CPI measurement issues with a special focus on owner-occupied housing

Alexandra Heath

Bank for International Settlements[1]

"Problems in measuring inflation"

National Bank of the Republic of Macedonia, Tuesday 22 May 2007

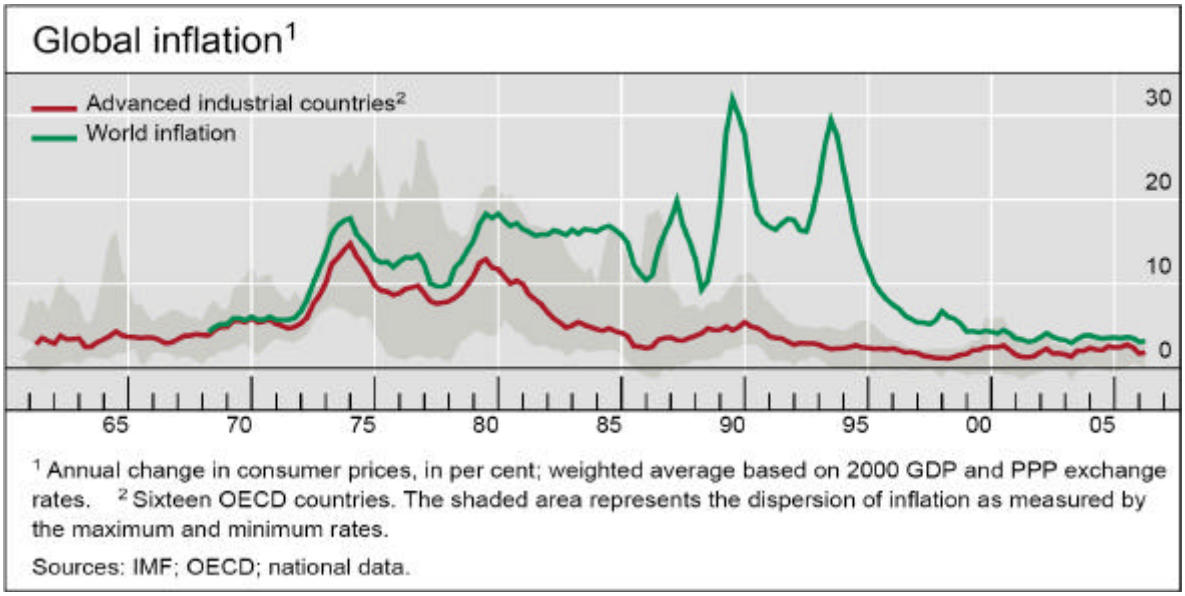
1. Introduction

CPI inflation is one of the most important indicators of macroeconomic conditions and, as such, there has been a long history of collecting price data with a view to getting a clear picture of how aggregate price dynamics are evolving. Even so, there appears to have been more interest in the measurement issues associated with the CPI and the impact these might have on macroeconomic policy making in the past decade or so. There are several reasons for this.

First, there has been a step down in the level of inflation in most countries, both industrialised and emerging (see Graph). This means that some of the biases that can arise from measurement issues are proportionately larger than they used to be. For example, the recent rebasing of the Japanese CPI led to a revision of 0.5 percentage point, which had significant policy implications given the low level of Japanese inflation. At the same time, there has been a trend towards inflation targeting, which by its very nature leads central bank analysts to spend more time thinking about the underlying drivers of aggregate price developments. Measurement issues are also likely to receive considerable attention at the time a central bank is thinking about defining an appropriate inflation target.

Another very important motivation for increased attention to measurement issues is that there have been significant changes in economic structure over the past decade or so. This is true even in industrialised economies that have not experienced major transitions like those experienced in Macedonia. First, there have been significant changes in the way things are sold to consumers. For example, there has been a trend in many countries to larger stores that can offer the same goods for less cost, and the internet has revolutionised the marketing and distribution of other goods and services, such as airline travel. Second, there has been an increase in the rate at which new products are introduced and in the rate at which they evolve. Innovation has been particularly rapid for IT based products such as computers and mobile phones, but also extends to products such as cars and white goods. Finally, globalisation and increased integration across economies have also increased the range of products that are available to consumers.

[1] Any opinions expressed in this paper are those of the author and do not necessarily reflect the views of the Bank for International Settlements.



In addition, a number of papers investigating the nature and magnitude of measurement issues have highlighted the importance of various measurement issues. The Boskin Report, commissioned by the US Congress in 1996, published estimates that the US CPI was upwardly biased by about 1.1 percentage points a year. These findings spurred several methodological changes in the US and elsewhere. Studies for other countries and more recent estimates for the US indicate biases with smaller magnitude. The challenges that have arisen in the process of designing and publishing the HICP for the euro area have also highlighted the importance of understanding and dealing with measurement issues.

At the end of the day, CPI measurement issues are likely to be economy-specific to some extent: different countries will face different constraints when trying to correctly measure the evolution of consumer prices. For example, the ability to deal with quality adjustment issues will depend on the availability of detailed product-level data, and sampling issues will potentially create bigger biases in economies undergoing more significant structural changes in terms of retail distribution networks and product innovation.

In 2006, the Irving Fisher Committee, hosted by the BIS, held a workshop focussing on CPI measurement issues from the perspective of central bankers. In what follows I have used material from this workshop extensively. In addition to detailed discussions on specific topics, there were several broad messages from this workshop. The first is that each economy has different features and the compilers of CPI data, optimally an agency independent of the central bank especially in the inflation targeting context, face different constraints and degrees of data availability. As such, there is no one-size-fits-all solution to the measurement issues and trade-offs that arise when compiling a CPI. Having said that, it is important for each compiling agency to continue to work towards improving their CPI and, to this end, an appreciation of the trade-offs is important for deciding where effort is best spent. The second message is that macroeconomic policy makers rely on the quality of CPI data and on information about the collection of CPI data. As such it is important for policy makers to form strong working relationships with their national statistical agency, and in this regard, workshops like this one, which involves central bankers as well as statisticians, was very positive.

The first part of the paper discusses the main biases that arise in the measurement of the CPI. In particular, the paper briefly describes and discusses some of the important issues associated with substitution, new product, outlet and quality bias. The second section looks at the measurement issues associated with owner-occupied housing, and makes a few

comments on the measurement of services. Finally, there is a brief discussion about measuring underlying inflation for macroeconomic policy purposes.

2. Biases in CPI measurement

There are three kinds of measurement problem that can affect the final CPI number. First, there are issues associated with aggregating individual price series that result in *substitution bias*. Second, there are several issues associated with the sampling and collection of individual price series. *New product bias* arises when the prices of new consumption items are not sampled, whereas *outlet bias* may be introduced if price collection does not take into account changes to the structure of retail distribution. Third, concerns about *quality bias* raise questions about how price movements should be adjusted for changes in the characteristics of a product over time. Issues associated with product innovation that underlie possible new product bias are also relevant here.

The Irving Fisher Committee on Central-Bank Statistics conducted a survey as background for its workshop on "CPI measures: central bank views and concerns" in April last year, which included a question on which of these measurement issues was regarded as the most problematic. For central banks which provided a ranking, two thirds ranked quality bias and substitution bias as the top two biases. Within this group, a slightly larger number ranked quality bias above substitution bias. New product bias was also recorded in the top two ranking for around one quarter of responding central banks, mostly from Europe.

Estimates of these biases, and their relative importance, are available for some countries: they vary from less than 0.5 percentage points to well over 1 percentage point, but are not generally reliably estimated. The tone of caution is well expressed in the following quote from Rudd (2006):

"[T]his approach is uncomfortably subjective. (Under the circumstances, that might be the best we can do.) Related to this is a concern that quantification of a bias estimate (and its confidence interval) might imply a degree of certainty that we simply do not have. Finally, this method tells us little or nothing about the nature of time variation in overall CPI bias, which might be of particular interest to a monetary policy maker."

2.1 Substitution bias

Substitution bias refers to the bias in price indices that results from consumers switching their consumption patterns. This is probably the best researched and understood of the biases discussed here, and in many cases changes made to CPI compilation methods have reduced this source of bias. Substitution bias can enter into a price index at two levels.

First, it can occur at the expenditure class level, usually in response to changes in tastes and income levels. For example, as income levels change, there may be a tendency for consumer to buy rather than to make clothes. The only real way of minimising this bias is to update expenditure weights regularly. Theoretically, superlative indices, eg the Tornqvist index offer the best solution, as they utilise information on base and current period weights. From a practical perspective, however, it is difficult to obtain current period weights in a sufficiently timely manner to make these indices useful for policy purposes. A practical compromise is generally considered to be a chained Laspeyres index that updates weights annually.

Second, substitution can occur below the expenditure class level, for example, a switch from one brand of biscuit to another. The Consumer Price Index Manual recommends that using a geometric mean to aggregate sub-expenditure class price movements is the best way of mitigating this problem because it allows for some substitution, whereas other averaging methods do not. It is also important in this context to ensure that the range of items sampled

at the sub-expenditure class level is sufficiently large to ensure that the unavailability of a particular product, such as a particular variety of fruit, does not affect the integrity of the CPI calculations.

2.2 New product bias

New product bias arises when there is rapid product innovation and prices for new items, eg MP3 players, are not sampled or given a weight in the consumption bundle. As with substitution bias at the expenditure class level, the only real solution for new product bias is to update weights on a regular basis. Assigning an appropriate weight in a timely fashion to a completely new item can be difficult if there is no recent information from a comprehensive household expenditure survey. Problems stemming from product innovation are also relevant for the question of quality bias (see below).

2.3 Outlet bias

Outlet bias also relates to sampling issues, but occurs when there is a significant change in retail distribution networks. For example, outlet bias could arise if there is a substitution away from small corner stores to large malls, or if there is a change from buying airline tickets at travel agencies towards buying them over the internet. In this case, the question of what weight to give to different possible sources for a given consumption item is relevant. It is likely that the question can only be answered by doing a targeted analysis of the consumption item in question because this information is not generally available from household expenditure surveys. The factors driving changes in retail distribution networks, such as cost minimisation, may also lead to quality changes in the products being sold.

2.4 Quality bias

Quality bias is arguably the most important measurement issue associated with compiling a CPI. There are several reasons why quality adjustment may have become a more important issue over time. The first is that the rate of product innovation and the rate at which new products are being introduced have increased. The second is that there has generally been an increase in the variety of products available of any given type. This makes sample design more complicated and makes differentiating products by quality as well as price more important. Bias arising from not adjusting prices for quality changes, or not correctly adjusting for quality changes could work in either direction. Available estimates suggest that quality adjustments that generate downward pressure on prices, eg those arising from technological change, are most likely to have been neglected, resulting in an upward bias.

For homogenous items, the issue of quality adjustment is relatively straight forward. If there is a change to the number of sheets of toilet paper in a roll, or the weight of a packet of cornflakes, it is easy to measure the change in quality and to make the appropriate price adjustment. Quality adjustment for other, less homogenous items, such as fruit and vegetables, is often done using the *expert judgement* of those collecting the prices first hand. The price collector may not only make subjective judgements about quality changes, but may also decide to replace a particular item in the sample that has been discontinued, or has otherwise become unreliable, with a comparable item.

Implicit methods of quality adjustment include the *overlap method* and the *class mean method*. The overlap method relies on the prices of the old item and its replacement to both be available in two periods, which enables the price series to be spliced. This method implicitly assumes that all the price difference was due to a quality difference. In light of evidence that retailers wait for the introduction of new products to change prices, this assumption may not be a good one to start with. The class mean method calculates the price changes for all the items within a given group that did not have any changes in

characteristics. This method benefits from having large samples as this increases the observations for which no quality change occurred and prices in both periods are available. This method also deals with the removal of old products and the introduction of new ones, although, if the rate of change in the characteristics of the consumption item is too high, there may not be a sufficient number of matching prices for the overall estimate to be reliable.

A more explicit adjustment for a quality change can be made when prices are available for the individual product features. For example, if the price of a car increases because a feature such as airbags has become a standard part of the price, the individual value of an airbag can be used to make the quality adjustment. This is known as the *production* or *option cost approach*. To the extent that individual features are more expensive when sold separately, this will result in a downward bias in prices because the quality adjustment is too large. This method of quality adjustment is sometimes combined with expert judgement to mitigate this effect, and is frequently used for motor vehicles and computers.

Another explicit method of quality adjusting price data, which has become increasingly popular, is the *hedonic pricing* method. This method regresses the price of an item on information about specific features that contribute to its value. Using the estimated parameters, it is then possible to calculate the change in the price between two varieties of the same good due to changes in specific features. Based on the survey carried out for the IFC workshop, the use of hedonic pricing for quality adjustment appears to be more pervasive than in the past, and is most likely to be used for audio visual, computing and other electrical appliances. In some cases, these methods are also used to quality adjust clothing and book prices.

The decision about whether or not hedonic pricing methods are justified relies on several factors. Working in favour of hedonic pricing methods for quality adjustment is the fact that they are relatively objective and have the potential to be more transparent. This benefit can be enhanced by the organisation doing the statistical collection and quality adjustment releasing details of the methodology used, including information such as the assumed functional forms used. Having said that, hedonic pricing methods of quality adjustment rely on having large amounts of data on the prices and characteristics of individual items, and there are a number of significant methodological issues, such as the functional form assumption, which must be dealt with.

3. Individual price measurement

In addition to the biases associated with aggregating, sampling and correctly measuring constant quality price developments, there are some price series which have proved to be far more difficult to measure than others, both at a conceptual and a practical level. In particular, there has been some disagreement about the correct way to measure owner-occupied housing prices, and this has been reflected in the variety of methods used across countries. Indeed, the difficulties are so great in some cases that owner-occupied house prices have been excluded altogether. Measuring constant quality prices for services has also been challenging for CPI compilers. These issues may become more important over time as services become an even more significant share of the consumption bundle.

3.1 Owner-occupied housing

In terms of measurement issues associated with a specific price series, housing service prices, particularly the price of owner-occupied housing, is one of the most discussed. This is partly because housing services typically account for a large fraction of consumption expenditure. In addition, there are several approaches to measuring owner-occupied housing services, each based on different assumptions, which can yield quite different readings for inflation. At this stage, the difficulties associated with measuring owner-occupied housing

prices has meant that Eurostat has decided not to include these prices in the euro area HICP until some satisfactory resolution of these issues can be found. There are four main methods: the rental equivalence approach, the user-cost approach, the payment approach and the net acquisition approach.

The *rental equivalence* approach assumes that housing services from rental property and owner-occupied property are perfect substitutes. Over the past decade or so, a number of countries have experienced rapid growth in house prices accompanied by relatively moderate growth in rental prices, a combination which suggests that this assumption does not hold. In addition, the stocks of private (no price regulation) rental properties must be sufficiently large and comparable to owner-occupied housing to allow enumeration. Within the EU there are six countries for which the rental market makes up less than 10% of the housing market (Eiglsperger, 2006). Another potential problem with the rental equivalence method is that it assigns a relatively large weight to housing services, and so any biases in the measurement of the price index are magnified. Despite these issues, the rental equivalence method is widely used in a range of countries including Germany, Hong Kong, Japan, Mexico, the Netherlands, Poland, Singapore, Thailand and the United States (Table 1).

The *user cost* approach tries to estimate the actual and imputed costs of using the owner-occupied housing stock in any given period. This is more complicated for housing that some other consumption items because it is a durable good that delivers services over many periods, well beyond the date of purchase. Typically, the costs that are measured include interest payments of mortgages, depreciation and other recurring costs such as taxes and minor maintenance costs. The opportunity costs of investing financial assets in a house and capital gains (with a negative sign) could also be included. However, these factors are often excluded based on the argument that capital gains and foregone interest earnings are important for the investment component of the house-purchase decision, but are not relevant for the consumption of housing services, eg Statistics Canada. In an environment where there is perfect substitutability between the rental and owner-occupied markets, the user cost and the rental equivalence methods should yield similar results.

There are several issues with the user cost approach. First, the inclusion of mortgage interest payments in the CPI is problematic if monetary policy uses interest rates as an instrument and targets CPI inflation. In this situation, an increase in interest rates designed to curb inflation, will automatically increase CPI inflation through this component. In cases where this is a potential problem, the target inflation rate is usually defined as the CPI excluding interest payments. Another issue with the user cost approach is that several of the imputed costs are a function of house prices, which include an asset (investment) as well as a consumption component. It is important, for example, that the value of land is removed from the calculation because land is not consumed. Even if this can be done, it is necessary that the data on house sales is representative of the whole owner-occupied housing stock especially in situations where there are rapidly rising house prices. The user cost approach is used in Canada, China, South Africa and Sweden.

The *payment approach* reflects actual outlays made by households over the period in which the consumer good is used. In general, this includes mortgage interest payments, taxes and maintenance and repair costs, and is therefore closely related to the simplified user cost approach. The main distinctions are the size of maintenance and repair costs that will be included, and the fact that depreciation is excluded. The weight given to owner occupied housing is also generally smaller in the payment approach. The issue of interest rates being the instrument of monetary policy as well as one of the components of the monetary policy target will continue to be present.

Table 1: CPI treatment of housing services

	Method of measuring owner-occupied housing (OOH)	Housing Weights		
		OOH	Rental	Total
Argentina	Not included	-	12.70	12.70
Australia	Net acquisition	11.21	5.22	16.43
Belgium	Not included	-	6.39	6.39
Brazil	Not included	-	3.80	3.80
Canada	User cost	15.98	6.14	22.12
China	User cost	na	13.00	13.00
Euro area	Not included	-	6.32	6.32
France	Not included	-	6.14	6.14
Germany	Rental equivalence	na	na	21.17
Hong Kong SAR	Rental equivalence	na	na	29.86
Indonesia	Not included	-	26.25	26.25
Italy	Not included	-	2.70	2.70
Japan	Rental equivalence	13.60	3.48	17.08
Korea	Not included	-	13.14	13.14
Mexico	Rental equivalence	11.97	2.52	14.49
Netherlands	Rental equivalence	9.13	6.69	15.82
Poland	Rental equivalence	11.97	4.54	16.51
Portugal	Not included	-	2.02	2.02
Russia	Not included	-	12.53	12.53
Singapore	Rental equivalence	na	14	14
South Africa	User cost	13.94	4.27	18.21
Spain	Not included	-	2.31	2.31
Sweden	User cost	16.07	10.75	26.82
Switzerland	Not included	-	18.72	18.72
United States	Rental equivalence	23.44	5.83	29.27
Source: IFC (2006), pp 63-64.				

The final method of measuring prices for owner-occupied housing is the *net acquisitions* approach. This approach aims to measure the expenditures made on new dwellings for the purposes of owner-occupation, as well as other current expenditures, eg maintenance costs. This method has the advantage of measuring the prices of owner-occupied housing consistently with the prices of other consumer durables, such as motor vehicles and white goods: ie purchases of housing from other private households are excluded, and the expenditure is recorded at the time it is made rather than as the value of the flow of services being imputed over the life of the asset. As discussed earlier in the context of the user cost approach, it is important to separate out the asset price component of a new dwelling. In practice, net acquisition measures exclude the value of land from the value of new dwellings to this end. One consequence of these assumptions is that the weight assigned to owner-

occupied housing under the net acquisition concept is relatively small compared with those discussed earlier. Diewert (2002) estimated that the expenditure weight is likely to be around half the size of the expenditure weights calculated under user cost methodologies.

Eiglsperger (2006) points to several data problems with the net acquisitions methodology: it can be difficult to obtain data on recent sales of new dwellings in a timely manner, and removing the value of land can be complicated. In the case of Australia, obtaining data has not proved to be so problematic. The cost of constructing a new dwelling is primarily based on data provided by project home builders. As such the cost of purchasing new dwellings reflects factors such as the cost of building materials and labour costs. Maintenance costs are assumed to reflect the same influences, and are included by increasing the weight of this component rather than by measuring these price movements separately.

Some assessment of the relative merits of the different approaches and the consequences of their underlying assumptions can be made by looking at some preliminary estimates of what the HICP for the euro area would have looked like had different measures of owner-occupied housing prices been included (Eiglsperger, 2006). The user cost estimates, based on OECD calculations result in the most volatile CPI readings, mainly because they are strongly influenced by changes in interest rates and house prices. The rental equivalence and net acquisition methods result in somewhat similar outcomes that are somewhat similar to the HICP as currently measured. The net acquisition method is the one preferred by Eurostat from a conceptual stand point, and data are now being collected to enable a more formal calculation of owner-occupied housing prices on this basis. The IMF Manual is neutral on this topic.

3.2 Services

Services have become increasingly important as a part of the consumption basket over time and accurately measuring service prices presents one of the most difficult challenges to CPI compliers. There are several reasons for this. First, services are rarely homogenous, and tend to be delivered in quite a disaggregated way, so designing the sampling is often less straight forward than it is for goods. A good example of these problems is offered by health services. It would be difficult to sample the full range of services and providers that can be purchased, so often statistical collections rely on ‘representative’ data, such as standard recommended fees and data collected by government departments for policy purposes. Second, quality adjustment is particularly challenging for services: it is rarely straightforward to quantify the characteristics of services, and even when it is possible, the sampling issues make it hard to get sufficiently detailed data to perform an adequate quality adjustment.

One example where the significant sampling issues have been overcome is the measurement of financial services in the Australian CPI. Statisticians at the Australian Bureau of Statistics have arranged to collect details on large numbers of individual bank accounts, which enables them to calculate changes to the fees paid and changes in interest rate margins, which is also another way of charging customers for financial intermediation services (Australian Bureau of Statistics, 2004).

4. Measuring underlying inflation

In general, macroeconomic policy should not be responding to volatile, temporary relative price movements, but rather to lower frequency price developments that are more widely spread. This is certainly true for central banks and monetary authorities with a mandate to target inflation, but also for other central banks for whom consumer price inflation is still an important part of the macroeconomic environment within which monetary policy decisions are being made. Understanding the way in which the CPI is compiled and the extent to which

price developments can be affected by measurement issues are both issues that are important for being able to make this distinction.

From a policy perspective, it is also often helpful to have aggregate measures of underlying inflation. For example, a complete and consistent time series for underlying inflation can be used for analytical exercises such as forecasting. Underlying inflation measures may also be useful for communicating with the public. However, making the statement that it is useful to have measures of underlying inflation that capture the persistent trends in aggregate prices does not give very precise guidance on how to go about calculating these measures.

There are several approaches to measuring underlying inflation. The first is to exclude specific price series that tend to be more volatile or unrepresentative of aggregate price movements. Standard core inflation measure that exclude food and fuel in all time periods fall into this category. The danger with this method is that some price movements, such as the ongoing price pressures from rising oil and food prices, may be excluded even when they are potentially important for understanding underlying inflation trends. Exclusion-type measures are probably the most common.

The second is to use statistical measures, such as a trimmed mean, to down-weight information coming from large relative price movements. These methods have the advantage that they use all the information available, although the influence of large relative price movements will be quite limited. However, the optimal, or even an acceptable, degree of trim is somewhat arbitrary. In addition, it is not clear that some of the assumptions underlying the justification for these measures hold when there are individual price series with large weights, eg housing in the United States. Although not as wide-spread as exclusion based measures, statistical measures of this sort are used in a number of countries including Australia, Germany, Poland, Sweden and Switzerland.

A third method is to reweight items according to characteristics of the individual price series, such as their persistence. An example is the weighted variance CPI measure used in Canada. The fourth method uses assumptions about the time series properties of a measure of underlying inflation, possibly in conjunction with some structure from an economic model. Examples include dynamic factor models, used in Italy and Switzerland, and theoretical models, such as the one under development at the Bank of Italy.

In general, there are no precise criteria for judging which measure of underlying inflation is the ‘best’. However, there are several features that the literature has agreed are important for a measure of underlying inflation to be ‘good’ for the purposes of monetary policy. In particular, underlying inflation should: not be more variable than headline inflation; have limited bias with respect to headline inflation; not be subject to revision; be replicable; be timely; and improve forecasting performance. Another feature that is generally regarded as desirable is that it is internationally comparable, but this is probably of second-order importance for the domestic monetary policy maker per se.

In practice, most central banks are likely to use more than one measure of underlying inflation for analytical purposes, recognising that no one measure is optimal. Many central banks also use more than one underlying inflation measure as sufficient statistics for a more detailed analysis of individual prices series in their communication. Publishing more than one measure of underlying inflation can potentially complicate the central bank’s communication with the public about underlying inflationary pressures, particularly if these measures diverge. However, the experience of several central banks, including the Reserve Bank of Australia, suggests that this is not necessarily a problem.

References

Australian Bureau of Statistics (2004), Experimental price indexes for financial services: 1998-2003, Information Paper, ABS Cat. No. 6413.0.

Diewert, E (2003), "The treatment of owner-occupied housing and other durables in a consumer prices index".

Eiglsperger, Martin (2006), "The treatment of owner-occupied housing in the harmonised index of consumer prices", IFC Bulletin 24, August 2006, pp 68-79.

IFC Bulletin (2006), "Proceedings of the IFC workshop on CPI measures: central bank views and concerns", IFC Bulletin 24, August 2006.

Rudd, Jeremy (2006), "Measurement error in the US CPI", IFC Bulletin 24, August 2006, pp 152-155.