

THE PERFORMANCE OF TRIMMED MEAN MEASURES OF UNDERLYING INFLATION

Andrea Brischetto and Anthony Richards*
Reserve Bank of Australia

Revised draft: May 2007

Abstract

This paper uses data for Australia, the United States, Japan and the euro area to examine the relative performance of the headline CPI, exclusion-based ‘cores’, and trimmed means as measures of underlying inflation. We find that trimmed means tend to outperform headline and exclusion measures on a range of different criteria, indicating that they can be thought of as having better signal-to-noise ratios. We also find that there is a wide range of trims that perform well. One innovation for the United States is to break up the large implicit rent component in the US CPI into four regional components, which improves the performance of trimmed means, especially large trims such as the weighted median.

The results lend support to the use of trimmed means as useful measures of underlying inflation at the current juncture where the growth of China and other emerging markets is having two offsetting effects on global inflation. Whereas some central banks have tended to focus on headline inflation and others have focused more on exclusion measures, our results provide some justification for a middle path, namely using trimmed mean measures which deal with outliers at both ends of the distribution of price changes in a symmetric manner.

*This is an updated and revised version of RBA Research Discussion Paper 2006-10. We thank a number of colleagues including Christian Gillitzer, Jonathan Kearns and John Simon for helpful discussions. The views expressed here are those of the authors and do not necessarily reflect those of the RBA. Address for both authors is Economic Group, Reserve Bank of Australia, PO Box 3947, Sydney 2001, Australia. Emails: BrischettoA@rba.gov.au, RichardsA@rba.gov.au.

THE PERFORMANCE OF TRIMMED MEAN MEASURES OF UNDERLYING INFLATION

Andrea Brischetto and Anthony Richards

1. Introduction

Headline consumer price inflation can be quite noisy, especially over short periods, and sometimes will not give a good measure of the current trend in inflation, which is a key input for any inflation forecast and for monetary policy decision making. Accordingly, many central banks focus on various measures of underlying inflation which they expect will contain less noise and may thus be helpful for meeting their ultimate objectives. Exclusion measures, which remove the effects of price changes of a fixed set of items such as food and automotive fuel, are widely used in a range of countries. Statistical measures, in particular trimmed mean measures¹, represent another approach and have received significant support from academic studies and are used by some central banks. In contrast, some other central banks downplay the use of underlying measures in their analysis or discussion of inflationary trends, focusing more on the headline consumer price index (CPI), often on the rationale that it is the ultimate goal of policy.

This study assesses the relative benefits of headline CPI, exclusion measures, and trimmed mean measures as indicators of the underlying trend in inflation. We use a number of criteria, but they mostly come down to a search for measures which appear to contain a high signal-to-noise ratio. One aspect of this is that if underlying measures of inflation are successful in removing noise from CPI data, they should also be helpful in predicting near-term inflationary outcomes. Hence, an important test for various measures of underlying inflation will be the extent to which they have some predictive power for near-term inflation. Accordingly, our criteria correspond closely to Blinder's often-cited comments on his time watching month-to-month inflation as a central banker: 'The name of the game then was distinguishing the signal from the noise, which was often difficult. The key question on my mind was typically: What part of each monthly observation on inflation is durable and what part is fleeting ... To me, the durable part of the information in each monthly inflation report was the part that was useful in medium- and near-term inflation forecasting' (Blinder 1997, p 157).

There are a number of earlier studies that assess different measures of underlying inflation, although these mostly use data for a single country.² However, the recent run-up in oil prices has shown that there is still a wide range of views as to the relative merits of different measures of underlying

1 We use the term 'trimmed mean measures' to also encompass the weighted median, which can be viewed as the 50 per cent symmetrical trimmed mean.

2 For example, see Bryan and Cecchetti (1994) and Smith (2004) for the United States, Vega and Wynne (2003) for the euro area, Shiratsuka (2006) for Japan, and Kearns (1998) and Roberts (2005) for Australia. We are aware of only one other multi-country study of different measures of inflation, by Catte and Sløk (2005).

inflation – witness the different focus of some central banks on the headline CPI, others on inflation excluding the effect of fuel and food, and still others on statistical measures. Accordingly, there is scope for further work in this area.

We use data for four different economies – the United States, the euro area, Japan and Australia – to look for evidence on the performance of different measures of underlying inflation that is reasonably robust across these economies and thus is likely to also apply to other countries. We consider the performance of a wide range of trims, as we move from the headline CPI (0 per cent trim) to the weighted median (50 per cent trim), and also compare these against the performance of exclusion measures of underlying inflation. We also address several technical issues, including the effect of items with very large CPI weights on measures of underlying inflation. Our approach is unashamedly empirical and atheoretical, focusing not on debates over how to define underlying inflation or on the theoretical case for using trimmed mean measures³, but rather assessing which of the measures of inflation are likely to be useful indicators for analysts both inside and outside central banks.

To preview our conclusions, we find that trimmed mean measures offer substantial advantages over headline inflation in terms of providing measures of inflation which have a higher signal-to-noise ratio. Our analysis of different degrees of trimming suggests that there is a wide range of trims that outperform exclusion measures of inflation. In addition, we find that the presence of large expenditure items has implications for the degree of trimming that will be optimal. In the case of the US CPI, we highlight the benefits of breaking up the large imputed rent of home-owners component into regional subcomponents, which yields improved measures of underlying inflation. Overall, while there are some important technical issues which must be considered when constructing trimmed mean measures, we conclude that these measures are quite robust in the sense that there is a wide range of trims that perform well for all the economies we consider.

2. Background on Underlying Inflation

2.1 Different Measures of Underlying Inflation

There are a number of different views as to what constitutes underlying inflation. At a theoretical level, it is sometimes said to correspond to the ‘monetary’ component of inflation. At a more practical level it is frequently described as either the ‘persistent’ or ‘generalised’ component of inflation. Behind both of these latter notions is the idea that in any month or quarter, there can be significant noise in the CPI (or other standard price indices) which may not be indicative of the underlying trend in inflation. The noise in short-horizon movements in the CPI reflects a range of relative price movements due to developments in commodity markets, weather, infrequent resetting of prices or taxes, and so on. Of course, not all changes in relative prices are noise, but much of this high-frequency movement in the CPI will indeed be relatively temporary.

³ Numerous other papers, including Roger (1998) and Wynne (1999), provide a discussion of the definition of core or underlying inflation; Bakhshi and Yates (1999) discuss the theoretical case for trimmed mean measures.

Policy-makers and other analysts often make use of measures of core or underlying inflation, which should be subject to less noise than headline inflation and should provide a better measure of the trend in inflation. Most measures of underlying inflation can be distinguished by whether they use reweighting across items or smoothing over time to reduce the noise in the headline CPI (Table 1).

Table 1: Classification of Measures Of Underlying Inflation

	No reweighting of CPI items	Smoothed by reweighting of CPI items
Not smoothed across time	Headline CPI inflation	Exclusion-based ‘cores’ Trimmed means Volatility weighted Persistence weighted
Smoothed across time	Exponentially smoothed Component smoothed	Dynamic factor models

One method of removing noise is to use an average of recent monthly or quarterly rates, for example using annual rates of inflation rather than just the price movement in the latest month or quarter. While this will generally be somewhat effective in creating a smoother measure, it will be a backward-looking one and will probably place insufficient weight on the information in the data for the most recent month or quarter. Indeed, as was seen in a number of countries in late 2006, annual rates of inflation can change rapidly, not only because of some movement in prices in the most recent period, but because of movements in prices (notably for oil) that happened a year earlier.

Of course, more sophisticated approaches to smoothing are possible. For example, Cogley (2002) proposes an exponentially weighted average of past headline rates as a measure of underlying inflation. And a recent paper by Gillitzer and Simon (2006) suggests a varying degree of time averaging for different items in the CPI basket. Their ‘component-smoothed inflation’ measure proposes only minimal averaging for items with prices that are relatively smooth but a higher degree of averaging for items with more volatile prices.

However, most measures of underlying inflation typically do not use time averaging but instead use some form of reweighting of CPI expenditure items. The most common approach is to calculate the inflation rate for the CPI basket excluding a few items (typically food and/or energy) which historically have had particularly volatile prices. Such exclusion-based measures – frequently simply called ‘core’ inflation – remove the direct effect of movements in the prices of some items on the rationale that they tend to be volatile and often not reflective of the underlying or persistent inflation pressures in the economy. However, exclusion measures of underlying inflation may not always be appropriate. For example, there may be large temporary movements in other components of the CPI that are not excluded from such measures. In addition, there may indeed be information about underlying inflation pressures in the excluded components, but such information will be lost in an exclusion measure.

Trimmed mean measures, which are discussed further in Section 3, represent an alternative approach and are used in a number of central banks. These are measures of the mean or average price change for the CPI basket after taking away, or trimming, the more extreme price movements in any period. The items trimmed are not prespecified but are based on the distribution of price changes in the period in question. These measures provide a measure of inflation that is not excessively affected by large price changes – either increases or decreases – in individual items.

Whereas exclusion-based measures and trimmed means can be said to preserve CPI weights to a significant extent, some other underlying measures reweight CPI items based on the time-series properties of prices for the individual CPI components. For example, Cutler (2001) proposes reweighting the CPI items according to their persistence. And a number of authors have proposed weighting items by the variance of their price changes or relative price changes. Laflèche (1997) suggests modifying such weights by multiplying them with CPI weights to get a ‘double-weighted’ measure with fixed weights based partly on expenditure patterns (as in the CPI) and partly on the relative noisiness of prices.

Finally, measures based on dynamic factor models are an additional type of underlying measure which employs both reweighting of CPI items and smoothing over time (see e.g., Cristadoro *et al* 2005). In addition, these measures typically also include information from outside the CPI. In some cases, they may include hundreds of additional variables (including real sector variables, monetary variables, and financial variables such as the exchange rate) which might have some economic relationship with inflation.⁴

The above measures can also be distinguished by a number of other criteria. These include the ease of communication, degree of revision, and the extent to which they use information from outside the CPI. In this paper we focus on three measures – headline, ‘core’ and trimmed means – that use only the information in the current CPI release, are not subject to revision (with the exception of revisions to seasonal factors for individual CPI items), and are relatively simple in their calculation. We do not argue that these are necessarily superior to the other measures, and indeed would suggest that central banks should calculate a range of measures. However, the three that we focus on can be easily calculated and communicated to the public, so they may be especially well suited for use both inside and outside of central banks.

2.2 Criteria for Assessing Measures of Underlying Inflation

Our approach in this paper is an empirical one, but to help motivate the work in Section 5, we begin by presenting a simple framework for thinking about candidate measures of underlying inflation. We denote the unobservable true rate of underlying inflation in each month or quarter by π^* . We can view any candidate measure of underlying inflation (π_i) as equal to the actual unobservable underlying rate, plus an idiosyncratic (or noise) term:

⁴ Model-based approaches (Quah and Vahey, 1995) also use variables from outside the CPI, with more economic structure than the data-driven approach of dynamic factor models.

$$\pi_{i,t} = \pi_t^* + e_{i,t} \quad (1)$$

We can think of monthly or quarterly headline inflation, π_{CPI} as being one of the candidate measures of underlying inflation and also represented in this way. The idiosyncratic term for any candidate measure should have an expected value close to zero, and the usefulness of the measure will be inversely related to the variance of the idiosyncratic component.

Based on this framework, there are several criteria we could use to assess different candidate measures of underlying inflation.⁵ First, if the true level of unobservable underlying inflation is expected to be fairly persistent and to evolve only slowly based on variables such as inflation expectations and the output gap, we would expect a good measure of underlying inflation to be relatively smooth. Alternatively, a volatile measure is unlikely to be a good estimate of underlying inflation. Hence, we might consider the size of month-to-month or quarter-to-quarter changes in short-horizon inflation rates as a simple measure of how much noise there is in a particular proxy for underlying inflation.⁶ We implement this in Section 5.1.

Another criterion by which measures of underlying inflation are often judged is how closely they mirror movements in some measure of the longer-run trend in CPI inflation. This test can be done in two dimensions: in terms of long-run average inflation rates for different measures, and the ability to match movements in some measure of trend inflation over time.

Comparisons of long-run average rates of growth are typically done relative to the long-run rate of growth of the headline CPI. Although institutional arrangements differ across countries, most central banks share the explicit or implicit goal of ensuring low and stable inflation, usually as measured by the CPI (or the CPI excluding effects from factors such as mortgage interest charges or large tax changes). Hence for practical policy purposes, we would expect that a candidate measure of underlying inflation should have a similar long-run mean to the target variable. Accordingly, in Section 5.2 we calculate the long-run mean rates of inflation for our various underlying measures, and compare them with average headline inflation.

An assessment of the ability of underlying measures to match movements in trend inflation requires the choice of a reference measure for the longer-run trend. This is somewhat arbitrary, but researchers such as Bryan and Cecchetti (1994) have proposed a two- or three-year-centred moving

5 Another approach would be to assess how different measures of underlying inflation accord with various economic relationships, such as Phillips curve models of inflation. We do not attempt this in the current paper, given the relatively short data samples that we study and because it may be difficult to choose models that are both relatively standardised across these countries but also reflective of the way that inflation is typically modelled in each country. However, it would be surprising if measures that do well against the criteria in this paper did not also perform well in other contexts. Indeed, the Australian experience is that it is less difficult to model trimmed mean inflation than either headline or exclusion measures.

6 Of course, we would not overstate the benefits of smoothness. However, a greater degree of smoothness in exclusion or trimmed mean measures is presumably a positive thing, especially given that – unlike some other measures – they are not subject to any intertemporal smoothing in their construction.

average of monthly headline CPI inflation rates. In Section 5.3, we calculate the root mean squared error (RMSE) of the candidate measures with respect to similar measures of trend inflation.

Finally, an additional criterion that is frequently suggested, including in the quote from Blinder (1997) in our introduction, is that a good measure of underlying inflation should have some predictive power for inflation: we implement this in Section 5.4. However, it is important to consider what types of predictive power one should expect to see. In particular, we do not expect that there exist measures of underlying inflation that lead headline inflation in any deep sense or that there are any significant lead or lag relationships between headline inflation and various candidate measures of underlying inflation.⁷ That is, like most other researchers in this area, we view the search for good measures of underlying inflation as a search for measures which have a high signal-to-noise ratio in terms of current underlying inflation, rather than for leading indicators of inflation: analysts looking for the latter are likely to find it more fruitful to estimate forecasting models based on the variables that drive inflation.

Accordingly, we think it is more plausible to think that any predictability is more a statistical, and fairly short-term, phenomenon. We view true unobservable underlying inflation as being a relatively persistent variable, typically changing only slowly as the fundamental drivers of inflation evolve. So unobservable underlying inflation in period t should be a very good predictor of itself in period $t+1$. And, consistent with Equation (1), headline inflation in period $t+1$ will be given by unobservable underlying inflation in period $t+1$ plus an error term. So a good estimate of period t underlying inflation should be a reasonably good predictor of headline inflation in period $t+1$. However, this will be somewhat dependent on the variance of e_{CPI} : if headline inflation is extremely noisy then even a very good measure of underlying inflation will have trouble forecasting it. In summary, we do not expect that measures of underlying inflation should necessarily have any predictive power for headline inflation at any deep level: rather because true underlying inflation is fairly persistent, a good estimate of underlying inflation will also be persistent and will have some predictive power for other measures of inflation in the near-term.

Similar insights hold for Granger causality tests involving the CPI and various candidates for underlying inflation. Suppose we have two noisy measures of underlying inflation that each contain some independent information about true unobservable underlying inflation. In general, if underlying inflation is quite persistent and we have a sufficiently long data sample, we would expect *both* to Granger cause each other.⁸ In reality, the sample period may not be long enough to observe

7 It seems unlikely that there exist a significant number of CPI expenditure items which are true leading indicators of the overall trend in inflation. One way to shed light on this is to calculate correlations between inflation in individual expenditure items and some measure of underlying inflation. For Australia, there are some items for which the highest correlation appears to be leading or lagging, rather than contemporaneous. However, one would expect to find some sample variation in such correlations. When we divide the sample and ask if the leads or lags are stable, we find that many of the apparent lead or lag relationships from the first half of the sample disappear (or reverse) in the second half.

8 Our argument is clearly at odds with a significant amount of earlier work (see, for example, Marques, Neves and Sarmento 2003) which maintains that Granger causality should go only from underlying measures to headline inflation. Our point of departure is that we do not assume that there is any candidate measure that is so close to true

Granger causality in both directions. If so, it is likely that the presence or absence of Granger causality between the two measures of underlying inflation will depend on how close each is to true underlying inflation (that is, on the relative variance of the e_i terms). If one measure tends to be a significantly better proxy for underlying inflation than the other, then we would expect Granger causality to be more likely to go from the former to the latter than vice versa. But if one measure is a sufficiently poor proxy for underlying inflation then we might not observe Granger causality in either direction. In particular, if CPI inflation is a particularly noisy measure of underlying inflation, in small samples we might find that it is not Granger caused by even very good measures of underlying inflation.

3. Background on Trimmed Mean Measures of Underlying Inflation

3.1 Definition of Trimmed Mean Measures

Trimmed mean measures of underlying inflation were first proposed by Bryan and Pike (1991) and Bryan and Cecchetti (1994). The rationale for trimmed means is that there are sometimes very large changes in prices for particular items that have a significant effect on a conventional average of all price changes, but which are quite unrepresentative of the rest of the price changes that are observed. Whereas exclusion measures remove some pre-specified items in every period regardless of whether or not their price changes are extreme, trimmed mean measures exclude – or more correctly down-weight⁹ – the impact of items based on whether or not their price changes appear to be outliers in the period in question. These measures represent an attempt to estimate the central part of the distribution of price changes, and to provide a measure of inflation that is not excessively affected by large price changes – either increases or decreases – in individual items.

Formally, the trimmed mean rate of inflation is defined as the average rate of inflation after trimming away a certain percentage of the distribution of price changes at both ends of that distribution. In all the work that follows, we focus on symmetric trimmed means. The first step in calculation is to order the distribution of price changes of CPI components for any period from lowest to highest. The n per cent trimmed mean rate of inflation for that period is calculated by excluding n per cent (taking into account the weights of the components in the CPI) from the top and bottom of the ordered distribution of price changes and then taking the weighted average of

unobservable underlying inflation as to have essentially no error term, and we allow for the possibility that the noise in the headline measure may be so large that its error term dominates any regression involving headline inflation.

⁹ We say ‘down-weight’, because even when an expenditure item is trimmed and ‘excluded’ it still affects the trimmed mean. In particular, the fact that an item has been trimmed in any period means that some other item which experienced a relatively high or low price change will not be trimmed. Trimmed mean measures are accordingly sometimes referred to as limited influence measures.

prices in the remaining central $(100-2n)$ per cent of the distribution. A whole suite of such measures can be provided by trimming an increasing proportion of the distribution from either end. At one extreme, the 0 per cent trimmed mean corresponds to the conventional headline rate of inflation. At the other extreme is the weighted median or the 50 per cent trimmed mean which is the price change at the 50th percentile (by weight) of the ordered distribution of price changes.

The statistical rationale for trimmed mean measures is usually based on the evidence that distributions of price changes for individual CPI components tend to show excess kurtosis, or fat tails. Fat-tailed distributions are more likely to yield extreme observations from one of the tails of the distribution, and it is possible that these observations are not balanced in any particular period by an offsetting movement in the price of another CPI component. Under such circumstances, trimmed means are likely to be better estimators of the central part of the distribution than are sample means. Roger (2000) provides a further discussion of the case for using trimmed means rather than simple means.¹⁰

3.2 Technical Issues in Calculating Trimmed Mean Measures

There are a number of technical, but nonetheless important, issues that present themselves in the calculation of trimmed mean measures of underlying inflation. These relate to particular characteristics of the actual distribution of price changes. In general, they are concerned with improving the performance of trimmed mean measures by making the underlying distribution of price changes smoother and more continuous, or correcting for possible average bias (relative to overall CPI inflation).

3.2.1 *Symmetric or asymmetric trims*

In this paper we limit our analysis to symmetric trimmed means, which remove the same proportion of the distribution at the high and low ends. However, some other applications have suggested the use of asymmetric trims. For example, Roger (1997) suggested the use of the 57th percentile for the New Zealand CPI, while Dolmas (2005) suggests trimming 25 per cent off the top and 19 per cent off the bottom of the distribution for the US personal consumption expenditures (PCE) chain price index.

Our use of symmetric trims in this paper is based both on *a priori* preferences and concerns about the communication issues surrounding underlying measures. In particular, intuitively it seems more desirable that outliers at both ends should be treated similarly. Perhaps more importantly, it may be easier to explain symmetric trims to the public, without raising concerns that the central bank is trying to manipulate measures of the inflation rate.

¹⁰ More complex versions of trimmed means are also feasible (see Roger 2000 and Aucremanne 2000), but we limit our analysis to the more standard applications.

The rationale for asymmetric trimming has typically been based on the finding that distributions of price changes are skewed so that a symmetric trimmed mean may be biased with respect to CPI (mean) inflation. However, the potential for bias in trimmed means will depend not just on the shape of the distribution, but also on the relationship between the relative volatility of prices and their average rate of growth. For example, if prices for services tend to be less volatile but to rise faster than those for goods then they will tend to be trimmed less often, which would result in a tendency for trimmed means to be biased upwards.

In the end, the decision as to whether asymmetric trimming should be employed requires a judgment that any apparent bias in some historical period is systematic and likely to persist into subsequent periods. Our own analysis in Section 5.2 finds little evidence of bias for the four economies we study, suggesting that asymmetric trims may typically not be necessary. In addition, it is worth noting that the performance of trimmed mean measures against the criteria in Sections 5.3 and 5.4 would be penalised to the extent that the measures were biased, so the strong performance of these measures suggests that bias is not a major issue.

3.2.2 *Choice of frequency of price changes*

Annual trimmed mean measures of inflation are often calculated by using the distribution of monthly or quarterly price changes to calculate measures of monthly or quarterly underlying inflation, which are then cumulated to get an underlying price index which can be used to generate the annual rate of underlying inflation. However, an alternative is to calculate the annual rate of underlying inflation by taking the distribution of year-ended price changes for all the CPI sub-components and calculating a trimmed mean of these annual rates.¹¹

In general, it might be expected that these should yield very similar estimates of annual underlying rates. However, consider the extreme case where the CPI could be calculated on a daily basis. In this case, we would expect that most individual prices in the economy would remain unchanged, so a trimmed mean of price changes across expenditure classes might also be zero, and cumulating these daily rates over a full year would suggest a zero annual rate of underlying inflation. By contrast, a trimmed mean based on the distribution of annual changes would indicate a very different rate, no doubt much closer to the true underlying rate. Of course, this example is extreme because CPIs are not calculated on a daily basis. So any bias from using monthly or quarterly price changes is likely to be much smaller. Nevertheless, given that there will typically be some items with prices changed (or sampled) irregularly, there may be a case for using price change distributions over longer horizons, and the two methodologies may sometimes give slightly different pictures of underlying inflation pressures over the prior year.

¹¹ For example, the measures calculated by the Reserve Bank of Australia, Bank of Japan, and the Federal Reserve Banks of Cleveland and Dallas are calculated at the quarterly or monthly level and then cumulated to get annual rates. In contrast, the Bank of Canada, Swiss National Bank, and Reserve Bank of New Zealand calculate trimmed means from the distribution of year-ended price changes.

In work not presented here, we have attempted to assess the relative merits of annual rates calculated via the two different approaches. The conclusions must be somewhat tentative given the relatively small number of independent annual observations in our sample period. However, the results do suggest that when the year-ended distribution of price changes is used, it may be best to avoid the use of weighted medians or other large trims.¹² Overall, weighing up results against some different criteria, there do not seem to be any strong empirical reasons to prefer one approach to calculating annual trimmed mean inflation over the other. However, one clear benefit to using trimmed means calculated from higher frequency price changes is that they also provide measures of underlying inflation for the most recent months or quarters, allowing policy-makers or analysts to make assessments about the current trend in inflation as opposed to relying on backward-looking annual rates of inflation.¹³

3.2.3 *Seasonality*

A related issue is the case of irregular price changes due to seasonal price adjustment, including for some administered prices. In such cases, a typical pattern will be for a price to be unchanged for most of the year, then to jump in one particular month or quarter by a relatively large amount. Many countries do not seasonally adjust the raw price data in their CPIs. In such cases the infrequent price adjustment may have a noticeable effect on the headline CPI, and a more significant effect on the calculation of trimmed mean measures. In particular, for a range of trims we would expect items with seasonal prices to be trimmed in the period when their prices are increased but not trimmed in the periods when their prices are unchanged. If there are a sufficient number of items with such characteristics, trimmed mean estimators based on non-seasonally adjusted data can be downwardly biased, as has been shown by Roberts (2005) using Australian data. However, a simple response which should eliminate this potential downward bias is to use seasonally adjusted prices for individual expenditure items which appear to have seasonal prices. This is the approach we follow below.

3.2.4 *Treatment of housing or other large expenditure items*

Another important issue in calculating trimmed mean measures for some countries is that such measures, in particular the weighted median, can be significantly affected by the presence of items that have a large weight in the CPI basket. This is most relevant in the case of housing costs. In the United States the component for ‘owners’ equivalent rent of primary residence’ accounts for around 24 per cent of the total CPI, compared with an average weight of 1.1 per cent for the 69 other expenditure items in our dataset. One consequence is that this one item is the median item fully 58 per cent of the time in our sample over 1993-2007. This component is also relatively large in Japan,

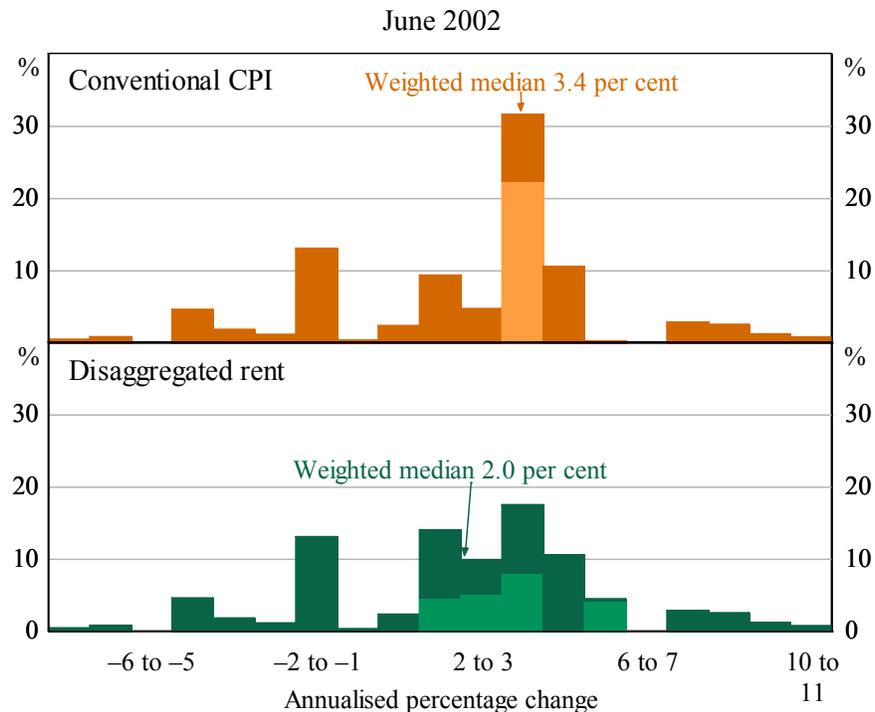
12 One problem with the weighted median based on the distribution of annual price changes is that – even at the centre of the ordered distribution – there can sometimes be noticeable gaps in the annual inflation rates of adjacent items, so the annual weighted median can be quite volatile: these effects are obviously much smaller for the cumulated monthly weighted median rates.

13 An additional drawback to the use of the year ended distribution is that it presents challenges at times of CPI reweightings (especially involving introduction or removal of expenditure items).

at about 14 per cent of the CPI. The largest items in the CPI are significantly smaller in the other two economies, accounting for less than 8 per cent of the total CPI in Australia and the euro area.¹⁴

We can use US CPI data to illustrate the consequences for trimmed means of having an item with a very large expenditure weight. The top panel of Figure 1 shows the distribution of annualised monthly price changes and the weighted median for June 2002, a month when this issue was particularly important. In the bottom panel, we show the distribution of price changes when more disaggregated data for implicit home-owners' rent are used. This component cannot be disaggregated further into finer expenditure items, so instead we break it into regional subcomponents. Regional price changes for this item are published by the Bureau of Labor Statistics (BLS), and we estimate appropriate weights by using dwelling stock data from the US Census Bureau. This gives weights in the total CPI of around 4 per cent for implicit rent in the North-east, 9 per cent for the South, 5 per cent for the Mid-west and 5 per cent for the West. The adjustment results in a distribution that is somewhat smoother and much less peaked.

Figure 1: US Distribution of Monthly Price Changes



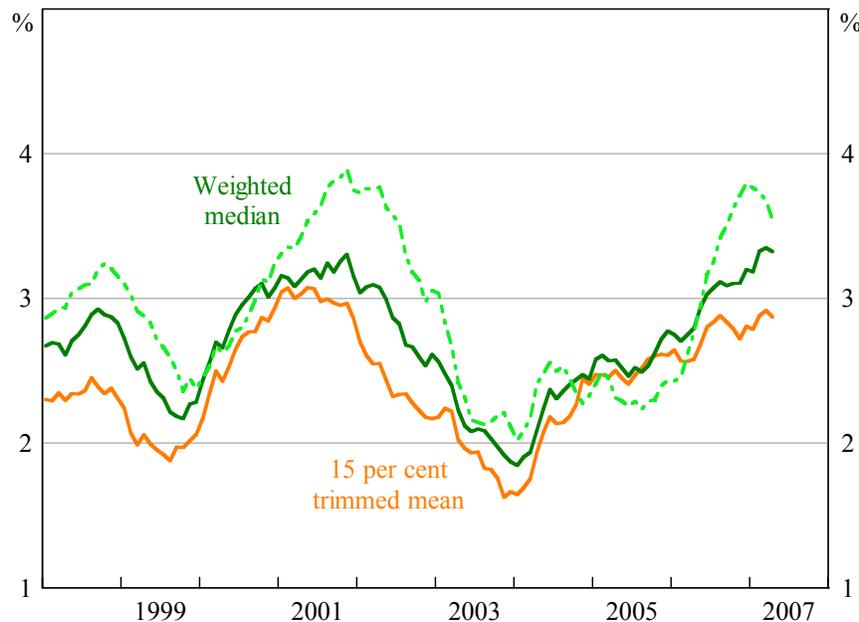
Notes: Bars do not sum to 100 because a small proportion of distribution at both extremes is not shown. The lighter shading shows the implicit rent components.

¹⁴ This reflects different methodological approaches to the treatment of housing costs. For example, Australia uses the 'net acquisitions approach' by including only the construction cost (but not land component) of additions to the owner-occupier dwelling stock, while the United States and Japan follow the 'consumption approach' by including an implicit cost (the estimated rental value) of housing for all owner occupiers. All three (and the euro area) include an item capturing the cost of renting for that part of the population that rents.

The conventionally calculated rate of mean CPI inflation is obviously unaffected by this change. However, the shift in the shape of the distribution results in a large shift in the weighted median inflation rate, from an annualised rate of 3.4 per cent using the aggregated data to 2.0 per cent using the less aggregated data. In contrast, the effect on smaller trims is much smaller: for example the 15 per cent trimmed mean rate shifts only from 1.8 per cent to 1.7 per cent when disaggregated data are used.

As is shown in Figure 2, this issue is particularly important during 2001 and 2002, and again in late 2006 and early 2007 when our adjustment reduces the cumulated annual inflation rate for the weighted median by as much as 0.6-0.7 percentage points, moving it closer to other measures of inflation at those times. Furthermore, the analysis in Section 5 suggests that the disaggregation of the rent component results in an improvement in the statistical performance of trimmed mean measures. This accords with the notion that these measures are likely to yield more accurate indications of the central tendency of the distribution when reasonably disaggregated price data are available and the distribution of the underlying changes is smoother.^{15, 16} Unfortunately, data constraints prevent us conducting a similar disaggregation for Japan.

Figure 2: US – Consumer Prices
Year-ended percentage change



Note: The dashed line is calculated without disaggregating the implicit rent component, while the solid lines disaggregate it into 4 regional sub-components.

¹⁵ One consequence of breaking up implicit home-owners' rent into four components is that one of these is the median item 'only' 37 per cent of the time, versus 58 per cent when it is a single item.

¹⁶ See also Bryan, Cecchetti and Wiggins (1997) for additional evidence that the gains from trimmed mean measures increase as more disaggregated data are used.

4. Data

Our analysis uses disaggregated price data for the CPI for Australia, the United States and Japan, and the harmonised index of consumer prices (HICP, hereafter referred to as the CPI) for the euro area. Our focus on the US CPI rather than on the PCE chain price index is mostly for consistency with the other economies in our study.

The disaggregated data used in our empirical work are monthly for the United States, euro area and Japan, and quarterly for Australia, as published by the statistical authorities in each economy.¹⁷ For Australia, we choose to focus on the low-inflation period starting from the March quarter of 1993, and for consistency we present results for the United States from January 1993 onwards.¹⁸ For the euro area we were constrained by data availability and present results for the period from January 1996. For Japan, we present results from April 1997, so that the effects of the 1997 VAT increase are excluded. The sample period for our assessment of the performance of the underlying measures ends in March 2007 for the United States and Australia, and February 2007 for the euro area and Japan.

Our analysis uses expenditure items at a relatively fine level of disaggregation.¹⁹ In the case of Australia, Japan and the euro area, the CPI and its components are published in non-seasonally adjusted terms only. In the case of the United States, the calculation of the overall CPI uses seasonally adjusted data for some components and unadjusted data for the rest (although a comprehensive list of which components are seasonally adjusted over time was not readily available). Given the importance of seasonal adjustment in calculating trimmed mean measures, we have seasonally adjusted the data for all four economies using the Census X12 method in EViews, adjusting all components where a regular seasonal pattern could be identified.

Based on these data, we construct seasonally adjusted CPIs for the four economies. We also calculate seasonally adjusted exclusion measures of core inflation.²⁰ And we calculate a range of

17 Namely the Australian Bureau of Statistics for Australia, Bureau of Labor Statistics for the US, Eurostat for the euro area and Statistics Bureau, Ministry of Internal Affairs and Communications for Japan. For Australia, we exclude interest charges up to 1998, and we make one specific adjustment to remove the effect of the tax changes of 1999–2000 from individual expenditure items. In the case of Japan, CPI weights prior to 2000 were not available and so are approximated using data from the Family Income and Expenditure Survey (on which the actual CPI weights are based). Otherwise, the data are as used by the statistical authorities in the construction of CPIs. Nevertheless, there were some minor discrepancies between our replication of CPI inflation and the CPI data published by statistical authorities, presumably reflecting rounding in input data. For consistency, the trimmed mean, exclusion (except in the case of Japan – see Footnote 20) and headline inflation measures used in this paper are all calculated from the same set of disaggregated prices and weights data.

18 However, there is every reason to expect that the results of the paper would apply also to earlier time periods when inflation was higher and more variable: indeed, we have confirmed this for data back to 1982 for Australia and 1987 for the United States.

19 The number of CPI components varies slightly over time as the composition of the CPIs changes, but in the most recent period is 90 for Australia, 70 for the United States, 50 for Japan and 93 for the euro area.

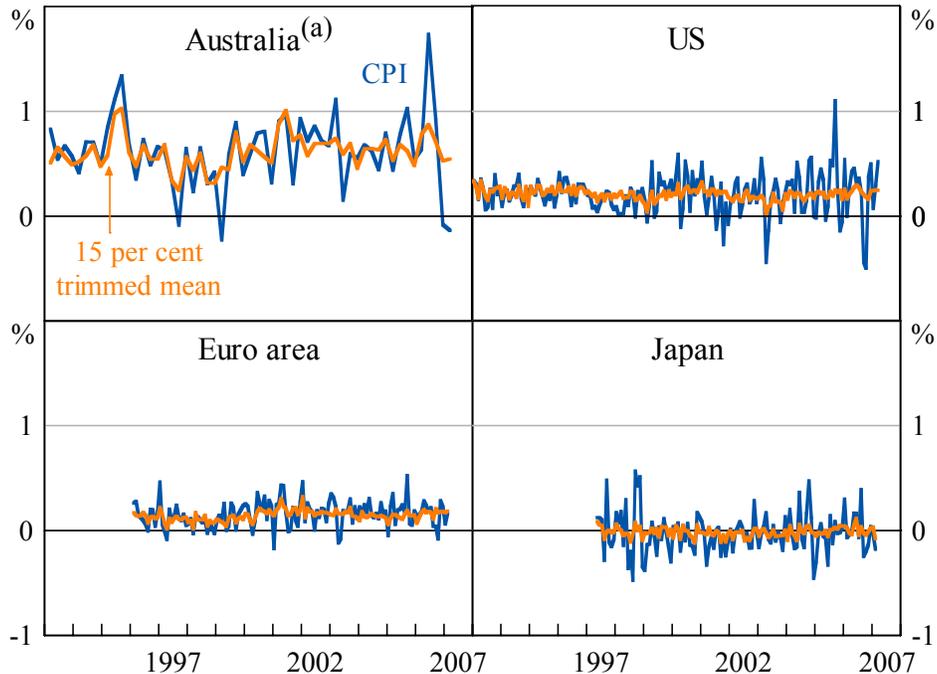
20 Items excluded are: for Australia, petrol and fruit & vegetables; for the US and Japan, food and energy; and for the euro area, food, energy, alcohol and tobacco. In the case of Australia, the US and euro area, we construct exclusion

trimmed mean measures using symmetric trims gradually trimming from 1 to 50 per cent of the price distribution from both ends.

We provide a first look at the trimmed mean measures of underlying inflation in Figure 3, along with headline CPI inflation (that is, the 0 per cent trimmed mean constructed from our seasonally adjusted data) for each economy. There are 50 different trims that we could illustrate, and in an earlier draft we used the middle of all feasible trims – the 25 per cent trim – for illustrative purposes. However, in Section 5.5 we suggest that smaller trims may sometimes be preferred, so we instead show the 15 per cent trim. It is immediately obvious that monthly or quarterly trimmed mean inflation is much less noisy than CPI inflation, suggesting that trimmed means may be promising measures of underlying inflation. We examine this proposition in more detail in the next section.

Figure 3: Consumer Prices

Monthly percentage change



Note: (a) Quarterly percentage change

measures using seasonally adjusted components where appropriate. For Japan, we do not have sufficiently disaggregated data to do this so we seasonally adjust the published exclusion measure.

5. Empirical Analysis

A number of criteria have been proposed as desirable for measures of underlying inflation. Some of these relate to the production and dissemination of measures, for example that they should be timely and verifiable (Roger 1997), understandable by the general public, computable in real time, and not subject to revision (Wynne 1999). Both trimmed mean measures and exclusion measures satisfy these criteria. Other criteria are more empirical: for example, Roger suggests that measures should be ‘robust and unbiased’ while Wynne suggests they should be ‘forward looking in some sense’. In this section we focus on some empirical criteria, as discussed in Section 2.2. For the most part, these can be viewed as various ways of searching for underlying measures that meet Blinder’s (1997) goal of better identifying the signal from the noise.²¹

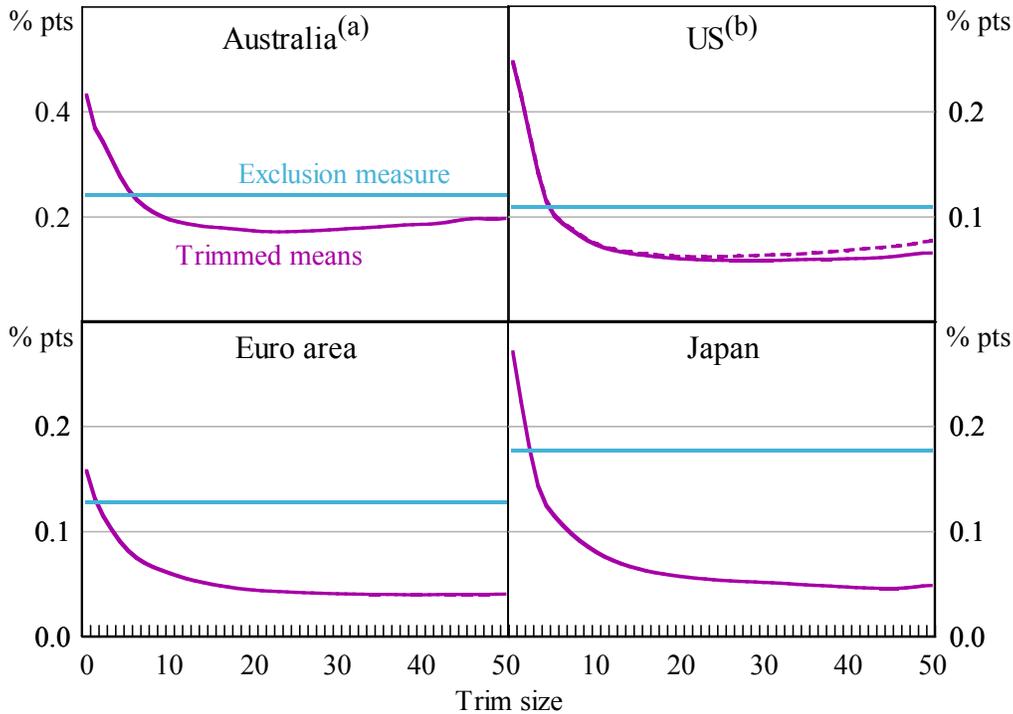
5.1 Smoothness of Underlying Inflation Measures

One way to assess smoothness is to examine which measures are more persistent. The alternative which we calculate is the standard deviations of period-to-period *changes* in monthly or quarterly inflation rates. Figure 4 presents these data for the headline CPI (represented by the 0 per cent trim), the exclusion measure (shown as a horizontal line), and for the various trimmed means for trims from 1 to 50 per cent. The results are fairly consistent across economies. They suggest that as we increasingly trim price changes from both ends of the distribution we obtain a much smoother measure of high-frequency inflation. Most of the gains in smoothness have been obtained after about 10 per cent of the distribution has been trimmed from either end. The lowest point for each line corresponds to the smoothest measure, which represents a reduction in volatility of around 60–85 per cent, with larger reductions for the economies with monthly data. The trims that would appear ‘optimal’ (on this criterion) differ but are all in the range of about 25–45 per cent. However, for all economies there is a fairly wide range of trims that offer broadly similar improvements in smoothness. The exclusion measures are also smoother than the headline CPI measures, although in all cases they offer much less reduction in noise than most of the trimmed means.

²¹ It is easily shown, based on the framework in Equation (1) and plausible assumptions about some variances and covariances, that the criteria in Sections 5.1, 5.3 and 5.4 are all driven by the magnitude of the variance of the idiosyncratic term in Equation (1).

Figure 4: Smoothness of Underlying Inflation

Standard deviation of change in monthly inflation



Notes: (a) Quarterly inflation.

(b) Dashed line is calculated without disaggregating the implicit rent component.

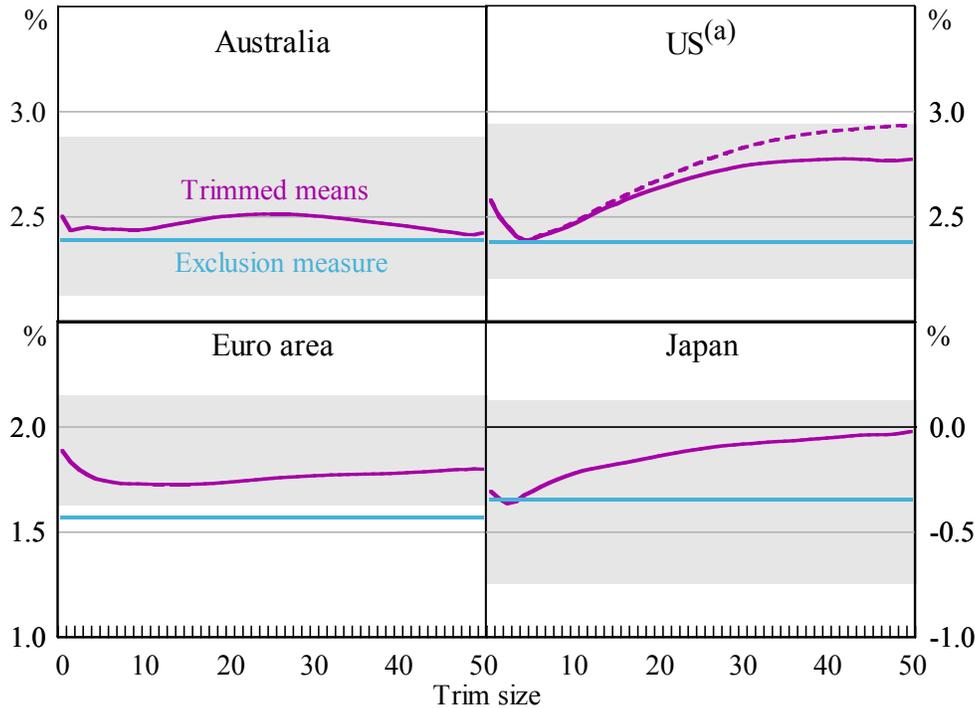
5.2 Bias with Respect to Headline Inflation

Figure 5 shows average annualised rates of underlying inflation for the four economies. To provide some perspective on the observed differences in inflation rates, we also include a simple estimate of the 95 per cent confidence interval around the mean growth for the headline CPI.²² The results for all four economies indicate that the trimmed mean measures typically have a fairly similar mean inflation rate to the headline CPI (again, this is the 0 per cent trim measure). The largest difference is for the standard weighted median for the United States which has a mean annualised growth rate which is 0.35 percentage points higher than the rate for the headline CPI. However, when we disaggregate the implicit rent component the difference falls to 0.20 percentage points. There is some tendency for the absolute differences in mean growth rates relative to the CPI tend to be lower for smaller trims than for the weighted median. In addition, the average growth rates for the trimmed means are sometimes above (the United States and Japan) and sometimes below (the euro area) the average growth of the headline CPI, whereas the average growth rate of the exclusion measure is in

²² The confidence interval is the standard one under the assumption that all monthly/quarterly inflation rates are drawn from a normal distribution, namely $\mu \pm (t\text{-stat} \cdot \sigma) / \sqrt{N}$, where the standard deviation is calculated using high-frequency inflation rates then annualised. We would not suggest that monthly inflation rates are literally *i.i.d.* but use this measure to give readers some common simple metric of whether deviations in growth rates are relatively large or small.

all cases below the growth of the headline CPI. The latter is not surprising given the increase in real energy prices over the sample period.

Figure 5: Mean Underlying Inflation Rates
Annualised



Notes: Shaded areas provide an estimate of 95% confidence interval around mean CPI (0 per cent trim) inflation.
(a) Dashed line is calculated without disaggregating the implicit rent component.

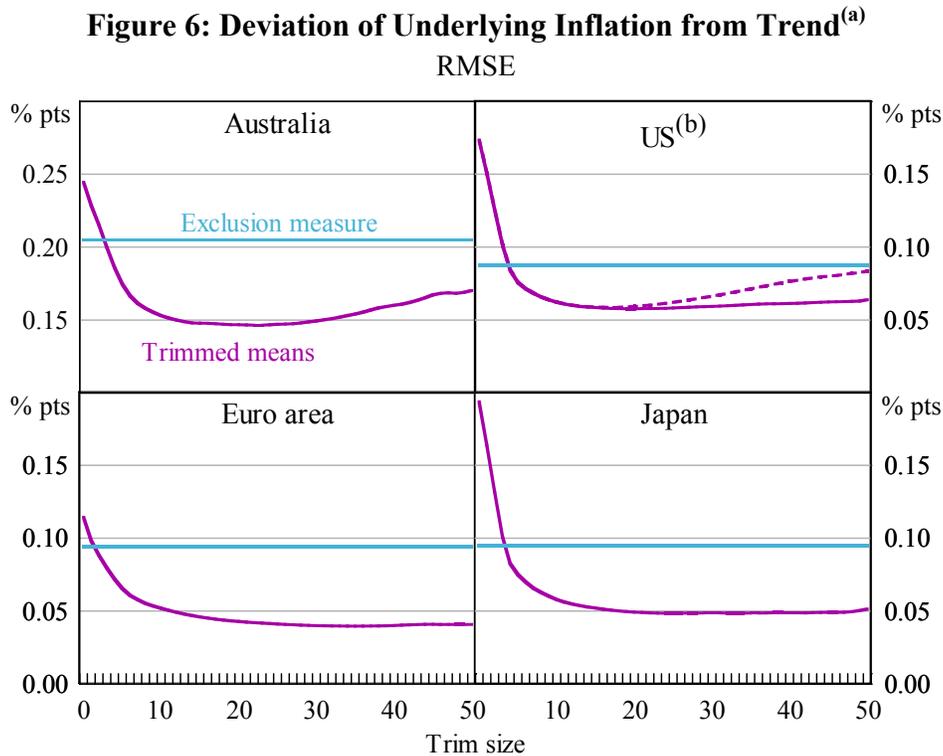
There would seem to be a number of ways that one could do formal statistical tests for bias with respect to the headline CPI.²³ Such tests provide mixed evidence regarding the hypothesis of unbiasedness for the trimmed mean measures (and for the exclusion measures). However, formal statistical tests for unbiasedness must be somewhat tentative over relatively short periods of 10–15 years, especially at times of large movements in fuel prices at the end of the sample period. In cases such as these, reasonable people might differ as to how to interpret a difference in average growth rates between an underlying measure and the headline CPI: perhaps additional data will suggest that

²³ There is relatively little discussion in earlier research as to how to test for bias. At least three possibilities suggest themselves. First, one could calculate simple confidence intervals as in Footnote 22 and shown in Figure 5. Second, there is the standard approach of regressing underlying inflation on a constant and the headline rate (or vice versa) and testing the hypothesis that the constant is zero and the slope coefficient is unity. Third, one could test for cointegration and a unit elasticity between underlying and headline price indices. The first may not be entirely appropriate as the time series properties of headline inflation may render the assumption of normality invalid. The second seems inappropriate because it is less a test of similarity of long-run growth rates and more a test of period-by-period correlations. The third seems most appropriate in principle, but may suffer from the usual problems of running cointegration tests over relatively short samples.

any differences in growth rates reflected anomalous movements in the headline CPI rather than in the underlying measures. Accordingly, given that the differences in average rates of growth appear relatively small (sometimes positive and sometimes negative), we would characterise the trimmed means for these economies as appearing to be unbiased measures.²⁴

5.3 Closeness to Trend Inflation

In this section, we calculate the root mean squared error (RMSE) of the candidate measures versus a measure of trend inflation. The trend is defined as the centred 25-month moving average of CPI inflation for the United States, Japan and the euro area, and the 9-quarter moving average for Australia. The results are shown in Figure 6 and are reasonably consistent with the earlier analysis of smoothness. In all four cases, even small trims result in a substantial reduction in the RMSE



Notes: (a) 25-month- or 9-quarter-centred moving average of CPI inflation.
(b) Dashed line is calculated without disaggregating the implicit rent component.

²⁴ However, in analysis not presented in this paper we have also calculated trimmed mean measures for the United Kingdom and found evidence of substantial bias as we move towards the weighted median. As Catte and Sløk (2005) have noted, this may be because services prices tend to be less volatile than goods prices and are trimmed less frequently, and because the gap between goods and services inflation has been somewhat larger in the UK than elsewhere. Of course, a biased measure can still provide useful information; if the bias is relatively constant, the measure may still be a useful indicator of *changes* in underlying inflation. Indeed, trimmed mean measures for the UK appear to generally perform reasonably well in other tests.

relative to trend inflation. The maximum reductions in RMSE range from 65–75 per cent for the economies with monthly data and about 40 per cent for Australia’s quarterly data. For each economy there is a wide range of trims offering substantial improvements relative to both the headline CPI and the exclusion-based measure. The ‘optimal’ trims (according to this measure) are around 15–35 per cent. Again, the US trimmed means calculated using disaggregated data for implicit rent outperform the aggregated measure, especially as we move towards the weighted median.

5.4 Ability to Predict Near-term Inflation

A number of researchers (for example, Cogley 2002) have assessed measures of underlying inflation by how well they forecast inflation over periods as long as three years (or even five years in the case of Bryan and Cecchetti 1994). The rationale is often that this is the horizon of ‘greatest relevance to monetary policy’. However, we do not consider predictability over such long horizons to be a sensible test.²⁵ Consider the case where a central bank had a good measure of underlying inflation to help inform its policy process. Hence when the measure indicated a build-up in underlying inflationary pressures the central bank would tend to respond by tightening monetary policy, placing downward pressure on inflation over the next few years. So we would not expect to see any predictability from underlying inflation to headline inflation at relatively long horizons. More generally, over periods of two or three years, developments in inflation will be determined by movements in the fundamental economic determinants of inflation, including the influence of monetary policy.

Instead, as discussed in Section 2.2, we think it is more plausible that any predictability is likely to be a fairly short-term phenomenon. Accordingly, we instead examine whether CPI inflation over three or six months (or one or two quarters) can be predicted by various measures of underlying inflation over the corresponding preceding period. The first exercise we conduct is to assess the performance of a rule that simply uses various measures of lagged inflation as the prediction for future headline inflation, so that we can calculate the RMSE for each underlying measure i as given by:

$$RMSE_i = \sqrt{\frac{\sum_{t=1}^N (\pi_{i,t}^k - \pi_{CPI,t+k}^k)^2}{N}} \quad (2)$$

where $\pi_{i,t}^k$ denotes inflation in measure i over the k periods to time t .

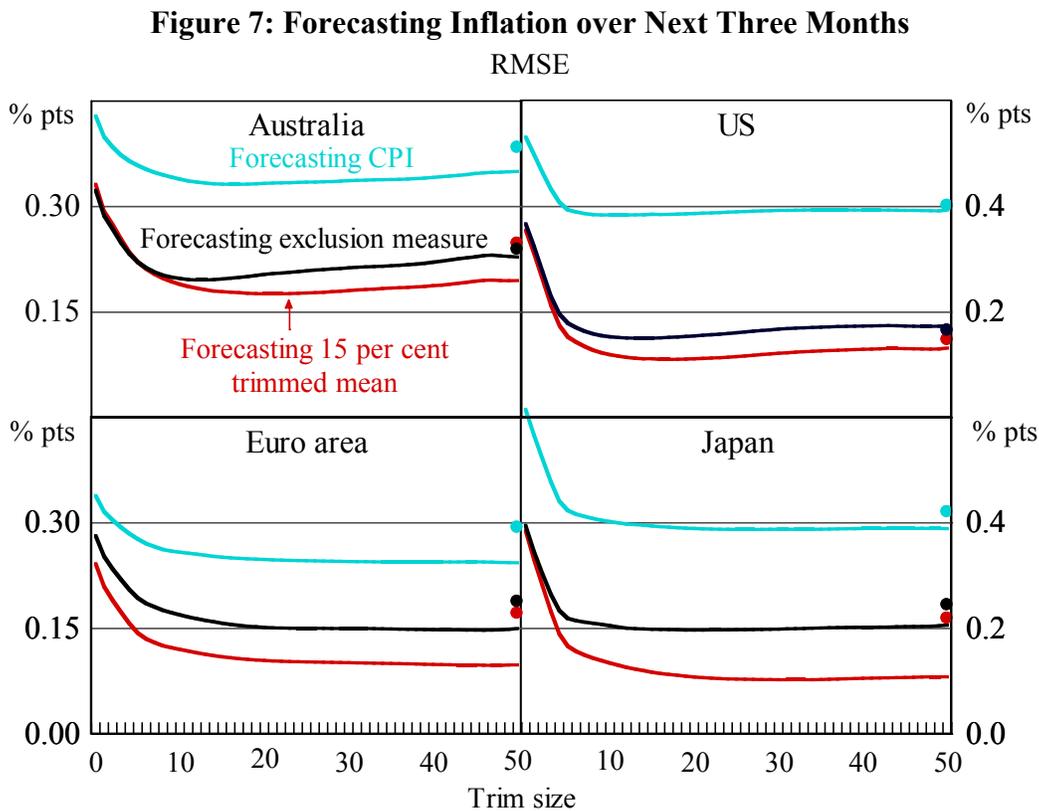
Some researchers, starting with Bryan and Cecchetti (1994), have conducted similar tests using a regression framework – that is, allowing for the possibility of a coefficient on the lagged underlying inflation rate that is different to one (and a non-zero constant term). However, other researchers (for example, Rich and Steindel 2005) have found that the results of such exercises can differ significantly between in- and out-of-sample tests, with some instability in coefficient estimates. Our

²⁵ Some of our reservations appear to be also shared by Wynne (1999).

estimates may be more robust in the sense that there is no difference between in- and out-of-sample tests. Equivalently, our analysis corresponds to feasible forecasting rules that might be quite close to those followed by real-world analysts, who might well forecast that near-term CPI outcomes will be similar to recent underlying inflation outcomes.²⁶

Of course, if headline inflation is very noisy it raises the question of whether the ability to forecast it is a particularly useful test of underlying measures. Instead, if there exists another measure that seems to better reflect the trend in inflation, a better test might be whether candidate measures of underlying inflation can predict that measure. Accordingly, we also assess how well our various measures of underlying inflation can predict either the exclusion measure of underlying inflation or the 15 per cent trimmed mean.

The results for the three-month (or one-quarter) horizon are shown in Figure 7 and indicate that the use of trimmed means for forecasting headline inflation result in reductions in RMSE of 25–40 per cent relative to forecasting with headline inflation itself. As was also the case for the other criteria,



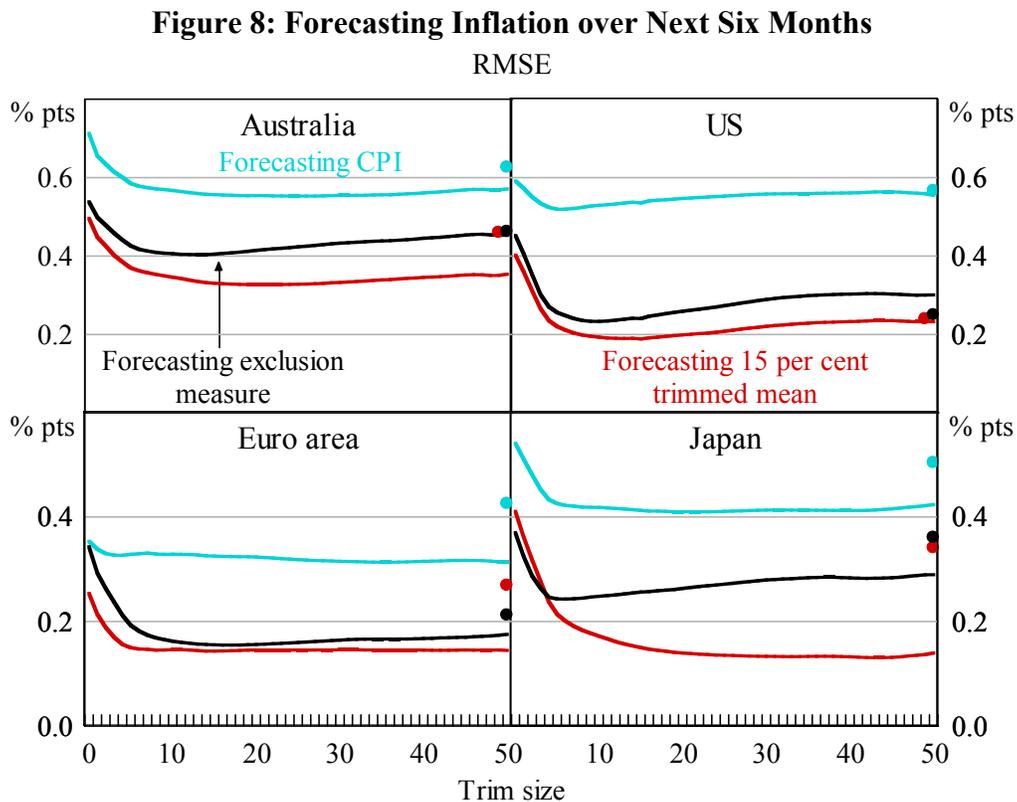
Note: Dots show forecast performance of exclusion measures

²⁶ Others who have used this approach include Cutler (2001) and Vega and Wynne (2003).

most of the gains accrue after only a relatively small amount of trimming. ‘Optimal’ trims according to this criterion vary from around 10 per cent for the United States to 50 per cent for the euro area. As with other criteria, for each economy there is a wide range of trimmed means that perform substantially better than either the headline or exclusion measures.²⁷

The results show a substantially lower RMSE for forecasting the two measures of underlying inflation. Interestingly, the trimmed mean measures tend to outperform the exclusion measure in forecasting the exclusion measure. The smallest forecast error occurs when the trimmed mean measures are used to forecast the 15 per cent trimmed mean inflation rate. The RMSEs from this exercise are up to 60–85 per cent lower than the RMSEs when the CPI is used to forecast itself.

The results for the six-month (or two-quarter) horizon are shown in Figure 8 and indicate smaller gains from using trimmed means to forecast the CPI. Nevertheless, the use of trims results in reductions of up to about 10–25 per cent in RMSE terms. However, the overall modest improvements in RMSEs suggest that central banks wanting to improve forecasts for the CPI at



Note: Dots show forecast performance of exclusion measures

²⁷ Of course, the similarity of these results with those in Section 5.3 is not surprising. As Vega and Wynne (2003) have noted, a measure that does well in tracking a centred moving-average trend should also do well in explaining a forward-looking trend.

horizons of six months or more should rely more on economic models and other information than simply on the current level of underlying inflation.

The second assessment of predictability uses Granger causality tests between headline CPI inflation, exclusion measures, and 15 per cent trimmed mean inflation. For Australia, we test for Granger causality in quarterly inflation. For the other three economies, we look through the noise in monthly data by using inflation over three-month periods (period t to $t+3$) and we use lagged three-month inflation ($t-3$ to t) as the explanatory variable.²⁸ The results are shown in Table 2.²⁹

Table 2: Testing for Granger Causality between Inflation Measures

Direction of Granger causality test	Australia	Euro area	Japan	US
Headline to trimmed mean	No	No	No	No
Headline to exclusion	No	No	No	No
Trimmed mean to headline	Yes	Yes	No	No
Trimmed mean to exclusion	Yes	Yes	Yes	Yes
Exclusion to headline	No	No	No	No
Exclusion to trimmed mean	No	No	No	No

The tests show that there is usually no evidence for Granger causality, with the exception that trimmed mean inflation Granger causes headline inflation for two of the four economies and Granger causes the exclusion measure in all four economies. Our interpretation of the results, based on the framework discussed in Section 2.2, is as follows. In an ideal world, since these variables can all be viewed as proxies for true unobservable underlying inflation, we might expect to see bi-directional Granger causality in most cases. However, in short samples and with significant amounts of noise, especially in the headline CPI, it would not be surprising if we frequently fail to find Granger causality. Cases where we do find Granger causality are likely to be where a relatively good estimate of underlying inflation has information that helps predict a more noisy (but not excessively noisy) measure. This is consistent with the results that show that trimmed mean inflation often Granger causes other measures of inflation, but that there is no Granger causality from either of those other two measures.

It is noteworthy that the goodness of fit in the equations for forecasting the CPI is typically quite low. One interpretation might be that underlying measures – both trimmed mean and exclusion

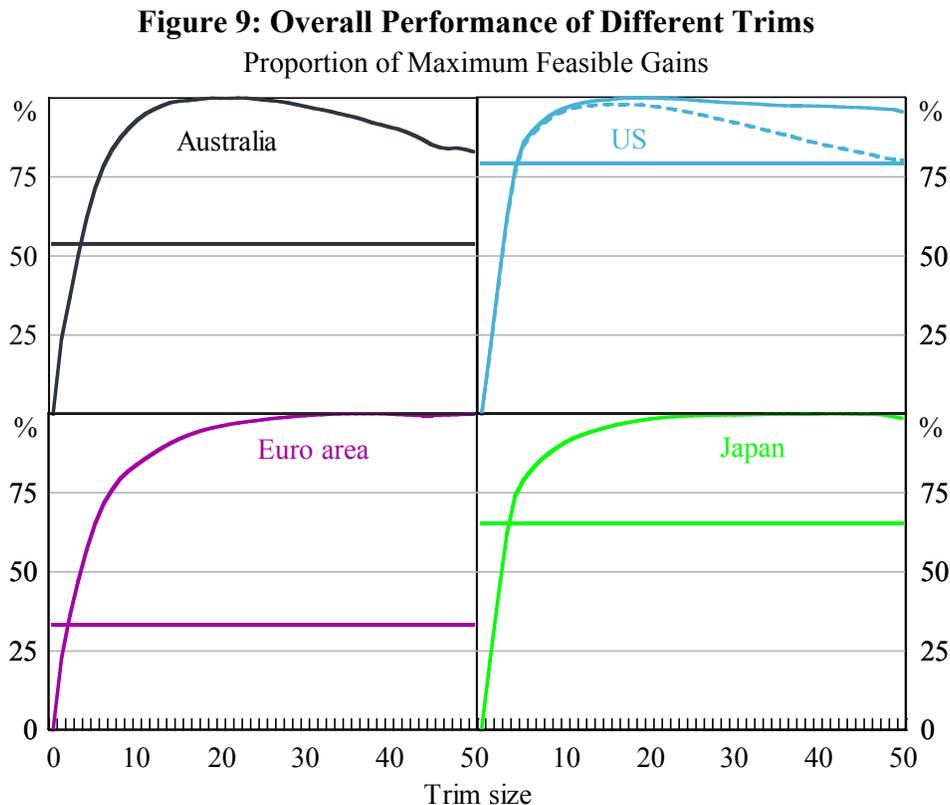
²⁸ For the monthly data, the tests use overlapping data to improve the efficiency of the parameter estimates: we use Newey-West errors to control for the moving average error term that is introduced. The equations all contain just one lag, the maximum that is suggested by the data.

²⁹ All results are valid at the 5 and 10 per cent levels, with the exception of Australian trimmed mean causing CPI inflation at only the 10 per cent level. In three cases, we make a judgmental adjustment where the addition of another lagged inflation variable raises the explanatory power of the equation, but the parameter estimate on the added variable is negative. We do not take this as evidence of Granger causality based on the view that Granger causality tests should not just look at F-statistics but – where there is some expectation as to the sign of the relationship – also at the sign of the implied effect.

measures – are of limited value because their ability to forecast headline inflation is only modest. An alternative interpretation would be that it is exactly because of the noise in headline inflation that we need measures of underlying inflation, and that their primary function is as guides to the true level of current inflationary pressures rather than as leading indicators of future headline inflation. We tend to favour the latter interpretation, and suggest that the predictability results and earlier exercises in this section present a powerful case for trimmed mean measures being much closer to an optimal measure of underlying inflation than either the CPI or the exclusion measures.

5.5 Is There an Optimal Trim?

Based on the analysis above, we now consider the question of which of the feasible trims might be preferred over others. Although many different criteria might be considered, we simply take the results from each of Figures 4, 6 and 7 and construct a simple summary measure of how close each trim is to an overall ‘optimal’ one.³⁰ The results are shown in Figure 9.



Notes: Horizontal lines represent performance of exclusion measure and dashed line for US is calculated without disaggregating the implicit rent component.

³⁰ Specifically, for the data in each of the three figures we calculate the maximal feasible improvement according to that criterion (in the case of Figure 7 we use the CPI prediction results), and the proportion of feasible improvement that is captured by each trim and by the exclusion measure. We then take the mean of these three proportions for each trim, rescaled so that the ‘best trim’ is equal to 100 percent.

The optimised trims (according to this measure) are around 20 per cent for the United States and Australia and around 40 per cent for Japan and the euro area. However, if we define trims that capture 80 per cent of the feasible gains as being close to optimal, the range of ‘close to optimal’ trims starts at between the 5 and 9 per cent trims and extends all the way to the 50 per cent trim (weighted median) for all countries. In none of these four economies would the exclusion measure have been included among the ‘close to optimal’ measures of underlying inflation (although the US core measure would come close). The key point from this is that there is a wide range of trims that offer substantial improvements over the headline or exclusion measures.

Yet if central banks are to calculate and publish a trimmed mean measure of inflation, it will be necessary to choose between the range of trims that seem to perform well.³¹ In considering this question there will inevitably be trade-offs between various goals for underlying measures. For example, if the goal is to arrive at a series which is very smooth and gives good signals of *changes* in the trend level of underlying inflation then it may be optimal to trim a significant proportion of the distribution of monthly or quarterly price changes – say by choosing the optimal trims of 20-40 per cent suggested by Figure 9. Yet the more that is trimmed, the greater the likelihood that the remaining part of the distribution may no longer give a good reading of the *level* of overall inflation: that is, a small amount of bias may be introduced. The latter point would suggest that if there is a choice between a range of trims that appear to be close to optimal, it may be best to err towards the lower end of the range to minimise the possibility of bias. This might suggest trims closer to those published by the Federal Reserve Bank of Cleveland (8 per cent from each end of the distribution) and the Reserve Bank of Australia (15 per cent).

6. Assessment of Recent Underlying Inflation

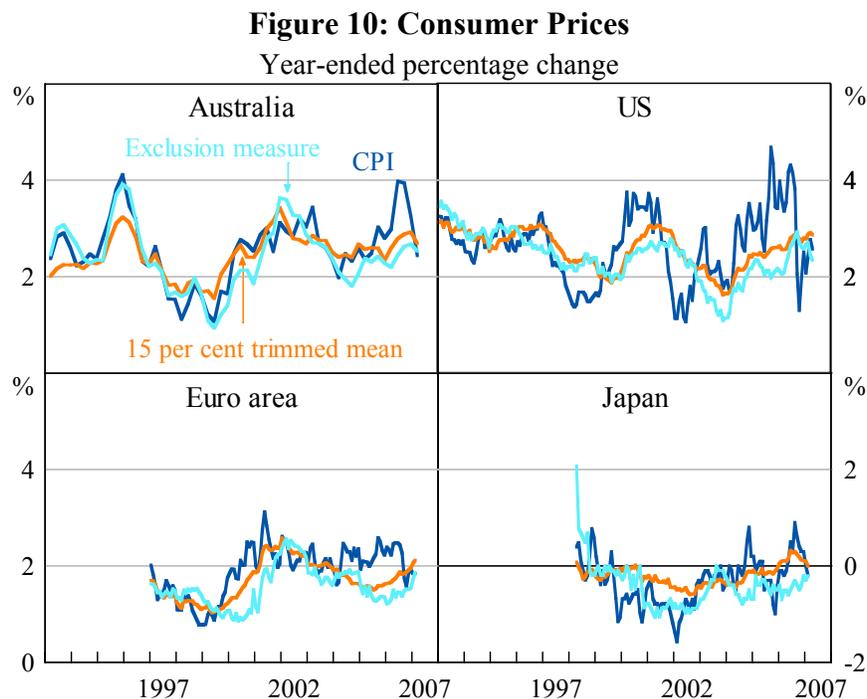
Given that 10 or more years of monthly and quarterly data suggest that trimmed means perform well against the various criteria in Section 5, we now look at the picture they provide about recent developments in inflation in the period when movements in oil prices have significantly affected headline CPI inflation.

There has been significant debate over how to think about the effect of rising oil prices, with some central banks focusing mostly on exclusion measures, but others looking more at headline inflation. One justification for the latter stance is that rising oil prices have been caused in large part by the growth in demand for oil from China and other emerging markets. Yet the growth of these countries has also resulted in falls in prices of a range of manufactured goods, and it would seem

31 At one level, it might seem that the weighted median is the least arbitrary choice, because it focuses on the centre of the distribution, which seems quite a natural choice. Indeed, this may be the reason that it was used in many of the early applications of trimmed means. However, smaller trims tend to perform slightly better on most of the criteria used in this paper, and the possibility of bias might also suggest against such a large trim. One potential problem with the weighted median is that it is the rate of inflation for a single item, and there can sometimes be noticeable ‘gaps’ even in the middle of the ordered distribution of price changes. This has prompted some authors (e.g. Laflèche 1997, Kearns 1998) to suggest various smoothing techniques that give some weight to the ‘adjacent’ price changes, but the more straightforward solution is simply to use a smaller trim.

inappropriate to leave out one effect that is boosting headline inflation while leaving in the effect that is reducing inflation. Fortunately, trimmed mean measures can deal even-handedly with these two effects, by down-weighting potential outliers at both ends of the distribution.

In Figure 10 three measures of annual inflation are shown: headline CPI inflation, the traditional exclusion measure and a trimmed mean measure based on cumulated monthly or quarterly changes. Although we have demonstrated that a broad range of trims performs similarly well as measures of underlying inflation, we use the 15 per cent trim, based on the discussion in Section 5.5. For the United States, the trimmed mean measure is based on the disaggregated regional data for implicit rents.



It is noteworthy that in all four economies the trimmed mean measures tend to provide somewhat less noisy measures of annual inflation than the traditional core measure. Both of those measures are much less volatile than the headline measure, which was boosted by rising oil prices for much of 2004-06 and then fell sharply in late 2006 when oil prices fell. Looking at the most recent period, the trimmed means suggest that underlying inflation has been running at a higher rate than is suggested by the more traditional exclusion measures of core inflation. Nevertheless, the trimmed mean measures suggest that, compared with the experience of the 1970s, inflationary pressures have remained generally well-contained in the face of a major shock to oil prices.

7. Conclusion

This paper has examined the relative performance of the headline CPI, exclusion-based cores, and trimmed means as measures of underlying inflation. Although there are many other candidates for measures of underlying inflation, the measures studied here share the benefit of being easily calculated and communicated. Based on data for four economies, we find that trimmed means tend to outperform headline and exclusion-based core measures on a range of different criteria, which indicate that trimmed mean measures can be thought of as having a higher signal-to-noise ratio than either of the other measures. This makes trimmed means more useful for extracting information about the current trend in underlying inflation from the relatively noisy monthly or quarterly CPI data.

The results lend support to the use of trimmed means as useful measures of underlying inflation in the circumstances seen in recent years when the growth of China and other emerging markets has been having two offsetting effects on global inflation. Whereas some central banks have tended to focus on headline inflation and others have focused more on exclusion measures, our results provide some justification for a middle path, namely using trimmed mean measures which deal with outliers at both ends of the distribution of price changes in a symmetric manner.

Our results suggest a few important technical points regarding the calculation of trimmed means. First, the performance of trimmed means (especially large trims such as the weighted median) can be improved by breaking up expenditure items which have a large weight in the CPI. We show that breaking up the large implicit rent component in the US CPI into four regional indices sometimes has a substantial effect on estimates of weighted median inflation. Second, when calculating trimmed means based on the distribution of annual price changes it may be preferable to avoid using the weighted median but rather to use significantly smaller trims. Third, when using measures based on cumulated monthly or quarterly trimmed means, there is a fairly wide range of trims that appear to perform well, though on balance we would suggest erring towards smaller trims.

Although our results suggest that trimmed mean measures are useful indicators of underlying inflation pressures, we cannot rule out the possibility that further research will suggest additional improvements to the calculation of these measures. In addition, it is unlikely that any single measure of underlying inflation can be held up as the ‘best’ measure at all times and in all countries. The relative usefulness of different measures may change depending on the nature of prevailing shocks, suggesting that central banks and other analysts should look at a range of measures of inflation in assessing inflation developments. Our results suggest that trimmed mean measures should have a reasonably prominent place in any suite of measures, although there will be no substitute for detailed analysis of all the factors driving the CPI at any point in time.

References

- Armour J (2006), 'An Evaluation of Core Inflation Measures', Bank of Canada Working Paper No 2006-10.
- Aucremanne L (2000), 'The Use of Robust Estimators as Measures of Core Inflation', National Bank of Belgium Working Paper No 2.
- Bakhshi H and T Yates (1999), 'To Trim or Not to Trim? An Application of a Trimmed Mean Inflation Estimator to the United Kingdom', Bank of England Working Paper No 97.
- Blinder AS (1997), 'Commentary on "Measuring Short-Run Inflation for Central Bankers"', Federal Reserve Bank of St. Louis *Review*, 79(3), pp 157–160.
- Bryan MF and CJ Pike (1991), 'Median Price Changes: An Alternative Approach to Measuring Current Monetary Inflation', Federal Reserve Bank of Cleveland *Economic Commentary*, Dec 1.
- Bryan MF and SG Cecchetti (1994), 'Measuring Core Inflation', in NG Mankiw (ed), *Monetary Policy*, University of Chicago Press, Chicago, pp 195–215.
- Bryan MF, SG Cecchetti and RL Wiggins II (1997), 'Efficient Inflation Estimation', NBER Working Paper No 6183.
- Catte P and T Sløk (2005), 'Assessing the Value of Indicators of Underlying Inflation for Monetary Policy', OECD Economics Department Working Paper No 461.
- Cristadoro R, M Forni, L Reichlin and G Veronese (2005), 'A Core Inflation Indicator for the Euro Area', *Journal of Money, Credit and Banking*, 37(3), pp 539–560.
- Cogley T (2002), 'A Simple Adaptive Measure of Core Inflation', *Journal of Money, Credit, and Banking*, 34(1), pp 94–113.
- Cutler J (2001), 'Core Inflation in the UK', Bank of England External MPC Unit Discussion Paper No 3.
- Dolmas J (2005), 'Trimmed Mean PCE Inflation', Federal Reserve Bank of Dallas Research Department Working Paper No 0506.
- Gillitzer C and J Simon (2006), 'Component-smoothed Inflation: Estimating the Persistent Component of Inflation in Real Time', Reserve Bank of Australia Research Discussion Paper No 2006-11.
- Kearns J (1998), 'The Distribution and Measurement of Inflation', Reserve Bank of Australia Research Discussion Paper No 9810.

- Laflèche T (1997), 'Statistical Measures of the Trend Rate of Inflation', *Bank of Canada Review*, Autumn, pp 29–47.
- Marques CR, PD Neves and LM Sarmiento (2003), 'Evaluating Core Inflation Indicators', *Economic Modelling*, 20(4), pp 765–775.
- Quah D and SP Vahey (1995), 'Measuring Core Inflation', *Economic Journal*, 105(432), pp 1130–1144.
- Rich R and C Steindel (2005), 'A Review of Core Inflation and an Evaluation of its Measures', Federal Reserve Bank of New York Staff Reports No 236.
- Roberts I (2005), 'Underlying Inflation: Concepts, Measurement and Performance', Reserve Bank of Australia Research Discussion Paper No 2005-05.
- Roger S (1997), 'A Robust Measure of Core Inflation in New Zealand, 1949–96', Reserve Bank of New Zealand Discussion Paper No G97/7.
- Roger S (1998), 'Core Inflation: Concepts, Uses and Measurement', Reserve Bank of New Zealand Discussion Paper No G98/9.
- Roger S (2000), 'Relative Prices, Inflation and Core Inflation', IMF Working Paper No WP/00/58.
- Shiratsuka S (2006), 'Core Indicators of Japan's Consumer Price Index', *Bank of Japan Review*, E-7 (November), pp 1–9.
- Smith JK (2004), 'Weighted Median Inflation: Is This Core Inflation?', *Journal of Money, Credit and Banking*, 36(2), pp 253–263.
- Vega J-L and MA Wynne (2003), 'A First Assessment of Some Measures of Core Inflation for the Euro Area', *German Economic Review*, 4(3), pp 269–306.
- Wynne MA (1999), 'Core Inflation: A Review of Some Conceptual Issues', European Central Bank Working Paper Series No 5.