

---

---

**MEASUREMENT BIAS IN THE HICP:  
WHAT DO WE KNOW AND  
WHAT DO WE NEED TO KNOW?**

Mark A. Wynne  
Diego Rodriguez-Palenzuela

**Research Department  
Working Paper 0206**



**FEDERAL RESERVE BANK OF DALLAS**

---

---

**MEASUREMENT BIAS IN THE HICP:  
WHAT DO WE KNOW AND  
WHAT DO WE NEED TO KNOW?**

Mark A. Wynne  
Diego Rodriguez-Palenzuela

**Research Department  
Working Paper 0206**



**FEDERAL RESERVE BANK OF DALLAS**

# MEASUREMENT BIAS IN THE HICP: WHAT DO WE KNOW AND WHAT DO WE NEED TO KNOW?

Mark A. Wynne  
*Federal Reserve Bank of Dallas*

Diego Rodriguez-Palenzuela  
*European Central Bank*

**Abstract:** The Harmonized Index of Consumer Prices (HICP) is the primary measure of inflation in the euro area, and plays a central role in the policy deliberations of the European Central Bank (ECB). The ECB defines its Treaty mandate of price stability as “...a year-on-year increase in the Harmonised Index of Consumer Prices (HICP) for the euro area of below 2% [...] to be maintained over the medium term.” Among the rationales given for defining price stability as prevailing at some positive measured inflation rate is the possibility that the HICP as published incorporates measurement errors of one sort or another that may cause it to systematically overstate the true rate of inflation in the euro area. This paper reviews what currently is known about the scope of measurement error in the HICP. We conclude that given the vague conceptual framework of the HICP, the scant research on price measurement issues in the EU and the ongoing improvements in the HICP, there is very little scientific basis at this time for a point (or even an interval) estimate of a positive bias in the HICP.

**Keywords:** Harmonized Index of Consumer Prices, HICP, price measurement.

**JEL Codes:** C43, E31

## 1. Introduction

Fifteen or twenty years ago, few economists outside the community of scholars specialising in these matters were interested in the question of how well national statistical agencies measure inflation. At that time inflation in most countries was still running at rates that clearly could not be attributed to measurement error, but rather to errors on the part of economic and monetary policy makers. Since then, a number of (not

completely unrelated) developments have moved the question of measurement to centre stage.

The first was the remarkable success of central banks in all of the industrialised countries in bringing inflation rates down to levels not seen in decades. The transition to a low inflation environment, one in which measurement error could possibly account for a significant fraction of observed inflation, once again made the question of how well inflation statistics measure what they are supposed to measure of central importance.<sup>1</sup> The second key development was the shift of many central banks around the world towards inflation targeting as a strategy or framework for monetary policy. Inflation targeting shifts the question of inflation measurement to centre stage in the formulation of monetary policy. Indeed, some central banks have eschewed the adoption of formal inflation targets because of the difficulty of accurately measuring inflation in a dynamic economy.<sup>2</sup> The third development was the remarkable rate of productivity growth in the manufacture of computers and related high-tech equipment, and the extraordinary rate of improvement in the quality of the output of this and related sectors. The rapid rate of improvement in the performance of computers and other IT equipment raised the question of how well statistical agencies dealt with quality change in measuring inflation. In some countries (in particular the US) the decision was made to make greater use of hedonic methods to try to control for these quality improvements to ensure that the raw data used to construct aggregate price statistics were capturing pure price changes. And the dramatic changes in quality-adjusted prices that hedonic methods revealed lead to renewed concerns about the possibility of significant substitution bias in traditional fixed-weight Laspeyres-type measures of prices and output, and ultimately to a major overhaul of the US national accounts with the fixed-weight Laspeyres measures of output being replaced by chain-weighted Fisher measures.<sup>3</sup> Recently the failure to account fully for quality improvements in high-tech equipment has been advanced as a possible explanation of sluggish measured investment growth in some European countries.<sup>4</sup> Finally, there has been technical progress in the measurement of inflation itself. This includes the development of large scale price-scanner databases, the drastic reduction in computing costs facilitating the implementation of hedonic methods and analytical

advancements in the economic theory of price indices, like the derivation of exact index numbers under decreasingly restrictive assumptions<sup>5</sup>.

Our objective in this paper is to assess the state of play on price measurement in the EU. In particular we want to pose and start answering the question of whether the Harmonized Index of Consumer Prices (HICP), which is central to the monetary policy deliberations of the European Central Bank (ECB), overstates (or understates) the rate of inflation in the euro area.<sup>6</sup> The ECB has defined its mandate of price stability as “...a year-on-year increase in the Harmonised Index of Consumer Prices (HICP) for the euro area of below 2% ... to be maintained over the medium term.”<sup>7</sup> Among the rationales given for defining price stability as prevailing at some positive measured rate of inflation is the possibility that the HICP as published incorporates measurement errors of one sort or another that may cause it to systematically overstate the true rate of inflation in the euro area.<sup>8</sup> After the publication of the Boskin Report in the US (Boskin et. al. (1996)), several studies were conducted for European countries to assess the extent of mismeasurement in national CPIs. These studies reached different conclusions, but there was and there persists a vague consensus among economists and policy makers that inflation statistics persistently overstate the true rate of inflation by an unspecified amount.

In this paper we will do two things. First, we will review the current state of knowledge about possible measurement biases in consumer price inflation statistics in the EU, and specifically the HICP. We will start by briefly summarising the conceptual framework of the HICP and then sketch the two broad strategies that have been implemented for assessing the extent of measurement bias. We will review the studies that were done for various EU countries following the publication of the Boskin Report in the US, and review the work that has been done since. Second, we will propose some priorities for future research on the accuracy of the HICP. We will argue that the paucity of studies on which to base an assessment of the accuracy of the HICP indicates that there is a lot of work that still needs to be done. Against this background, we will conclude by outlining our view of the research priorities.

Before proceeding, it is important to note that any attempt to assess the extent of measurement bias in a price index inherently entails some attempt to hit a moving target.

The practices of statistical agencies may deliver better or worse results at different points in time depending on economic conditions. For example, substitution bias is less of a problem when relative prices are less volatile, which in turn depends in part on the level of inflation.<sup>9</sup> Measurement problems are also likely to be more severe during episodes of rapid technological change. Even in an economy that was not continually changing, the scope for measurement error varies over time as statistical agencies learn of problems and move to correct them. These considerations apply with even greater force to any attempt to assess measurement bias in the HICP. In many ways, the HICP is a work in progress.

## **2. The Harmonized Index of Consumer Prices (HICP)**

The Harmonized Index of Consumer Prices (HICP) has its origin in the requirement of Article 109j of the Treaty on European Union (the Maastricht Treaty) that “a high degree of price stability” be among the criteria to be used in assessing whether Member States had achieved the necessary high degree of sustainable convergence criterion prior to the launch of EMU.<sup>10</sup> Protocol No. 6 of the Treaty further clarified the price stability criterion, stipulating that convergence consisted of having an inflation rate that “...does not exceed by more than 1½ percentage points that of, at most, the three best performing Member States in terms of price stability.” Furthermore, Protocol No. 6 required that “Inflation shall be measured by means of the consumer price index on a *comparable* basis, taking into account differences in national definitions” (emphasis added). It was this requirement that inflation be measured on a comparable basis that led to the development of the HICP programme. The consumer price indexes of the Member States were deemed inadequate for this purpose, as they differ greatly in terms of their coverage and even in terms of their conceptual frameworks. The creation of the HICP is therefore intimately linked to the single currency project within the EMU and this fact has unavoidably conditioned the properties of the HICP as a price index. In particular, the creation of the HICP was driven by the need to have a measure of consumer price inflation available in time for the assessment of convergence prior to the launch of EMU, and the compromises this process necessitated reflected the need for comparability and an index that would mainly be used for monetary policy purposes. The imperatives that drove the creation of the HICP may explain some of the shortcomings of its conceptual

framework. Diewert (2002) identified at least four approaches to the construction of price index numbers, namely the fixed basket approach, the test or axiomatic approach, the stochastic or statistical approach, and the economic approach, and examines the implications of each for the choice of HICP index number concept.<sup>11</sup> The fixed-basket approach entails comparing the prices of a fixed-basket of goods in two periods to arrive at a measure of the change in the price level. The best-known examples of this approach are the well-known Laspeyres and Paasche price indexes. The test or axiomatic approach seeks an index number formula that satisfies a number of tests (such as invariance to changes in units, monotonicity and so on) that have been proposed over the years. The stochastic or statistical approach views the index number problem as an estimation problem, specifically that of finding the best estimate of the central tendency of the sample of individual price changes observed between two periods. The economic approach is rooted in the neoclassical theory of consumer behaviour, and the price index or cost of living index is the change in the cost of attaining some reference level of utility between two periods. How does the HICP fit into this classification scheme?

According to Eurostat (2001) the HICP is “...a Laspeyres-type price index that is based on the prices of goods and services available for purchase in the economic territory of the Member State for the purpose of directly satisfying consumer needs.” (p. 19, emphasis in original) Given this pricing concept, the coverage of the HICP is defined as “household final monetary consumption expenditure.” Eurostat (2001) further explains “The HICP is not a cost of living index. That is, it is not a measure of the change in the minimum cost for achieving the same ‘standard of living’ (i.e. constant utility) from two different consumption patterns realised in the two periods compared and where factors other than pure price changes may enter the index” (p.19, emphasis in original).<sup>12</sup> The weights used to calculate the HICP may relate to a reference period that is up to seven years prior to the current year<sup>13</sup>, but must be updated each year for “...especially large changes in the expenditure pattern” (p. 20). The HICP was first published in March 1997 and is published each month about two weeks after the end of the month to which it refers. Since November 2001, Eurostat has released a “flash” estimate of the HICP shortly (a couple of days) after the end of the reference month.

Where does this place the HICP in the classification scheme proposed by Diewert? Eurostat states quite explicitly that the HICP is not a cost of living index, so its conceptual framework presumably lies in either the fixed-basket, axiomatic or statistical approaches to index number construction. However, Diewert (2002) shows quite convincingly that this cannot be the case, since all three of these approaches (as well as the economic approach) would rule out the use of the Laspeyres formula for the calculation of the HICP. We will return to the issue of the conceptual framework of the HICP below. For now, it suffices to note that the HICP does not easily fit into any existing approach to the index number problem.

As already noted, the HICP is defined as a Laspeyres-type index. Thus for a country  $j$  the HICP in month  $t$  is defined as

$$HICP_{j,t} = \sum_i \omega_{i,j,b} \left( \frac{p_{i,j,t}}{p_{i,j,r}} \right) \quad (1)$$

where  $p_{i,j,t}$  is the price of the  $i$ 'th commodity in country  $j$  in the comparison month  $t$ ,  $p_{i,j,r}$  is the price in the reference month  $r$ , and  $\omega_{i,j,b}$  is the weight assigned to the  $i$ 'th commodity based on expenditure shares in the expenditure base period  $b$ . As already noted, the weights used to aggregate the elementary aggregates can be based on expenditure patterns up to seven years prior to the year for which the HICP is being compiled, although some countries update the weights annually. The reference period for the HICP sub indexes is 1996=100; for the new sub-indexes introduced in January 2000, the reference period is December 1999=100.

The elementary aggregate indexes, the individual  $p_{i,j,t}$ , or price relatives  $(p_{i,j,t} / p_{i,j,b})$  are calculated as either arithmetic or geometric means<sup>14</sup>. Specifically, the price relatives should be calculated as:

$$(p_{i,j,t} / p_{i,j,b}) = \left( \frac{1}{n} \right) \sum_k p_{k,i,j,t} / \left( \frac{1}{n} \right) \sum_k p_{k,i,j,b}$$

or equivalently as:

$$(p_{i,j,t} / p_{i,j,b}) = \left( \prod_k (p_{k,i,j,t} / p_{k,i,j,b}) \right)^{\left( \frac{1}{n} \right)}$$



HICP regulations allow the use of arithmetic means of price relatives under special circumstances.<sup>15</sup> Note that one of the Boskin Commission’s recommendations was that the US Bureau of Labor Statistics switch to the use of geometric means at the elementary aggregate level to eliminate the problem of formula or lower level substitution bias, which they estimated as contributing as much as 0.25 percentage points to the overstatement of inflation in the US. However, Triplett (1998) argues that the undesirable properties of the arithmetic mean of price relatives that led to its being banned in the computation of elementary aggregates in the HICP are not necessarily eliminated by the geometric mean.

The classification of goods and services in the HICP is based on the classification of consumer expenditures devised for the 1993 System of National Accounts. The COICOP/HICP (Classification of Individual Consumption by Purpose adapted for the HICPs) omits some COICOP categories (e.g. narcotics, prostitution, imputed rentals for owner-occupied housing) and combines others to ensure that the weight of each category exceeds 0.001 in most of the member states. It is worth noting here that Triplett (forthcoming) has criticised the adoption of the COICOP for classifying the components of consumer price indexes as “wholly inappropriate” because it is not rooted in economic theory.

As noted earlier the HICP plays a central role in the policy deliberations of the ECB. The Monetary Union Index of Consumer Prices (MUICP) is used to assess inflation developments within the euro area. The MUICP is a simple weighted average of the HICPs of the Member States participating in EMU, with country weights equal to the country’s share of household final monetary consumption. The MUICP is calculated as an annual chain index with country weights  $w_{j,t}$  for country  $j$  at date  $t$  defined as the share of the country’s household final monetary consumption expenditure in the EMU total:

$$MUICP_t = \sum_{j \in MU} w_{j,t} HICP_{j,t} \quad (2)$$

For example, the country weights used in 2000 are derived from national accounts data for 1998, price updated to December 1999.<sup>16</sup>

Compilation of the HICP remains a highly decentralised affair, with responsibility for compiling national HICPs resting with national statistical institutes. At the time of writing, almost no detailed information on the methods and practices of the various national statistical was readily available. While the harmonization programme appears to have been very successful at increasing the degree of comparability of consumer price indices across EU countries and has achieved harmonisation on best practice in most areas, there is very little documentation currently available to outsiders explaining the details of how national HICPs are constructed. Eurostat's comprehensive compendium of HICP reference documents (Eurostat (2001), which includes the two reports of the Commission on the harmonisation process (Commission of the European Communities 1998 and 2000)) and its plan to make available a handbook on the elaboration of the HICP are welcome developments, which should be further complemented with additional releases of documents and information on current practices to implement the HICP in member countries.<sup>17</sup>

### **3. What do we know about bias in measures of consumer price inflation in Europe?**

Lets start with the question of what we currently know about the accuracy of measures of consumer price inflation in the EU. The answer is: Very little. To date there have been a number of attempts to assess the accuracy of the national CPIs in individual member states using the direct or "brute force" approach of looking at components of the indexes and then aggregating to obtain an estimate of the overall bias. This was the approach of the high-profile Boskin Commission in the US, and in some circles remains the preferred approach to studying measurement problems. A drawback of this approach is that it can be difficult to be sure that all potential biases have been found. There is the risk of paying too much attention to narrow categories of goods or services where there may have been a tendency to overstate inflation, and neglect broader categories of goods and services where there may be a tendency to understate inflation. An alternative, indirect, approach is to try to assess the extent of measurement bias indirectly by comparing measured consumer price inflation with either survey data of one sort or another, or by trying to estimate quality or other types of bias using alternative data on consumer spending patterns. The drawback of the indirect approaches to assessing the

extent of measurement error is that we usually need to make a variety of strong assumptions to draw inferences about measurement bias, although it is a matter of debate whether these assumptions are any stronger than those needed to implement the ground up direct approach.

The reference period for almost all of the existing studies of the accuracy of measures of consumer price inflation preceded the launch of the HICP, so they typically did not consider the extent to which their conclusions applied to the HICPs of the member states. The principle of harmonisation on best practice that was and continues to be a guiding principle of the HICP program suggests that the results of these studies may be of limited applicability to the HICP, although this will vary from country to country to the extent that some countries were already employing best practice and to the extent that the differences between national CPIs and HICPs are not all that great for some countries. **Table 1** gives some sense of the degree of overlap between national CPIs and national HICPs. Note that for some countries the coverage of the two is almost identical, which would in principle allow us to directly infer from estimates of bias in the national CPI the magnitude of the likely bias in the HICP. There does not appear to have been any attempt to perform a comprehensive audit of the CPI or HICP in any of the member states comparable to the audits of the US CPI carried out by the Stigler (1961) and Boskin (1996) Commissions<sup>18</sup>. Indeed many of the studies of bias in European CPIs seem to have been prompted by the Boskin Commission report rather than undertaken independently. Nor do there appear to have been any independent attempts to carry out a review of possible measurement problems comparable to Gordon's (1990) study of the measurement of durable goods prices. In assessing the state of knowledge, we can draw a distinction between the surveys that were prepared around the same time as the publication of the Boskin report in the US, and the work that has been done since.

### **3.1 Post-Boskin reviews.**

Of the studies to date, the most comprehensive by far, and a model for what is needed for other countries, and ultimately for the EU (or euro area) as a whole, is that by Hoffman (1998) for Germany. In addition to pulling together existing studies of potential problems in the German CPI, Hoffman carried out a number of new studies, significantly

adding to the body of knowledge on the potential for quality bias in the German CPI. However, Hoffman notes in his conclusions that the paucity of detailed case studies of mismeasurement of individual prices severely limits the ability to come up with a reliable estimate of the overall bias.<sup>19</sup> The number he provides as an estimate of the bias is 0.75 per cent a year, but he notes that the bias may well be larger in certain circumstances, especially during periods of falling prices<sup>20</sup>.

Lequiller (1997) provides an overview of the measurement problems in the French CPI. Taking the Boskin Commission report as his point of departure, he notes that some of the Commission's suggestions for improvements in the US CPI were already in place in France, specifically more frequent updating of the weights at the upper level and the use of the geometric mean to combine prices at the lower level, thereby reducing the potential for substitution bias. He concedes the possibility of some outlet substitution bias of between 0.05 and 0.15 percentage points a year, comparable to the figures cited by the Boskin Commission for the US, but notes that this estimate is based on a single study. On the subject of new goods and quality change, he notes the paucity of studies on which to base a firm conclusion. He cites one study of PC prices in France that shows a discrepancy of as much as 4.4 percentage points a year between the official index for PCs and a quality-adjusted index based on hedonic methods.<sup>21</sup> His bottom line, however is that there is simply not enough evidence to warrant a strong conclusion about the extent of the overall bias.

Two studies have looked at the possibility of measurement error in the UK Retail Price Index (RPI). Oulton (1995) reviews the traditional sources of bias and concludes that "...substitution and outlet bias are probably not significant sources of error in the UK. The two other sources of bias [quality change and new goods] most probably do lead to significant overstatement, but the size of the upward bias cannot at the moment be quantified." (Oulton, 1995, p. 60) Oulton based this conclusion on the fact that the RPI was at the time subject to considerably more frequent (annual) re-basing than the US CPI. He also cites the findings of Manser and McDonald (1988) for the US that the difference between a chained Laspeyres index and a true cost of living index was negligible (of the order of 0.01 percentage points per annum). However, subsequent research by Blow and Crawford (1999) call this conclusion into question. In his discussion of outlet substitution

bias, Oulton notes that planning controls in the UK might have impeded (perhaps indefinitely) the emergence of the type of low-cost high-volume outlets that revolutionised US retailing in the 1980s. Again, the vast majority of the references in his paper are to studies that have been done using US data. Only (ten of fifty six references are to studies of the UK or the RPI, and the bulk of these are to publications describing the construction of the RPI rather than independent studies of the accuracy of its components. Cunningham (1996) guesstimates that the extent of the bias in the RPI is 0.35 to 0.8 percent a year, although again almost exclusively on the basis of studies for the US.

The only other study of the overall bias in national CPIs for Europe of which we are aware is that by Folkertsma (1998) for the Netherlands, which concludes that "...the CPI in the Netherlands is probably biased but because of the lack of empirical research in this area the size of the measurement bias cannot be assessed." Once again the bulk of the references in Folkertsma's study (twelve of sixteen) are to studies of US data.

**Table 2** summarises the findings of these surveys. The coverage of the studies is about half of the EU or EMU. Crudely adding the estimates together, we would conclude that bias in the HICP for the EU might be of the order of magnitude of 0.28 to 0.37 percentage points per annum, or 0.28 to 0.30 percentage points per annum for the MU. These estimates are based on the assumptions that nothing has been done to correct the problems identified by these studies, and that the bias in all of the other member states of the EU or MU is zero. We believe that both of these assumptions are too strong to allow us to place much confidence in these numbers.

### **3.2 Ongoing research.**

Unlike the studies reviewed in the previous section, the more recent research has rarely aimed at providing estimates of the overall bias in CPI or HICP inflation. Instead, it has focused on deepening our understanding about the different sources of biases in the main sub-indices of national CPIs or HICPs. In this respect the more recent contributions provide a firm basis for future progress in assessing the accuracy of the HICP, in that they identify areas where problems have been found and oftentimes also propose methodological innovations for dealing with these problems.

The more recent work can be grouped under two broad (not disjoint) categories. First, the papers that have put forward some form of methodological innovation in the measurement of price indices. Areas where innovation has been particularly intense are the exploration of alternative data sources and research on new methods for quality adjustment. Second, papers that have scrutinised the scope for measurement bias in specific sectors of the economy. Sectors that have received particularly close attention are information processing equipment, rental housing markets and means for private transportation.

### **Alternative data sources**

Increasing attention is being paid to data sources that could complement the traditional direct sampling by the national statistical institutes, like scanner data, household surveys and databases provided by manufacturers' and consumers' associations.<sup>22</sup>

Of these possibilities, the avenue of research that has already proved to be particularly fruitful and to have a notable potential for further HICP improvements is the use of comprehensive transaction data available in scanner databases. The volume by Feenstra and Shapiro (forthcoming) is a pioneering exploration of the potential for scanner data to improve the measurement of prices. Silver and Heravi (2002) is a prominent example of an attempt to exploit the potential of scanner data in a European context. Scanner data are electronic records of transactions collected by business establishments in the course of doing business. From the perspective of measuring consumer prices, the most relevant scanner data is that collected by retail outlets by scanning barcodes at checkout lines. Silver and Heravi (2002) use data on transactions from a scanner database that includes data on purchases of washing machines, dishwashers, television sets, cameras and vacuum cleaners in the UK in 1998, with a view to assessing the potential bias arising under the matched-models approach to quality adjustment. They conclude that unless the matched-models approach is used in a context of frequently updated sampling and is implemented as a relatively high frequency chained index, it will be subject to important biases, particularly in settings with rapid quality change. Specifically, comparing hedonic-based quality-adjusted indexes based on

scanner (which virtually reflect the universe of transactions in that year) with quality adjusted indexes based on the matched-models method, they find that the latter suffers from a *downside* bias of 3.2 percentage points per year. Silver and Heravi provide a comprehensive argument as to why the matched-models method is likely to induce a bias in consumer price indices. Specifically, the matched-models approach tends to introduce a sample selection bias in the price index. Requiring that a given item should be present in the sample both in the former and the latter period when the price comparison is made implies two cases of sample selection. *New* models that are present in the latter but not in the former period, as well as *old* items that are present in the former but not in the latter, are more frequently dropped out of the sample. Then, if the pricing of old and new products is systematically different than pricing of matched models (as it is frequently the case since product prices tend to change along the product life cycle) then the matched-models method introduces a sample selection bias in price measurement. They conclude that, although increasing chaining and sampling frequency alleviates the sample selection bias of the matched-models method, sectors with rapid technological and quality change may call for the combined use of scanner data and hedonic-based quality adjustment.

The findings of Silver and Heravi are an important caution to those who assume that the use of hedonic methods for quality adjustment will automatically lower measured inflation rates. Bascher and Lacroix (1998) report estimates of hedonic models for dishwashers and apparel (women's suits and men's shirts) in the French CPI. Unfortunately they do not compare the results of quality adjustment using their hedonic models with traditional methods. However, their hedonic model has been used since September 1997 to make quality adjustments to dishwashers. **Figure 1** shows the annual rate of change on the relative price of dishwashers in France over the course of the 1990s. Note that since the introduction of the hedonic index, the rate of change of the dishwashers index increased.

A second source of information that could become a useful complement to sampling by the national statistical institutes are the databases collected by consumers and manufacturers' associations. For example, Hoffman (1998) used data on price and quality characteristics of a number of consumer durables from the magazine *Test* published by Stiftung Warentest to evaluate quality adjustments in the German CPI. Also

in this vein, Licandro *et al.* (2001) make use of the more detailed information (relative to data available to the Spanish National Statistical Institute) on traded new automobiles' characteristics, to gauge the possible bias in the corresponding HICP category for Spain (see below for a summary of the results).

### **Research on methodological issues**

Recent contributions pertaining to measurement methods have clustered around three main topics: alternatives to traditional hedonic regressions for quality-adjusted prices, sampling issues (in particular, outlet substitution bias), and the use of non-parametric methods to assess the scope of different sources of measurement bias<sup>23</sup>. Accurately adjusting prices for changes in product quality remains one of the thorniest issues in measurement. Many critics of the practices of national statistical agencies have argued that greater use of hedonic methods to deal with quality change (whether for the purposes of assessing when product substitutions are comparable or for making direct quality adjustments) would enhance the accuracy of price indexes. Statistical agencies have resisted the use of hedonic methods for a variety of reasons (see Triplett (1990) for a good review), including the sensitivity of the estimated quality adjustments to model specification, and the difficulty of employing them in real time. Methodological innovations that address these concerns are therefore to be welcomed. Sampling issues remain of central concern in determining the accuracy of measures of inflation at the consumer level. The development of superlative price indexes by Diewert (1976) facilitated the derivation of estimates of substitution bias that were not conditional on assumptions about functional forms for utility or demand functions. Further innovations that allow less restrictive assumptions to be employed in assessing the accuracy of price indexes are thus greatly to be welcomed.

Starting with alternatives to the traditional hedonic regression, Bover and Velilla (2002) apply panel data methods to analyse prices of dwellings in a sample of newly constructed multi-unit sites for the period 1993-1997. They compare price increases in dwellings based on a number of quality adjustment methods with those from a naïve indicator<sup>24</sup>, which only takes into account the size of the dwelling. The latter might be seen as a proxy for an owner occupied house prices index that follows a cost-of-



acquisition approach and that does not introduce adjustment for quality change. The comparison of the best quality-adjusted specification with the naïve indicator reveals that the latter suffers an upside bias from quality adjustment in the range of 0.75 to 1.2 percentage points per year. An important contribution of Bover and Velilla (2002) is that they show that panel data methods can be used to construct parsimonious quality adjusted prices of dwellings. These can be derived from a sample that includes information only about the price, the size (i.e. number of square meters) of the dwelling, and a site indicator where the dwelling belongs. Specifically, they show that their parsimonious quality-adjusted prices turn out to be very similar to those derived from a conventional hedonic regression which takes into account a large number of individual characteristics of the dwellings.

As regards recent progress on the measurement of the outlet substitution bias, Covas and Silva (2001) revisit this issue for the Portuguese CPI. They make use of data from the Portuguese household budget surveys to complement the sample on which the CPI is based. Allegedly, the latter may have failed to take into account the rapid expansion of large retailers in the 1990s. They conclude that the outlet substitution bias in the Portuguese CPI may have been close to 0.5 percentage points per year, i.e. larger than reported in earlier studies. They indicate however that the outlet substitution bias is likely to have decreased to about 0.25 percentage points per year in the later years, although the size of the bias in the later years is particularly difficult to gauge.

Finally, a recent strand of the literature has highlighted the strong potential of non-parametric methods to tackle a number of measurement issues. In particular, Blow and Crawford (1999) use non-parametric methods to examine the problems of substitution bias, new goods bias and quality bias in the UK RPI. They conclude that substitution bias in the RPI caused it to overstate the increase in the true cost of living by up to 3.2% over the period 1976-1997.<sup>25</sup> They also look at quality change in the audio-visual equipment component of the RPI and find that failure to make appropriate adjustments for quality change in this category caused the RPI to overstate inflation by about 1% over the period 1974-1996, or about 0.05% per annum. Finally, they estimate that the gain in welfare associated with the introduction of a single new good (the UK National Lottery) reduced the measured rate of inflation from 2.14 percent per annum to

1.97 percent<sup>26</sup>. Of particular note for the debate over research priorities in the area of price measurement is their concluding observation that

*“The bounds on the biases caused by a single instance of a new good and quality change in one section of the RPI are comparable to the overall bias caused by commodity substitution. We therefore conclude that, although substitution bias in the RPI formula is significant, it is likely to be much smaller than the biases that can be caused by continual product innovation in the form of new goods and quality change.”* (Blow and Crawford 1999, p. xii, emphasis added)]

### **Research on sectoral HICP items**

Regarding the sectoral aspect of price measurement, the bulk of the more recent work has concentrated in three areas. First, and not surprisingly, high technology goods, particularly information processing equipment, have attracted a considerable amount of attention given the rapid pace of quality upgrading within this class of goods. Problems of quality bias are believed to be particularly acute within this class of goods, but the implications for overall inflation at the consumer level (although not at the producer level) are less obvious since these goods make up a relatively small part of the HICP.<sup>27</sup> Second, housing services, which represent a substantial part of the HICP, but exhibit limited technological change.<sup>28</sup> And third, goods related to private transportation, which represent an intermediate case in terms of bias and weight in the HICP relative to the other two sets of items.<sup>29</sup> Overall, research findings tend to confirm that a positive and considerable bias exists in the HICP for goods related to information processing equipment and private transportation and suggests that price increases for rental housing services have tended to be underestimated, at least in the case of Germany.

Starting with prices of information processing equipment, Izquierdo and Matea (2001) exploit a database on desktop and laptop personal computers' prices and characteristics for the period 1990-2000 to come up with an estimate of the measurement bias in the Spanish price sub-index for information processing equipment.<sup>30</sup> Their quality-adjusted estimates based on the standard hedonic method point to the possibility of a substantial upside bias the official measure. While the latter reports an average decrease of 9 percent per year in information processing equipment prices, their estimates

indicate an average price decline of 35 percent per year, i.e. a bias of 26 percentage points on average per year. At the same time, their results need to be interpreted with caution, since certain important quality characteristics of PCs, like the type of microprocessor and weight, are missing from the sample.

Hoffmann and Kurz (2002) explore the potential for measurement error in the CPI for rental housing services for West Germany.<sup>31</sup> Their results suggest that the year to year rate of increase in this component of the CPI for West Germany could be reflecting a *downside* bias of 0.5 percentage points since approximately 1992 (i.e. coinciding with German re-unification). This result is based on a comparison of the official CPI for rental housing services with an index of rental housing services derived from an alternative survey-based data base, the German Socio-Economic Panel (GSOEP), for the period 1985-1998. Although the exact nature of their estimate of the bias in the CPI is difficult to pin down, their discussion suggests that the downside bias could be related to the fact that the German CPI sub-index for rental housing services is a matched-models index (i.e. the sample of rents results from following dwellings, and not households, over time). They show in particular that rent adjustments in dwellings tend to occur upon tenant turnover. Since overall tenant turnover in the former West Germany increased after German re-unification, a possible explanation of the emergence of a bias in the official series since re-unification is that the German CPI sub-index for rental housing services under-represents tenant turnover in the underlying sample.

Finally, as regards the latest evidence on bias stemming from imperfect quality adjustment methods for automobiles, two studies are particularly salient: Bode and van Dalen (2001) and Licandro, Izquierdo and Maydeu (2001). These papers use data from different countries (the Netherlands and Spain) and very different methods to account for goods' heterogeneity (the user-cost approach and principal components methods respectively). Nonetheless, they yield broadly similar results, suggesting the existence a substantial upward measurement bias in the respective official HICP item.

Bode and van Dalen (2001) report the results of estimating hedonic regressions for new cars in the Netherlands under a user-cost approach<sup>32</sup> and compare the effects of quality adjustments based on these regressions with the published indexes for new cars. Over the 1990-1999 sample period they find that while the official CPI for new cars

increased 11.2 percent over this period, the hedonic indexes show that quality adjusted prices were roughly constant, recording changes of between 0.1 percent and -3.6 percent depending on specification and implying a cumulative upward bias in the CPI of between 11.1 and 14.8 percentage points.

As regards quality adjustment for prices of cars in Spain, Licandro, Izquierdo and Maydeu (2001) exploit a very detailed database of sold cars' prices and characteristics provided by a manufacturers association for the period 1997-2000. Building on these data, they construct a price index comparable to the one provided by the Spanish National Statistics Institute (SMSI) and compare it to their own hedonic-type of quality adjusted indicator. Their results indicate that the average difference (in percentage rates of increase) between the indicator constructed to follow the official one and indicator that they propose is 3.5 percentage points per year. The estimator of quality adjusted car prices of Licandro, Izquierdo and Maydeu (2001) is based on a two-stage extension of the conventional hedonic methodology. In a first stage, car characteristics are divided in a number of groups, each of them is then aggregated into a "quality characteristic" of the car. This reduction in the dimension of the space of characteristics is based on the use of principal-components techniques. Once the quality characteristics of the cars have been constructed, they are used as input in the conventional hedonic regression (i.e. instead of directly using the raw characteristics in the right-hand-side of the hedonic regression). However, Licandro, Izquierdo and Maydeu (2001) do not provide an assessment of the costs and benefits of this two-stage procedure, relative to conventional hedonic regressions, either in the context of their sample, or in more general theoretical terms.

Two conclusions are warranted based on this review of recent research on the accuracy of price indexes at the consumer level in Europe. First, and most importantly, the recent work has shown that the common presumption that official inflation statistics routinely overstate the true rate of price increase seems to be wrong. Recent research has uncovered important examples of understatement of inflation which would need to be offset against examples of overstatement of inflation in arriving at an estimate of overall bias. Second, while the recent research has started to fill in some of the gaps identified in the earlier surveys by Cunningham (1996), Folkertsma (1998), Hoffman (1998), Lequiller (1997), and Oulton (1995), we are nowhere near having as the kind of

comprehensive information that would be needed to arrive at a reasonably precise estimate of the overall accuracy of the HICP. Thus some of the indirect approaches to estimating bias are worth considering, as they typically yield an estimate of the accuracy of the overall index without having to rely on detailed studies of individual categories of goods.

### 3.3 Indirect approaches to inferring measurement error

An alternative to the direct approach to estimating the overall bias in the HICP is to try to infer it indirectly using alternative data. One ingenious suggestion put forward by Nordhaus (1998) involves comparing median household income deflated by some measure of consumer prices with self-reported measures of well being. Taking data from the University of Michigan Household Survey (specifically the response to the question asking how the household's financial situation had changed over the past year) and regressing it on the change in median household income deflated by the CPI allows one to back out an estimate of the bias in the CPI. For example, a finding that more households reported their financial situation as having improved over the past year than report it as having deteriorated, at a time when a CPI deflated measure of median household income shows no change, suggests that the deflator used to deflate income may be overstating the inflation experienced by the average household. **Figure 2** replicates Nordhaus' scatter plot.

A simple regression of the form

$$\psi_t = a((\varpi_t - \pi_t) - b) + \varepsilon_t$$

where  $\psi_t$  is the percentage of survey respondents reporting themselves as being better off less the percentage of respondents reporting themselves as being worse off than they were 12 months ago,  $\varpi_t$  is the growth of median household income over the past year, and  $\pi_t$  is measured inflation, allows one to estimate the extent of bias in the CPI.<sup>33</sup> If the CPI correctly measures inflation, then the estimated parameter  $b$  should be zero. The simple regression line shown in Figure 2 indicates that on average equal numbers of people report themselves as being better off as report themselves as being worse off when

median household income deflated with the All Items CPI-U was declining at 1.1 percent a year.<sup>34</sup>

Unfortunately, a comparable exercise at the EU or MU level is not feasible. While the European Commission's monthly Harmonized Consumer Survey does provide usable measures of households self assessed well being similar to those in the Michigan Survey (specifically the responses to question 1 "How does the financial situation of your household now compare with what it was 12 months ago?"), we do not have EU or MU-wide data on income distribution or median household income. There are very limited time series data available at the level of the individual member states that would allow us to employ this approach to the problem of bias estimation: we were only able to obtain data for France, Italy, Sweden and the UK.

Statistics Sweden reports data on median income from work for persons employed full time over full year on an annual basis from 1990 through 1998.<sup>35</sup> However, the European Commission Harmonised Consumer Survey for Sweden only begins in 1995, the date of Sweden's accession to the EU, leaving little in the way of usable data. We have not been able to find a comparable survey of household well being that could be used instead.

We had more success with France, Italy and the UK. For France we were able to obtain data on median income from Casaccia and Seroussi (2000) for the period 1950-1998, with some gaps. The Harmonized Consumer Survey included the question on household finances over the past 12 months from 1985 on, so we have a useable sample of about fifteen annual observations. However, the period of overlap with the HICP is only three years, too short to allow any meaningful comparisons. For the UK we are able to obtain a long time series (1961-1998) on median household income drawn from the Family Resources Survey and the Family Expenditure Survey which we used for comparison with the responses to the Harmonized Consumer Survey.<sup>36</sup> Finally, for Italy Brandolini (1999) reports various measures of the distribution of post-tax household incomes distribution for the period 1965-1998.

**Figures 3-5** show scatter plots of growth in real median household income against the balance statistic from the European Commission's monthly survey for the UK, France and Italy. What is striking is that on average more households report themselves as being

worse off than as being better off in every single year of the sample in both France and Italy. This despite the fact that real median income appears to be growing in most years. Crude application of Nordhaus' methodology to these data would suggest that far from overstating inflation at the consumer level in these countries, the national CPIs drastically *understate* it! The only country for which there was any meaningful degree of overlap between the various data sources was France. The broken line in Figure 4 is the Nordhaus regression line fitted to three years of data on HICP deflated median household income growth. The qualitative impression is the same as one gets from the CPI-deflated numbers, namely that the official measures of inflation appear to understate inflation as experienced by the average household.

#### **4. What do we need to know?**

Based on the evidence reviewed in the previous section, we believe that there are major gaps in our knowledge about the accuracy of consumer price inflation statistics in Europe. We think that the focus of future research ought to be on assessing the accuracy of the HICP rather than on national CPIs, given the central role that this index now plays in economic and monetary policy deliberations in Europe. Two important issues will need to be addressed before it will be possible to attempt to provide a comprehensive answer to the question of whether the HICP overstates the true rate of inflation in the euro area. The crucial first step is to define what the true rate of inflation is.<sup>37</sup> The second is to understand in detail how exactly the HICP is computed on a month-to-month basis.

The mainstream approach to evaluating measures of consumer price inflation involves comparing official measures with the cost-of-living index of economic theory. If, as Eurostat has stated on numerous occasions, the HICP is not based on the theory of the cost of living index, we need to know what theory it is based on, and whether, how and to what extent this theory differs from the theory of the cost of living index. The HICP is of course not unique in eschewing the cost of living as its conceptual framework. However, as Oulton (1995) noted in his review of the UK RPI “Unless the notion of a true index is to be purely subjective, existing indexes must be assessed in the light of the only existing objective standard, which is provided by the economic theory of index numbers. In the case of the RPI, the comparison must be with a COL index.” (Oulton,

1995, p. 61). Likewise Blow and Crawford (1999) argue in taking the cost of living index as the benchmark against which to compare the RPI that "...in very many cases, the uses to which the RPI is put require it to be interpreted as a cost-of-living index. So asking if it is a good approximation to a true cost-of-living index is a legitimate question and an important one for many users." (Blow and Crawford, 1999, p. vii) The argument that the HICP cannot be biased because it is by definition measuring what it is supposed to measure is not very compelling.

In section 2 above we sketched out the conceptual framework of the HICP as documented in Eurostat (2001). Diewert (2002) provides a comprehensive critique of the HICP's conceptual framework as it currently stands. He concludes that "The "theory" of the Harmonized Index of Consumer Prices seems to lack an underlying firm theoretical basis. Evidently, its primary purpose is as a measure of inflation that is based on actual transactions that use money. However, as we have argued...above, a measure based on "monetary" transactions is too broad to be useful...when the inflation measurement goal of the harmonized index is narrowed down to focus on purchases of consumer goods and services...the "general theory" of the HICP does not constrain the index as much as an explicit producer or consumer theory approach would." (Diewert, 2002, 42)

One key difference between the HICP and the cost of living index has to do with the appropriate treatment of durable goods. In the cost of living framework it is the service flow yielded by durable goods that ought to be priced, and not the purchase price of the good. (Although in reality this principle only seems to be followed when it comes to pricing the services of owner-occupied housing). However, the situation is exactly the reverse under the HICP concept.

Perhaps the price concept that forms the basis of the HICP is the right one for a measure of inflation for monetary policy, and we need to make a stronger distinction between increases in the cost of living and monetary inflation. On the face of it, the concept of household final monetary consumption expenditures that forms the basis of the HICP is appealing from the stand point of monetary policy. The biggest drawback of this concept is that the measure of inflation that it gives rise to does not have a rigorous foundation in economic theory. This is in stark contrast to the cost of living concept, which is familiar to anyone who has ever studied intermediate microeconomics.<sup>38</sup> Some



detailed analysis of the differences between the two measures is also probably warranted. Under what circumstances would the different treatment of durable goods in the two indexes cause them to diverge?<sup>39</sup> For example, would a revised HICP that included the net acquisition cost of owner occupied homes be more sensitive to housing market booms than a cost of living index? Is this desirable from the perspective of monetary policymakers?

At present the HICP does not include the costs of owner-occupied housing. Eurostat is working towards the inclusion of housing prices using a net acquisition cost approach ( on an experimental basis initially, later as part of the index). How this will be done in practice remains to be seen, but will raise a whole new set of questions. For example, how will changes in the quality of houses be handled?<sup>40</sup> Is sample representivity a problem (in smaller countries the number of new houses on sale each month available to the price surveyor may be quite limited)? More generally, the rationale for excluding the options of imputed rental value and user cost was that those payments depend on interest rates in a rather mechanical fashion. But interest rates (through the cost of mortgages) may have sizeable effects also on the net acquisition value. How large are those effects precisely and could they pose problems for the HICP as an indicator for monetary policy?

The second crucial issue that needs to be addressed is the paucity of information available to outsiders about how the HICP is constructed. There is a very limited amount of documentation available about the construction of national CPIs, and in those cases where it is available it is only in the language of the state concerned. It would be extremely useful to the international community of researchers working on measurement issues related to the HICP to have, as early as possible, documentation available for all countries in English. Note that Eurostat has made available a significant amount of extremely useful information about the HICP program through their reports to the Council (see Commission of the European Communities (1998, 2000), both of which are included in Eurostat (2001)). However, it is still rather difficult to get a good sense of how the different national statistical agencies go about compiling the raw data that go into the HICP. For example, to what extent do national statistical agencies rely on sampling of goods and outlets when deciding what to price and where? How are samples

selected and updated? When identical new model year cars are being priced for the HICP, are they adjusted for quality exactly the same way in all countries in which they are priced?<sup>41</sup> Knowing what the practices of the statistical agencies are would greatly help in setting priorities for future research. For example, if we were reasonably confident that statistical agencies followed best practice when it came to making quality adjustments, but followed more informal procedures as regards outlet selection, it would make more sense to devote more resources to figuring out the potential for outlet substitution bias than to second guessing the agencies' quality adjustments.

**Figure 6** gives some sense of the likely differences in practice in terms of quality adjustment.<sup>42</sup> Information processing equipment is one of the most traded and tradable items in the HICP. We conjecture that across all of the product categories in the HICP, this one is most likely to have identical products being priced in all countries. While it would not be surprising to find minor differences in the levels of prices of information processing equipment in different countries (due to say differences in taxes), it is remarkable that the rates of change are so dramatically different. The figure shows that in Austria, France and the Netherlands, information processing equipment relative to the overall HICPs for those countries cost only about a quarter of what it did in 1996, while in Belgium, Greece and Luxembourg it cost a bit more than half of what it did in 1996.

The problems that arise in accurately measuring inflation are usually grouped under three headings. The first type of problem relates to the sampling procedure (including the choice of when to sample products) used to select the goods in the consumption basket, as well as to record their price and quality characteristics. The second broad class of problems arises due to substitution on the part of consumers away from relatively expensive products toward relatively cheaper products. These substitution problems arise as a result of the choice of index number formula and weighting scheme. We can also include in this category the problem of outlet substitution as consumers shift their spending from traditional retail outlets to newer discount stores. The third broad class of problems arises due to changes in the quality of goods and services, and the arrival of new goods.

The limited availability documentation on the implementation of the HICPs at national level makes it difficult to assess the extent to which potential heterogeneity in

sampling timing and methods (including the frequency with which weights are updated and the basis on which the updating in weights is made –i.e., whether weights are updated on the basis of new consumers’ survey or just on changed prices) gives rise to measurement bias. We include under this heading the issue of list versus transactions prices. For almost all uses to which a price index is put, it is important that, to the extent that is feasible, the prices that enter into the index are those at which actual transactions take place, and not list prices at which relatively few transactions occur.<sup>43</sup> The two reports from the Commission to the Council on the HICP do not appear to address the issue of transactions versus list prices explicitly, although the guidelines on the treatment of sales prices or price reductions are relevant.

Arguably, substitution bias in the traditional sense is *by definition*<sup>44</sup> not a problem in the HICP. The coverage of the HICP is defined as “final monetary consumption expenditures of households” and the HICP is intended to measure the average price change experienced by households attempting to maintain a given pattern of these final monetary consumption expenditures. Given that discussions of substitution bias are based on a comparison of a calculated index with a theoretical ideal of a cost of living index based on household expenditure functions, any discussion of the substitution bias in the HICP might need to be re-formulated in terms of the fundamentally different concept of the so-called representativity bias. However, this simply brings us back to the question of what is the theoretical basis for the HICP; representative of what?

Even if we were to take the theory of the cost of living index as the appropriate benchmark for the HICP, there are other reasons to believe that there may be relatively little substitution bias in the HICP. At the highest levels of aggregation, this reflects the fact that the weights are updated annually (albeit frequently on the basis of “price updating” rather than annual consumer expenditure surveys, which in turn raises interesting questions) rather than being held fixed for long periods of time. At the lower level it appears that the lower-level substitution bias that was of such concern to the Boskin Commission is probably not a major issue in the HICP because of widespread use of geometric means.

As for outlet substitution bias, evidence is too limited to attempt a clear assessment. Hoffman (1998) notes that traditional high-cost outlets are probably over

represented in the German CPI, but concludes that based on observed trends in market shares outlet substitution bias is unlikely to amount to more than 0.1 percent annually. We would be surprised if outlet substitution bias turned out to be a major source of bias in the HICP, since the retail sector in many (but not all) European countries seems to be a lot more regulated than the retail sector in the US, where the phenomenon was first noted.<sup>45</sup> Arguably in the presence of such regulation, it may make little difference where the statistical agencies collect their raw data.

The big unanswered questions have to do with quality change and the introduction of new goods. There is some reason to believe that the rules put in place governing the treatment of new goods are such that this is might be less of a problem with the HICP than has been the case with some national CPIs. The HICP program requires that new goods be included when they achieve a sales volume of over one part per thousand of consumers spending. As noted above, to date there have been remarkably few studies of the problem of quality change in European price statistics. Many of the existing estimates of the extent of measurement bias in national CPIs rely heavily on studies for the United States. For example, Hoffmann (1998) notes that, at the time he was writing, there were only three studies of the quality problem for Germany. At the outset we cannot rule out the possibility that the paucity of studies of quality change in Europe reflects fundamental differences between Europe and the US, specifically a less innovative environment and a slower pace of product innovation and technical change. We are inclined to discount this possibility. It seems to us that the electronic goods purchased by European households, to take but one example, have experienced improvements in quality over the past two decades at a rate comparable to the US. By way of illustration, **Figure 7** shows the recent behaviour of computer prices in the US CPI and the HICP for the EU (EICP). Note that the US series, which has been quality adjusted using a hedonic model since January 1998, shows a much more rapid rate of decline (an average of -29.4 percent per annum) than the EU series (an average of -13.5 percent per annum).

We also caution against a presumption that the methods employed by national statistical agencies to make quality adjustments are inherently susceptible to producing an upward bias in a measure of consumer prices. It has long been argued by practitioners that the difficulty of disentangling changes in “fashion” from changes in quality in the

apparel component of a consumer price index make it as likely that the consumer price index understates inflation as overstates inflation. Wynne and Sigalla (1996) cite an example pointed out by Jack Triplett in an unpublished conference paper wherein he noted that the infant's and toddler's component of the US CPI (which was presumably less influenced by fashions cycles) showed a much more rapid rate of increase than the men's or women's components.<sup>46</sup>

The classification of goods and services in the HICP does not distinguish between men's, women's and children's apparel, but it does distinguish between "clothing materials" (COICCOP/HICP code 03.1.1), "garments" (COICOP/HICP code 03.1.2) and "other articles of clothing/clothing accessory" (COICOP/HICP code 03.1.3). **Figure 8** shows the rate of change of these three components of the HICP over the past five years. We see that the rate of change of the garments component of the index (which we would expect to be the most susceptible to a fashion cycle induced understatement of inflation) does indeed grow at a slower rate than the other components. And as Astin (1999) points out, the weight of apparel in the HICP is on the order of 25 times the weight of personal computers and other high tech goods where quality change is commonly thought to impart an upward bias to the price index. Some additional, albeit more ambiguous, evidence is presented in Table 3, which reports the rate of change of detailed components of the French CPI.

For the purposes of assessing the extent of potential bias due to, for example, inadequate quality adjustment we would ideally want to obtain the exact specifications of the products priced by the agencies and then construct an alternative price index based on, say, hedonic quality adjustment. In reality, however, this is likely to be infeasible for a variety of reasons (either the agencies do not have the necessary detailed product descriptions or they cannot release them due to confidentiality requirements).

Alternatively we could gather price information and product specifications from technical publications, mail order catalogues, consumer magazines and the like, and use these data to construct alternative quality adjusted price indexes. This is the strategy adopted by Gordon (1990) in his monumental study of the prices of producers' durable equipment in the United States. The primary drawback of this approach is that the data used to construct the alternative series may not be representative or strictly comparable to

the data used by the national statistical agency. A discrepancy between the two series could be due to the methods of quality adjustment or the use of different data (which may be more or less representative of the prices paid by households). More importantly from the perspective of obtaining a quick answer to the question of how well the HICP measures inflation, the resources and time needed to produce a scholarly study such as Gordon's should not be underestimated.

It may therefore be useful to explore alternative indirect strategies for estimating the extent of quality bias in the HICP other than by estimating hedonic price indexes. In particular the approach developed by Bils and Klenow (2001) based on quality Engel curves may have some potential. The revealed preference approach of Blow and Crawford (1999) may also hold some promise in this regard, as may the statistical approach of Lichtenberg and Griliches (1989). What we do know is that within the realm of traditional (non-hedonic) methods to account for quality change that are used in member countries, there is great variety in outcomes (mainly, markedly different patterns of price decrease in high-tech sectors).

On the issue of new goods, we note that the rule that new goods be included in the HICP once they achieve a sales importance of 1/1000 may help alleviate the potential for an upward bias from this source. However, it is still an open question as to how well the 1/1000 rule does in capturing the gains on consumer surplus from the introduction of new goods? At this point it would appear that all we can do is some suggestive numerical exercises. It probably won't be until after a major new product has arrived and appeared to have been overlooked by the HICP that we will be able to do more. However, under this heading we might also want to include expanded product variety. Some of the products suggested by national statistical agencies as new goods (lamb, discos) have been around for a while, but are only now becoming significant at the level of some member states. Included under this heading are the greater variety of products that are available in many categories of goods. Here the work of Hausman (1997, 1999) provides a useful starting point, and a reminder that greater variety is widespread in all categories of spending.

## **5. Concluding observations.**

In this paper, we have attempted to summarise the current state of knowledge about the potential for measurement bias or error in the HICP. This question is important to the European Central Bank, as it has noted that the possible presence of measurement error in the HICP justifies a definition of price stability as prevailing at small positive rates of measured inflation. Based on the evidence reviewed above, we believe that at this point in time there is very little scientific basis for putting a point or even an interval estimate on the likely magnitude of the overall bias in the HICP. There are three reasons for this. First the HICP is still an evolving measure of inflation. Practices are being refined and improved on an ongoing basis. Problems that may have been identified in national CPIs may have been corrected in the HICP. Second, assessing the potential for bias requires a detailed knowledge of how the HICP is actually constructed. The various Regulations set certain standards, but at this time we do not have a definitive handbook of HICP methods documenting actual (as opposed to aspired to practices) that can be studied to see whether there is room for improvement. And finally, to date there has been little in the way of independent research assessing the accuracy of the various components of the HICP (or national CPIs for that matter).<sup>47</sup> Some of the research that has appeared in recent years and was reviewed above has shown that the widespread belief that official inflation statistics routinely overstate the true rate of inflation is misplaced. We have also shown that calculations of the sort Nordhaus (1998) used to illustrate the presence of a significant upward bias in the US CPI suggest that measures of consumer price inflation in Europe understate inflation. However, we caution against reading too much into these results, given the short samples of data on which they are based.

<b>Table 1</b> Coverage of HICP and CPI (percent)		
	Expenditure covered by HICP but excluded from CPI	Expenditure covered by CPI but excluded from HICP
Belgium	0.0	1.1
Denmark	8.2	11.4
Germany	2.0	11.4
Greece	6.9	0.0
Spain	0.0	0.5
France	8.7	0.0
Ireland		
Italy	0.0	4.8
Luxembourg		
Netherlands	8.4	17.3
Austria	13.2	8.3
Portugal	0.0	7.7
Finland	0.0	0.0
Sweden	0.0	16.5
UK	12.4	9.6

**Notes to Table 1:** Source: Commission of the European Communities (1998).



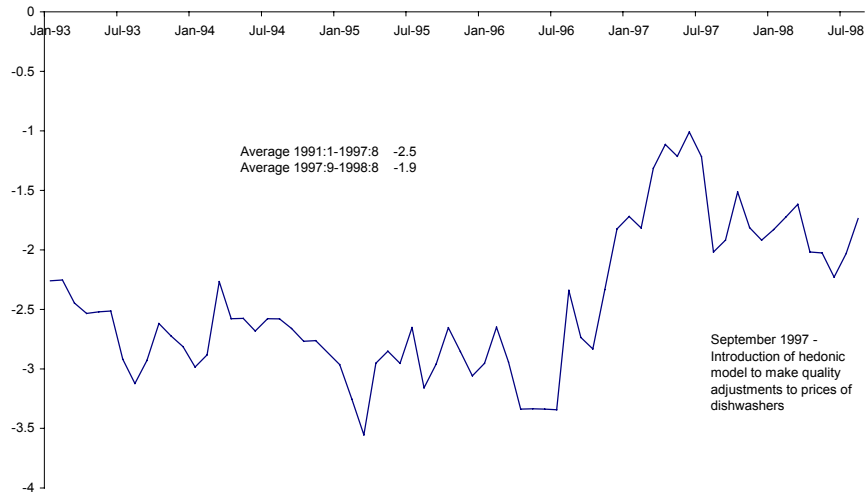
<b>Table 2</b> Summary of existing surveys of measurement error			
	Germany (Hoffman, 1998)	France Lequiller (1997)	UK Cunningham (1996)
Substitution bias - Upper level - Lower level	< 0.1	0.05 - 0.10	0.05 – 0.1
Outlet substitution bias	< 0.1	0.05 - 0.15	0.1 – 0.25
Quality change	< 0.5	?	0.2 – 0.3
New goods	< 0.1	?	0.0 – 0.15
Total	≈ 0.75	> 0.1 - 0.25 ?	0.35 - 0.8
Expenditure covered by CPI but excluded from HICP	11.4	0.0	9.6
Share of			
- EICP	27.149	16.380	16.34
- MUICP	34.651	20.907	NA
Implied bias in			
- EICP	≈ 0.20	> 0.02 - 0.04 ?	0.06 – 0.13
- MUICP	≈ 0.26	> 0.02 – 0.04	NA

**Notes to Table 2:** Source: Hoffman (1998), Lequiller (1997), Cunningham (1996). Shares of EICP and MUICP are country weights for 2000, price updated to December 1999 weights, from Commission of the European Communities (2000).

<b>Table 3</b>			
Rate of change of selected components of French CPI 1991-1998			
	Men's	Women's	Children's
Coats & jackets	0.8	1.0	0.5
Pants	0.9	0.7	0.0
Casual attire	0.1	0.4	0.8
Dress shoes	1.0	0.7	1.9

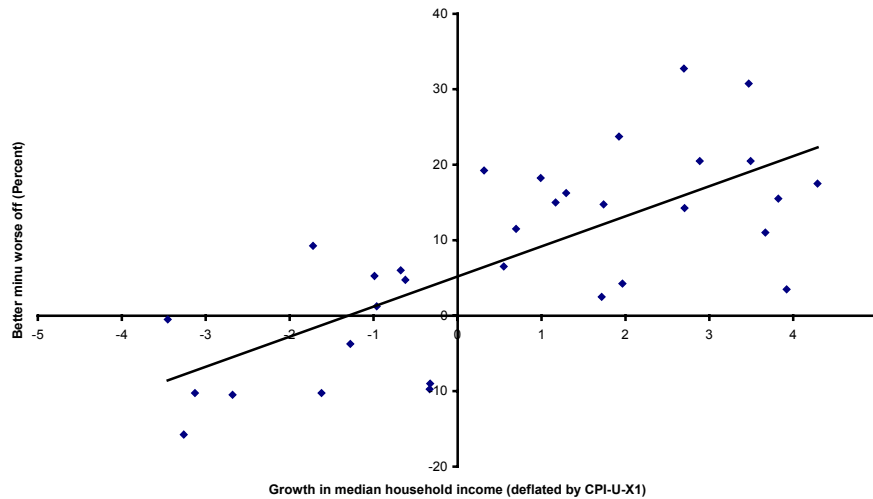
**Notes to Table:** Source INSEE.

**Figure 1**  
**Change in relative price of dishwashers**  
**France 1993-1998**

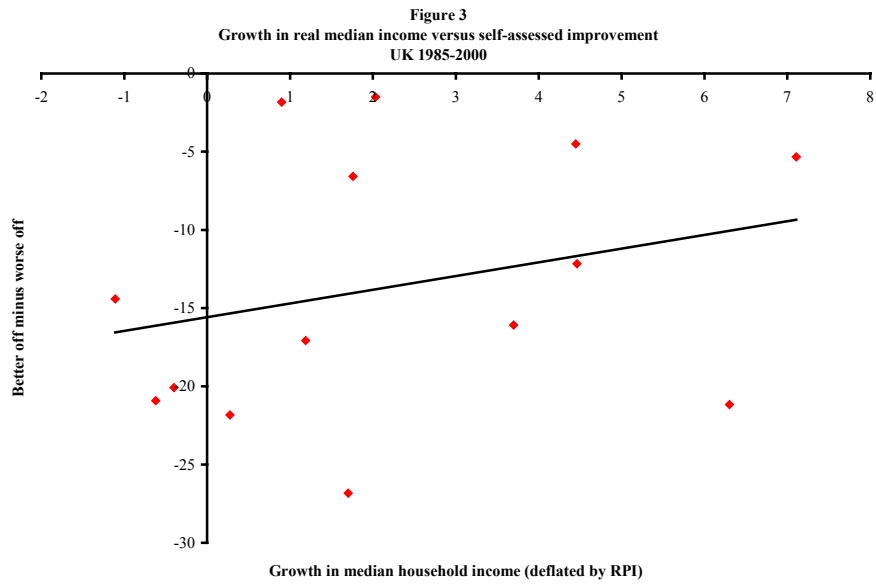


**Notes to Figure 1:** Source: INSEE. Annual change in CPI series “Lave-vaisselle” 4315 divided by CPI.

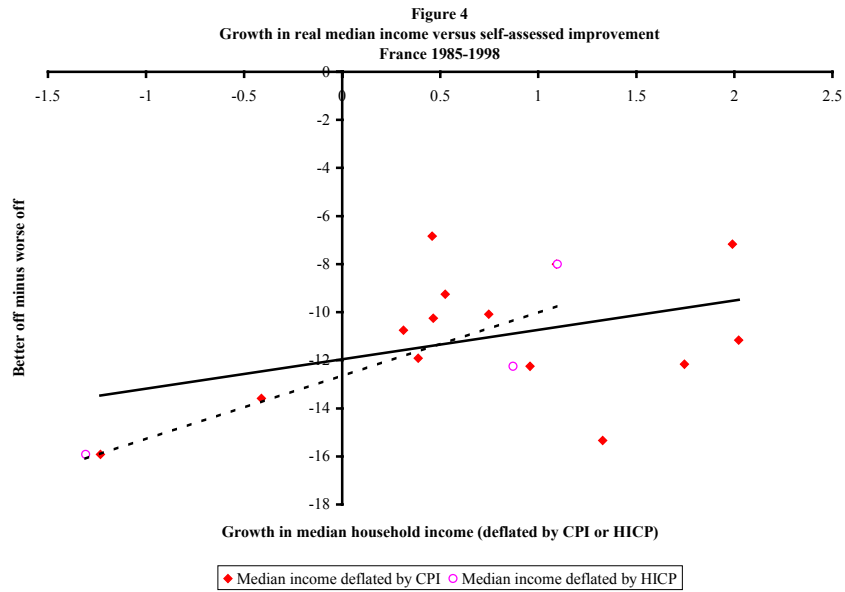
Figure 2  
Growth in real median income versus self-assessed improvement  
US data 1967-1998



**Notes to Figure 2:** Scatter plot of growth in real median income against change in self assessed well being. Real median household income is defined as Nominal median household income from Table A-1 of U.S. Census Bureau (2001) deflated by the CPI-U-X1 research series. The measure of self-assessed well being is from the University of Michigan’s monthly Survey of Consumers. Question A2 of the Survey asks: “We are interested in how people are getting along financially these days. Would you say that you (and your family living there) are better off or worse of financially than you were a year ago?” We used the responses to this question (reported in Table 4 of the monthly reports) to calculate a balance statistic (i.e. percent reporting themselves better off minus those reporting themselves worse off). The Figure shows the average over the course of the year of the quarterly balances.

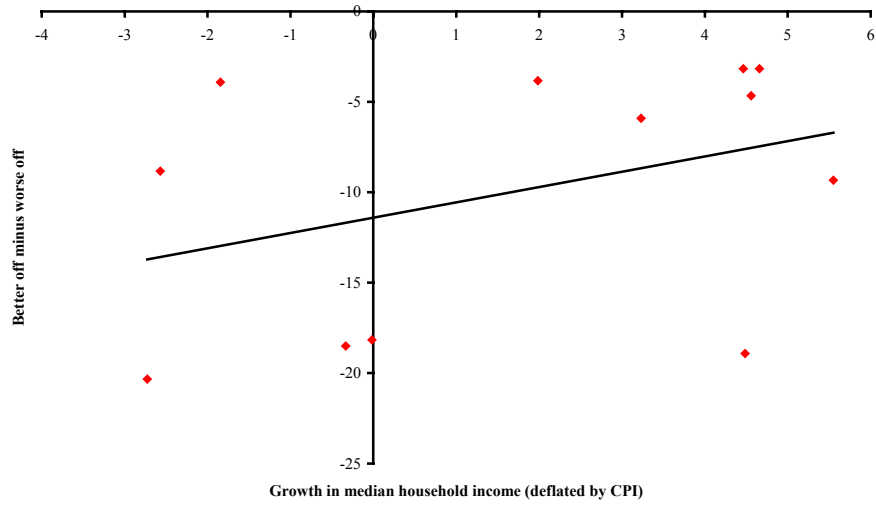


**Notes to Figure 3:** Balance statistic from the European Commission’s Harmonized Consumer Survey. Median income from Institute for Fiscal Studies. RPI from Haver Analytics.



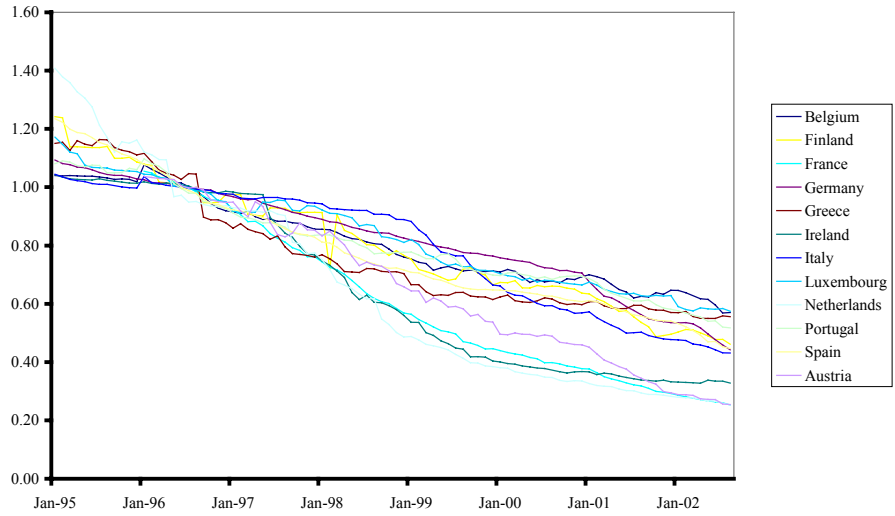
**Notes to Figure 4:** Data on median household income from Casaccia and Seroussi (2000). Balance statistic from the European Commission's Harmonized Consumer Survey. CPI data from Haver Analytics.

Figure 5  
Growth in real median household income versus self-assessed improvement  
Italy 1985-1996



**Notes to Figure 5:** Data on median household income from Brandolini (1999). Balance statistic from European Commission's Harmonized Consumer Survey. CPI for Italy from Haver Analytics.

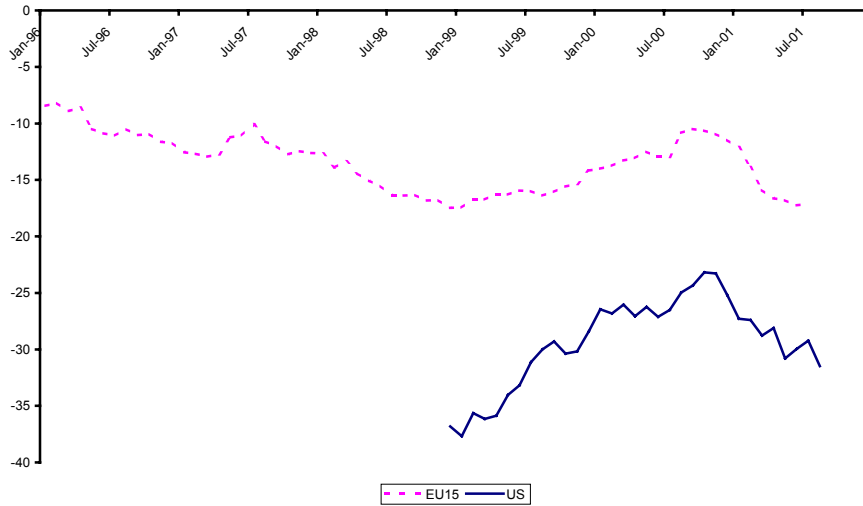
**Figure 6**  
Relative price of information processing equipment in the euro area



**Notes to Figure 6:** HICP for Information processing equipment (COICOP/HICP code 09.1.3) divided by the aggregate HICP for each country.

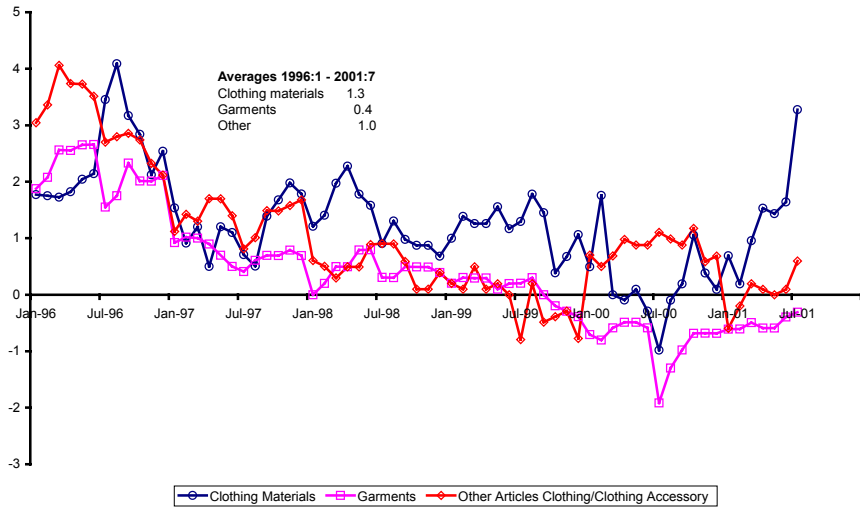


**Figure 6**  
Information processing equipment  
EU15 and US



**Notes to Figure 7:** Annual change in relative prices of personal computers. The US series is the CPI for Personal computers and peripheral equipment divided by the All Items CPI-U. The EU15 series is the HICP for Information processing equipment (COICOP/HICP code 09.1.3) divided by the EICP. Information processing equipment had a weight of 0.32 percent in the EICP in 2000. The US CPI series for personal computers has been quality adjusted using hedonic methods since January 1998, and had a relative importance of 0.079 percent in the CPI-U as of December 2000. Source: Haver Analytics.

**Figure 7**  
**Components of apparel inflation**  
**EU15 level**



**Notes to Figure 8:** Source: Haver Analytics.

## Acknowledgements

The views expressed are those of the authors and do not necessarily reflect the views of the European Central Bank, the Federal Reserve Bank of Dallas or the Federal Reserve System. Eric Millis provided outstanding research assistance. We thank the editor, two anonymous referees, and participants in the November 2001 CEPR/ECB Workshop on Measurement Issues for comments.

## Notes

---

<sup>1</sup> It is probably no accident that prior to the appointment of the Boskin Commission in 1995, the last time the problem of price measurement had attracted the attention of lawmakers in the US was at the end of the 1950s, a period of low inflation comparable to what the US experienced in the late 1990s. The Price Statistics Review Committee was created in July 1959 and issued its report in 1960. The report and a series of staff studies were published as Stigler (1961).

<sup>2</sup> See in particular Greenspan (2001b) who noted that “When industrial product was the centrepiece of the economy during the first two-thirds of the twentieth century, our overall price indexes served us well....But in our new century, the simple notion of price has turned decidedly ambiguous...how will we measure inflation....in the twenty-first century when our data -- using current techniques -- could become increasingly less adequate for tracing price trends over time?...For all these conceptual uncertainties and measurement problems, a specific numerical inflation target would represent an unhelpful and false sense of precision.”

<sup>3</sup> The change is explained in Landefeld and Parker (1995). Triplett (1992) and Young (1992) provide background.

<sup>4</sup> See in particular the June 2000 *OECD Economic Outlook*, the August 2000 *Monthly Bulletin* of Deutsche Bundesbank, the December 2000 issue of Sveriges Riksbank's *Inflation Report* and the July 2001 *Monthly Bulletin* of the European Central Bank.

<sup>5</sup> Diewert (2001) gives a historical overview of the link between theory and practice of price measurement issues. He concludes that the challenges to inflation measurement will only increase in the future, as technological change will make inflation measurement increasingly complex. He lists current analytical developments that should facilitate meeting those challenges in the future.

<sup>6</sup> In the press release announcing its strategy for monetary policy, no mention is made of the possibility of a measurement bias in the HICP, and the only feature of the HICP that the ECB cited as justifying the choice of this index for defining price stability was the fact that it was (and remains) the only price index that was sufficiently harmonised across the euro area at the launch of EMU. See European Central Bank (1998) and (2001).

<sup>7</sup> Eurostat uses the “American” spelling Harmonized Index of Consumer Prices in its publications, while the European Central Bank seems to prefer the “British” spelling Harmonised Index of Consumer Prices. We will use the spelling preferred by Eurostat.

---

<sup>8</sup> See for example European Central Bank (2001) p.39.

<sup>9</sup> See for example Balke and Wynne (2000) and Silver and Ioannidis (2001).

<sup>10</sup> Article 121 of the Consolidated Version of the Treaty on European Union (Amsterdam Treaty).

<sup>11</sup> He also mentions the approach of Divisia, but dismisses it as being unhelpful in terms of providing practical guidance to the construction of index numbers.

<sup>12</sup> There is a potential inconsistency here with how the HICP treats quality change. In the same document the Commission notes ‘ “Quality change” occurs whenever ...a change in [product] specification has resulted in a significant difference in *utility* to the consumer between a new variety or model of a good or service and a good or service previously selected for pricing in the HICP for which it is substituted.’ (Commission of the European Communities, 2000, p. 62, emphasis added)

<sup>13</sup> Table 8 in page 58 of Eurostat (2001) lists the weights reference period for all EU countries as of January 1999 and as projected in the next years. It also provides information on the weight updating frequency by those countries using chained indices.

<sup>14</sup> The formulae applied for the computation of the elementary aggregates by some the EU national statistical offices is summarised in Eurostat (2001), Table 9 (p. 60).

<sup>15</sup> According to Eurostat (2001) p. 59 “The arithmetic mean of price relatives may be applied in exceptional cases where it can be shown to meet the comparability requirement.” Eurostat does not give any indication as to how often these exceptional circumstances arise.

<sup>16</sup> Commission of the European Communities, 2000, p.14.

<sup>17</sup> See Astin (1999).

<sup>18</sup> The Stigler Commission reviewed all of the major price indexes published by the US government, not just the CPI.

<sup>19</sup> The vast majority of the some 160+ references in Hoffman’s paper were to studies for other countries or general problems of index number construction.

<sup>20</sup> Following up on his overview of potential measurement error in the German CPI, Hoffmann (1999) provided further evidence of the potential for quality bias in the German CPI. Citing his own earlier work and research by Harhoff (1999) he suggested, on a somewhat more tentative basis, that given the rules employed by the German Federal Statistical Office to make quality adjustments, the potential for upward bias increases at lower inflation rates and as of 1998 was around 0.25 percent per annum.

<sup>21</sup> See Moreau (1996).

<sup>22</sup> See Feenstra and Shapiro (forthcoming). De Haan (2001) reports that scanner data will be used in the compilation of the CPI for the Netherlands from May 2001.

---

<sup>23</sup> Other methodological issues in addition to these three topics highlighted in the main text have been tackled in the recent literature. In particular, Poinat (1997) examines the potential for bias in the French CPI due to annual chain linking of December indexes. Ruiz-Castillo et. al. (2000b) examine the distributional consequences of quality change bias in the Spanish CPI. Importantly, they do not have independent evidence of the extent of the quality bias in the Spanish CPI, but rather rely on the figures arrived at by the Boskin Commission, suitably adjusted for Spain. However, this paper addresses a source of bias that might be of interest to monetary policy makers. This is the bias that arises when statistical agencies fail to allow for price changes that occur between the period when household consumption patterns are investigated to obtain the weights to be used to aggregate prices and the reference period for the price index. Ruiz-Castillo et. al. (2000a) term this the Laspeyres bias, and (interestingly enough) show that it was *negative* on average over their sample period.<sup>23</sup> More recently, a number of sectoral studies on price measurement issues have been undertaken by Banco de España, focusing on owner housing prices, automobiles and personal computers. Finally, Neves and Sarmento (1997) report estimates of the size of the substitution bias in the Portuguese CPI of between 0.05 and 0.1 percentage points a year. Santos and Coimbra (1995) present estimates of bias due to the failure to fully account for quality improvements in cars.

<sup>24</sup> They also compare their quality-adjusted series to the National Accounts deflator for residential constructor. This comparison suggests that the latter features a significant downside bias of 3.5 percentage points per year. However, the new residential house deflator is constructed as an index of the costs of the construction inputs. The discrepancy reflects partly a potential bias from insufficient quality adjustment together with declining profit margins for residential house developers.

<sup>25</sup> They note that the RPI is not designed to be a cost of living index, but argue that given that the many uses to which the RPI is put require that it be interpreted as a cost of living index it is legitimate to ask how well it approximates this theoretical ideal.

<sup>26</sup> Blow and Crawford (2001b) revisit this question using further methodological innovations based on non-parametric methods. The size of the bias from excluding the National Lottery as a new good is increased somewhat in this case, when the Törnqvist index is used to compare the change in prices under inclusion and exclusion of the new good.

<sup>27</sup> The series “Information processing equipment” COICOP/HICP category 09.1.3 had a weight of 0.338 percent in the MUICP in 2001.

<sup>28</sup> The series “Actual rents for housing” COICOP/HICP category 04.1 had a weight of 5.82 percent in the MUICP in 2001.

<sup>29</sup> The series “ Motor cars” COICOP/HICP category 07.1.1 had a weight of 4.581 percent in the MUICP in 2001.

<sup>30</sup> About 90% of which corresponds to personal computers.

<sup>31</sup> Additional research on housing services, for the case of Ireland, has been undertaken by Conniffe and Duffy (1999), that focuses on the problems of correcting for differences in the characteristics of different houses to compute an index of average house prices. Since the Irish CPI measures the costs of owner-occupied housing using mortgage interest costs, and the HICP

---

does not yet include the costs of owner-occupied housing, the results of this research are of limited applicability for our purposes.

<sup>32</sup> This is particularly relevant for the case of the CPI in the Netherlands, which employs the theory of the cost of living index as the conceptual framework for its CPI. The user-cost approach aims at measuring the total cost per unit of service from a commodity to its owner.

<sup>33</sup> The precision of this approach may be enhanced if the distribution of responses (instead of the summary statistic “% better off - % worse off”) is made available by the statistical agency and if this distribution can be compared with the distribution household incomes deflated by the CPI.

<sup>34</sup> Nordhaus’ estimate is 1.5 percent but we were unable to replicate his results exactly.

<sup>35</sup> At <http://www.scb.se/eng/befovalfard/inkomster/hink/hinktab2.asp>

<sup>36</sup> We thank Tom Clark of the UK’s Institute of Fiscal Studies for helping us track down this data.

<sup>37</sup> Commission of the European Communities (1998) notes that “The HICPs can all be said to meet their purposes of ‘measuring inflation by consumers’ to a degree which is unknown (and perhaps unknowable) because there is no reference by which to determine the extent of any bias.” (p.11)

<sup>38</sup> There is a curious inconsistency in the position of advocates of the cost of living approach. They argue that the service flow of durables should be priced because consumers acquire durables with a view to enjoying their services over several years. However, the standard theory of choice that underlies the construction of the traditional cost of living index ignores the fact that consumers maximise not just current utility but also expected future utility. Ideally we should then focus on an intertemporal cost of living index as advocated by Alchian and Klein (1973) and Pollak (1983), and operationalized by Shibuya (1992). However, this raises a whole new set of intractable issues. Some of these issues are addressed in Diewert (2002).

<sup>39</sup> Diewert (2000) reports some illustrative calculations along these lines.

<sup>40</sup> Recall that failure to properly account for quality change in the stock of owner occupied housing in the US prior to 1978, specifically, deterioration in quality due to depreciation, imparted a downward bias to the US CPI for a long time. See Randolph (1988).

<sup>41</sup> Commission of the European Communities (2000) suggests not.

<sup>42</sup> We only show the data for the euro area countries to reduce clutter.

<sup>43</sup> Lequiller (1997) acknowledges that the French CPI may have overstated inflation during the 1993 recession by failing to take full account of retailers greater use of unadvertised discounts during this period. Note that Hoven (1999) states explicitly that list prices are used for the compilation of the CPI for new cars in the Netherlands

<sup>44</sup> In the sense that “substitution” captures an aspect of economic behaviour that cannot be accounted for by a purely statistical index. In practice, national statistical institutes routinely take into account the issue of substitution bias through the update of the relevant basket whenever the effect of the update amounts at least to 0.1% of the index.

---

<sup>45</sup> The US-based retailer Wal-Mart entered the German market in 1998, opening a chain of superstores similar to those it operates in the US offering a range of discount department store goods and groceries. In May 2000, it lowered its prices on a number of staples below their purchase cost in an attempt to gain market share. The Aldi and Lidl chains followed suit, prompting the German Cartel Office to open an investigation under Section 20(IV)(2) of the Act Against Restraints of Competition (Gesetz gegen Wettbewerbsbeschränkungen) which prohibits businesses with superior market power from pricing below cost (except in very limited circumstances). In September 2000, the Cartel Office found that all three retailers were indeed engaging in illegal price cutting, and ordered all three to raise prices. See, for example, the Associated Press story by Stephen Graham of September 9, 2000.

<sup>46</sup> See also Astin (1999) and Gordon (2001).

<sup>46</sup> Camba-Mendez, Gaspar and Wynne (2002) summarise the proceedings of a workshop organised by the CEPR and ECB to address issue of price measurement in Europe.

## References.

Alchian, A. A., and B. Klein (1973) On a correct measure of inflation, *Journal of Money, Credit and Banking* 5, 173-191.

Astin, J. (1999) The European Union Harmonized Indices of Consumer Prices (HICP). *Statistical Journal of the UN Economic Commission for Europe* 16, 123-136.

Balke, N. and M. Wynne (2000) An equilibrium analysis of relative price changes and aggregate inflation. *Journal of Monetary Economics* 45, 269-292.

Bascher, J. and T. Lacroix (1998). On the use of hedonic methods in the CPI: An application to consumer durables and apparel. Paper presented to fourth meeting of Ottawa Group, Washington DC. Available at <http://www.statcan.ca/secure/english/ottawagroup/pdf/24B&L98.pdf>

Bils, M. and P. Klenow (2001) Quantifying Quality Growth. *American Economic Review* 91, 1006-1030.

Blow, L., and I. Crawford (1998) A quality-constant price index for new cars in the UK, 1986 to 1995. Institute for Fiscal Studies Working Paper No. W98/12.

----- and ----- (1999) *Cost-of-living indices and revealed preference*. London: Institute for Fiscal Studies.

---

----- and ----- (2001a) The cost of living with the RPI: Substitution bias in the UK retail prices index. *Economic Journal*, 111, F357-382.

----- and ----- (2001b) A non-parametric method for valuing new goods. Paper presented at the CEPR/ECB workshop “Issues in the measurement of price indices”, 16/17 November 2001.

Bode, B., and J. van Dalen (2001). Quality corrected price indexes of new passenger cars in the Netherlands, 1990-1999. Working Paper, Rotterdam School of Management, Erasmus University, Rotterdam. Presented to sixth meeting of Ottawa Group, Canberra, Australia, April 2001.

Boskin, M. J., E. R. Dulberger, R. J. Gordon, Z. Griliches and D. W. Jorgenson (1996) Towards a More Accurate Measure of the Cost of Living. Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index.

Bover, O. and P. Velilla (2002) Hedonic house prices without characteristics: the case of new multiunit housing. European Central Bank Working Paper No. 117.

Brandolini, A. (1999) The distribution of personal income in post-war Italy: Source description, data quality and the time pattern of income inequality. Banca D'Italia Temi di Discussione del Servizio Studi 350.

Camba-Mendez, G., V. Gaspar, and M. Wynne (2002) *Measurement Issues in European Consumer Price Indices and the Conceptual Framework of the HICP*. Frankfurt am Main: European Central Bank.

Casaccia, M. and G. Seroussi (2000) *Séries longues sur les salaires: Édition 2000*. INSEE Resultats No. 735. Paris: INSEE.

Conniffe, D. and D. Duffy (1999) Irish house price indices: Methodological issue. Paper presented to annual meeting of Irish Economics Association.

Commission of the European Communities (1998) *On the Harmonization of Consumer Price Indices in the European Union*. Report from the Commission to the Council. Catalogue number: CB-CO-98-133-EN-C. COM (1998) 104 Final. Brussels: European Commission.

----- (2000) *On the Harmonization of Consumer Price Indices in the European Union*. Report from the Commission to the Council. COM (2000) 742 final. Brussels: European Commission.

Covas, F. and J. S. Silva (1999) Outlet substitution bias. *Banco de Portugal Economic Bulletin* September, 77-85.



---

Cunningham, A. W. F. (1996) Measurement Biases in Price Indexes: An Application to the UK's RPI. Bank of England Working Paper Series 47.

De Haan, J. (2001) Generalized Fisher price indexes and the use of scanner data in the CPI. Paper presented to sixth meeting of the Ottawa Group.

Deutsche Bundesbank (2000) *Monthly Report August 2000*. Frankfurt am Main: Deutsche Bundesbank.

Diewert, W. E. (1976) Exact and superlative index numbers. *Journal of Econometrics* 4, 114-45.

----- (2000) The Consumer Price Index and Index Number Purpose. University of British Columbia Department of Economics Working Paper No. 00-02.

----- (2001) Research in Price Measurement for the Next Twenty Years. University of British Columbia Department of Economics Working Paper No. 01-11.

Diewert (2002) Harmonized Indexes of Consumer Prices: their Conceptual Foundations, European Central bank Working Paper No. 130.

European Central Bank (1998) A stability oriented monetary policy strategy for the ESCB. ECB Press Release, October 13.

----- (2001) *The Monetary Policy of the ECB*. Frankfurt am Main: European Central Bank.

Eurostat (2001) *Compendium of HICP reference documents (2/2001/B/5)- December 2001*. Luxembourg: Office for Official Publications of the European Communities. Also available at  
:<http://forum.europa.eu.int/Public/irc/dsis/hicop/library?l=/public&vm=detailed&sb=Title>

Feenstra, R. C., and M. D. Shapiro (forthcoming). *Scanner data and Price Indexes*. Chicago: University of Chicago Press.

Folkertsma, C. K. (1998) Measurement bias of the CPI. *Economisch Statistische Berichten* 83, 189-191.

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

----- (2001) Apparel prices and the Hulten/Bruegel Paradox. Department of Economics, Northwestern University.

Greenspan, A. (2001) Transparency in monetary policy. Remarks at the Federal Reserve Bank of St. Louis Economic Policy Conference, October 11.

---

Harhoff, D. (1999) A hedonic price index for automobiles in Germany. Draft, University of Munich.

Hausman, J. A. (1997) Valuation of new goods under perfect and imperfect competition. In Timothy F. Bresnahan and Robert J. Gordon (eds.) *The Economics of New Goods*. Chicago: University of Chicago Press.

----- (1999) Cellular telephone, new products and the CPI. *Journal of Business and Economic Statistics*, 17, 188-194.

Hoffman, J. (1998) Problems of Inflation Measurement in Germany. Discussion Paper 1/98. Economic Research Group of the Deutsche Bundesbank.

----- (1999) The treatment of quality changes in the German Consumer Price Index. Paper presented to fifth meeting of Ottawa Group, Reykjavik, Iceland. Available at <http://www.statice.is/ottawa/hoffmann.pdf>

Hoffmann, J. and C. Kurz (2001) Rent indices in West Germany: 1985 to 1998. Discussion Paper 01/02. Economic Research Centre of the Deutsche Bundesbank. Also European Central Bank Working paper No. 116.

Hoven, L. (1999) Some observations on quality adjustment in the Netherlands. Paper presented to fifth meeting of Ottawa Group, Reykjavik, Iceland. Available at <http://www.statcan.ca/secure/english/ottawagroup/ottawa/hoven.pdf>

Izquierdo, M. and M<sup>a</sup> de los Ángeles Matea (2001). Hedonic Prices for Personal Computers in Spain during the 1990s. Banco de España ‘ Servicio de Estudios Economicos, n<sup>o</sup> 74, 2001.

Landefeld, J. S. and R. P. Parker (1995) Preview of the comprehensive revision of the national income and product accounts: BEA’s new featured measures of output and prices. *Survey of Current Business*, 75, 31-38.

Lequiller, François (1997) Does the French Consumer Price Index Overstate Inflation? *INSEE Document de Travail G 9714*.

Licandro, O., M. Izquierdo and A. Maydeu (2001) Car quality improvements and price indices in Spain. Banco de España ‘ Servicio de Estudios Economicos, n<sup>o</sup> 72, 2001.

Lichtenberg, F. and Z. Griliches (1989) Errors of measurement in output deflators. *Journal of Business and Economic Statistics*, 7, 1-9.

Manser, M. E. and McDonald, R. (1988) An analysis of substitution bias in measuring inflation, 1959-1985. *Econometrica* 56, 909-30.

---

Moreau, A. (1996) Methodology of the price index for microcomputers and printers in France. In OECD (ed.) *Industry Productivity: International Comparison and Measurement Issues*. OECD Proceedings. Paris: OECD.

Neves, P. D. and L. M. Sarmento (1997) The substitution bias of the consumer price index. *Banco de Portugal Economic Bulletin*, June, 25-33.

Nordhaus, W. D. (1998) Quality changes in price indexes. *Journal of Economic Perspectives*, Winter, 59-68.

National Statistics (1997) Is the Retail Prices Index Biased? *Consumer Price Indices*, December 1997. Available at [http://www.statistics.gov.uk/themes/economy/Articles/PricesAndInflation/downloads/Is\\_the\\_RPI\\_biased.pdf](http://www.statistics.gov.uk/themes/economy/Articles/PricesAndInflation/downloads/Is_the_RPI_biased.pdf)

OECD (2000) OECD Economic Outlook, 67. Paris: Organisation for Economic Cooperation and Development.

Oulton, N. (1995) Do UK Price Indexes Overstate Inflation? *National Institute Economic Review* 152, 60-75.

Poinat, F. (1997) Estimation of the potential bias of the French CPI due to annual linking of intermediary aggregates. Paper presented to third meeting of Ottawa Group, Voorburg, Netherlands. Available at <http://www.statcan.ca/secure/english/ottawagroup/pdf/07POI3.pdf>

Pollak, R. A. (1983) *The Theory of the Cost of Living Index*. Oxford: Oxford University Press.

Randolph, W. C. (1988) Housing depreciation and aging bias in the consumer price index. *Journal of Business and Economic Statistics* 6, 359-71.

Ruiz-Castillo, J., E. Ley and M. Izquierdo (2000a) The Laspeyres bias in the Spanish consumer price index. Fedea Documento de Trabajo 2000-05.

-----, ----- and ----- (2000b) Distributional aspects of the quality change bias in the CPI: Evidence from Spain. Photocopy, Departamento de Economía Universidad Carlos III.

Santos, E. A. and C. Coimbra (1995). Preços de automóveis e alterações de qualidade. Uma abordagem em termos de preços hedónicos. *Boletim Económico*, Banco de Portugal, 1, September.

Shibuya, H. (1992) Dynamic Equilibrium Price Index: Asset Prices and Inflation, *Bank of Japan Monetary and Economic Studies* 10, 95-109.

---

Silver, M. and S. Heravi (2002) A failure in the measurement of inflation: Results from a hedonic and matched experiment using scanner data. European Central Bank Working Paper No. 144.

Silver, M. and C. Ioannidis (2001) Intercountry differences in the relationship between relative price variability and average prices. *Journal of Political Economy* 109, 355-374.

Stigler, G. J., ed. (1961) *The Price Statistics of the Federal Government: Report to the Office of Statistical Standards. Bureau of the Budget*. New York: National Bureau of Economic Research.

Sveriges Riskbank (2001) *Inflation Report*. December 2000.

Triplett, J. (1990) Hedonic methods in statistical agencies: An intellectual biopsy. In E. R. Berndt and J. E. Triplett (eds.), *Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth, NBER Studies in Income and Wealth Volume 54* (pp. 207-33). Chicago: University of Chicago Press.

----- (1992) Economic theory and the BEA's alternative quantity and price indexes. *Survey of Current Business* 72, 49-52.

U.S. Census Bureau (2001) *Money Income in the United States: 2000, Current Population Reports, P60-213*. Washington D.C.: U.S. Government Printing Office.

Wynne, M. A., and F. Sigalla (1996) A survey of measurement biases in price indexes, *Journal of Economic Surveys* 10, 55-89.

Young, A. H. (1992). Alternative measures of change in real output and prices. *Survey of Current Business* 72, 32-48.