sources: CSO, 1989; CSO, 1997b; CSO, 2001.

2.3 The Implications of Using the CPI as a Cost-of-Living Index

To express the index of the cost of living symbolically, we begin with the consumer’s expenditure function $e(\mathbf{p}, u)$, which specifies the least expenditure required to obtain a given level u of utility at prices \mathbf{p} : i.e. the “cost of living”. Given u , and base prices \mathbf{p}_0 , the index at prices \mathbf{p} is then defined as $C(\mathbf{p}_0, \mathbf{p}, u) = 100 \times e(\mathbf{p}, u) / e(\mathbf{p}_0, u)$, which is sometimes referred to as the Konüs cost of living index. $C(\mathbf{p}_0, \mathbf{p}, u)$ is closely related to the Hicksian compensating variation (CV) of a price change, where $CV = e(\mathbf{p}_0, u) - e(\mathbf{p}, u)$. If a change in \mathbf{p} raises C , then it has a negative CV, indicating a fall in utility relative to the base period at constant income. In general the true index is dependent on the level of utility u , except when preferences are homothetic, which is the necessary and sufficient condition for $e(\mathbf{p}, u)$ to take the form $uf(\mathbf{p})$, for some function f . For this paper, it is natural to specify $C(\mathbf{p}_0, \mathbf{p}, u)$ in terms of utility in the base-period for prices.

The Laspeyres index relates to the relative cost of the base consumption vector $\mathbf{x}(\mathbf{p}_0, u)$ (see Appendix A1), and may be written $L(\mathbf{p}_0, \mathbf{p}, u) = 100 \times \mathbf{p} \cdot \mathbf{x}(\mathbf{p}_0, u) / e(\mathbf{p}_0, u)$, bearing in mind that $e(\mathbf{p}_0, u)$ is the cost of $\mathbf{x}(\mathbf{p}_0, u)$ at

prices \mathbf{p}_0 . While $\mathbf{x}(\mathbf{p}_0, u)$ is one way of getting utility u at prices \mathbf{p} , it is not necessarily least-cost, so that $L(\mathbf{p}_0, \mathbf{p}, u) \geq C(\mathbf{p}_0, \mathbf{p}, u)$.³

Whether or not the cost-of-living framework is adopted in principle, it has yet to be fully implemented anywhere, and because many practical uses of the CPI carry cost-of-living overtones, it is important to know the extent to which the CPI deviates from a cost-of-living index. The seminal recent work is the report (Boskin *et al.*, 1996) of the Boskin Commission, whose findings are extensively discussed by the commissioners and others in a symposium in *Journal of Economic Perspectives*, 1998.⁴ Boskin *et al.* (1998, p. 11, Table 1) estimate the annual bias in the US CPI to be 1.1 percentage points per annum, within a range of plausible values of 0.8 to 1.6 points. Of the 1.1 points, they attribute 0.4 points to substitution bias and 0.6 to new product/quality change, with the residual attributed to “outlet substitution” i.e. to the fact that the sample of outlets from which prices are sampled is only rotated slowly. Using estimates from the Congressional Budget Office, the Boskin Report concludes that over-indexation of federal tax and expenditure programmes could add \$1.07 trillion to US national debt by 2008 (Boskin *et al.*, 1996, p. 8).

A provisional estimate puts total spending by the Government of Ireland in 2002 on social insurance and social assistance payments at €9,520m, and in the December 2002 budget the increase for 2003 was put at €530m, i.e. a rise of 5.6 per cent. The budget estimate for the annual rise in the CPI in 2002 was 4.7 per cent, while the forecast rise in 2003 was 4.8 per cent (Government of Ireland, 2003, p. 3; 2002, pp. A.11 and E.11). If we were to transfer Boskin’s estimate of bias to Ireland, then it could raise questions concerning about €100m of the increase in the social welfare budget for 2003.

2.4 From Fixed Weight Index to Cost-of-Living Index

There is a continuing debate among economists and official statisticians over the desirability and practicability of developing cost-of-living indices, as alternatives to fixed-weight series: for example, see Schultze (2003), National Research Council (2002), and Triplett (2001), who discusses the positions taken by official statistical and other agencies in some of the major industrial countries. In some cases, notably the USA, the cost of living index is explicitly adopted by official statisticians as the correct framework for the CPI. In others

³ This conclusion relates primarily to levels, not to rates of change. A Laspeyres price index overstates the true cost of living, but it is not generally true that a fixed-base Laspeyres index overstates the rise in the cost of living between *any* pair of dates. Such an index overstates the rise between the base point $t=0$ and any $t_i>0$, but it does not necessarily overstate the rise between any pair of dates $t_i>0$ and $t_j> t_i$.

⁴ *Journal of Economic Perspectives*, Vol. 12, No. 1, Winter 1998, pp. 3-78; also see three related papers in *Journal of Economic Perspectives*, Vol. 17, No. 1, Winter 2003, pp. 3-58.

(e.g. the UK, Ireland, and within the EU at Commission level), the framework has been rejected, while in yet others (e.g. Canada) the position is ambiguous (Triplett, 2001, p. F311, and CSO, 2003, p. 37).

Boskin *et al.* make a number of recommendations, all within the context of their overall proposal that the US Bureau of Labor Statistics (BLS) should establish a cost-of-living index. They propose moving from fixed-weight indices to indices that allow for substitution when prices are aggregated, and they recommend frequent reweighting. They urge that entry lags for new products should be shortened, but do not propose the more radical step of including an estimate of the consumer surplus (i.e. the compensating variation) that accrues when new goods appear. They propose that quality change should be handled by expanded use of hedonic regression techniques.

In 1997, the BLS announced that its objective was for the CPI to become a cost-of-living index, and that its decisions had reflected that intention for some time (Schultze, 2003, p. 4). Subsequent to the Boskin report, the BLS commissioned a panel to investigate the construction of a cost-of-living index (see National Research Council, 2002). Abraham (2003) describes recent progress, including the introduction of weighted geometric averaging of price-relatives at the lowest level of aggregation, to tackle substitution bias. Compilation of the US CPI begins with about 200 item-area components. For a subset accounting for about 60 per cent of expenditure, the new procedure generates a subindex for each component that is, in effect, a weighted geometric mean of the individual price-relatives within the components, using base expenditure shares as weights (Moulton, 1993, and Dalton *et al.*, 1998). Geometric averaging is intended to allow for substitution within each component or stratum, with an implicit assumption of constant expenditure shares, whereas the former Laspeyres procedure reflected an implicit assumption of constant quantities. There is no particular reason to assume constant expenditure shares, but as Abraham (2003, p. 50) suggests, it “is almost certainly preferable to assuming no substitution at all”. Combination of the subindices continues to involve the Laspeyres procedure.

In Ireland, some recent initiatives by the CSO are consistent with the Boskin recommendations for the US. From December 2001 simple (i.e. unweighted) geometric averaging replaces arithmetic averaging of prices at the lowest level of aggregation for about half of the “basket”, i.e. where prices are taken directly from retail outlets (CSO, 2003, p. 45). This should take some account of substitution between identical or near-identical goods as relative prices change. The current (December 2001) series, based on the 1999-00 Survey, shortens the re-weighting cycle from seven years to five, so that in future CPI weights should on average be closer than hitherto to actual expenditure proportions. Other relevant recent changes include the adoption

of two techniques of adjustment for changes in quality (CSO, 2003, p. 46).

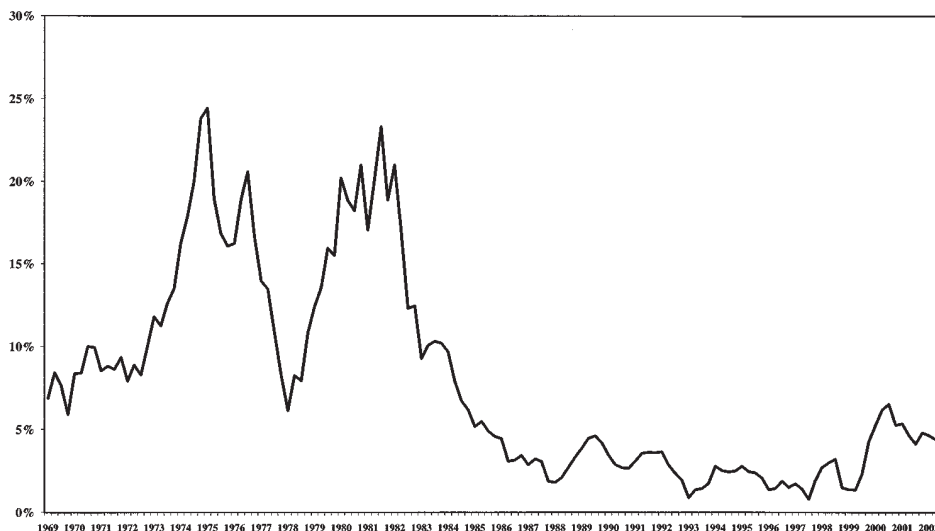
Despite these developments, the Irish CPI continues in essence to be a fixed-weight Laspeyres index, and therefore, compared with a cost-of-living index, it remains prone to bias. The following analysis is concerned with just one of three major biases discussed by Boskin: the substitution bias that arises because the base quantity vector, as used in the Laspeyres index, is no longer least-cost when relative prices change.

III VARIATION IN RELATIVE CONSUMER PRICES: IRELAND, 1985-2001

3.1 Inflation in Ireland

Inflation was comparatively low in the 1990s, and the annual average rate of increase in the CPI was 2.9 per cent in the sample period, 1985-2001. A longer series captures the high-inflation periods following the two oil-shocks of the 1970s, so that from 1968 to 2002 the annual average rate of increase was 7.6 per cent. Throughout those years there was considerable variation in the annual rate as revealed by Figure 2, and in this section we explore the extent to which this was accompanied by variation in relative prices. Evidence of changing expenditure proportions (see Figure 1) is not necessarily evidence of

Figure 2: *CPI, Annual Inflation Rate from 1969 (first quarter) to 2002 (third quarter)*



Data sources: CSO online and Statistical Bulletin.

variation in relative prices: if preferences are non-homothetic, then expenditure proportions will change over time as income varies at constant relative prices, and we should therefore examine prices directly.

3.2 *Measuring Changes in Relative Prices: Methodological Issues*

In a study of the connection between price dispersion and inflation, Parks (1978; and see also Theil, 1967, Chapter 5) introduced a measure of relative price variation that has since been widely used (e.g., Jaramillo, 1999, Fielding and Mizen, 2000). It measures the dispersion of inflation rates for different commodities about the mean, on a one-period basis, and is defined as follows: $RPV_t = \sum_i w_i (Dp_{it} - \bar{Dp}_t)^2$, where $Dp_{it} = \ln(p_{it}) - \ln(p_{i,t-1})$ and $\bar{Dp}_t = \sum_i w_i Dp_{it}$. The p_{it} are prices or indices for category i at time t , and the w_i are the expenditure shares in some base-year. Differencing ensures that RPV is invariant to equiproportionate changes in all prices and is unaffected by choice of base-year when applied to an index, whereas in levels the variance, the coefficient of variation and Braithwait's measure (1980, p. 71) all fail to satisfy either or both of these conditions.

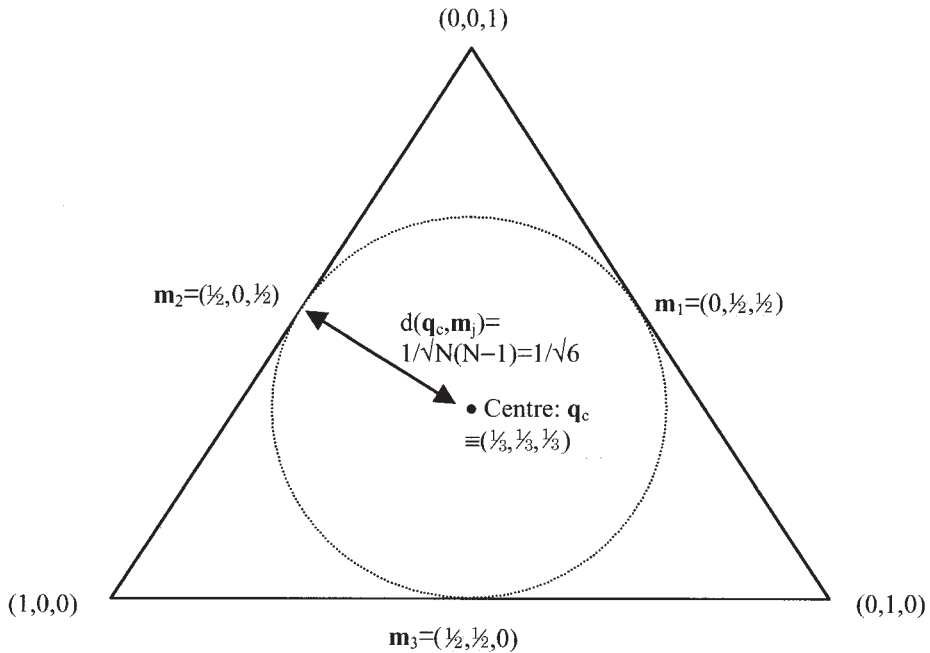
An alternative approach, using levels, is to normalise by mapping each price-vector \mathbf{p}_t into the unit simplex $S_N = \{\mathbf{q}_t > \mathbf{0} \mid \sum_j q_{tj} = 1\}$.⁵ The transformation $\mathbf{q}_t = (1/\sum_j p_{tj})\mathbf{p}_t$ maps any \mathbf{p}_t into S_N , and we identify changes in relative prices with changes in the distance between normalized price-vectors. The distance between two points in S_N is $d(\mathbf{q}_t, \mathbf{q}_r) = \sqrt{\sum_j [(q_{tj} - \bar{q}_t) - (q_{rj} - \bar{q}_r)]^2}$, or equivalently $d(\mathbf{q}_t, \mathbf{q}_r) = \sqrt{[(N-1)(s_t^2 + s_r^2 - 2s_{tr})]}$ where the s terms are the sample variances and covariance computed from the components of the two vectors. A large value of $d(\mathbf{q}_t, \mathbf{q}_r)$ reflects some combination of high variances with a small covariance. (Further technical detail is given in Appendix A.2.)

3.3 *Measuring Changes in Relative Prices: Empirical Analysis, 1985-2001*

First, we compute RPV for Madden's ten commodity groups using the appropriate CPI series, with $t=2001$ and $t-1=1985$. We find $\sqrt{RPV_t} = 23.12$ per cent compared with a rise in the CPI of 58.4 per cent, which suggests a significant change in relative prices.

Using the alternative approach, we condense quarterly or monthly CPI data into annual averages, to assist interpretation. $T=17$ (i.e. 1985-2001), and Table 1 sets out the distances between all possible pairs of vectors in the set of averages, after normalisation to the unit simplex S_{10} . No pair of points coincides, and with three small exceptions, the distance away from any year always grows as we move down the table.

⁵ For vector inequalities, this paper uses the notation \succeq , $>$, \gg , whereby $\mathbf{x} \succeq \mathbf{y}$ denotes $x_i \geq y_i$, each i ; $\mathbf{x} > \mathbf{y}$ denotes $x_i > y_i$, each i with $x_i > y_i$, some i ; $\mathbf{x} \gg \mathbf{y}$ denotes $x_i > y_i$, each i .

Figure 3: Illustration of the Unit Simplex for $N = 3$ 

The triangle represents S_3 , the unit simplex in three dimensions ($N = 3$): i.e. the set (in this case a plane) of all price vectors $\mathbf{q}_t > \mathbf{0}$, after normalisation to $\sum_j q_{tj} = 1$.

The base-year price-vector maps to the centre, \mathbf{q}_c .

Given an arbitrary price-vector \mathbf{q}_t , a sufficient condition for $\mathbf{q}_t \geq \mathbf{0}$ is that \mathbf{q}_t lie within the inscribed circle, or equivalently $|\mathbf{q}_t - \mathbf{q}_c| < \sqrt{N(N-1)}$. A necessary condition for $q_{ti} = 0$, some i , is that $|\mathbf{q}_t - \mathbf{q}_c| \geq \sqrt{N(N-1)}$.

The average value of all cells in Table 1 is $\bar{d} = 0.0332$, which must be scaled to be interpreted. The base year (1987) maps to the centre of the simplex, because all of its coordinates are equal. Consider the distance from the centre of the simplex to the mid-point of any side, where one price equals zero and all others equal $1/(N-1)$. This is illustrated in Figure 3 for $N=3$, and it is the least distance at which any point must lie from the centre to permit the possibility of a zero price for at least one commodity-group. Such a possibility is an extreme improbability at the given level of aggregation, so we interpret this distance, equal to 0.1054 for the case $N=10$, as a very large movement in relative prices. During 1985-2001 the average distance between normalised

Table 1: Euclidean Distances $d(\mathbf{q}_t, \mathbf{q}_r)$ Between Pairs of Price Vectors $\mathbf{q}_t, \mathbf{q}_r$ on the Unit Simplex S_{10} , at Times t, r , Annual Data, 1985-2001

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
1986	0.0184															
1987	0.0252	0.0078														
1988	0.0297	0.0119	0.0051													
1989	0.0288	0.0124	0.0059	0.0041												
1990	0.0299	0.0147	0.0088	0.0085	0.0062											
1991	0.0351	0.0190	0.0125	0.0103	0.0102	0.0071										
1992	0.0444	0.0277	0.0207	0.0176	0.0182	0.0167	0.0106									
1993	0.0484	0.0315	0.0244	0.0209	0.0219	0.0214	0.0155	0.0064								
1994	0.0528	0.0357	0.0286	0.0248	0.0258	0.0262	0.0209	0.0120	0.0063							
1995	0.0574	0.0406	0.0335	0.0297	0.0304	0.0307	0.0254	0.0162	0.0108	0.0056						
1996	0.0582	0.0421	0.0352	0.0315	0.0320	0.0329	0.0281	0.0197	0.0144	0.0095	0.0054					
1997	0.0626	0.0475	0.0409	0.0375	0.0376	0.0381	0.0336	0.0258	0.0210	0.0170	0.0126	0.0083				
1998	0.0713	0.0563	0.0497	0.0462	0.0462	0.0466	0.0419	0.0339	0.0292	0.0248	0.0198	0.0165	0.0096			
1999	0.0768	0.0625	0.0563	0.0529	0.0528	0.0537	0.0493	0.0422	0.0374	0.0329	0.0283	0.0246	0.0177	0.0099		
2000	0.0813	0.0696	0.0643	0.0617	0.0613	0.0627	0.0592	0.0533	0.0487	0.0449	0.0408	0.0361	0.0291	0.0245	0.0176	
2001	0.0879	0.0752	0.0694	0.0665	0.0662	0.0668	0.0627	0.0559	0.0514	0.0474	0.0427	0.0388	0.0313	0.0238	0.0163	0.0120

Note: Within each year, the CPI price-series for each commodity group is based on 1987=100. The price-vector for each year is mapped into the unit simplex, and the Euclidean distances are computed. Bold-face type indicates the three cases where $d(\mathbf{q}_{t+k+1}) < d(\mathbf{q}_t, \mathbf{q}_{t+k})$, some $k > 0$.

price-vectors (i.e. 0.0332) is about one-third of this critical value, and again we conclude that significant changes in relative prices have occurred. The conclusion is identical for NIE price data, where \bar{d} is only marginally larger at 0.0355.

IV CALCULATION OF SUBSTITUTION BIAS: METHODOLOGY

4.1 Calculation of the Cost-of-Living Index: (i) Using the Expenditure Function

The first method of measuring bias begins with a system of estimated demand equations, from which we derive the consumer's expenditure function. From this we construct the cost-of-living index for the base-period level u of utility, which is then compared with a Laspeyres index L derived from the base-period consumption vector predicted by the same system of equations (here, the base and reference periods are both 1987). That vector is one way, not necessarily the cheapest, of getting utility u at any price-vector \mathbf{p} . Thus the cost-of-living index should never exceed the Laspeyres index, and the difference between them measures substitution bias. This methodology is used in Braithwait (1980) for example, and also in Irvine and McCarthy (1978), which is the only published empirical investigation of the topic using Irish data.

A potential difficulty is that the technique involves a high level of aggregation because of the impracticability of estimating the very large number of parameters of a highly disaggregated system. Bias may be missed if much substitution occurs at a more disaggregated level, although Boskin *et al.* (1998, p. 7) suggest that "estimates of substitution bias [produced by this approach] in, for example, Jorgenson and Slesnick [1983], are quite similar to those in numerous far more disaggregated studies by the BLS, which use the second method [i.e. based on a superlative index: see Section 4.4]".

4.2 Calculation of the Expenditure Function: Choice of Estimated Demand System

Early investigations of demand in Ireland, including Casey (1973), O'Riordan (1976), and McCarthy (1977), were based on the linear expenditure system (LES), and McCarthy's paper led to the pioneering work on substitution bias of Irvine and McCarthy (1978). The LES, while simple and tractable, is a rather inflexible functional form. For example, all goods must be normal with respect to income, and net complementarity is ruled out. Later research uses less restrictive functional forms: Conniffe and Hegarty (1980) and Madden (1993) use the Rotterdam model, and Madden also uses the CBS

model and the Almost Ideal Demand System (AIDS). Conniffe and Eakins (2003) return to the LES, in an investigation of stochastic specification; because it uses only five commodity groups their model is unsuitable for use here.

The AIDS expenditure function is a “flexible functional form”: that is, it may act as a second-order approximation to an unknown underlying function, and consequently it avoids the inflexibility of the LES. It “gives an arbitrary first-order approximation to any demand system; it satisfies the axioms of choice exactly; it aggregates perfectly over consumers without invoking parallel linear Engel curves; it has a functional form which is consistent with known household-budget data...” (Deaton and Muellbauer, 1980a, p. 312). Finally, the expenditure function may be extracted relatively easily from an estimated AIDS model, estimated in levels. Madden’s *AIDS model 3*, estimated in levels, is used here (see Madden, 1993, pp. 102-103).⁶ The investigation of substitution bias rests on a maintained hypothesis of maximising behaviour, and consequently it is appropriate to use this version of the model, which has homogeneity and Slutsky-symmetry imposed.

4.3 Application of Madden’s Parameter Estimates

A series for the expenditure function $e(\mathbf{p}, u)$, February 1985 to December 2001, may be calculated by combining CPI price data with Madden’s parameter estimates. The expenditure function is as follows, with the indices i and j running from 1 to N :

$$\ln e(\mathbf{p}, u) = \alpha_0 + \sum_i \alpha_i \ln(p_i) + \sum_{ij} \gamma_{ij}^* \ln(p_i) \ln(p_j) + \beta_0 u \prod_i p_i^{\beta_i}.$$

The parameters α_i and β_i (for all $i > 0$) appear in the demand equations, which also contain parameters γ_{ij} that are related to the γ_{ij}^* as follows: $\gamma_{ij} = \frac{1}{2}(\gamma_{ij}^* + \gamma_{ji}^*) = \gamma_{ji}$. Madden’s regression results thus provide estimates of all the parameters in $e(\mathbf{p}, u)$, other than α_0 and $\beta_0 u$. We assume that $u=0$ at very low expenditure levels, and α_0 is evaluated by setting $e(\mathbf{p}, u)=1$, $\ln[e(\mathbf{p}, u)]=0$, and letting $u=0$ with 1987 prices. Deaton and Muellbauer (1980a, p. 316) note that “in many examples the practical identification of α_0 is likely to be problematical” and continue: “[α_0] can be interpreted as the outlay for a minimal standard of living when prices are unity (usually in the base year...)”, and the procedure adopted here is consistent with this. Given α_0 , $\beta_0 u$ is evaluated using the reported mean expenditure (annualised) from the 1987

⁶ Madden’s regression estimates are based on price series set at 100 in 1985 (personal communication from Dr Madden), so a 1985 base is used when utilising them. For all other purposes, the base year is 1987 as previously described. The regression estimates are not given in Madden (1993), but have been made available by Dr Madden and are used with his permission. For expenditure categories see Note 2.

Household Budget Survey (CSO, 1989, Table 2). This normalisation of utility is necessary to satisfy the condition that $e(\mathbf{p}, u)$ should equal actual expenditure in the 1987 base-period. Finally, the base-utility true index of the cost of living is $C=100 \times e(\mathbf{p}, u)/e(\mathbf{p}_0, u)$.

4.4 Calculation of the Cost-of-Living Index: (ii) the Superlative Index

Diewert's (1976) concept of the "superlative price index" is fundamental to developing a practical cost-of-living index, and it also underpins the alternative approach to measuring substitution bias, using a superlative index instead of an index derived from an expenditure function: for example, see Kokowski (1987); Manser and McDonald (1988); Aizcorbe and Jackman (1993).

The superlative index is defined as any index that is "exact for (that is, consistent with) a homothetic preference function that can approximate arbitrary homothetic preferences" (Diewert, 1998, p. 48). Diewert shows that certain "superlative" statistical price indices, such as Fisher's index and Törnqvist's index, may be used to approximate a true index of the cost of living. Fisher's index is the geometric mean of the Laspeyres and Paasche indices, and is an exact true index if the underlying preference ordering is quadratic (Diewert, 1981, p. 184). Törnqvist's index \mathcal{T} is defined by $\ln(\mathcal{T}) = \sum \bar{w}_i \ln(p_i/p_{0i})$,⁷ where the \bar{w}_i are the arithmetic means of base- and comparison-period expenditure weights, and p_{0i} and p_i are base- and comparison-period prices, respectively. Törnqvist's index is an exact true index if $\ln[e(\mathbf{p}, u)]$ is quadratic in utility and the logs of prices (Diewert, 1981, p. 191). For further details see Appendix A1.

The Fisher and Törnqvist indices have a number of desirable properties, in addition to those possessed by all superlative indices. Notably, they measure the rate of inflation in an intuitively appealing manner: the Fisher measures it as the arithmetic mean of inflation in the two constituent indices, and for the Törnqvist it is the arithmetic mean of inflation in each of the constituent price-series, weighted by the \bar{w}_i . These indices are discussed further in Diewert (1998, pp. 48-49), and in references contained therein.

Despite their desirable properties, superlative indices have two problematical features. First, Diewert's approximation result is based on the assumption of homotheticity, which implies that all income elasticities of demand are unity, so that the fraction of income spent on each good remains constant as income varies. This is inconsistent with most expenditure data: for

⁷ Törnqvist's index is a weighted geometric mean of price-relatives. Since January 1999, the BLS has used a similar formula to combine prices of individual items at the most disaggregated level, which differs from the Törnqvist index in weighting by base rather than average expenditure shares (Moulton, 1993, Dalton *et al.*, 1998, and Abraham, 2003). In Ireland, the CSO uses unweighted geometric averaging.

example, see Deaton and Muellbauer (1980b, p. 144, and also pp. 78 and 173). Secondly, there is the practical difficulty that computation of superlative indices requires information on quantity (or expenditure) data for each commodity at each date, alongside the price data: it is the incorporation of comparison-period consumption data that allows superlative indices to capture substitution effects. These data may not be available at all – which is why Laspeyres indices are so widely used. Alternatively, the data may be available after a lag, but an essential requirement of a consumer price index is that it should be published in a timely manner. To the extent that a price index is used to index other economic variables, later revision will be unhelpful and perhaps unwelcome to many users (see Abraham, 2003, pp. 50-51). For example, from July 2002 the BLS has published an index in which the lower-level (geometric) stratum indices are aggregated using a Törnqvist index, but the final version of this index is only available after a two-year lag (Schultze, 2003, p. 6).⁸

V SUBSTITUTION BIAS: EMPIRICAL RESULTS

We focus primarily on the first approach to constructing a cost-of-living index, using the expenditure function. This allows us to exploit the quarterly and (from 1997) monthly frequency of CPI data, and to make valid comparisons with the CPI itself, based on the same data. Furthermore, the approach allows us to explore cost-of-living indices for different income groups. However, we also present annual estimates of superlative indices based on NIE data: for comparison, and also because of the potential importance of these indices as alternatives to fixed-weight indices.

5.1 Calculation of Substitution Bias for Household at Average Expenditure Level

Table 2 sets out annual values for the cost-of-living index $C=100 \times e(\mathbf{p}, u) / e(\mathbf{p}_0, u)$, along with a Laspeyres index L based on predicted AIDS expenditure weights for 1987.⁹ These relate to a household at average expenditure as reported in the Household Budget Survey, 1987. The indices are close: i.e. substitution bias is generally small, as shown in Table 2. The bias reaches a maximum of 0.47 per cent in 1998, and except in 1997-1999 the

⁸ For a critical discussion of superlative indices, see Deaton (1998, pp. 41-42). See also Shapiro and Wilcox (1997) on the use of the constant elasticity of substitution price index, to produce an early estimate of a superlative index.

⁹ The upper bound on the cost-of-living index is the Laspeyres index based on the AIDS model's predicted weights. See Braithwait (1980, p. 70, note 10).

bias is always below 0.3 per cent in any year. Overall growth multiples of C and L are respectively 46.4 per cent and 46.8 per cent, 1987-2001, which translate into annual average (geometric mean) growth rates of 2.76 per cent and 2.78 per cent respectively. Consequently, the average bias is 0.02 per cent per annum, 1987-2001.¹⁰ On an annual average basis the bias is always positive, as expected, and as Table 2 reveals. The only exceptional month is November 1990, when the computed Laspeyres value (L=110.7) lies below the true index value (C=110.8). Presumably this arises through a departure of the computed expenditure function from concavity (see Appendix A3, and Madden, 1993, p. 117).

Table 2: *True Cost-of-Living Index C_t Versus Computed Laspeyres Index L_t , Annual Data, 1985-2001^a*

t	Cost-of-Living Index C_t	Computed Laspeyres Index L_t	Consumer Price Index CPI_t	Substitution bias $\frac{L_t - C_t}{C_t} \times 100\%$	Annual average % bias, 1987 to t	Excess of CPI $\frac{CPI_t - C_t}{C_t} \times 100\%$
1985	93.3	93.4	93.4	0.11	0.054	0.15
1986	97.0	97.0	97.0	0.02	0.021	-0.02
Base 1987	100.0	100.0	100.0			
1988	102.3	102.3	102.1	0.01	0.009	-0.12
1989	106.5	106.5	106.3	0.01	0.007	-0.14
1990	109.6	109.7	109.9	0.02	0.008	0.19
1991	113.0	113.1	113.4	0.04	0.009	0.29
1992	116.4	116.6	116.9	0.12	0.025	0.41
1993	118.1	118.3	118.6	0.15	0.024	0.38
1994	120.9	121.2	121.3	0.19	0.027	0.28
1995	123.8	124.2	124.4	0.27	0.034	0.43
1996	126.1	126.4	126.5	0.27	0.030	0.32
1997	127.6	128.1	128.3	0.36	0.036	0.57
1998	130.4	131.0	131.4	0.47	0.042	0.82
1999	132.7	133.1	133.6	0.31	0.026	0.68
2000	140.1	140.4	141.0	0.16	0.013	0.64
2001	146.4	146.8	147.9	0.28	0.020	1.01

^a $C_t = 100 \times e^{(p,u)/e(p_0,u)}$. C_t and L_t are derived from Madden's AIDS Model 3, using CPI price data.

¹⁰ For comparisons such as these, the usual summary statistic is the annual average percentage bias, i.e. the excess annual average growth in L relative to C over the period of length T. If $L_0 = C_0 = 100$, then the annual average bias is equivalent to $(L_T/C_T)^{1/T} - 1$, or $(\ln(L_T) - \ln(C_T))/T$.

For comparison, Table 2 also shows the official CPI series, rebased to 1987. It tracks the computed Laspeyres index closely, the annual difference being below one index point in each year except 2001, when it reaches 1.1 points (0.75 per cent). The percentage excess of the CPI over the true cost-of-living index is always below one per cent in absolute value except in 2001, and is always positive after 1989.

5.2 Superlative Indices

NIE data provide expenditure weights for each year for seventeen commodity groups (after excluding housing), and a price-series may be derived in each case as described in Section I. From these data Fisher and Törnqvist price indices F_t and \mathcal{T}_t may be calculated for 1985 to 2001. They are shown in Table 3, along with Laspeyres and Paasche indices L_t^N and P_t derived from the same data. The superscript N distinguishes this Laspeyres index from the index computed in Section 5.1.

Table 3 shows that $F < L^N$ and $\mathcal{T} < L^N$ in 2001, with annual average biases in L^N of 0.153 and 0.102 per cent relative to F and \mathcal{T} respectively, 1987-2001. From Table 2 the annual average bias for the derived Laspeyres is 0.020 per cent, relative to the true index C of the cost of living, so if we were to look only at 2001, we might suspect the “demand system” method of underestimating substitution bias. However, a second striking feature of the data in Table 3 conflicts with this conclusion. After 1987 $P_t > L_t^N$ in six out of ten years up to 1999. This is unusual, and it follows that $F_t > L_t^N$ in those years. Similarly, in eight of twelve years to 1999, $\mathcal{T}_t > L_t^N$. Thus, neither of these superlative indices is a satisfactory approximation to a cost of living index for these data.¹¹

5.3 Comparisons with Earlier Studies

Section 5.1 reports an annual average bias of 0.02 per cent. This compares closely with Irvine and McCarthy (1978), from which a figure of 0.02 per cent to 0.03 per cent per annum, 1968-1974, may be extracted from their Table 2 (Irvine and McCarthy, 1978, p. 160). These are small numbers compared with the results reported in a number of UK and US studies, using various methodological approaches, no doubt partly because of differences in aggregation levels.

At a lower level of aggregation, we should expect higher estimates of substitution bias, reflecting greater possibilities of substitution within commodity groups than between them. For example, Braithwait (1980, p. 70),

¹¹ Similar results emerge when we merge some of the expenditure categories to arrive at the ten groups mentioned in Section I, and apply either CPI data or NIE price data. On the relation between the Laspeyres and and Törnqvist indices, see Appendix A1.

Table 3: *Price indices based on NIE expenditure data. Annual, 1985-2001.*^a

<i>t</i>	<i>Laspeyres</i> <i>Index</i> L_t^N	<i>Paasche</i> <i>Index</i> P_t	<i>Fisher's</i> <i>Index</i> F_t	<i>Törnqvist's</i> <i>Index</i> \mathcal{T}_t	<i>Annual</i> <i>Average %</i> <i>bias,^b 1987-t</i> $L_t^N \vee F_t$	<i>Annual</i> <i>Average %</i> <i>bias,^b 1987-t</i> $L_t^N \vee \mathcal{T}_t$
1985	88.6	88.5	88.5	88.5	0.029	0.032
1986	92.6	92.6	92.6	92.6	-0.006	-0.003
Base 1987	100.0	100.0	100.0	100.0		
1988	103.7	103.7	103.7	103.7	0.001	0.001
1989	107.4	107.4	107.4	107.4	0.016	0.017
1990	114.5	114.4	114.5	114.5	0.020	0.003
1991	117.3 ^c	117.3^c	117.3^c	117.4	-0.003	-0.019
1992	120.1	119.9	120.0	120.1	0.017	0.002
1993	122.3	122.4	122.3	122.4	-0.007	-0.022
1994	125.6	126.1	125.9	126.0	-0.027	-0.041
1995	128.3	128.9	128.6	128.8	-0.030	-0.044
1996	130.7	131.0	130.9	131.1	-0.013	-0.028
1997	132.4	133.0	132.7	132.9	-0.024	-0.041
1998	135.7	135.7	135.7	136.0	0.001	-0.019
1999	138.9	138.6	138.8	139.2	0.012	-0.016
2000	145.5	141.4	143.4	144.1	0.110	0.076
2001	151.5	145.2	148.3	149.4	0.153	0.102

^aIndices computed from seventeen expenditure categories: 1. *Food*; 2. *Non-alcoholic beverages*; 3. *Alcohol*; 4. *Tobacco*; 5. *Clothing & footwear*; 6. *Fuel & power*; 7. *Household durables*; 8. *Household non-durables*; 9. *Personal transport eqpt.*; 10. *Operation of personal transport eqpt.*; 11. *Public transport*; 12. *Communication*; Recreation, entertainment and education: 13. *Eqpt.* and 14. *Services*; Misc. goods & services: 15. *Professional services*, 16. *Goods*, and 17. *Other services*. *Housing* is excluded – see Section I.

^bFor the method of calculating annual average percentage bias, see footnote 10.

^cTo two places of decimals, in 1991 $L^N < F < P$.

After 1987, bold type indicates those cases where the Paasche or Fisher or Törnqvist index exceeds the Laspeyres index in value.

using the methodology of Section 5.1 with a different demand specification and less aggregation, reports an average substitution bias of 0.07 per cent per annum for the USA during 1958-1973. Other studies using the “superlative index” approach produce estimates of annual average substitution bias close to 0.2 per cent for the USA. These include Kokowski (1987, cited in Aizcorbe and Jackman, 1993, p. 25): 0.16 per cent per year, 1972-1980; Manser and McDonald (1988, pp. 909 and 921): 0.18 per cent per year, 1959-1985; Aizcorbe and Jackman (1993, p. 29): 0.2 per cent per year, 1982-1991). Boskin *et al.* (1996, p. 53; 1998 pp. 9-11) estimate the “upper-level substitution bias” as 0.15

per cent per annum,¹² basing their conclusion on a number of studies that draw on both methodologies. Moulton (1996, p. 160) lists the results of fourteen recent US studies, showing estimates of overall bias typically of at least 0.5 per cent. Finally, using revealed preference methods, Blow and Crawford (2001) calculate nonparametric bounds to the true index for the UK, 1976-1997. They estimate (p. F378) that the Retail Price Index overstates inflation by up to 3.2 per cent over that period, which corresponds to an annual average bias of up to 0.15 per cent.

5.4 Implications for Consumer Demand

The estimate of substitution bias presented here is remarkably similar to that of Irvine and McCarthy, despite the separation in time and the difference in functional form. Moreover, at the given level of aggregation (nine or ten commodity groups), both estimates suggest low elasticities of substitution. To take this much further we should need estimates of the Morishima elasticities (Blackorby and Russell, 1989), and to compute these we should need values for compensated own-price and cross elasticities of demand. However, some of the evidence in Section 5.2 is illuminating. There we found that the Paasche index lies above the Laspeyres in six out of ten years up to 1999 for the seventeen-group NIE data, and the same is true during a substantial part of 1997-1999 for ten-group CPI or NIE price data. If the reference-period (1987) and comparison-period ($t > 1987$) weights were identical, so would be the Laspeyres and Paasche values in the comparison period. Now consider the case of homothetic preferences, where there are no income effects on expenditure shares. If substitution effects are relatively large, then we expect the Laspeyres index for the comparison period to give excess weight, relative to the Paasche, to groups that have experienced relatively large price increases, so that $L > P$. When we observe $P > L$, one possibility is a combination of homothetic preferences with small substitution effects: i.e. small own-price substitution effects, low levels of net substitutability, with perhaps significant net complementarity.¹³ An alternative and likely possibility is that preferences may be nonhomothetic and that, whatever may be the substitution effects, these are swamped by income effects, arising both from changes in income and from price changes.

¹² "Upper level bias": at a fairly high level of aggregation, as in this paper.

¹³ The condition for the relative expenditure share w_i/w_j to fall when p_i rises is that the Morishima elasticity $M_{ij}(p,u) > 1$. The statement in the text follows from the relation $M_{ij} = \varepsilon_{ij} - \varepsilon_{ji}$ where ε_{ij} and ε_{ji} are compensated price elasticities (Blackorby and Russell, 1989, pp. 885-886). On the empirical relationship between Paasche and Laspeyres indices, see National Research Council, 2002, pp. 22-23.

5.5 *The Impact of Price-Changes on Different Income Groups*

The report of the Household Budget Survey for 1987 includes estimates of average weekly expenditure for each income decile in the sample (CSO, 1989, Table 2). These data may be used to construct true cost-of-living indices for each decile, and the index values for selected months are set out in Table 4. From 1987 until the end of 1994, the indices rise approximately in step across the income range, with slightly less growth at the bottom and more at the top. From that point, there is a negative correlation between income-rank and the index, and the spread grows over time. By December 2001, the range between the indices for the first and tenth deciles represents 6.1 per cent of the index for the mean household, with the first decile, at 154.6, lying 3.8 per cent above the mean, and the tenth, at 145.5, lying 2.3 per cent below it. This represents an excess annual average growth rate for the first decile's cost-of-living index of about 0.4 per cent, 1987-2001, compared with the tenth decile.

The differential impact of inflation across the income range arises from disproportionate exposure of poor households to the more rapidly inflating commodity prices. For the six commodity groups that have the greatest overall growth in prices,¹⁴ the cumulative predicted expenditure weights (1987-2001) are in the range [0.82, 0.90] for households in the first decile, and only [0.73, 0.81] for the tenth decile. The main factors are food and tobacco. During 1987-2001, the ranges for the predicted weights on food (the sixth group in order of price increase) for first and tenth decile households are [0.40, 0.49] and [0.09, 0.20] respectively. Tobacco prices grew far faster than any others, and while the tenth decile's predicted weight on tobacco is 0.02 at most, and often zero, the first decile's is in the range [0.10, 0.12].

In contrast with these results, Irvine and McCarthy (1978, p. 160) find growth of the cost of living to be related *positively* to income, but the excess growth rate for the highest income group, compared with the lowest, is only 0.14 per cent per annum 1968-1974, on average. There are several possible explanations for this. First, Irvine and McCarthy use NIE price data, versus CPI data here. However, if Table 4 is re-computed using NIE data, the pattern across the deciles is the same, albeit with a smaller dispersion. Second, relative growth of prices in their sample period was very different from 1987-2001. Fuel prices grew fastest in 1968-74, and were third from slowest in 1987-2001. Tobacco prices accelerated, from last to first place, and drink prices accelerated from sixth to second place. Thirdly, Irvine and McCarthy begin with a different demand model from this paper, with parameter estimates

¹⁴ Overall growth in consumer price indices, 1987 (average) to December 2001: Tobacco: +109%; alcohol: +70%; services: +66%; transport (excl. petrol) and equipment: +57%; other goods: +54%; food: +52%. The remaining groups are:- durables: +33%; fuel and power: +29%; petrol: +10%; clothing and footwear: -19%.

from a different although overlapping sample period (1953-74, compared with 1958-88 for Madden).

Table 5: *Variation of True Index C_t Across Seven Income Groups Used in Irvine and McCarthy (1978)*

t	<i>Lowest income group</i> C_{1t}	<i>2nd</i> C_{2t}	<i>3rd</i> C_{3t}	<i>4th</i> C_{4t}	<i>5th</i> C_{5t}	<i>6th</i> C_{6t}	<i>Highest income group</i> C_{7t}
1968	100	100	100	100	100	100	100
1974, from Irvine & McCarthy (1978, p. 160). ^a	179.5	180.1	180.3	180.4	180.8	181.0	181.0
1974: based on Madden's AIDS model, Irvine & McCarthy's data.	173.1	173.9	174.5	174.9	176.4	178.3	179.7

^a These data relate to the base-year utility level in Irvine and McCarthy, who also provide (p. 162) results based on the current-year utility level, showing a similar pattern across the income range.

Table 5 sets out the result of using Madden's AIDS model to construct cost-of-living indices for Irvine and McCarthy's seven income groups, using their data. This produces a similar pattern to Irvine and McCarthy: the index values for each group are lower, and the dispersion is wider, but the exercise replicates the positive correlation between income rank and growth in the cost of living, 1968-74. This suggests that the ranking differences between Irvine and McCarthy's findings for 1968-74 and those reported in Table 4 are due largely to differences in the behaviour of prices in the two periods, rather than to differences between the estimated demand models.

VI CONCLUSIONS

From 1985 to 2001, when annual average growth of the CPI was 2.9 per cent, there was significant relative movement among the price indices for ten main categories of consumption expenditure. Nevertheless, we find only a small substitution bias, of 0.02 per cent per annum on average, which is consistent with the conclusions of Irvine and McCarthy (1978) for 1968-1974, and which is much smaller than the estimate of 0.15 per cent for upper-level bias that is presented in the Boskin report. The estimate of substitution bias is based on a comparison between a true index of the cost of living and a Laspeyres price index (not the CPI), both derived from Madden's estimated

AIDS demand system. However, except in 2001, the annual average values of the computed Laspeyres index and the CPI are within one index point of each other.

The Paasche index estimated on NIE data exceeds the Laspeyres index for almost half of the period after 1987, which supports the contention that substitution effects and substitution biases are small at this level of aggregation. Substitution bias of the measured size will have a limited economic significance, but further research, which would be facilitated by the use of scanner data, is necessary to establish the extent of lower-level substitution bias. To reach any conclusions for Ireland about the overall impact of bias along the lines pioneered by Boskin for the US, we should also need estimates of its other components, i.e. from quality change and new goods.

Superlative price indices are estimated here using expenditure data from NIE, but these exceed the associated Laspeyres index L^N for a significant part of the sample period, and are therefore not good approximations to a cost of living index for these data. This suggests that the superlative index approach might not provide a satisfactory cost-of-living index for Ireland if it were applied only at a high level of aggregation.

Irvine and McCarthy report a positive relation between income-rank and the impact of inflation, during a six-year period (1968-1974) in which the CPI rose by about 80 per cent. However, across the income spectrum the dispersion of changes in the cost of living is quite small. That finding for 1968-74 is replicated by applying the *AIDS*-based cost-of-living index to Irvine and McCarthy's data. By comparison, this paper finds a negative relation between income-rank and the impact of inflation after 1994, with a comparatively wide dispersion towards the end of a fourteen-year period (1987-2001) in which the rise in the CPI was below 50 per cent. The explanation for this turnaround in the impact of inflation lies in changes between the two periods in the relative growth-rates of prices for the different expenditure groups, with for example tobacco prices moving from slowest-growing to fastest. This disproportionate impact across the range of income clearly raises important policy issues for indirect taxation.

REFERENCES

- ABRAHAM, KATHERINE G., 2003. "Toward a Cost-of-Living Index: Progress and Prospects", *Journal of Economic Perspectives*, Vol. 17, No. 1, Winter, pp. 45-58.
- AIZCORBE, ANA M., and PATRICK C. JACKMAN, 1993. "The Commodity Substitution Effect in CPI Data, 1982-91", *Monthly Labor Review*, Vol. 116, No. 12, December, pp. 25-33.
- BLACKORBY, CHARLES, and R. ROBERT RUSSELL, 1989. "Will the Real Elasticity

- of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities)", *American Economic Review*, Vol. 79, No. 4, September, pp. 882-888.
- BLOW, LAURA, and IAN CRAWFORD, 2001. "The Cost of Living with the RPI: Substitution Bias in the UK Retail Prices Index", *The Economic Journal*, Vol. 111, No. 472, June, pp. F357-F382.
- BOSKIN, MICHAEL. J., ELLEN R. DULBERGER, ROBERT J. GORDON, ZVI GRILICHES, and DALE W. JORGENSON, 1996. *Towards a More Accurate Measure of the Cost of Living: Final Report to the US Senate Finance Committee*, Washington: United States Senate, 4 December. Reprinted in Dean Baker (ed.), *Getting Prices Right*, London: M.E. Sharpe for The Economic Policy Institute, 1998, pp. 5-77.
- BOSKIN, MICHAEL. J., ELLEN R. DULBERGER, ROBERT J. GORDON, ZVI GRILICHES, and DALE W. JORGENSON, 1998. "Consumer Prices, the Consumer Price Index, and the Cost of Living", *Journal of Economic Perspectives*, Vol. 12, No. 1, Winter, pp. 3-26.
- BRAITHWAIT, STEVEN D., 1980. "Substitution Bias of the Laspeyres Price Index: An Analysis Using Estimated Cost-of-Living Indexes", *American Economic Review*, Vol. 70, No. 1, March, pp. 64-77.
- CASEY, MICHAEL G., 1973. "An application of the Samuelson-Stone Linear Expenditure System to Food Consumption in Ireland", *The Economic and Social Review*, Vol. 4, No. 3, April, pp. 309-342.
- CENTRAL STATISTICS OFFICE (CSO), online. Data for *Consumer Price Index and National Income and Expenditure*, available from http://www.eirestat.cso.ie/6_filelist.html [last accessed September 2003].
- CENTRAL STATISTICS OFFICE (CSO), 1989. *Household Budget Survey 1987, Volume I: Detailed Results for All Households*, Dublin: The Stationery Office.
- CENTRAL STATISTICS OFFICE (CSO), 1997a. "Consumer Price Index (CPI): Introduction of Updated Monthly Series (Base: November 1996=100)", *Statistical Bulletin*, Vol. 72, No. 2, June, pp. 257-278.
- CENTRAL STATISTICS OFFICE (CSO), 1997b. *Household Budget Survey 1994-1995, Volume I: Detailed Results for All Households*, Dublin: The Stationery Office.
- CENTRAL STATISTICS OFFICE (CSO), 2001. *Household Budget Survey 1999-2000 – Preliminary Results*, Dublin: The Stationery Office.
- CENTRAL STATISTICS OFFICE (CSO), 2003. "Consumer Price Index (CPI): Introduction of Updated Series", *Statistical Bulletin 2002*, Vol. 77, pp. 37-64.
- CONNIFFE, DENIS, and JOHN EAKINS, 2003. "Does the Stochastic Specification of the Linear Expenditure System Matter?", *The Economic and Social Review*, Vol. 34, No. 1, Spring, pp. 23-32.
- CONNIFFE, DENIS, and A. HEGARTY, 1980. "The Rotterdam System and Irish Models of Consumer Demand", *The Economic and Social Review*, Vol. 11, No. 2, January, pp. 99-112.
- DALTON, KENNETH V., JOHN S. GREENLEES, and KENNETH J. STEWART, 1998. "Incorporating a Geometric Mean Formula into the CPI", *Monthly Labor Review*, Vol. 121, No. 10, October, pp. 3-7.
- DEATON, ANGUS, 1998. "Getting Prices Right: What Should be Done?", *Journal of Economic Perspectives*, Vol. 12, No. 1, Winter, pp. 37-46.
- DEATON, ANGUS, and JOHN MUELLBAUER, 1980a. "An Almost Ideal Demand System", *The American Economic Review*, Vol. 70, No. 3, June, pp. 312-326.

- DEATON, ANGUS, and JOHN MUELLBAUER, 1980b. *Economics and Consumer Behaviour*, Cambridge: Cambridge University Press.
- DIEWERT, W. ERWIN, 1976. "Exact and Superlative Index Numbers", *Journal of Econometrics*, Vol. 4, No. 2, May, pp. 115-145.
- DIEWERT, W. ERWIN, 1981. "The Economic Theory of Index Numbers: A Survey", in Angus Deaton (ed.), *Essays in the Theory and Measurement of Consumer Behaviour, in Honour of Sir Richard Stone*, Cambridge: Cambridge University Press, pp. 163-208.
- DIEWERT, W. ERWIN, 1998. "Index Number Issues in the Consumer Price Index", *Journal of Economic Perspectives*, Vol. 12, No. 1, Winter, pp. 47-78.
- EUROSTAT, 2001. *Compendium of HICP Reference Documents*, (2/2001/B/5), Brussels: Commission of the European Communities.
- FIELDING, DAVID P., and PAUL MIZEN, 2000. "Relative Price Variability and Inflation in Europe", *Economica*, Vol. 67, No. 265, February, pp. 57-78.
- GOVERNMENT OF IRELAND, 1997. *Budget 1997*, Dublin: The Stationery Office.
- GOVERNMENT OF IRELAND, 2002. *Budget 2003*, Dublin: The Stationery Office.
- GOVERNMENT OF IRELAND, 2003. *2002: Statistical Information on Social Welfare Services*, Dublin: The Stationery Office.
- IRVINE, IAN, and COLM MCCARTHY, 1978. "Expenditure Patterns and the Welfare Effects of Inflation: Estimates of a 'True' Cost-of-Living Index", *The Economic and Social Review*, Vol. 9, No. 2, January, pp. 157-165.
- JARAMILLO, CARLOS F., 1999. "Inflation and Relative Price Variability: Reinstating Parks's Results", *Journal of Money Credit and Banking*, Vol. 31, No. 3, Part 1, August, pp. 375-385.
- JORGENSON, DALE, and DANIEL T. SLESNICK, 1983. "Individual and Social Cost-of-Living Indexes", in W.E. Diewert and C. Montmarquette (eds.), *Price Level Measurement*, Ottawa: Statistics Canada, pp. 241-336.
- KOKOWSKI, MARY F., 1987. "Consumer Price Indexes by Demographic Groups", Bureau of Labor Statistics (Washington DC) Working Paper No. 167.
- KONÜS, A.A., 1939. "The Problem of the True Index of the Cost of Living", *Econometrica*, Vol. 7, No. 1, January, pp. 10-29, reprinted in translation from the Russian. First published in *The Economic Bulletin of the Institute of Economic Conjecture* (Moscow), No. 9-10 (36-37), September-October 1924, pp. 64-71.
- MADDEN, DAVID, 1993. "A New Set of Consumer Demand Estimates for Ireland", *The Economic and Social Review*, Vol. 24, No. 2, January, pp. 101-123.
- MADSEN, JAKOB B., 1996. "Formation of Inflation Expectations: From the Simple to the Rational Expectations Hypothesis", *Applied Economics*, Vol. 28, No. 10, October, pp. 1331-1337.
- MCCARTHY, COLM, 1977. "Estimates of a System of Demand Equations Using Alternative Commodity Classifications of Irish Data, 1953-1974", *The Economic and Social Review*, Vol. 8, No. 3, April, pp. 201-211.
- MANSEY, MARILYN E., and RICHARD J. MCDONALD, 1988. "An Analysis of Substitution Bias in Measuring Inflation, 1959-85", *Econometrica*, Vol. 56, No. 4, July, pp. 909-930.
- MOULTON, BRENT R., 1993. "Basic Components of the CPI: Estimation of Price Changes", *Monthly Labor Review*, Vol. 116, No. 12, December, pp. 13-24.
- MOULTON, BRENT R., 1996. "Bias in the Consumer Price Index: What is the Evidence?", *Journal of Economic Perspectives*, Vol. 10, No. 4, Fall, pp. 159-177.

- NATIONAL RESEARCH COUNCIL, 2002. *At What Price? Conceptualising and Measuring Cost-of-Living and Price Indexes*. Charles Schultze and Christopher Mackie, (eds.), Washington DC: National Academy Press.
- O'RIORDAN, WILLIAM K., 1976. "The Demand for Food in Ireland 1947-1973", *The Economic and Social Review*, Vol. 7, No. 4, July, pp. 401-415.
- PARKS, RICHARD W., 1978. "Inflation and Relative Price Variability", *Journal of Political Economy*, Vol. 86, No. 1, February, pp. 79-95.
- SCHULTZE, CHARLES L., 2003. "The Consumer Price Index: Conceptual Issues and Practical Suggestions", *Journal of Economic Perspectives*, Vol. 17, No. 1, Winter, pp. 3-22.
- SHAPIRO, MATTHEW D., and DAVID W. WILCOX, 1997. "Alternative Strategies for Aggregating Prices in the CPI", *The Federal Reserve Bank of St. Louis Review*, Vol. 79, No. 3, May/June, pp. 113-125.
- THEIL, HENRI, 1967. *Economics and Information Theory*, Amsterdam: North-Holland.
- TRIPLETT, JACK E., 2001. "Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?", *The Economic Journal*, Vol. 111, No. 472, June, pp. F311-F334.

APPENDIX

A1. Price Indices: Laspeyres, Paasche, Fisher, Törnqvist

The Laspeyres price index is $L_\tau = \sum_i p_{\tau i} x_{0i} / \sum_i p_{0i} x_{0i} = \sum_i w_{0i} r_{\tau i}$. The first subscript indexes time t , i indexes goods, and p_{ti} , x_{ti} , w_{ti} and r_{ti} are respectively price, quantity, expenditure weight and the price-relative p_{ti}/p_{0i} . The base- and comparison-periods for prices are $t=0$ and $t=\tau$ respectively. In a modified Laspeyres index the weights are w_{bi} where the base period $t=b$ is not the reference period $t=0$. The Paasche price index is $P_\tau = \sum_i p_{\tau i} x_{\tau i} / \sum_i p_{0i} x_{\tau i} = 1 / (\sum_i w_{\tau i} / r_{\tau i})$ and the Fisher index is $F_\tau = \sqrt{(L_\tau P_\tau)}$, i.e. the geometric mean. Clearly F lies between P and L .

If $E_{\tau i}$ and $E_{\tau i}^0$ are NIE expenditure data for category i at time τ , for current ($t=\tau$) and base ($t=0$) prices respectively, then $E_{\tau i} = \sum_j p_{\tau ij} x_{\tau ij}$ and $E_{\tau i}^0 = \sum_j p_{0ij} x_{\tau ij}$ where j indexes the components of i . If we define $I_{\tau i}$ as $E_{\tau i} / E_{\tau i}^0$, then $I_{\tau i} = \sum_j p_{\tau ij} x_{\tau ij} / \sum_j p_{0ij} x_{\tau ij}$, which is a Paasche price index, as is $I_\tau = \sum_i E_{\tau i} / \sum_i E_{\tau i}^0 = \sum_{ij} p_{\tau ij} x_{\tau ij} / \sum_{ij} p_{0ij} x_{\tau ij}$. The Törnqvist index is defined by $\ln(\mathcal{T}_\tau) = \sum_i \bar{w}_{\tau i} \ln(r_{\tau i})$ where the $\bar{w}_{\tau i}$ are arithmetic means of the weights at $t=0$ and $t=\tau$. By strict concavity of $\ln()$, $\sum_i w_i \ln(r_{\tau i}) < \ln(\sum_i w_i r_{\tau i})$ for any set of positive fractional weights w_i with $\sum_i w_i = 1$. Applying the weights for $t=0$ and $t=\tau$, and adding the resulting inequalities with weights $1/2$ and $1/2$, we have

$$\begin{aligned} \ln(\mathcal{T}_\tau) &= \frac{1}{2} \sum w_{0i} \ln(r_{\tau i}) + \frac{1}{2} \sum w_{\tau i} \ln(r_{\tau i}) \\ &< \frac{1}{2} \ln(\sum w_{0i} r_{\tau i}) + \frac{1}{2} \ln(\sum w_{\tau i} r_{\tau i}) = \ln(\sqrt{L_\tau \sum w_{\tau i} r_{\tau i}}) \end{aligned}$$

so that $\mathcal{T}_\tau < \sqrt{L_\tau \sum w_{ti} r_{ti}}$. If the w_{ti} and w_{0i} are close, then $\sum w_{ti} r_{ti} \approx L_\tau$ and $\mathcal{T}_\tau < L_\tau$. However, this inequality does not hold in general, and it is possible to have $\mathcal{T}_\tau > L_\tau$.

A2. Measuring Changes in Relative Price Levels

In levels, we might use the variance $\sigma^2(p_{tj})$ among components of \mathbf{p}_t as a measure of changes in relative prices. However, σ^2 changes when all prices change by a scalar multiple. One solution is to normalise in each year relative to the price of an arbitrarily-chosen numeraire. We adopt the alternative, of mapping price-vectors into the unit simplex. This avoids an arbitrary choice of numeraire; it is standard in microeconomic theory; it has a natural geometric interpretation; and distances on the simplex have a ready statistical interpretation.

We need a measure of deviations between points in S_N . A difficulty is that individual series either consist of index numbers, and are unique only up to a scalar multiple; or they consist of prices, and reflect arbitrary choices of units of measurement, and again are unique only up to a scalar multiple. If \mathbf{P} is a $(T+1) \times N$ matrix of observations on price indices (or prices) for N goods in periods $0, 1, \dots, T$, then \mathbf{PB} is an equally valid *representation*, given an appropriate diagonal $N \times N$ matrix \mathbf{B} with each $b_{jj} > 0$: e.g. $b_{jj} = 1/p_{tj}$, each j , re-setting the base period at τ . Normalising at each date creates a matrix $\mathbf{Q} = \mathbf{APB}$, where \mathbf{A} is $\text{diag}\{a_{tt}\}$ and normalises to the simplex, and each row of \mathbf{Q} is a normalised price-vector. If $\mathbf{p}_{\tau+r}$ is a scalar multiple of \mathbf{p}_τ , these map to the same point in S_N . Finally, each a_{tt} depends on the b_{jj} as well as on the components of \mathbf{p}_t . We identify movements in relative prices with changes in the distances $d(\mathbf{q}_t, \mathbf{q}_r) = \sqrt{\sum_j (q_{tj} - q_{rj})^2}$ between successive rows of \mathbf{Q} . If $d=0$ in some representation \mathbf{PB} , then $d=0$ in any representation. Otherwise d depends on the representation of prices, i.e. on \mathbf{B} . Thus, interpretations of Euclidean distances should be made cautiously.

We may write $d(\mathbf{q}_t, \mathbf{q}_r) = \sqrt{\sum_j [(q_{tj} - \bar{q}_t) - (q_{rj} - \bar{q}_r)]^2}$, because $\bar{q}_t = \bar{q}_r = 1/N$, or $d(\mathbf{q}_t, \mathbf{q}_r) = \sqrt{[(N-1)(s_t^2 + s_r^2 - 2s_{tr})]}$. In particular $d(\mathbf{q}_t, \mathbf{q}_c) = \sqrt{[(N-1)s_t^2]}$ where one point is the centre \mathbf{q}_c of the simplex, i.e. the base-year point $(1/N, 1/N, \dots, 1/N)$. From \mathbf{q}_c to the mid-point of any of the $(N-2)$ -dimensional sides of the simplex, where one coordinate is zero and all others are $1/(N-1)$, the distance is $1/\sqrt{N(N-1)}$.

A3. Substitution Bias and the Concavity of $e(\mathbf{p}, \mathbf{u})$

At any prices the value of the true expenditure function $e(\mathbf{p}, \mathbf{u})$ cannot exceed the cost of the base-year consumption vector, which is one way to buy \mathbf{u} , but is not necessarily least-cost. However, it is possible for the estimated function to violate this condition.

Taylor-expanding $e(\mathbf{p}, u)$ about base prices \mathbf{p}_0 and using Shephard's lemma (i.e. $\frac{\partial e(\mathbf{p}_0, u)}{\partial \mathbf{p}} = \mathbf{x}(\mathbf{p}_0, u)$, the Hicksian demand function), we have: $e(\mathbf{p}, u) = \mathbf{p}'\mathbf{x}(\mathbf{p}_0, u) + \frac{1}{2}(\mathbf{p} - \mathbf{p}_0)'\mathbf{H}^\theta(\mathbf{p} - \mathbf{p}_0)$, where \mathbf{H}^θ is the Hessian of $e()$ evaluated at some $\theta \in (\mathbf{p}, \mathbf{p}_0)$. \mathbf{H} is also the Jacobian of $\mathbf{x}(\mathbf{p}_0, u)$, i.e. the substitution matrix. Because e is concave, \mathbf{H} is negative semi-definite everywhere, so that $e(\mathbf{p}, u) \leq \mathbf{p}'\mathbf{x}(\mathbf{p}_0, u)$ always. Then $C \leq L$, where these are indices relating to the underlying, and unobservable, expenditure and demand functions. However, while $e(\mathbf{p}, u)$ must be concave, the estimated expenditure function \hat{e} might not be, because of random error in parameter estimates, or because the form of the demand system is misspecified. Thus, there may exist prices at which $\hat{e} > \mathbf{p}'\hat{\mathbf{x}}(\mathbf{p}_0, u)$, in which case for the estimated indices we should have: $\hat{C} > \hat{L}$.

Madden (1993, p. 117) reports that the estimated substitution matrices for his various *AIDS* models (estimated in levels) have six negative eigenvalues, and that "In most cases, those eigenvalues which are positive are quite small in size, suggesting that the matrices may not be far from being negative semi-definite'.