To trim or not to trim?

An application of a trimmed mean inflation estimator to the United Kingdom

Hasan Bakhshi and Tony Yates

We are very grateful to Andrew Brigden for the work he did developing computer programs used in this paper. Comments from Bill Allen, Michael Bryan, Steve Cecchetti, Mark Cornelius, Alec Chrystal, Alastair Cunningham, Mervyn King, Danny Quah and Scott Roger materially improved this paper. All remaining errors are ours alone. The views expressed in this paper are those of the authors and not necessarily those of the Bank of England.

Issued by the Bank of England, London, EC2R 8AH to which requests for individual copies should be addressed: envelopes should be marked for the attention of the Publications Group (Telephone 0171- 601 4030). Working papers are also available on the Bank's Internet site at http://www.bankofengland.co.uk.

ISSN number 1368-5562

Contents

Abstract

Although the *target* of monetary policy is clear, there have been suggestions that the *conduct* of monetary policy is improved by monitoring ‘trimmed mean’ inflation rates, the mean of some central portion of the distribution of price changes. This paper assesses critically the theoretical and empirical arguments for trimming, and applies Bryan *et al*’s (1997) concept of the ‘optimal trim’ to the United Kingdom.

To trim or not to trim?

An application of a trimmed mean inflation estimator to the United Kingdom

1 Introduction

Monetary policy aims to control inflation. A key problem in the design of monetary policy is what measure of inflation to use. The Government's inflation target stipulates that monetary policy should aim to achieve an annual rate of increase in retail prices (excluding mortgage interest payments) — RPIX — of 2 1/2%. Although the *target* of monetary policy is clear, there have been suggestions that the *conduct* of policy might be improved by monitoring other measures of inflation as a means to achieving the RPIX target.

In this paper we comment on one such proposal by Bryan and Cecchetti (and most recently with co-author Wiggins) that policy-makers might gain from calculating what is called a 'trimmed mean', the mean of some central portion of the distribution of price changes.(1)

Trimmed means are motivated by two sets of arguments: economic and statistical. The economic argument is that in some models of sticky prices, relative price shocks temporarily affect the aggregate price level, even though the long-run effect, when all prices have adjusted, is zero. Trimmed means help alleviate this problem. We argue that the case is not so clear-cut, as there are other models that suggest that trimming may make things worse, not better. The statistical argument runs as follows: the measured inflation rate is the mean of a sample of price changes drawn

(1) Bryan, M and Cecchetti, S (1994), 'Measuring core inflation', in Mankiw, N G (ed) *Monetary Policy*, Bryan, M *et al* (1997), 'Efficient inflation estimation', *mimeo*. The methodology has been applied to the New Zealand inflation rate by Roger (1997).

from an unknown population. If this population is normally distributed, then the sample mean is the most efficient estimator of the population mean. But if the population has fatter tails, then means calculated over sample distributions with trimmed tails will be more efficient estimators of the population mean than the mean of the untrimmed distribution. We point out that the weights implied by the trimmed mean ('one' on the centre and 'zero' on the tails) are only an approximation to the weights implied by the theory of small samples.

Bryan *et al* evaluate different trimmed means according to their ability to proxy for long moving averages of the inflation rate in the United States. We conduct a comparable exercise for the United Kingdom. We conclude that the 'optimal trim' — the trim that best approximates the benchmark inflation rate — is not as robust an estimator of core inflation as appears to be the case for the United States. The optimal trim appears to be sensitive to the statistical criterion used to assess efficiency and to the precise way in which the underlying inflation rates are calculated. We also comment on some problems in choosing the appropriate benchmark inflation rate.

We conclude that the case for trimming inflation rates on statistical grounds is convincing. What is less clear is just how much trimming is optimal.

2 The theory of the trimmed mean

(i) The case in favour

The theoretical motivation for using trimmed means, articulated in Bryan and Cecchetti (1994), stems from the menu cost model of Ball and Mankiw (1995). The intuition is as follows. Suppose firms can review prices costlessly at regular intervals. Suppose too that they may change prices within this interval but in doing so incur a fixed cost.[2] The price-setting rule therefore has elements of 'time-dependence' and 'state-dependence'. Now assume that the economy is hit, for example, by a supply shock that causes some relative prices, but not the aggregate price level, to rise. The desired prices of a few firms rise by a considerable amount, and those of the rest of the firms in the economy fall by a small amount (leaving the aggregate of desired prices unchanged). Only the few firms with large positive changes in desired price will find it worthwhile to pay the menu cost and change actual prices. In this case, a measure of the aggregate price level which averages all price changes across firms will show an increase, even though the aggregate price level implied by desired prices has not changed. In the long run, once all firms reach their regular price-review dates, actual prices will converge on desired prices, and the mean of those desired prices will also equal the mean of actual prices.

In such an economy, a more accurate measure of the aggregate price level implied by desired prices (and therefore of the price level to which actual prices will converge in the long run) will be obtained by ignoring the few price changes that are large enough to warrant firms paying the menu cost associated with a change in their prices. Provided that sufficiently few firms find it worthwhile to pay the menu cost, then the median and the mode price change will be the best estimators of the long-run change in the price level (which is zero).

(2) This cost could come from a variety of sources. It could represent the risk of antagonising customers if prices change at unexpected junctures; or the cost of bringing forward pricing reviews and disrupting other management activities; or even the cost of being sufficiently aware of the optimal prices at all times, as against collecting the information to calculate optimal prices at discrete intervals.

(ii) The case against

The first problem with the case 'for' is that — taken literally — it assumes that only supply shocks(3) can cause skewed changes to desired prices. Our first comment is to point out that aggregate demand shocks can also generate skewed distributions of actual price changes.

Consider first an economy where price-setting is purely time-dependent (or where shocks are such that the state-dependence in prices never manifests itself). Suppose too that there is constant money supply growth (a series of positive aggregate demand shocks) but that there are no supply shocks, and so desired relative prices are not changing. This growth is never large enough to warrant firms paying the menu cost associated with discretionary price changes (if indeed there is any state-dependence at all), and firms simply update nominal prices at their review dates in line with this constant growth rate. If the price-review dates are staggered across firms, then although *desired* relative prices will not change, *actual* relative prices will change each time a firm reaches its review date. But the trimmed mean will discard the price changes of those firms who have reached their review dates. If, for example, at any one time there is only one firm changing prices, then the whole distribution of price changes will be concentrated at zero, with one firm forming the tail of the distribution at a value equal to the rate of increase in the money supply. A trimmed mean would discard the outlying firm. In this extreme case, the trimmed mean would never record any increase in prices at all, even though the aggregate price level over a long run would increase by the rate of growth of money!

In this economy, one should put more, not less, weight on observations at the tail of the distribution, since these contain more, not less, information about the underlying path of the aggregate price level. In sum, if there is staggered, time-dependent price-setting, and there are demand shocks but

(3) Strictly speaking, this should be relative price shocks, as distinct from aggregate demand shocks; a relative demand shock would do just as well in this model.

no supply shocks, we would adjust the distribution of prices in the opposite way to that suggested in Bryan and Cecchetti.

We can think of a similar result when price-setting is purely state-dependent, ie there are no regular price-review dates. In this economy nominal prices are left unchanged until the desired price moves more than some predetermined distance from the actual price. (In the literature this distance is defined by what is known as the Ss bound, where S denotes the upper value of the desired price and s denotes the lower value of this price). If either (i) the distance between S and s is not common amongst firms, or (ii) firms are at different points in the Ss interval immediately before the impact of the demand shock (perhaps because firms have different histories of shocks to relative prices), then a demand shock of a given size will mean that for some firms desired prices breach the Ss bound, while for others they do not.(4) Let us consider an extreme example:

Suppose an aggregate demand shock hits the economy and firms' Ss bounds are such that only one firm finds it optimal to change prices in response. A trimmed mean estimator would discard this one change and take the average of all other prices, leading us to conclude that the aggregate of all desired prices was unchanged. But this is clearly not so: all desired prices have increased. And the simple average of actual prices (which in this example would record an increase) would be a better estimate than the trimmed mean. Better still would be to 'trim' all other prices from the measure except those at the tail.(5)

In sum, aggregate demand shocks can cause skewness in relative price changes. We cannot dismiss this as a mere theoretical possibility: it may not be too far-fetched to imagine that the systematic positive skewness we observe in the distribution of UK prices — as Chart 11 later in the paper

(4) An example of a model where this might occur is presented in Caballero and Engel (1993).
(5) Note that these examples are consistent with rational expectations.

shows — is the result of the repeated positive nominal aggregate demand shocks experienced during a regime of positive inflation.

These examples of situations where demand shocks cause skewness in the price distribution give us insights into how to improve the estimation of inflation in desired or long-run prices. In these particular examples, price changes at the extremes of the distribution should get *more* weight in the estimator than those in the centre of the distribution, the reverse of the Bryan and Cecchetti trimmed mean. In general, in a world of menu-cost price-setting, the weight that is attached to each percentile of the distribution will depend on the exact balance of demand and supply shocks. Since this will vary over time, then so will the weights we attach to percentiles of the price distribution. But of course, if we knew the exact magnitude of demand and supply shocks at the time the shocks hit the economy then we would not need to calculate the trimmed mean, which is designed to adjust for the temporary influence of supply shocks on the aggregate price level!

In addition, the examples we have cited are ones where demand shocks are large enough to push only some firms into making state-dependent price changes. If they are sufficiently large that, when their effects on desired prices are added to those of any subsequent supply shock, all firms find it optimal to pay menu costs and make a price change, then that supply shock will have no distorting influence on the measured inflation rate in any case.

One final remark on the theoretical motivation for using the trimmed mean: it rests on menu costs being a significant economic phenomenon in the first place, and one that gives rise to the mixed time and state-dependent rules that we have talked about (of course this is also true of our counter-examples). But direct evidence on menu costs is not too supportive. For example, Kashyap (1995) reported that there were quite a few examples of small price changes in retail catalogues in the United States (recall that if there are menu costs, then firms are likely to ‘save up’ price increases and make big changes). Also in the United States, Blinder (1994) asked companies whether they think menu costs are a significant factor in deterring them from changing prices. It ranked sixth out of twelve different price stickiness theories.

In the United Kingdom, Hall *et al* (1997) asked around 650 companies the same question; only 7% of their sample thought menu costs were significant. Moreover, this same UK survey found that the mixed time and state-dependent pricing rule that underlies the Bryan *et al* trimmed mean is not that common: only 10% of firms operated this kind of pricing rule. 79% said they used a purely time-dependent rule. Now it must be the case that even these 79% of firms would change prices if the state shifted by a large enough amount, but if, over the sample period we are looking at, no such event took place, then it is reasonable to think of the United Kingdom as a predominantly time-dependent price-setting economy. And recall that, with continual positive demand shocks and time-dependent pricing, the trimmed mean may not be the best measure of core inflation.

3 Statistical theory

Bryan *et al* (1997) present the statistical arguments for using the trimmed mean measure. The intuition of the case they present is as follows. Recorded price changes are samples from a population of price changes whose distribution is unknown. If the population is normally distributed the best (most efficient) unbiased estimator of the population mean is the sample mean. But if the population distribution is leptokurtic (ie has fatter tails than the normal distribution), this result no longer holds. Sampling errors in the draws from the tails will cause the sample distribution to be skewed even if the underlying population is symmetric. The trimmed mean estimator — the average of some central portion of the distribution — is a more efficient estimator of the population mean in this case, since it is less affected by idiosyncratic draws of extreme price changes at either end of the distribution. Bryan *et al* conduct Monte Carlo tests to demonstrate this result: the more leptokurtic a population distribution, the less efficient an estimator of the population mean is the sample mean. Importantly, they also show that the more leptokurtic a population distribution the more of the tails of the sample distribution it is optimal to trim before averaging.

We can think of another motivation for using the trimmed mean that Bryan *et al* do not mention. The fact that the Office for National Statistics (ONS) in the United Kingdom uses expenditure weights in measuring prices that are only updated annually also bears on the question of trimming. Suppose the price of a particular good rises. Leaving the expenditure weights unchanged will mean that the expenditure component of the largest positive price changes will be overstated: unless goods are completely non-substitutable, consumers will substitute away from goods whose prices rise. In a symmetric distribution of sampled prices that accurately reflects the population, this would not be a problem. The expenditure on goods whose price change is at the top of the distribution would be overstated (large price rise times unchanged quantity), but the expenditure on goods whose price change is at the bottom would be understated (large price fall times unchanged quantity).

But, as Bryan *et al* point out, these distributional conditions are not in general met. Specifically, we might face a representative sample from an asymmetric population distribution, or an unrepresentative (ie asymmetric) sample from a symmetric population. So taking trimmed means could be motivated by infrequent adjustment of expenditure weights by the statisticians. The case might even in principle be stronger than that of Bryan and Cecchetti's based on infrequent adjustment of prices by firms. First, infrequent expenditure weight adjustment is a fact, not a theoretical proposition. Second, weight adjustment is simultaneous for all goods, and is purely time-dependent, so there is no danger of discarding 'leading indicator' price series, as there is when thinking about price adjustment. Against this, in the United Kingdom at least, the published price sub-components are still aggregated across some goods: it is likely that substitution between such goods at this level would be less than between individual goods.

The statistical case for trimming appears to be convincing. The question is how much of the distribution should be trimmed. Bryan *et al* show that trimmed mean estimators are better approximations of measures of 'core' inflation in the United States than simple mean inflation rates, where core inflation is defined as a centred moving average of inflation (of varying lengths). They calculate an 'optimal trim' by searching for that which best proxies these long-run averages of inflation. In Section 4 we conduct a similar exercise for the United Kingdom, and then reflect on what we can learn from the results.

4 Application to the United Kingdom

The remainder of this paper does two things. First, we replicate the Bryan *et al* analysis on UK data. Second, we discuss the methodology they use and the problems in moving from the statistical motivation for the trimmed mean to its application in practice.

(i) The optimal trim for the United Kingdom

We look for the optimal trimmed mean estimator of Bryan *et al*'s measure of core inflation using the entire 1974 to 1997 sample for which we have a sufficiently disaggregated distribution of price changes across goods. Specifically, following Bryan *et al*, we look to see which portion of the distribution of price changes, when averaged, is the closest approximation to a centred moving average of the mean of all price changes. In the baseline case, a 37-month centred moving average is used. We measure how close the approximation is by calculating the root mean square error (RMSE) and mean absolute deviation (MAD) for each trimmed mean estimator. The trimmed mean is calculated in each case by first computing annualised one-month changes in the non seasonally adjusted price of 81 individual RPIX components and second, averaging the central x% of the distribution of those price changes. We look at short-horizon inflation rates simply because the problems induced by menu-cost price-setting or sampling errors are likely to be more acute at the shortest horizons. The implication of the Bryan *et al* analysis is that, in time, all prices will change to reflect shocks to desired prices, and, in time, sampling errors will balance out through repeated draws from the underlying population of price changes.

In the baseline case we annualise the one-month change in prices by raising them to the power of 12. We then trim the resulting annualised series. An alternative would be first to trim the monthly inflation rates and then compound the resulting series to calculate annualised trimmed mean inflation rates. The latter is what Bryan *et al* appear to do in their baseline case. We also do this, in our robustness section. One reason for annualising first is that raising monthly price changes to the power of 12 accentuates the skewness of the distribution. This stacks the cards in favour of trimming: something we want to do given that our aim is to provide *critical* comment on some of Bryan *et al*'s arguments for trimming. A second reason is that the alternative of compounding trimmed monthly inflation rates means that the resulting annualised series conflates different goods in different months. The resulting annualised inflation rate for any given month cannot then be interpreted as a change in the value of a fixed basket of goods in the preceding twelve months.

We work with seasonally unadjusted data, unlike Bryan *et al*. The main reason for this is that the ONS in the United Kingdom does not publish seasonally adjusted price indices in the United Kingdom. In the United States, by contrast, the published data used by Bryan *et al* is seasonally adjusted by the Bureau of Labour Statistics. Given the many different methods of seasonal adjustment that exist in the literature we did not want to induce further degrees of uncertainty into our analysis.

The results for our baseline case are shown in Charts 1 and 2:

Chart 1

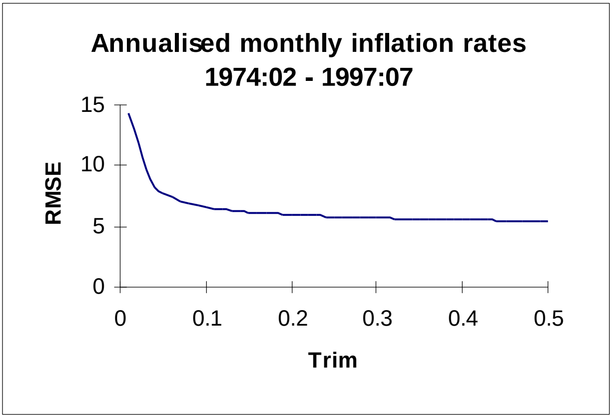
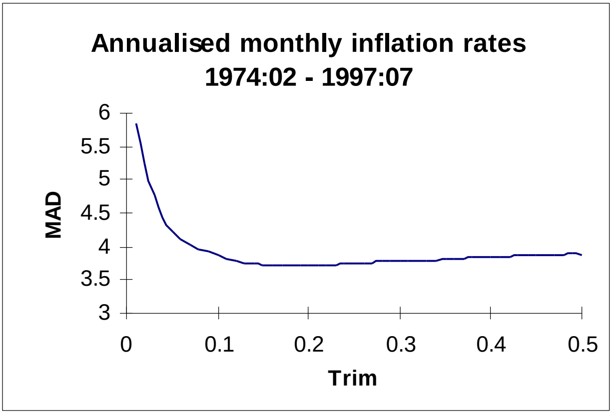


Chart 2



These charts show how the RMSE and MAD vary when we trim different portions of the distribution. We can find the 'optimal trim' — the trimmed mean estimator that best approximates the 37-month centred moving average of RPIX inflation — by picking out the trough of these charts. If we take the RMSE as the metric, the optimal trim is 0.47 (or 47% from each tail of the distribution). On the other hand using the MAD measure the optimal trim is only 0.17. It is not obvious which measure of efficiency is superior: note that the RMSE measure attaches a greater weight to large forecast errors (since in the RMSE the errors are squared first and summed before taking the square root).(6)

Table 1 The optimal trim for the United Kingdom

RPIX 1974:02 - 1997:07

	RMSE	MAD
Mean(a)	17.66	6.98
Median	5.51	3.88
Optimal trim	5.45	3.71
Trim at optimum	47%	17%

(a) We have used here a non chain-weighted RPIX to make it comparable with the trimmed mean.

Table 1 summarises the properties of different estimators of ‘core’ consumer inflation. The first point to note is that — just as Bryan *et al* found — trimming appears to help: trimming 47% of the cross-sectional distribution of consumer prices reduces the RMSE by just under 70%. Second, the RMSE and MAD give very different values for the optimal trim. Third, the RMSE and MAD associated with the untrimmed mean measure of inflation are both large (and larger than in Bryan *et al*). This should not be surprising: we trim annualised monthly changes in the price of non-seasonally adjusted RPIX components. So we would expect larger forecast errors around the core rate. (Using annualised monthly rates also exaggerates the difference — as measured by the RMSE and the MAD — between the mean and the trimmed mean estimators. As explained in footnote 5, the RMSE penalises large errors more heavily, so annualising increases the difference between the RMSE and the MAD measures.)

(6) This is easy to see with a simple example. Consider first a set of three forecast errors taking the values 2,2 and 2. The MAD here equals 6/3 and the RMSE also equals $\sqrt{\frac{12}{3}}$ =2. Now consider a distribution of three forecast errors where there is an outlier: 1,4 and 1. The MAD is again 6/3 but the RMSE has increased to $\sqrt{\frac{18}{3}}$.

Fourth, on both the RMSE and MAD measures, the median is almost efficient as the optimal trim. But this is not surprising from Charts 1 and 2: it is clear that the changes in the RMSE and MAD as we move the trim about its optimum are not large (the two curves are very flat). This is not a feature of Bryan *et al*'s results, but we argue later that this result is explained by the different annualising methods used in our baseline cases.

Bryan *et al* (1997) argue that the optimal trim for the US CPI (0.09) is robust to changes in the definition of the core inflation rate and to changes in the sample period over which the whole experiment is carried out. They show that the optimal trim is *not* robust to the level of disaggregation of the price components, nor to the horizon over which inflation is measured. Specifically, they find that the optimal trim is higher the more disaggregated the price components (they argue that this is because the population distribution is more leptokurtic the more disaggregated the price components). And they find that measuring component price inflation at lower frequencies (eg annual rather than monthly) — what they call
'pre-trim averaging' — decreases the size of the optimal trim. Intuitively, lengthening the observation interval reduces transitory noise.

(ii) Robustness of the trimmed mean calculation

We turn now to investigate how robust this optimal trim is to a variety of changes in our baseline experiment.

The first change we examine is to the method of annualising monthly price changes. The baseline case calculates the trimmed mean by first annualising monthly changes in RPIX components, and then trimming. Here, as in Bryan *et al*, we trim monthly inflation rates, without annualising them first, and subsequently derive an annual inflation rate for each month by compounding the monthly inflation rates over the previous twelve months (we refer here to these as 'monthly inflation rates').

Charts 3 and 4 show that the optimal trim on both the RMSE and MAD measure is now much lower, at only 0.04. It is clear that annualising monthly inflation rates by raising them to the power of 12 — as in Charts 1 and 2 — magnifies the transitory volatility around 'core' inflation induced by the price changes observed at the tails of the distribution. Importantly, it is no longer the case that the efficiency losses from trimming beyond the optimum are flat. And neither do Bryan *et al*'s results for the United States that also use this alternative annualising procedure display this pattern.

The earlier result that the payoff from trimming is flat around the optimum depends very much on which method of annualising is used.

Chart 3

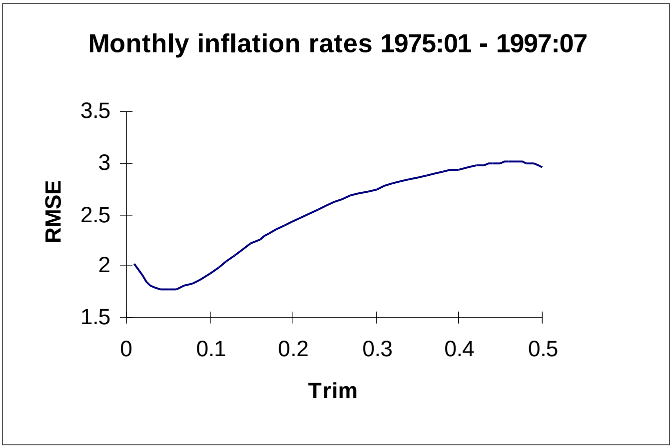
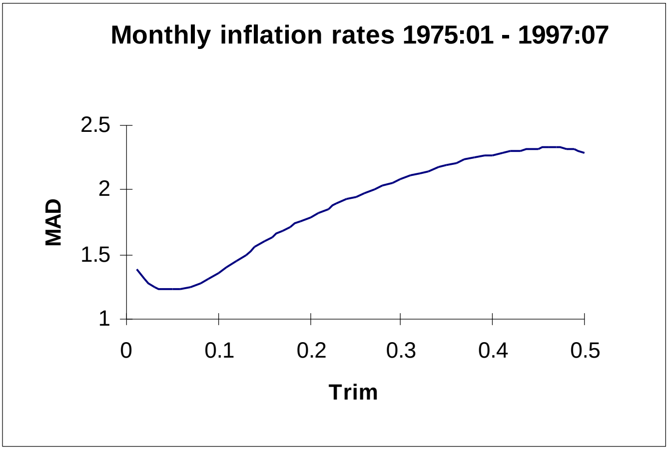


Chart 4



The second change to the baseline case we consider is to the level of aggregation of retail price components. The baseline case calculates trimmed means using the most disaggregated data that the ONS publishes. We also compute the optimal trim using data published at a higher level of aggregation, namely 29 components. The results are shown in Charts 5 and 6:

Chart 5

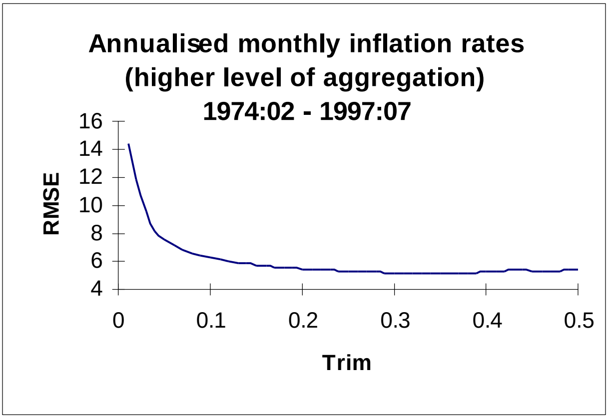
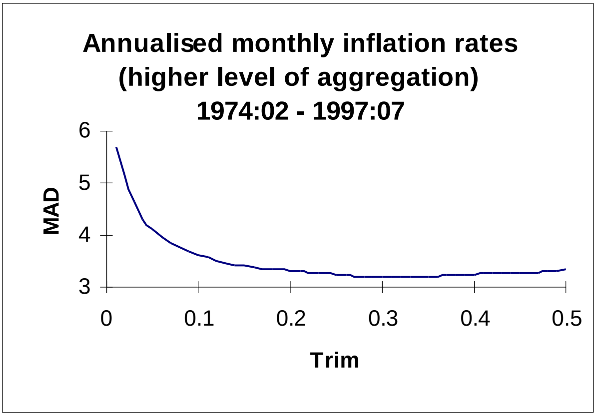


Chart 6



As Bryan *et al* report for the United States, the optimal trim falls on the RMSE criterion, from 47% to 36%. But it rises on the MAD criterion, from 17% to 33%.

In the baseline case, we annualised monthly changes in the rate of inflation of RPIX components before trimming. We get a handle on robustness across the horizon over which inflation is measured by trimming *12-month* changes in RPIX components and then computing the optimal trimmed mean. The optimal trim is now 30 on the RMSE basis and 28 using the MAD (see Charts 7 and 8).

Chart 7

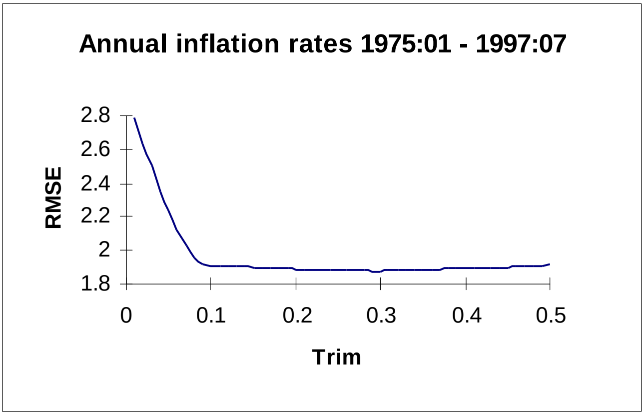
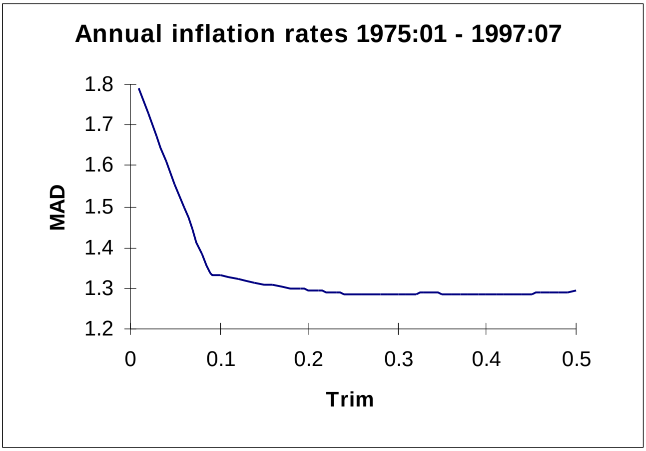


Chart 8



So, on the RMSE measure, the optimal trim is indeed lower for annual inflation rates compared with our baseline annualised monthly rates. But on the MAD measure it is again larger. It is also worth noting that the MAD and RMSE measures of the optimal trim are very similar for annual inflation rates.

We now take our baseline case and calculate the optimal trim for sub-samples of the period over which RPIX data has been published in the United Kingdom, namely 1974:01 - 1997:07. Bryan *et al* (1997) argue that the optimal trim for both the US CPI and PPI are robust to changes in the sample period. Our preliminary findings suggest that this may not be the case for the United Kingdom.

We calculate the optimal trim for ten-year rolling samples across the whole period 1974:01 - 1997:07.

Chart 9

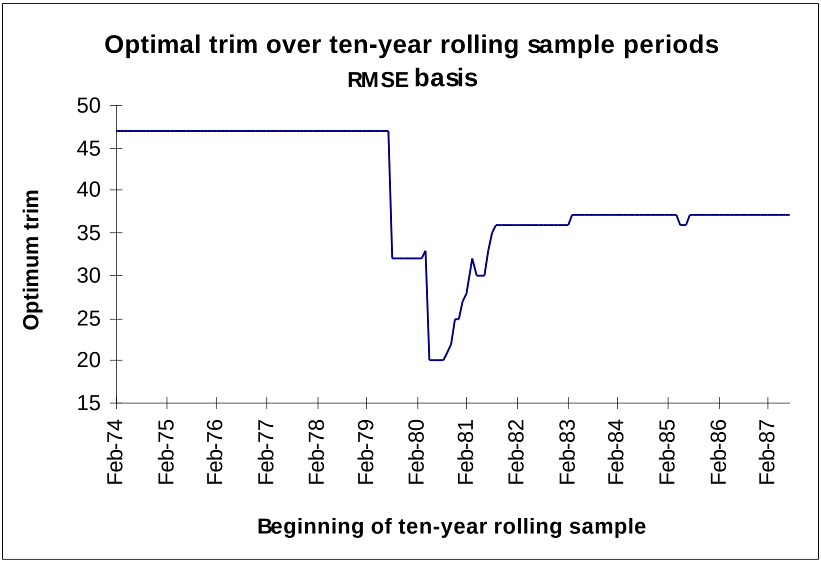
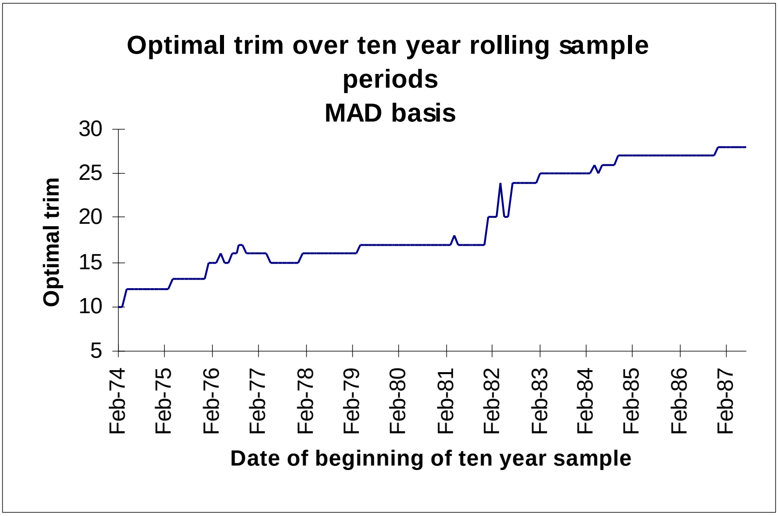


Chart 10



Charts 9 and 10 show that the optimal trim does not appear to be robust across sample periods. And the exact time profile of the optimal trim depends on which measure of efficiency is being used.

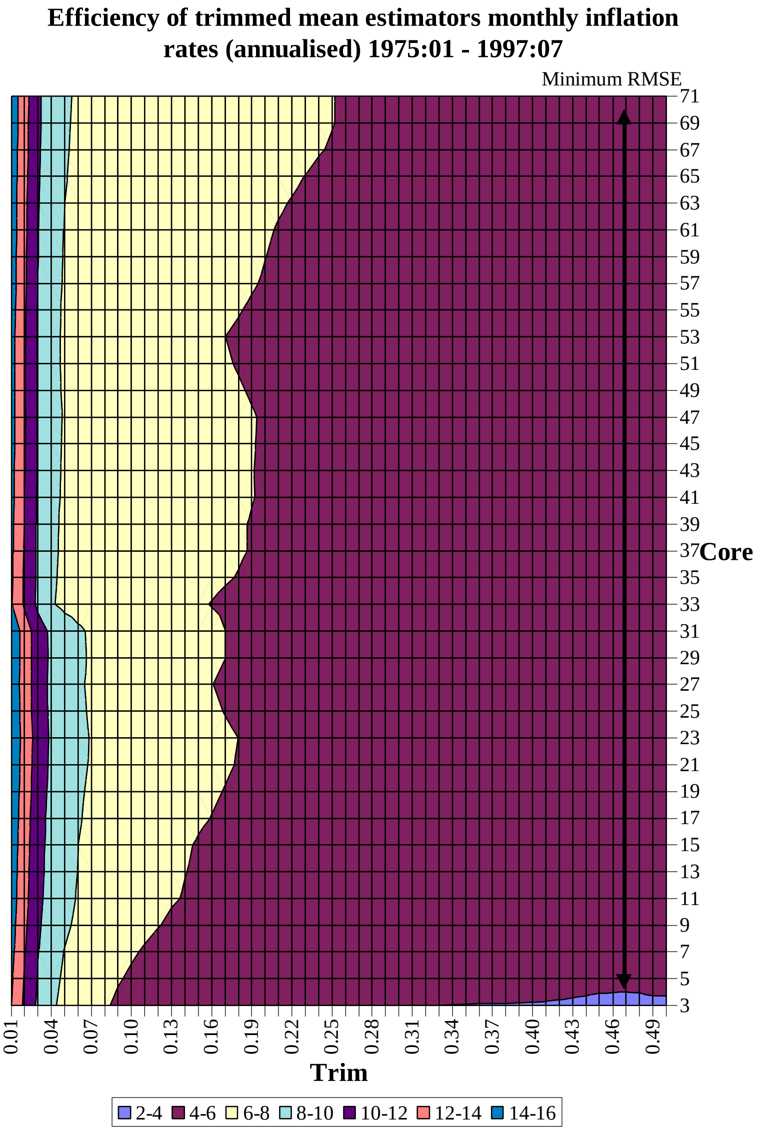
The baseline case defines the core inflation rate as the 37-month centred moving average of RPIX inflation. But how robust is the optimal trim to the way this core is measured? Table 2 reports optimal trims as a function of the length of the moving average specified for the benchmark, core inflation rate.

Table 2 Optimal trims monthly inflation rates (annualised) 1974:02 - 1997:07

Core inflation	Optimal trim using RMSE	Optimal trim using MAD
25	47	24
37	47	17
49	47	18
61	47	20

Table 2 shows that the optimal trim is quite robust to the definition of the benchmark inflation rate (we find that this is also the case for annual inflation rates, and for both annualised monthly and annual rates over the 1987:01 - 1997:07 sample period). This is displayed more generally in the contour chart below (see Chart 11). This plots RMSEs for different combinations of the core inflation rate (as measured by the moving average window) and the trim.   This shows that, at least for annualised monthly inflation rates, the optimal trim on the RMSE criterion, at 47, *never* changes as the core varies from a 3-month to 71-month centred moving average!

Chart 11



Efficiency of trimmed mean estimators monthly inflation rates (annualised) 1975:01 - 1997:07
Minimum RMSE
Core
71
69
67
65
63
61
59
57
55
53
51
49
47
45
43
41
39
37
35
33
31
29
27
25
23
21
19
17
15
13
11
9
7
5
3
0.01
0.04
0.07
0.10
0.13
0.16
0.19
0.22
0.25
0.28
0.31
0.34
0.37
0.40
0.43
0.46
0.49
Trim
2-4
4-6
6-8
8-10
10-12
12-14
14-16

Nevertheless, if we look across the range of our experiments, we conclude that the optimal trim is not as robust an estimator of core inflation for the RPIX as appears to be the case in the United States. The optimal trim appears to be sensitive to the statistical criterion used to assess efficiency — whether we measure efficiency using the root mean squared error (RMSE) or mean absolute deviation (MAD); to the precise way in which the trimmed mean inflation rates are calculated — whether or not monthly rates are annualised by raising to the power of 12 or whether component inflation rates are calculated at horizons longer than one-month inflation; as well as to the sample period. Though Table 1 and Charts 1-3 and 6-8 appear to show that beyond a certain cut-off point, roughly around 10%, there is little to be gained or lost by trimming any more of the distribution, this result seems to be related to the particular annualising method we have adopted in the baseline case. Charts 4 and 5, showing the RMSE and MAD for trimmed mean inflation rates using the alternative annualising method, do not share this feature. And neither do Bryan *et al*'s results for the United States that also use this alternative annualising procedure display this pattern. In sum, without a clear steer from statistical or economic theory, we are left without a way of deciding how much of the distribution to trim.

(iii) Problems with Bryan et al's methodology

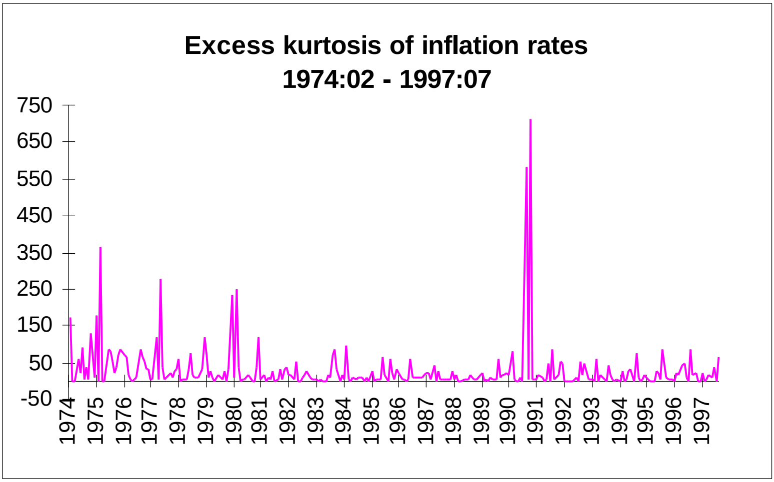
Here we raise a number of questions about the empirical exercise conducted in Bryan *et al*.

First, recall the statistical motivation for trimming. The argument was that the population of price changes was leptokurtic (had fat tails relative to the normal distribution), and that sampling errors in the tails meant that the observed large price increases (decreases) might not always be balanced out by observed large price decreases (increases) — a sample distribution could be skewed — even when the underlying population might be symmetric. Large price changes are 'trimmed' because they have a large moment about — or are a large distance from — the population mean. But 'trimming' as Bryan *et al* have used it gives a *zero* weight to these observations, when the intuition of the statistical argument for trimming is that the weight attached to a percentile in the distribution should be *inversely proportional* to its distance from the population mean. Of course we do not have the population mean to hand — this is the point of searching for an efficient inflation estimator — but it still should be possible to improve upon the zero-one weighting scheme adopted by Bryan *et al*, using the sample mean or the median as the benchmark inflation rate for calculating the weights attached to each percentile.

Second, Bryan *et al*'s Monte Carlo analysis showed that the optimal trim depends on the amount of kurtosis in the underlying distribution. Ideally, therefore, it is the kurtosis of actual price changes that should dictate the optimal trim in reality and not the ability of the resulting trimmed index to approximate some proxy for 'core inflation' (about which we will say something below).

Unfortunately, just as the exercise of trimming is designed to find an estimate of the population mean, we do not have precise information about the kurtosis of the population distribution. Bryan *et al* could have approximated this by the sample kurtosis, however, and in practice we think this would align the trimming exercise more closely to the statistical problem that motivated it. But presumably just as the population mean changes over time, so might the other moments. Even in the absence of menu costs, for example, a large supply shock may bring about large relative price changes in the distribution, and increase population kurtosis. During these periods, sampling errors at the tails would have a larger impact on the error in estimating the population mean, and so the optimal trim would be higher. In short, the optimal trim should depend on the population kurtosis, and will, in general, vary over time. Chart 12 suggests that a time-varying kurtosis is not implausible: the average 'excess' kurtosis over the sample period — ie kurtosis over and above that of the normal distribution — is 28.4 while the standard deviation around this mean is 67.4. However, we cannot rule out the possibility that we are simply observing variable samples drawn from a constant population.

Chart 12(7)



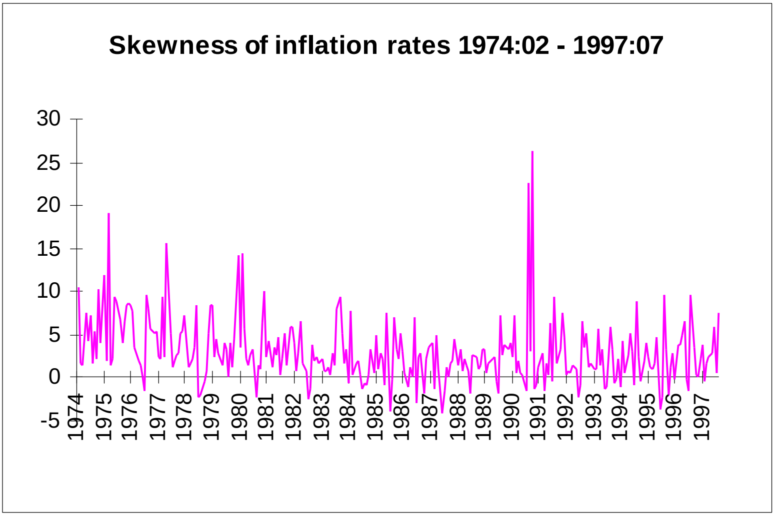
Third, applying an equal trim to the top and the bottom of the distribution — as Bryan *et al* discuss — is only valid when the underlying population distribution is symmetric. In a distribution that is positively skewed, the largest 5% of price changes will have a greater moment about the mean (will lie farther away from it in inflation units) than will the smallest 5% of price changes; in turn, sampling errors in the highest part of the distribution will also play a more important role in incorrectly estimating the mean than those in the lowest part. So in a positively skewed distribution, a larger proportion of the top of the distribution should be 'trimmed' in order that, over repeated draws, the expected impact of

(7) Charts 10 and 11 are in fact based on only those components for which inflation rates are available for the full sample period (64 out of 81).

sampling errors on the mean is zero. (Roger (1997) makes use of this wisdom in applying the trimmed mean to New Zealand, where the price distribution has a systematic, positive skewness.)

In other words, whether the ‘trim’ is symmetric — whether the amount of the distribution trimmed at the top is the same as that trimmed at the bottom — should depend on whether or not the underlying population is symmetric. The first problem is that we don’t have precise information about the population skewness: we only have the sample skewness. But a second possibility is that this population skewness could well vary over time: relative price shocks, even in the absence of menu costs, that result from supply shocks, could generate any distribution of desired price changes, and one that we would have no reason to expect would be symmetric, or time-invariant. Chart 13 shows that skewness is on average positive over the whole sample period, at 3.1, but that there is significant variation around this mean: the standard deviation around the mean is 3.8 over the sample period. This chart suggests that a time-varying population skewness may indeed be plausible.

Chart 13



To sum up these points: the optimal amount of these sample price-change distributions that we should trim depends on a possibly varying, and certainly unknown, population kurtosis. The ratio of the proportion to trim at the top to the proportion we should trim at the bottom depends on an equally variable and equally unknown population skewness. Finally, 'trimming' itself should ideally take the form of applying a distribution of weights to percentiles of price changes, where the weight is inversely proportional to the distance from the population mean. The zero-one weighting scheme implied by Bryan *et al*'s scheme can be improved upon, though this is not to say that zero-one trimming is not better than no trimming at all.

In the empirical part of Bryan *et al*'s analysis they let the optimal trim depend entirely on how well the trimmed estimator approximates a measure of core inflation, defined as a long centred moving-average inflation rate. The implication of the discussion above is that this may not be appropriate. Bryan *et al*'s idea here is that by taking long-window moving averages, the problems associated with estimating the population mean price change by the sample mean disappear. There are two possible reasons why this procedure might be justified.

First, recalling the theoretical motivation for the trimmed mean, if the moving average is calculated over a window that is longer than the typical length of time between price reviews then this measure will not be contaminated by supply shocks. But this requires that all supply shocks hitting the economy have worked themselves through by the end of the moving-average window. In a dynamic economy with supply shocks hitting each month, a core inflation rate calculated in this way may not suffice. The only situation in which it will be appropriate is one where all supply shocks that contaminate the inflation rate both appear *and* disappear within the centred moving-average window.

The second possible justification recalls the statistical motivation for using trimmed means. For a moving-average window of sample mean inflation rates to suffice as a measure of core inflation requires that sampling errors balance over time. Intuitively, if the population moments are stable, then calculating a moving average over a long enough window will mean that there have been enough draws from the tails of the distribution to ensure that extreme upward changes are balanced by enough changes in the opposite direction. The trouble is that the population moments — particularly the skewness and kurtosis — could well vary.

There may be a further problem with using the centred moving average as a core or benchmark rate against which to judge trimmed means. This is that the moving-average window may be long enough to contain aggregate demand shocks that are fully offset by the monetary policy authorities in the period. Yet these are genuine fluctuations in the inflation rate that would have persisted had the authorities not responded: it is not obvious why these shocks should be excluded from the benchmark rate.

It is not an easy task to find a benchmark against which to judge a new measure of inflation like the trimmed mean. For every such benchmark inflation rate, there is always the question: why not use that inflation rate, instead of the trimmed mean? For the centred moving average, the answer is straightforward. At each point in time we do not have the future inflation rates needed to calculate the centred moving average! But as we have discussed above, there are problems with using this measure: its usefulness rests on the time horizon being long enough to overcome problems relating to menu costs or sampling. If population moments change, or there are demand shocks, then (there will at least be periods when) the usefulness of the moving average is limited.

5 Conclusions

Some recent papers have argued that monetary policy-makers should use trimmed mean inflation rates in gauging inflationary pressure. At least two motivations have been given: the first is that supply shocks may induce temporary skewness of *actual* price changes if there are menu costs of adjusting prices, even though the *desired* (or long-run) price change is in fact zero. This induces transitory volatility into the conventionally measured inflation rate, the mean of all price changes. The second motivation is that in leptokurtic distributions of price changes (distributions with fat tails) sampling errors at the tails can induce a significant amount of volatility in the sample mean, even when the true population mean is unchanging. In both cases, it has been suggested by Bryan *et al* that these problems can be mitigated by taking the mean of some central portion of the distribution of observed price changes.

Rather than propose a new measure of inflation, we set ourselves the (far less noble!) task of commenting on one proposed by others. Our argument is not that trimming inflation rates is wrong. It is that while policy-makers can obviously gain useful information by looking at the cross-section distribution of price changes, the question of what information to discard is not at all straightforward. We have two broad points to make:

First, aggregate demand shocks can also bring about skewed changes in actual prices. This may occur if firms operate staggered, time-dependent pricing, or if there are heterogeneous state-dependent rules for setting prices. In all these cases, the implication is that in response to a demand shock, we may obtain a better estimate of inflation consistent with desired prices by putting more, not less weight on the extreme price changes in the distribution, since these contain the most, not the least, information about the future general price level. This compares with Bryan and Cecchetti's model of menu costs that predicts that the policy-maker should put zero weight on the extreme price changes of the distribution. So there is no *a priori* case for trimming on theoretical grounds.

Second, there are difficulties in moving from the statistical theory that motivated the trimmed mean, to a practical measure. The statistical logic behind trimming implies that the optimal trim — the amount of the tails we should chop off — depends on the amount of kurtosis in the population distribution. This kurtosis is unknown and possibly changing over time. In addition, whether we should trim the same amount from the top as from the bottom of the distribution depends on the population skewness. This is also unknown and possibly changing. Moreover, trimming assigns a zero-one weighting scheme to the price change observations. But intuitively, the percentiles should be weighted in inverse proportion to their distance from the population mean. Finally, we raise some objections to the use of a long centred moving average as a benchmark against which to judge how well trimmed means perform, not least that it assumes that all supply shocks that contaminate the inflation rate both appear and disappear within the centred moving-average window.

We conclude that the theoretical grounds for trimming are weak, but that the statistical arguments are more convincing. That said, there are many difficulties in applying the statistical logic when trying to calculate the optimal trim in practice. Given these difficulties it is perhaps unsurprising that we find a marked degree of non-robustness in our estimates of the optimal trim.

References

Balke, N and Wynne, M A (1996), 'An equilibrium analysis of relative price changes and aggregate inflation', *Federal Reserve Bank of Dallas Research Department Working Paper*, pages 96-9.

Ball, L and Mankiw, N G (1995), 'Relative price changes as aggregate supply shocks', *Quarterly Journal of Economics*, 110, pages 161-93.

Blinder, A (1994), 'On sticky prices: academic theories meet the real world', in Mankiw, N G (ed) *Monetary Policy*, NBER.

Bryan, M and Cecchetti, S G (1994), 'Measuring core inflation', in Mankiw, N G (ed) *Monetary Policy*, NBER.

Bryan, M, Cecchetti, S G and Wiggins, R (1997), 'Efficient inflation estimation', *mimeo*.

Caballero, R J and Engel, E M (1993), 'Heterogeneity and output fluctuations in a dynamic menu-cost economy', *Review of Economic Studies*, 60, pages 95-119.

Hall, S, Walsh, M and Yates, A (1997), 'How do UK companies set prices?' *Bank of England Working Paper* No 67.

Kashyap, A (1995), 'Sticky prices: new evidence from retail catalogs', *Quarterly Journal of Economics*, 110, February.

Quah, D and Vahey, S (1995), 'Measuring core inflation', *Economic Journal*.

Roger, S (1997), 'A robust measure of core inflation in New Zealand, 1949-96', Reserve Bank of New Zealand, *mimeo*.