RESERVE BANK OF NEW ZEALAND

DP2004/04

Estimates of the output gap in real time: how well have we been doing?

Michael Graff

May 2004

JEL classification: E37, E52, E58

Discussion Paper Series

DP2004/04

Estimates of the output gap in real time: how well have we been doing?

Abstract[1]

This paper addresses the real-time versus ex-post properties of the output gap as quantified by the Reserve Bank of New Zealand's multivariate (MV) filter, starting with the second quarter of 1997, when the current procedure was implemented.

There are three sources of revisions of the output gap: revisions of real GDP data, the end point problem of symmetric filters and changes to the calibration of the MV filter.

The performance of the output gap with respect to signalling inflationary pressure, as measured by future non-tradables inflation, has been reasonably good, both in real time and *ex post*. However, during the recorded history of the MV filter, the revisions to real-time output gap have been no smaller than had a standard Hodrick-Prescott (HP) filter been used. Moreover, the MV filter leads to permanently different levels of the output gap estimates if compared to a purely statistical trend.

The MV filter is a hybrid construct. The empirical reference to indicators of inflationary pressure distances it from the original concept of the output gap, where a deviation of observed from potential output is taken as a *cause* of inflationary pressure. There is some indication that a major recalibration of the MV filter in 2002 helped to maintain the correlation with a target variable that it is supposed to "explain".

1 I would like to thank David Archer, Nils Bjorksten, Tim Hampton, Ashley Lienert, Özer Karagedikli and James Twaddle for helpful comments on earlier versions of the paper. However, the views in this paper are mine and should not be attributed to the Reserve Bank of New Zealand.

1 Introduction

The output gap is a well-established theoretical concept in contemporary economics. Moreover, it plays a crucial role in many structural macro-models; it is frequently (and with some success) referred to in papers that are looking for a scheme to "explain" (reproduce) a historical path of central bank policy settings; and it is probably safe to assume that a substantial number of monetary policy makers as well as fiscal authorities[2] pay close attention to real-time estimates and forecasts of the output gap. It would be very hard to understand eg the behaviour of the US Federal Reserve without reference to the output gap or the employment gap.

In line with this, the Reserve Bank of New Zealand's medium-size structural macro-model FPS (Forecasting Policy System), which was introduced in the second quarter of 1997, puts heavy weight on the output gap as a key variable to describe and steer the state of the New Zealand economy. In FPS the output gap is quantified by the MV ("multivariate") filter.

Although the output gap plays such a prominent role in current economic theory and policy, it is still an inherently unmeasurable equilibrium based construct. The output gap refers to the deviation of realised from potential output, and this in a constantly changing economic environment. Successfully estimating the output gap requires that in addition to reasonably reliable data of current or near future *realisations* of economic activity (which are hard enough to get),[3] economic policy has to have access to reasonably reliable estimates or projections of *potential* (*equilibrium*) output.

Based on data from several countries, recent studies have cast some doubt on whether the prevailing techniques to estimate the output gap yield practically useful results. Particularly in *real time*, when information on the state of the economy is most important, estimates of the output gap appear to be uncomfortably unreliable.

2 For example, the recently introduced Swiss "debt break" specifies the budget deficit (surplus) as a function of an HP filtered real-time output gap.

3 Evidently, estimating potential output from unreliable estimates of real output is not likely to produce reliable estimates of output gaps.

Is this also true for the New Zealand output gap as quantified by the Reserve Bank's MV filter? To address this question, this paper assesses the *real time* properties of the New Zealand output gap *as actually estimated and forecasted by FPS* for the regular quarterly MPS (Monetary Policy Statements) since June 1997. With the latest data for 2004q1, the overall period is still uncomfortably short for statistical analysis, but – with this caveat in mind – it is informative to take a mainly descriptive look at the record.

Main findings are that while the MV filter has a limited potential to reduce future revisions of our real time output gap estimates, it ensures that the real time output gap reflects the amount of inflationary pressure in the economy. For the period under consideration, however, the real-time MV filtered output gap has not experienced less subsequent revision than had the HP filter been used. Furthermore, based on the starting values from the MV filter, the Reserve Bank of New Zealand's structural macro model has generated reasonably reliable forecasts of the output gap for only one quarter ahead.

The remainder of the paper is structured as follows: Section 2 discusses the prevailing approaches to derive operational definitions of the output gap, including the MV filter. Section 3 reviews the international evidence on output gap estimates in real time. Section 4 presents the results of our empirical evidence for New Zealand. Section 5 concludes.

2 Operational definitions of the output gap

Unsurprisingly, apart from a general understanding that the output gap g denotes the relative departure of empirical output Y from its "equilibrium" or "potential" Y^*,

$$g = (Y - Y^*) / Y^*, \qquad (1)$$

the current state of the art does not give a conclusive answer to how it should be conceptualised. We can distinguish (at least) three attempts at defining it: a substantive, a statistical and a functional.

(1) The *substantive approach* argues that potential output is a function of the amount of (a broad concept) of the factors of production voluntarily[4] available at the period under consideration and the (again: broadly defined) technology at hand to combine them to produce goods and services. This is clearly an economically meaningful concept,[5] and – presumably – this is the basis for its popularity.

To operationalise it, however, is a formidable task, and while a number of attempts to estimate full capacity production functions have been conducted, some, if not most, of the results have been rather disappointing. Consequently, the substantive approach, to which the output gap owes much of its credit, is not the one that practical economists usually refer to. (Neither does the MV filter.)

(2) According to the *statistical* approach, potential output is what you get when you send a real GDP series through a low pass filter (usually the HP filter).[6] Many of the practical methods nowadays in use to derive estimates of potential output nowadays rely on this statistical approach of extrapolation of a smooth trend from the historical path of the output series.

However, if these methods do not predict a constant growth rate for potential output but allow for some (but not total)[7] adaptation of potential to observed output, real-time gap estimates are imperfect in the sense that they are (1) prone to

4 This needs to be stressed, since potential labour is not just a linear function of a well defined demographic cohort, but intrinsically endogenous, replying to a wide array of economic incentives, regulatory interventions, changing tastes (eg for leisure, labour force participation of women, consumptional ingredients of prolonged education etc). In addition, effective working hours directly affect the intensity with which the stock of physical (and other) capital is used, so that the effects of fluctuations in effective labour are further amplified.

5 The neo-Keynesian representation for it neatly verbalised by Nelson and Nikolov (2003): "… economic theory suggests…that potential output corresponds to the output level that would prevail in the absence of nominal wage or price rigidity."

6 This is the answer one is likely to get from students in an examination. Of course, it is substantially flawed (like saying GDP is what is published by the Statistical Office), but in the end it is not too stupid, because this is how it is generally done.

7 In this trivial case the output gap would vanish (ie equal 0 for all observations).

revisions as new data keep coming in and (2) systematically biased in periods of structural change, since the trend is ultimately identified ex post by past *and* future realisations. Regrettably, this is true for linear time-invariant filters and band pass filters alike.[8] Moreover, the only theoretical notion behind this black box approach is that potential GDP is evolving along a path that shows a considerable amount of inertia.

(3) The *functional approach*: Potential output is the level of output at any point in time that results in zero inflationary pressure. This is sometimes labelled NAILO ("non-accelerating-inflation rate of output")[9] and is conceptually related to – but not identical with – the NAIRU ("non-accelerating-inflation rate of unemployment"). The difference between the two is that the first is based on the existence of an equilibrium potential output path, while the latter postulates an equilibrium rate of unemployment, but to the degree that there is a close relation between output and employment (which there certainly is, since the level of employment is the major causal determinant of output, be it factual or potential), the distinction between the two gets academic rather than practical.

Note that this is an elegant approach to overcome the practical difficulties with the substantive notion of the output gap. If theory tells you that a positive (negative) output gap creates inflationary (deflationary) pressure and/or over-employment (underemployment) of the factors of production, why not use this theoretical link to identify the output gap inductively by looking at inflationary and/or factor market pressures?[10] To start with, find the points in time when inflationary pressure was zero – eg realised inflation (π) equalled expected inflation (π^e) – and/or the points in time when unemployment/capacity utilisation was equal to "equilibrium" – eg some longer term average of their past realisations –, and you have identified periods where "functional" potential output equalled observed output. Then, specify functional relationships between the

8 For an elaboration of this point, cf. van Norden (2002).

9 Cf. Hirose and Kamada (2003).

10 Cf. Laxton and Tetlow (1992) for the seminal contribution for this approach.

output gap and inflationary pressure and/or unemployment/ excess capacity utilisation. Finally, get data on your indicators and refer to the functional relationships to derive a quantitative measure of the output gap.

However, there are two caveats. Firstly, to incorporate additional indicators for strain on resources into GDP centred estimates of the output gap, they themselves have to be formulated in gaps.[11] In other words, to help gauge the "unobservable" potential output a range of other "unobservables", eg the NAIRU and/or "desired" or "equilibrium" capacity utilisation are referred to. Hence, the problem of not being able to measure potential output directly translates into the problem of quantifying the NAIRU[12] and/or "equilibrium" capacity utilisation. The improvement in the augmented output gap measure is therefore subject to the validity of the approaches to get estimates of the "second order" unobservables.

Secondly, potential output is now partly endogenised. Specifically, to the extent that the additional information dominates the output gap estimate, this approach reverses the theoretical relationship

$$\textit{output gap} \rightarrow \textit{inflationary pressure}$$

into an inductive measurement model

$$\textit{inflationary pressure} \rightarrow \textit{output gap},$$

thereby depriving the output gap concept of some of its original substantive content. With potential output being identified contingent on observed inflation and/or inflationary pressure, one can no longer claim that the correlation between such an output gap measure and observed inflation represents a structural relationship. It

11 Laxton and Tetlow (1992), Butler (1996). The Reserve Bank of New Zealand's approach follows the same logic, cf. Conway and Hunt (1997).

12 For a fundamental critique of the NAIRU cf. Hagger and Groenewold (2003). Evidence for the practical usefulness of a Phillips curve relationship to forecast inflation is mixed. For example, Gruen et al. (2002) report encouraging evidence from Australia, whereas Robinson et al. (2003) point to difficulties with real-time estimates, and Lansing (2002) argues that it is of little or no use for the USA.

is there by construction.[13] Hence, with a functional measurement approach, the output gap loses some of its original sense and should properly be regarded as an econometric indicator of inflationary pressure.[14]

The output gap in FPS

The Reserve Bank of New Zealand's MV filtered output gap belongs to the aforementioned third group of practical output gap measures that rely on the *functional* approach. To pin down potential output, it refers to the standard HP filter augmented with an inflation gap, an employment gap, and a capacity utilisation gap. In particular, the MV filter minimises the following expression:

$$\Lambda = \sum_{t=1}^{T}(Y_t - Y^*_t)^2 + \lambda\sum_{t=2}^{T-1}[(Y^*_{t+1} - Y^*_t) - (Y_t - Y_{t-1})]^2 + \sum_{t=1}^{T}\theta_t\varepsilon_{\pi,t}^2 + \sum_{t=1}^{T}\mu_t\varepsilon_{U,t}^2 + \sum_{t=1}^{T}\rho_t\varepsilon_{CU,t}^2 \qquad (2)$$

The first two summands represent the standard HP filter ($\lambda = 1600$). The additional information is incorporated by the last three summands, which are the residuals from a "Phillips curve" relationship

$$\pi_t = \pi_t^e + F(L)(Y_t - Y^*_t) + \varepsilon_{\pi,t} \qquad (3)$$

13 This circularity can be traced to the very origins of multivariate filtering; cf. Laxton and Tetlow (1992: i): "… if movements of potential output have a different effect on inflation than do cyclical movements in output, then information on inflation may be useful in identifying potential output."

14 An algebraic representation may help to clarify this point: When potential output Y^* is made endogenous on inflation π and a vector of indicators for strain on resources s, we get $Y^* = Y - f(\pi, s)$. Recall the substantive definition of the output gap (in absolute terms): $g^{abs} = Y - Y^*$. Accordingly, in the limiting case that potential output is derived exclusively by correcting observed output by the deviation of potential from observed output, the magnitude of which is estimated through the indicator model f, both potential and observed output are cancelled from the "gap": $Y^* = Y - f(\pi, s) \Rightarrow g^{abs} = Y - Y^* = f(\pi, s)$.

and the residuals from the two "Okun's Law" relationships

$$\textit{Unemployment gap} = G(Y_t - Y^*_t) + \varepsilon_{U,t} \qquad (4)$$

$$\textit{Capacity utilisation gap} = H(Y_t - Y^*_t) + \varepsilon_{CU,t} \qquad (5)$$

The weights placed on these relationships are θ, μ and ρ. These, the lag operator $F(L)$ as well as G and H are set exogenously to determine the mapping of these gaps into the output gap space. The MV filter's Phillips curve inflation gap is $\pi - \pi^e$, the employment gap refers to a NAIRU measured by an HP filtered unemployment series,[15] and the capacity utilisation gap is defined as the deviation of surveyed capacity utilisation to "equilibrium" capacity utilisation, which is exogenously fixed at 89 per cent. Furthermore, the MV filter is augmented by a stiffener that imposes an exogenously chosen potential GDP growth rate for the last three years.[16]

Accordingly, the MV filter is not only an approach to mitigate the *end point* problem; its *final* output gap values will diverge from the *final* HP filtered output gap when the additional information in the MV filter is giving a diverging signal.[17]

The twofold reason for augmenting the HP filter into an MV filter is spelled out in a number of Reserve Bank of New Zealand papers, beginning with Conway and Hunt (1997, p. 2), who state that in comparison to conventional measures of the output gap, the MV filter (1) "should provide a more reliable gauge of inflationary pressure in New Zealand" and (2) display "improved updating properties in that the measure is less prone to revision as new data

15 Since the latter suffers from an end point problem itself, a recent modification is to stabilise it by a survey measure of reported skill shortages. Note that this too is incorporated as a gap, so that we now face a "third order" unobservable problem.

16 For a detailed technical exposition of the MV filter, cf. Conway and Hunt (1997).

17 This can be observed for the New Zealand output gap in the late 1980s as well as for the last few years, for which the MV filter produces an output gap that is considerably different from the HP filtered estimate (see section 4.3). The divergence tends to be more pronounced for periods of significant structural change, where univariate GDP smoothing and inflation or indicators for inflationary stress are more likely to give different signals. For a further illustration of this, cf. Benes and N'Diaye (2004) on the Czech Republic.

becomes available", to Citu and Twaddle (2003, p. 10), who recapitulate that since the MV filtered output gap incorporates "indicators of resource strain" it should (1) be assumed to be a better indicator of inflationary pressure than less sophisticated measures and (2) "reduce the severity of the end-point problem because the information from these indicators does not tend to be revised".

3 The output gap in real time

The performance of the Reserve Bank of New Zealand's output gap measure has repeatedly been analysed and evaluated,[18] and it is fair to summarise this research as basically supportive. While measurement problems at the boundary of the observable data are acknowledged, a strong correlation of the current output gap with near-term future non-tradables inflation is identified as a particularly useful characteristic, since it supplies the policy maker with a theoretically sound *and* practically useful intermediate target.

At the same time, it is commonly argued that it gives better and more structured guidance to monetary policy than alternative approaches which eg rely on indicator-based estimates of future inflation, on economic growth, or on monetary targets. Some scepticism, though, has been expressed, pointing to model uncertainty and recommending that the output gap should not be accorded spurious accuracy and be supplemented by other evidence on the state of the economy, possibly from rival models.[19]

Recently however, a new literature has emerged which expresses serious doubts as to whether the output gap is a *practically useful* concept. A tough challenge to the output gap is expressed by Orphanides and van Norden (2002, 2003), who, in a nutshell, argue that while the output gap might be a useful concept for theoretical thinking about inflationary pressures, and while in addition to this, this usefulness is empirically well-established *ex post*, its practical usefulness is severely impaired or even annihilated by the inherent difficulty to know with sufficient reliability the magnitude of the

18 For example Citu and Twaddle (2003), Claus et al. (2000), Gaiduch and Hunt (2002), McCaw and Ranchhod (2002), Razzak (2002), Twaddle (2002, 2003).

19 For example Claus et al (2000).

output gap at the time *when the policy maker needs to know* it, ie in *real time*. This, according to Orphanides and van Norden, is true for all currently applied GDP detrending methods used to come up with real-time output gap estimates. Specifically, the end point problem lies mainly in the fact that without knowledge of the future, it is impossible to distinguish between cycle and trend, so that when shifts of the latter are eventually discovered, prior estimates of potential output have to be revised. Moreover, Orphanides and van Norden's real-time simulations suggest that multivariate methods are no remedy to the end point problem: "Though the information from multivariate methods may be useful in principle, their added complexity introduces additional sources of parameter uncertainty and instability which may offset the potential improvement in real time." (2002, p 582).

This view is supported by empirical evidence presented in a couple of recent studies. Notably, Orphanides (2003, p 997) compares a reconstructed real-time output gap series for the US going back to 1951 with today's view and finds persistent underestimation through most of the period until the mid-eighties. In the mid-seventies, the misperception amounted to an incredible *ten percentage points* of potential output, which in a simulated real-time Taylor rule framework would suggest that the Fed's monetary policy during the "Great inflation" was by no means meant to be permissive. Similarly, Nelson and Nikolov (2003) reconstruct a real-time output gap series for the UK going back to 1965 and plug this into a standard monetary policy framework. They find that the Bank of England's failure to lean against inflation in the early 1970s can be attributed to a real-time perception of the output gap that was *seven percentage points* lower than what one would quantify it nowadays. Cayen and von Norden (2002) conduct similar analysis for Canada since 1981 and find revisions of up to *six percentage points* of potential GDP.[20] For Japan, Hirose and Kamada (2003) find that

20 While Cayen and van Norden evaluate a wide range of output gap estimation methodologies, they lamentably do not include the Bank of Canada's multivariate filter. They note (p. 58) that this would be "interesting". The reason for this omission is probably that the Bank of Canada's multivariate filter was only installed in the mid-nineties, and, whereas the other methodologies allow for "backcasts" to the beginning of the 1980s, the multivariate filter cannot easily be simulated. Note that the same is true for the Reserve Bank of New Zealand's MV filter.

since 1995 an output gap which is derived by an HP filter augmented with a Phillips curve relationship would have suffered revisions of the same magnitude. Finally, a recent analysis for Finland (Billmeier 2004) finds that out of nine output gap measures none would add significantly to a univariate autoregressive explanation of annual CPI inflation from 1980–2002 and attributes this to the fact that a "statistically satisfying measure of potential output" might not be feasible for a high volatility observed (yearly) output series like the Finnish one (p 27).

In other words, real-time uncertainty about the magnitude of the output gap is not merely a theoretical concern about problems with low pass filters. There is evidence that reliance on the output gap might be responsible for some of the gravest central bank mistakes of the last decades, when real-time output gap measures failed to take account of changes in the growth rate of potential output. These days, the ongoing discussion about the retarded effects of the IT revolution and the "jobless" recovery in the US points to the possibility of another major change in the growth rate of potential output.[21]

On the other hand, based on a simulated MV filtered real-time output gap series for Australia from 1971q4 to 2001q4, Gruen at al. (2002) report revisions below four percentage points of GDP.[22] Moreover, drawing on ex-post data for the Euro area from 1970q1 to 2000q4, Rünstler (2002) finds revisions to various real-time output gap estimates that do not exceed two percentage points of GDP throughout the 1980s and 1990s.[23] With respect to the Reserve Bank of New Zealand's macro model, the findings are even less

21 As Kahn and Rich (2003) have recently pointed out, accepting the "new economy" story and assuming a sustainable acceleration of potential output growth would significantly lower our present real-time estimates of the output gap which tend to attribute fast growth to cycle rather than trend.

22 Their estimation method corresponds to the Reserve Bank of New Zealand's MV filter with the HP filter and a time varying Phillips curve, but without the "stiffener", Okun's law and capacity utilisation relationship.

23 Apparently, apart from filtering technology, it also matters which series exactly is fed into the various filters. What would have resulted in tolerable (from a policy perspective) output gap revisions at one time and place might offer extremely misguiding policy advice at another time and place.

disturbing: the real-time misperception of the output gap has so far not exceeded *2.5 percentage points* of potential output.

However, FPS is in place since 1997 only, ie in a comparably stable economic environment, so that it has not been seriously tested so far. Moreover, though by far less spectacular than what is reported from the US, the UK and Canada, 2.5 percentage points is a considerable deviation from the ex-post estimate. Furthermore, McCaw and Ranchhod (2002, p. 13) estimate that, since FPS was installed, the "starting-point output gap estimation errors" account for "about half of the average CPI forecast bias".

As stated above, being aware of the end point problem, in FPS, the HP filter is modified by referring to conditioning real-time information vector plus the stiffener, so that these real-time estimates might be more reliable than those from a naïve HP filter. However, due to the short historical span since the implementation of FPS, the Orphanides critique has so far not been empirically applied to the output gap as projected by FPS.

While earlier assessments on pragmatic grounds –the lack of sufficiently long history of FPS to generate a documented output track – had to refer to *simulated* real-time data,[24] there is now a recorded track from FPS of nearly seven years, so that the opportunity arises to refer to the original projections rather than simulations. Accordingly, this paper follows a direct way and focuses on FPS's *de facto* real-time projections.

24 Gaiduch and Hunt (2000) report a standard error of 1.30 and a serial correlation of 0.97 for revisions to a simulated MV filtered output gap going back to 1971. However, no analysis with real-time data is attempted. The first and so far the only published comparison of the factual real-time MV filtered estimates with later vintages of the output gap that I am aware of can be found in McCaw and Ranchhod (2002, pp 13 ff). They decompose the revisions into "data" and "other" issues and find that data and GDP forecast revisions "have added unhelpful noise to our output gap starting-point estimates … they have not added bias" to the Reserve Bank of New Zealand's inflation forecasts.

In particular, we shall evaluate whether based on the historical track of FPS empirical support can be found for the following conjectures:

1 Real-time MV filtered output gap *estimates* (in quarter *t* for quarter *t*) are stable in the sense that they are reasonably close to the latest numbers produced by the same model.
2 The MV filter has mitigated the real-time problem, which is inherent to procedures relying on symmetric filters.
3 The real-time output gap is a useful indicator of inflation.

FPS has produced reasonably reliable forecasts of the output gap in the near future (at quarter *t* for quarter *t*+*x*).

Based on the empirical track of FPS so far some of the conjectures above are not supported by the available evidence.

4 Empirical evidence

4.1 Real-time versus ex-post output gap estimates

The first step of our analyses compares the historical record of documented real-time MV filtered output gap estimates (in quarter *t* for quarter *t*) with the latest numbers produced by the same model.

Before we proceed, some reflections on the appropriate reference series are in order. In FPS, historical output gaps are computed by sending the corresponding historical GDP and indicator series through the MV filter. The real-time estimate of the output gap for quarter *t* is the last one which is obtained by MV filtering. Given this "starting value", FPS produces its forecasts of the output gap in the future $t+x$ $(x \geq 1)$ through a production function based model. Furthermore, note that the resulting forecasts for realised output are not used to extend the MV filtered series.[25]

As one moves through time, this set-up implies two data related sources of revisions to the output gap. Firstly, what once was a

[25] This would require forecasts of the additional indicators as well. These, however, are not produced in the present forecasting environment.

monitoring quarter, later on no longer occupies a position at the very right margin, so that the filter has more information to separate trend from cycle or noise. This is the classical *end point problem*.

The second data related source for revisions to the output gap is revisions to the underlying data itself. Note that this refers mainly to the GDP series.[26] We can *a priori* assume that GDP data (be they estimates, provisional, or final official numbers) are improving as time goes by and are hence subject to revisions, since one can draw on a lager set of information and/or improved processing of the underlying data due to hindsight. Furthermore, the GDP series that are fed into the MV filter stem from two different sources. Given the publication lag of GDP data in New Zealand, at the time of running FPS for a given MPS round in quarter *t*, the latest published data is for quarter *t*–2. Thus, the last two GDP data of the series which is sent trough the MV filter are not from Statistics New Zealand, but Reserve Bank of New Zealand indicator based estimates for the "monitoring quarters" *t* and *t*–1.[27] Accordingly, each vintage of MV filtered GDP series undergoes a distinctive revision that is due to replacing the monitoring quarter GDP estimates by the first official numbers. This occurs exactly two quarters after the initial computation of the MV filtered output gap. If we now assume that the monitoring quarter forecasts are not only less reliable because they can refer to less information then subsequent estimates, but in addition to this, because Statistics New Zealand is better equipped to produce GDP estimates, this would correspond to a characteristic break in the MV filtered GDP series resulting in a deviation of the usual pattern of monotonically decreasing revisions to one with revisions peaking at the second round. Yet, even after that, idiosyncratic as well as conceptual revisions of the official New Zealand GDP numbers will continue to be a source of output gap revisions.

Revisions of quarterly GDP enter one by one into the real output path, and to the extent that potential output is affected less, this will

[26] The additional indicators are less prone to revisions, which is one of the reasons why they are referred to in the fist place.

[27] It is worth emphasising that in the end – referring to an unobservable aggregate – *all* GDP data are *estimates*. The difference between the monitoring quarters estimates and the "official" data is thus more a matter of degree rather than principle.

modify the HP filtered GDP component of the potential output series.

The literature argues that the instability of HP filter derived output caused by data revisions is minor in comparison to the end-point problem of the filter itself. Can we confirm this statement for New Zealand? To this end we have to construct two different real-time series: a real-time series *R* that results from a consecutive series of a MV filtered output end points with real-time data, and a "quasi real-time" series *Q* which repeats the same procedure, but refers to the final data (which, of course, were not at hand when the "real" real-time estimate was performed; hence: "quasi"). Now, let *F* stand for the last vintage of the HP filtered output gap and define total revision $T = F - R$ and $P = Q - R$. Referring to observations from 1985q1–2004q1 and regressing *T* on *P* we obtain

$$\underset{(-1.38)}{T = -0.26} + \underset{(2.38)}{0.62\,P}\,,$$

with an insignificant intercept and a significant coefficient for the data revision driven output gap changes. However, the coefficient of determination ($R^2 = 0.07$) reveals that data revisions explain less than 7 per cent of the total revisions of the HP filter derived New Zealand output gap.[28] Since the MV filter is drawing on additional indicators that – unlike GDP – are not subject to revisions, this should correspond to the upper limit of this source of revisions.

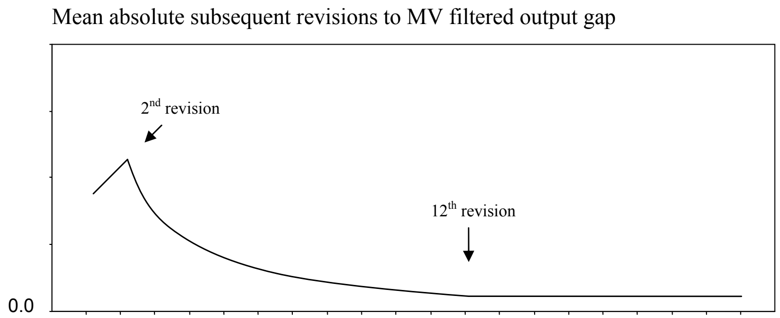
Other recent research on the Reserve Bank of New Zealand's macro model forecasting performance likewise confirms that for the MV filter, data revisions are the lesser of the "data related" problems. In particular, from a series of simulations run with the actual MV filter and different vintages of GDP data, Twaddle (2003, p. 6) concludes: "Practically all the 'data issues' related to output gap bias relates to not having the latest vintage of data rather than from poor monitoring quarter forecasts. If we replace the monitoring quarter

[28] To check for robustness, we re-ran this regression for a number of different samples. The results, however, remained qualitatively unchanged, and the coefficient of determination never exceeded 15 per cent.

GDP numbers with the first GDP outturns the bias is almost unchanged, though the standard deviation falls."

Given this, the reference series against which to evaluate the initial MV filtered output gaps should cover the domain of our last historical output gap series for which the classical asymmetric filter instability is no longer a (serious) problem. Now, focussing on the classical end point problem, how long should it take to see the output gap eventually to converge to a final number? From general experience with symmetric low pass filters applied to quarterly data, convergence should be roughly completed within three years, ie for the quarterly series at t+12. Moreover, this coincides with the MV filter domain to which the stiffener does not apply, so that changes in judgement on the sustainable growth rate of potential output do no longer affect the historical output gap estimates. According to these considerations, the stylised pattern output gap estimations for a given period through time should roughly resemble the representation given in figure 1.

Figure 1:
Expected impact of subsequent output gap revisions

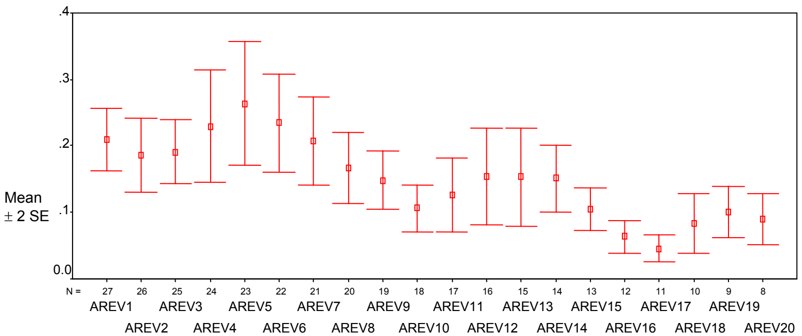


Let us now look at the historical data. Figure 2 plots the error bars (mean ± 2 standard deviations) for all observed revisions to output gaps referring to 1997q2 and later. In particular, AREV1 stands for all 27 hitherto recorded *first* revisions of the output gaps for 1997q2 to 2003q4, AREV2 for the 26 *second* revisions of the output gaps for 1997q2 to 2003q3, and so on until AREV20, which represents the eight recorded 20th revisions.

A comparison with figure 1 reveals that the empirical pattern is quite different from the stylised path. Firstly, there is no local maximum at AREV2; secondly, the revisions fail to converge to zero at about AREV12; and thirdly, they are not monotonically decreasing but rather following a cyclical pattern.

Figure 2:
Observed impact of subsequent output gap revisions

Mean absolute subsequent revisions to MV-filtered output gap



Percentage points of GDP, starting 1997q3, 1st to 20th revision

At this stage, we can conclude:

1. The substitution of official GDP data for RBNZ estimates, which goes along with the second output gap revision, has *not* been associated with remarkably large MV filter output gap revisions. In other words, the RBNZ monitoring quarter estimates have been reasonably precise on average in comparison to the subsequent first provisional official data.

2. The classical end point problem of not being able to separate trend from cycle in real time is reflected in a cyclical pattern of revisions. While this is due to the short sample period and might cancel out in data covering a large number of cycles, it reinforces the conjecture that output gap revisions are to a large extent a cyclical phenomenon.

3. Given the historical data, it is not safe to assume that revisions to our MV filtered output gap are negligible after 12 quarters.

For the following analyses, the third point poses a special problem. If the MV filtered output gap for a given quarter fails to converge to a constant within a finite period of time, we have to be aware of the fact that we do not possess a series of truly final numbers of historical output that we could compare with their real-time equivalents.

Let us now look at the paths of the *historical* output gaps for 1990q1 to 1996q4 as quantified ex-post by the MV filter from the time of its installation in 1997q2 to the present.

Figures 3.1 to 3.28 show how the MV filtered estimate for a given output gap – relating to the pre-FPS period, so that the end point problem should not be a concern, at least not for the earlier gaps – evolved through time.

This series of plots reveals that the MV filtered output gap for New Zealand seems to be an elusive phenomenon rather than a well quantifiable economic variable. Revisions keep occurring more than ten years after the quarter under consideration, with some time paths for a given output gap close to monotonically increasing/decreasing and others U- or S-shaped.

The two sources of output gap revisions that are responsible for the failure of the historical output gaps to converge to a final value, even after the classical end point problem is no longer a problem, are:

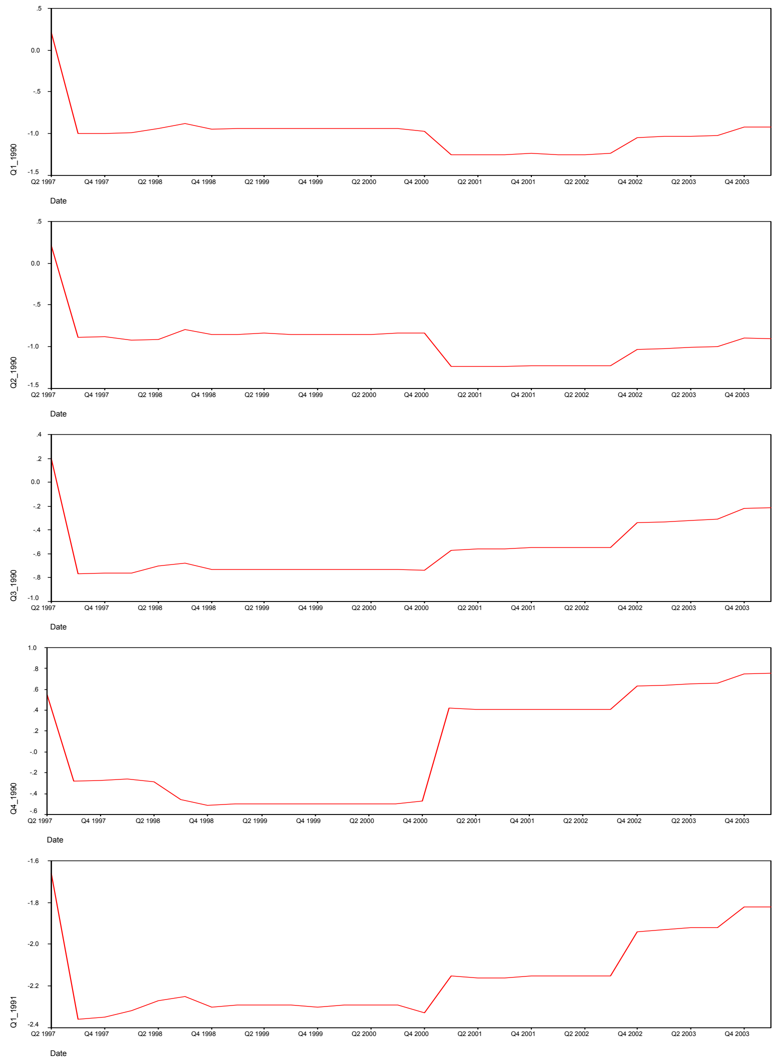
1. There are ongoing revisions of the official GDP data due to conceptual changes and rebasing of the data.

2. The definition of the MV filter changes over time.

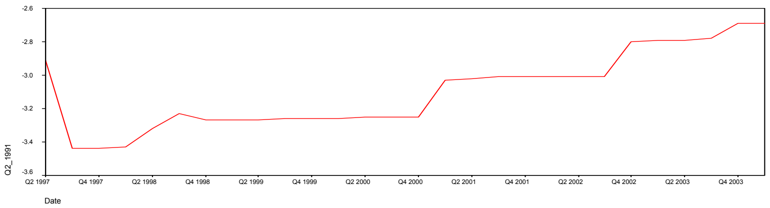
Hence, the MV filtered output gap does not converge to a final value not only because it is based on data that are subject to revisions, but because the procedure f that transforms the data input into an estimate of the output gap is itself a variable $f(t)$ that undergoes occasional changes. Thus, we cannot expect an output gap for a given period to ever converge to a final value, at least as long the changed formulas $f(t)$ are applied to the preceding periods.[29]

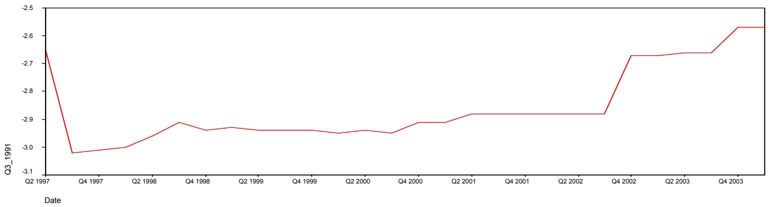
[29] Note that fixing previous gaps through not applying the current vintage of the MV filter is no viable remedy to this problem because it would create a worse one by destroying the internal consistency of any output gap time series.

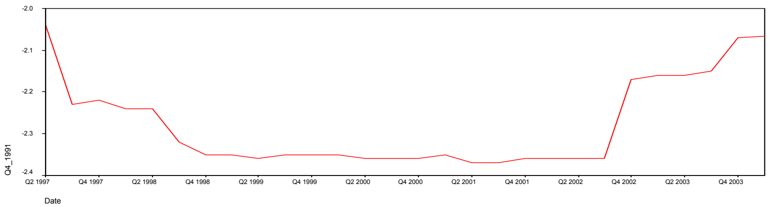
Figures 3.1–3.5: Stability of output gaps 1990q1–1991q1

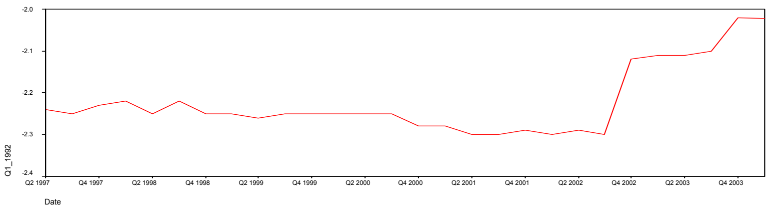


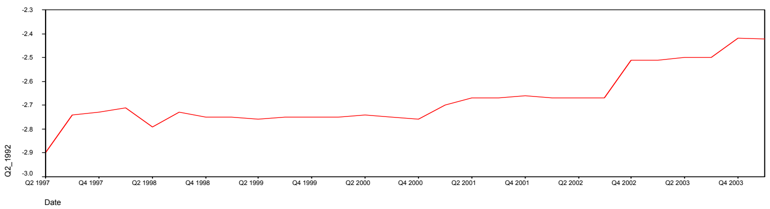
Figures 3.6–3.10:
Stability of output gaps 1991q2–1992q2



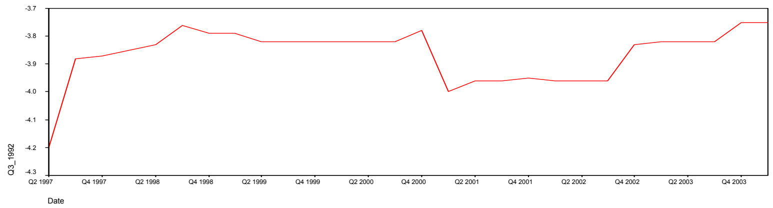


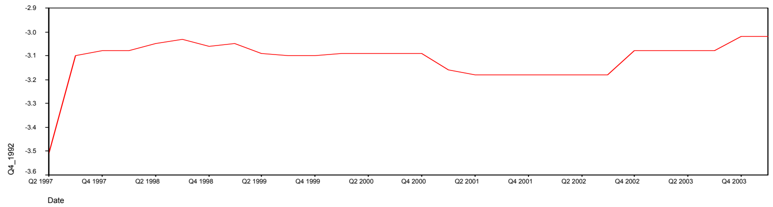


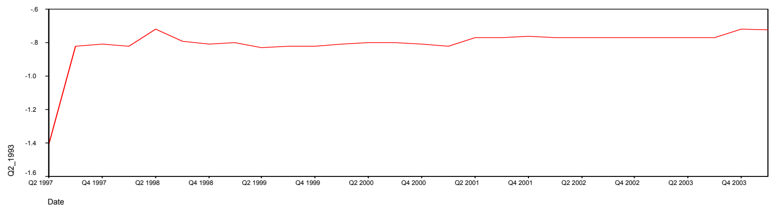


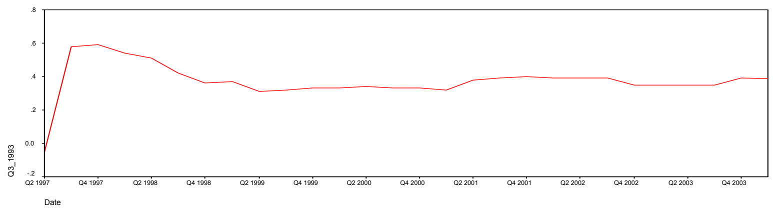


Figures 3.11–3.15:
Stability of output gaps 1992q3–1993q3 through time

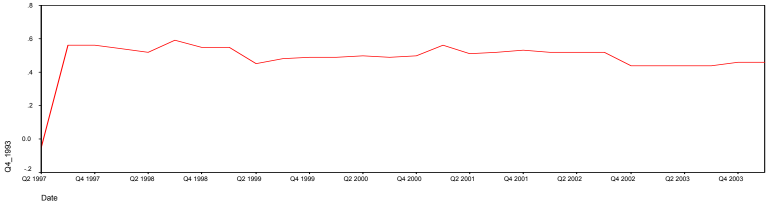


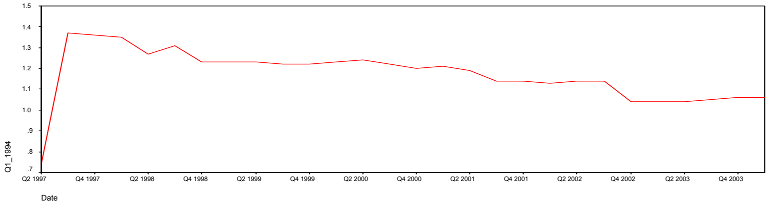


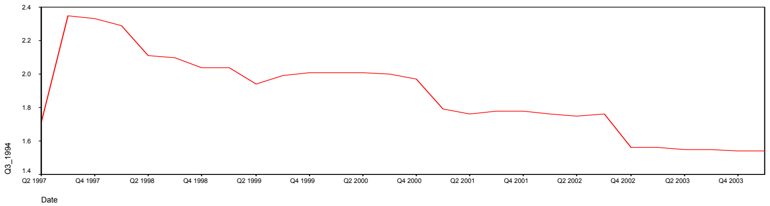




Figures 3.16–3.20:
Stability of output gaps 1993q4–1994q4







Figures 3.21–3.25:
Stability of output gaps 1995q1–1996q1

Figures 3.26–3.28:
Stability of output gaps 1996q2–1996q4

Consequently, we could stop here, claiming that where there is no real final series, it is futile to look for the reliability of a real-time series, since we do not have a yardstick for this kind of exercise. However, this would probably be overreacting, and in what follows, we shall pursue two alternative strategies to deal with this problem:

Firstly, we can relate the historical real-time output gap series to measures of inflationary pressure. Given that the MV filter produces an output gap measure belonging to the functional class, we have to be aware of some degree of circularity, but with this in mind, we shall follow this approach later in this paper.

Alternatively, we can pragmatically disregard the fact that substantial revisions keep occurring well beyond the twelfth revision and regard the ex-post series after an arbitrary number *r* of revisions as a sufficiently close approximation to the true numbers.

We shall start with this second approach. In particular, we refer to the above mentioned rule of thumb and let $r = 12$ quarters, so that with the latest data currently referring to 2004q1 and the twelfth revision referring to 2001q1 we can draw on a (pseudo-) final reference series until 2000q4. With the real-time MV filtered data starting in 1997q2 we thus have 15 data points for real-time versus ex-post comparisons. This is admittedly an uncomfortably short reference series. On the other hand, not much would be gained in terms of representativeness of the statistical basis if we added a few more quarters,[30] we stick to what we have and defer the comparison of longer de facto real time versus final output gaps to the future.

Figure 4 shows the MV filtered real-time output gap and its last vintage from the March 2004 model run. Obviously, the real-time and the ex-post series are correlated, and – with the notable exception of the intervals 1999q2–2000q2 and 2001q4–2002q4 – there are no strong signs of a phase shift. While to the right of the vertical line, which marks the end of what we regard as the reference series, the lines move closer together than in the left half of the graph, the period until the end of 2000 still sees the two series performing an obvious co-movement ($R^2 = 0.54$).

Averaging forecast errors rather than *absolute* forecast errors shows that the forecast bias for 1997q2–2000q4 has been negative with a mean upward revision of 0.76 percentage points. In other words, in real time, the MV filtered output gap signalled far less inflationary pressure than ex post.[31]

However, 0.54 is not a strong coefficient of determination, and in levels, the historical MV filtered output gap was uncomfortably inaccurate (see figure 5). At the beginning of the reference period the real-time output gap was below zero, while ex post it turned out

---

[30] For this, one would want to have time series that cover a fair number of cycles, ie some *decades* rather than years. Yet, as is demonstrated by the wide range of international findings on output gap revisions for comparable detrending techniques applied to different data sets, even very long time series are specific, so that caveats about generalisations pertain.

[31] Obviously, notwithstanding the real-time output gap bias, inflation in New Zealand did *not* get out of control. The reader may speculate whether this is because the policy maker was careful not to rely too much on the output gap, due to offsetting factors, or whether the output gap didn't matter after all.

to be close to two percentage points of potential GDP. Some quarters later, it underscored again (by more than one percentage point), and toward the end of 1999 the difference was again more than two percentage points. The average absolute real-time error as compared to the (pseudo-) final output gap for the 15 quarters from 1997q2–2000q4 amounted to 0.93 percentage points; the maximum was 2.2 percentage points. Though low in comparison with what has been found for other countries, these are still considerable magnitudes which should make us cautious about the uncertainty of our real-time output gaps.

Figure 4:
Ex-post vs. real-time output gap, 1997q2–2004q1

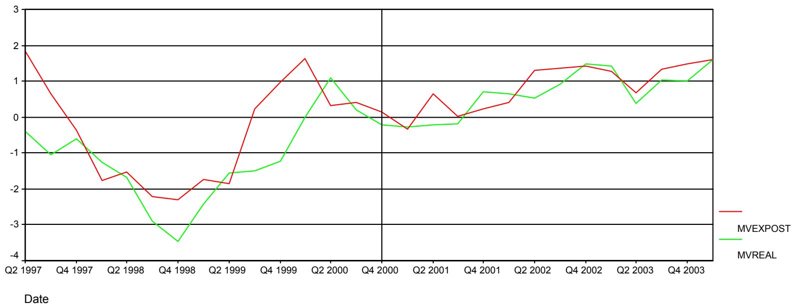
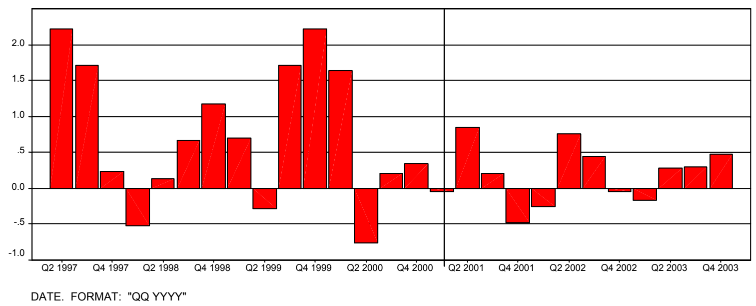


Figure 5:
Revisions of real-time output gaps (1997q2–2003q4)



4.2 The MV filter and the end point problem

The end point problems of low-pass filters are well known, and various remedies are suggested to deal with the resulting uncertainty with respect to real-time estimates of the output gap.

A common approach is to extend the underlying series with a couple of "future" values. The techniques encompass anything from purely statistical ARIMA forecasts[32] to expert "guesstimates", to full blown model based forecasts. With the additional data points thus obtained, the last factual real-time observation will no longer constitute the last data point, so that a symmetrical filter can be applied. However, due to the fact that the future is ultimately unknown, the end point problem is not solved, but rather transformed from a filter problem into a forecasting problem. Accordingly, the usefulness of this remedy depends on the accuracy of the forecasts that are fed into the filter. Given our limited ability to produces reliable forecasts of GDP, not too much might be expected from this method.

Other approaches within the HP filter framework are stiffeners that lower the sensitivity of the trend estimate with respect to the last data points. This prevents the filter from being too reactive to the last GDP data and hence boosts the estimated cycle in real time.[33]

Multivariate filters broaden the HP filter's minimisation problem in additional gaps, apart from the GDP gap. This addresses the HP filter's data related end point problem in two ways. Firstly, the additional indicators can be chosen to be less prone to data revisions than GDP. Secondly, the additional information can indicate how far actual output is deviating from the NAILO, thereby contributing to pin down a data based estimate of potential output at the margin of the GDP series, where the naïve HP filter is wanting.

The MV filter operates exactly along these lines. It augments the HP filter with an inflation gap, an employment gap, and a capacity utilisation gap, and modifies its last twelve data points with a stiffener. Accordingly, we might expect the MV filter to improve the

[32] Cf. Kaiser and Maravall (2001).

[33] An elaborate variant of this approach is explored in depth by Bruchez (2003).

correlation between real-time and final output gap estimates as compared to the un-augmented HP filters' performance. Previous analyses within the bank (notably Twaddle 2002, 2003) with simulated data for the 1990s come to the conclusion that to some extent this indeed appears to be true, though the superior revision properties of the MV filter are probably overstated because of the difficulties in simulating real-time estimates which do not refer to knowledge and experience which was not available in state space.

Being able to refer to a longer history of factual real-time estimates, our approach does not attempt to simulate real-time MV filter output gaps. Instead, it gives an exact, albeit short record of the Bank's factual real-time performance in estimating the output gap. The benchmark by which to evaluate the success of the MV filter in this respect is to compare the difference between the factual MV real-time estimates and the final MV filtered output gap series on the one hand with the difference between the real-time HP derived series and the corresponding final HP filter derived output gap series. This ensures an exact comparison of the revision properties. In particular, for this analysis, we need not be concerned with the question of which final reference series (MV versus HP) eventually comes closer to the "true" output gap, but with the amount of their respective revisions only. At the same time, we can test the extent to which, for the given time period, the revision properties of the real-time HP filtered output gap improve if we resort to the standard technique of extending the HP filtered series at the right margin by a number of forecasts. For the latter, we can refer to consistent data, namely the GDP forecasts to the right of the "monitoring quarters" from the corresponding RBNZ macro model runs. Hence, we refer to the following output gaps:

- an HP filtered real-time series based on the same GDP data that were underlying the Bank's real-time MV filter output gap estimates (ie with RBNZ estimates for the two last data points) and repeated filtering up to the quarter for which a real-time estimate is to be inferred (filtering from 1985q1);
- an HP filtered ex-post series, based on the GDP series from the March 2004 MPS round (1985q1–2004q1);

- an extended HP filtered real-time series based on the same GDP data that were underlying the Bank's real-time MV estimates plus twelve future GDP data points which are the forecasts from the corresponding model run, and repeated filtering up to the quarter for which a real-time estimate is to be inferred (filtering from 1985q1);
- an HP filtered ex-post series, based on the GDP series from the March 2004 MPS round including forecast for the next twelve quarters (1985q1–2007q1);
- the series of real-time MV filtered output gap estimates produced and finalised from the MPS June 1997 to the March 2004 forecast rounds, referring to the quarter that is about to come to an end (the second "monitoring quarter");

the MV filtered ex-post output gap series from the March 2004 MPS, where, as discussed above, the data points until 2000q4 (2001q2) will be denoted as the "reference" series.

Figures 6 to 8 show the record of real-time versus ex-post output gaps resulting from the naïve HP filter, the extended HP filter and the MV filter.

Would the problem of lacking reliability of real-time estimates of the output gap, which is inherent to all estimation procedures relying on symmetric filters and appears clearly in figure 6, have been alleviated by extending the HP filtered GDP series by twelve data points of the corresponding RBNZ forecasts? Has it successfully been mitigated by the MV filter? The graphs are not too conclusive and suggest considerable real-time uncertainty for all three output gaps.

To assess the comparative goodness of fit of the two real-time versus ex-post estimates, we thus refer to a simple and transparent numerical measure, the mean absolute real-time error (MAR), which is defined as $(1/t) * \Sigma|P_t - F_t|$, where P_t and F_t denote provisional and final values, respectively. In our context, this measure indicates how misleading (in absolute terms) the respective real-time estimates have been on average, under the assumption that the final series are the best proxy of the “true” number that the policy maker would have wanted to know. The results are given in table 1.

Figure 6:
Real-time output gap revisions: HP filtered

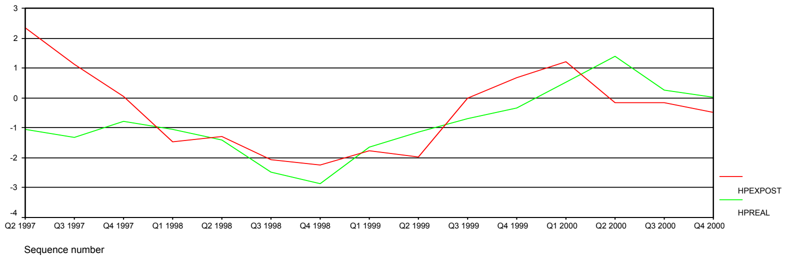


Figure 7:
Real-time output gap revisions: extended HP filtering

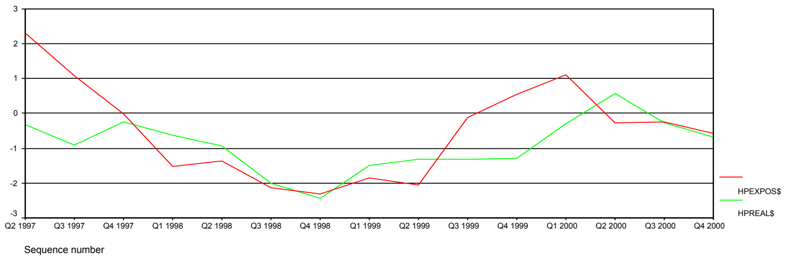


Figure 8:
Real-time output gap revisions: MV filtered

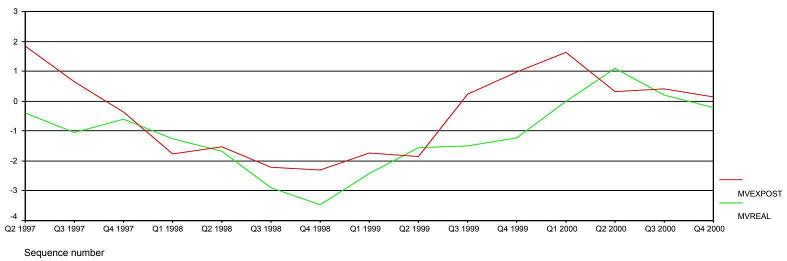


Table 1:
MAR of HP and MV filtered output gaps

	HP filter	extended HP filter	MV filter
1997q2–2000q4 (n = 15)	0.93	0.86	0.97
1997q2–2003q4 (n = 27)	0.69	0.66	0.70

While the HP filtered real-time output gap is stabilised by filtering a longer series which includes twelve GDP forecasts at the right boundary, the MV filter has not mitigated the end point problem; the subsequent output gap revisions during this reference period have been slightly *larger* than if the standard HP filter had been applied to the same data.

As a test of robustness of this unexpected finding, we extend the comparison to the last recorded revision in 2003q4. While this comparison must be regarded as preliminary, since the later ex-post data points are still close to the right margin and therefore prone to relatively large revisions, it should be informative as to whether the failure of the MV filter to produce less pronounced revisions in absolute terms is specific to our short reference period. The result is given in the second row of table 1, which confirms the same rank order with respect to revision properties as for the conservative reference period: The HP filter’s real-time revision properties are improved by referring to an extended forecasted GDP series, but not by the additions brought about by the MV filter.[34]

[34] That the mean absolute revisions during the extended comparison are lower for all filters is not evidence of any substantial improvement of either filtering technique, but simply due to the fact that the ex-post data points to the right have undergone fewer revisions, so that there are still errors in the pipeline.

This is an irritating result: Since it was introduced, the MV filter has not delivered the expected result.[35] Its revision properties should (but do not) outperform the HP filter, since it is bringing into the output gap estimation process variables that are related to the NAILO output gap concept, and theses are variables that do not get revised.

A possible explanation for this might be that the MV filter gains get swamped by problems in estimating the output gap when one does not have official GDP data to work with – estimates of GDP could be sufficiently noisy to do this. Is this is what we observe in the (relatively short) history of the MV filter in New Zealand? To check for this, we can look at the revisions induced by the *third* output gap estimate for a given quarter across the various filters. Though this is not a real-time estimate in chronological time, but rather an historical estimate referring to two quarters ago, it is the first that draws on official data rather than on estimates.

The results are given in table 2, which shows that the results from table 1 are essentially reproduced. The MV filtered output gap estimates for t–1 after the first revision as well as for t–2 after the second revision – ie the first estimate to draw on official data – have on average experienced the largest absolute revisions towards the March 2004 output gap estimate, while extended HP filtering would have yielded the best results. It is only after the fourth consecutive model run that the MV filtered output gap estimates has experienced less further revisions than the two HP filtered variants.

In other words, we can find vintages for which the remaining revisions towards the present "final" estimate are favouring the MV filter, but these do not coincide with the shift from the Reserve Bank's GDP estimates to the first official numbers but occur afterwards, and they are remote from being *real time* output gaps.

[35] This finding seemingly contradicts Twaddle's (2002) paper, which brings forward evidence that the MV filter is superior to the HP filter in reducing revisions when applied to historical data from New Zealand covering 1990–2001. However, Twaddle's analyses do not refer to the same data as and address a different question than this paper, and he is careful to emphasise that his simulated real-time data inevitably refer to information that was not available in real time and therefore tend to let the MV filter perform better than it would have done had it really been used in real time.

As the differences between the revisions due to the various filters are not very pronounced, we do not argue that there is a strong case to prefer the HP filter. Yet, the same applies vice versa. The MV filter may not have fared *worse*, but it certainly has not been an effective remedy to the end point problem, at least during the short period for which we have historical data. Applied to the data from this period, all filters considered here are found wanting.[36]

Table 2:
MAR of past output gaps, 1997q2–2000q4 (*n* = 15)

	HP filter	Extended HP filter	MV filter
t = q–1	0.78	0.73	0.79
t = q–2	0.57	0.54	0.56
t = q–3	0.65	0.67	0.59
t = q–4	0.65	0.64	0.49

Let us reflect on the implications: If the MV filter's raison d'être were to reduce the real time uncertainty of the end points of the HP filtered output gap, judged on its performance since its implementation, we have found no *empirical* reason to prefer it – at least not for purposes that rely on timely information, ie in quarter t on quarters t, t–1 and t–2.

[36] Note that the failure of the MV filter to reduce the end point problem in real time is not unique to the analysed time window in New Zealand; the same has been observed elsewhere: St-Amant and van Norden (1997: 55) state that "[T]he ability of the EMVF, the filter used to estimate the output gap in the Bank of Canada's staff projections of the Canadian economy to isolate business-cycle frequencies is worse than that of the HP filter … These problems seem to be the result of features introduced to improve the filter's estimate of current and recent output gaps." In this context, note that the Reserve Bank of New Zealand's MV filter is in essence modelled after the Bank of Canada's EMV filter. St-Amant and van Norden conjecture that this failure of "hybrid" methods to assure superior revision properties is due to the additional uncertainty which is introduced together with the additional information they are referring to.

This assessment might however be modified if the revisions from the real-time to the ex-post data were a minor concern compared to the correlation of the real-time data with what we are – at least from a Central Bank's perspective– ultimately interested in, namely inflationary pressure. Since there is no undisputable reference series for the output gap (potential output being inherently unobservable), the difference of real-time versus ex-post output gap estimates, interesting as they may be, may practically be of minor importance.

The crucial test for the usefulness of the output gap is then whether it helps predict inflation in real time. We turn to this issue now.

4.3 The output gap as an indicator of inflation

Following Orphanides and van Norden (2003, p 2), we shall distinguish between "suggested usefulness" of the output gap (which is implied by the established correlation between ex-post output gap measures and inflation) and "operational usefulness", which is based on output gap estimates in real time. According to the research undertaken by these authors, looking at the ex-post data will severely overstate the operational usefulness of the output gap. Moreover, when running simulations with various output gap estimation techniques, they find that some models perform well under certain circumstances, while others will be superior when feeding them with data from a different regime, which adds further uncertainty. Furthermore, the capacity of real-time output gap estimates to gauge inflationary pressure usually does not outperform simple bivariate models that refer to past inflation and output growth only.[37]

Turning to New Zealand, Claus et al (2000, pp 38 ff) compare the CPI inflation prediction potential of different ex-post output gap measures (HP filtered, MV filtered, SVAR, unobserved components) from 1997q1 to 1999q3. Since what their analysis addresses is "suggested usefulness", they rely on ex-post data and include the latest vintage of the MV filtered output gap. Applying a variety of simulated forecast models (Phillips curve, VAR, SVAR) they show that the considered output gap measures indeed convey information about the direction of the CPI one quarter ahead, for which they can predict the direction of inflation changes for roughly two thirds of the sample quarters. However, further ahead, the forecasting ability quickly deteriorates and converges to 50 per cent, which is what you would get with a random guess. Moreover, the MV filtered output gap outperforms the SVAR and unobserved components gaps on different measures of forecasting accuracy, but not the simple HP filtered gap.[38] Razzak (2002) finds that forecasting New Zealand's inflation after 1992 (ie since disinflation was accomplished) is a tricky matter, no matter what approach is used. In particular, neither the output gap nor other methods proved to add significant forecasting power to a benchmark model with only past inflation and inflation expectations as explanatory variables.

However, the usual benchmark models rely on autocorrelation within the inflation series, which implies a phase shift that makes these models inappropriate for forecasting purposes, where you would rather resort to a freezing the last observed value (the so "cold-deck" estimate), which is certainly no option for a forward looking monetary policy setup. Hence, we take a new look at the data and assess the operational usefulness of the output gap with respect to its forecast reliability rather than the performance of practically useless benchmark models.

Our starting point is the correlation between the MV filtered output gap and median non-tradables CPI inflation two quarters ahead.[39]

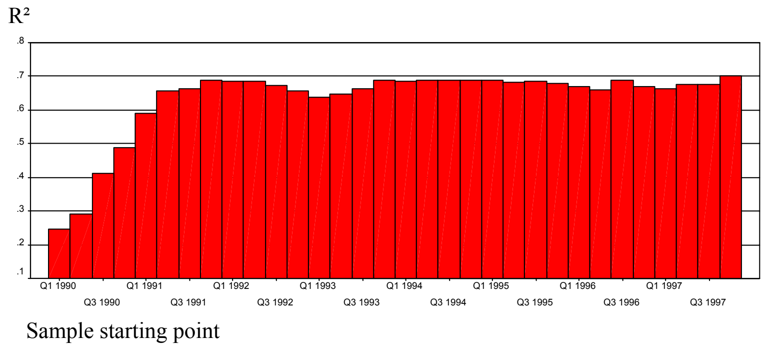
[37] Cf. Rünstler (2002) for an assessment of the "operational usefulness" of various output gap measures for the Euro area. His results are somewhat more favourable, suggesting that during the last two decades the output gap-inflation nexus might have been more stable in Europe than elsewhere.

[38] An earlier analysis with data for Australia (de Brouwer 1998) shows that from 1980q1 to 1997q4 – ex post, and along with other gap measures – the RBA's "MV HP output gap" (which is referring to the same filtering technology as the Reserve Bank of New Zealand's MV filter) adds significantly to forecast of changes of inflation in an error correction model that draws on a range of additional indicators of inflation.

[39] Other studies – eg Billmeier (2004), Cayen and von Norden (2002), de Brouwer (1998), Orphanides and van Norden (2993), Razzak (1997), Robinson et al. (2003), Rünstler (2003) – tend to prefer the correlation of the output gap with the GDP deflator, headline inflation or overall CPI inflation rather than non-tradables CPI inflation as the yardstick for real-time output gap estimates. Yet, the short term dynamics of inflationary pressure might affect the tradables sector more directly. As we shall demonstrate, this indeed is the case for New Zealand.

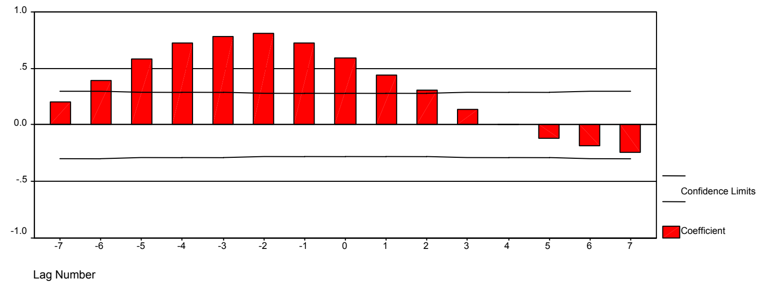
The reference series goes back to 1990, so that we can empirically determine the timing of this structural break with respect to the output gap-inflation nexus by repeatedly correlating the two series (the output gap advanced two quarters) with starting points moving from 1990 to the right. In figure 9, we plot the coefficients of determination resulting from the different sample sizes.

Figure 9:
Non-tradables π_{t+2} and g, R^2 for different samples



The structural break is quite obvious, and we can localise it at 1991q2. Dropping this observation lifts the R^2 to a level of 0.64 to 0.70 for all sample sizes. Consequently, we let the sample period start in 1991q2. Let us now look at the lead-lag structure. Figure 10 shows the cross correlation between the ex-post MV filtered output gap and median non-tradables CPI inflation.

Figure 10:
Lead-lag structure of non-tradables π_{t+2} and g



Obviously, as an indicator of inflation, the output has a lead of two quarters. Implying that the inflationary pressure indicated by the output gap needs some time to work itself through the economy, this lead is not implausible.

Figure 11 plots non-tradables CPI inflation along with the predicted series resulting from its linear regression on the ex-post MV filtered output gap series advanced two quarters:

$$\text{non-tradables } \pi_t = \underset{(26.1)}{2.75} + \underset{(9.90)}{0.66}\ g(\text{ex post})_{t-2}\,,$$

with the corresponding t-statistics in brackets and a coefficient of determination $R^2 = 0.66$. The "mechanical" interpretation of this regression is that since 1991q2, an ex-post output gap of zero would on average coincide with a non-tradables CPI inflation of roughly 2.75 per cent due to inertia, which might be attributed to inflation expectations.

Figure 11:
Non-tradables π_t and ex-post prediction by MV g_{t-2}

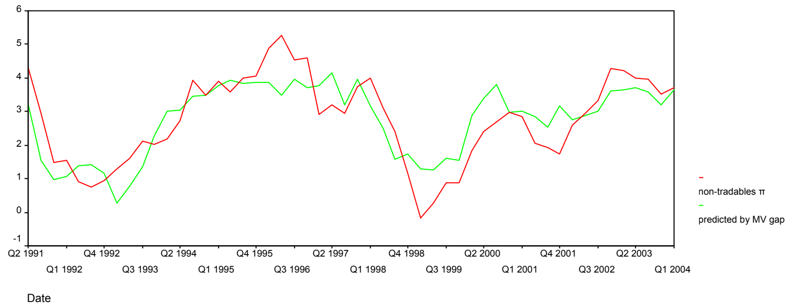
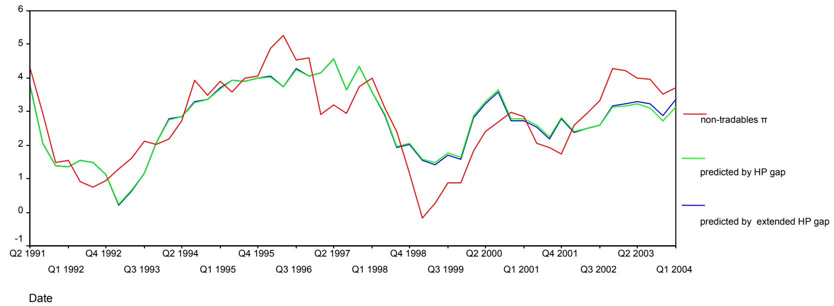


Figure 12 shows a roughly comparable prediction for the HP filter derived output gap inflation model. The coefficient of determination for this regression is 0.64.

Is median non-tradables CPI inflation the right reference series? Could it be that the output gap is performing better with traditional reference series such as headline inflation? Figure 13 shows that this is not the case. For 1991q2 to 2004q1, the output gap is not a useful indicator of future headline inflation; and this is true for all three filtering procedures considered in this paper (MV filter, HP filter, extended HP filter).[40]

[40] Razzak (1997) finds that from 1983q1 to 1995q3 the real time HP filtered output gap (the "filter") provides *better* CPI inflation forecasts than the corresponding ex-post series (the "smoother"). We can confirm that this holds until the end of 2000. Presuming that the output gap is driving inflation and that its ex-post estimates are generally more closely related to its "true" values than real-time estimates, this is a counter-intuitive result.

Figure 12:
Non-tradables π_t and ex-post prediction by HP g_{t-2}



We shall therefore stick to median non-tradables CPI inflation, where a significant correlation with the output gap is a proven fact. However, these result are so far based on ex-post data, ie on what we know now rather than on what we knew in real time. Let us therefore compare this with the predictive potential of the real-time output gaps. Given that the recorded real-time MV filtered output gap series commences in 1997q2, our comparison interval is reduced to 1997q4–2004q1. Hence we re-run the above regression for the shorter sample period with the ex-post versus the real-time output gap as alternative regressors. The result is plotted in figure 14.

Surprisingly, though the predicted paths are somewhat different, figure 14 does not suggest that either of them should be preferred. The same is true for the coefficients of determination: with the ex-post output gap, R^2 equals 0.71, whereas the real-time output gap yields an R^2 of 0.70, which is essentially equal. On the other hand, the HP filtered output gap for the same sample period produces an R^2 of only 0.47 in the real-time variant as compared to 0.64 ex post.

Figure 13:
Headline π_t and prediction by ex-post output gaps

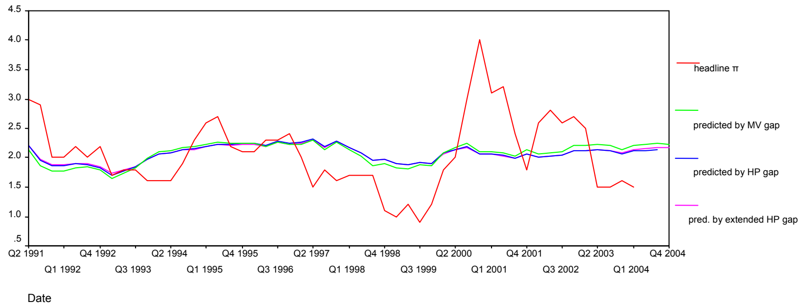


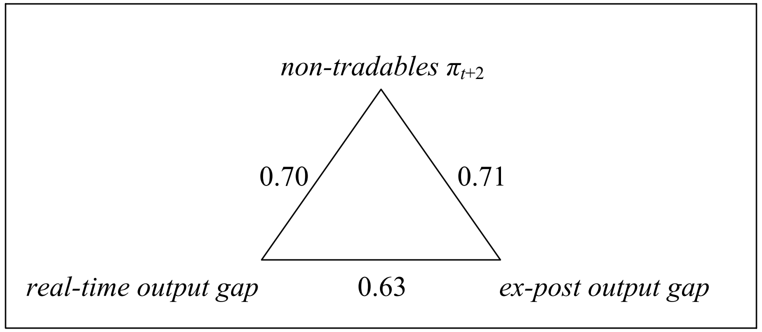
Figure 14:
Non-tradables π_t, prediction by g_{t-2}, ex post/real time

At first sight, this is an encouraging result. Though the real time MV filtered output gap is undergoing substantial revisions as time goes by, its very first estimate, where the end point problem is most severe, is correlated no less to the manifestation of inflationary pressure two quarters ahead – median non-tradables CPI inflation – than the (pseudo-) final ex-post series that can rely on hindsight to separate trend from cycle. Moreover, this finding cannot be

reproduced for the HP filter, where the real-time inflation prediction performance is markedly inferior.[41]

What do we take from this? Before we proceed, let us add one more bit to the picture: the correlation between the real-time and the ex-post output gap for the given sample period 1997q4–2003q1, ie the data range that underlies the inflation forecasts reported above, which amounts to $R^2 = 0.63$. Figure 15 summarises the results obtained so far.

Figure 15:
Correlations R^2 between π_{t+2} and g, (ex post/real time)



Interestingly, the correlation between the real-time output gap and future non-tradables CPI inflation is *higher* than the correlation between the real-time and the ex-post output gap. In other words, for the sample period under consideration, the real time MV filtered output gap is a *better* predictor of future median non-tradables inflation than of its own final estimate, and the final (ie March 2004) output gap estimate is *not* a better predictor of inflation than the very first real-time estimate.

41 We note in passing that the real-time versus ex-post performance of the output gaps derived from HP filtering a GDP series extended by twelve Reserve Bank forecasts is largely similar to that of the MV filter. The real-time series produces an inflation forecast with an R^2 of 0.68, which is quite similar to 0.67 ex post. (As a matter of fact, the real-time R^2 is even marginally higher.)

Hence, the MV filter is less a cure to the end point filter problem than a procedure that helps to bring about a reliable correlation between the real-time output gap estimate and inflationary pressure. However, let us recall that the MV filtered output gap – be it in real time or ex post – is contingent on several indicators of inflationary pressure, so that this correlation is there by construction.

Before we proceed, let us briefly reflect on terminology: the MV filter is *not* what the traditional usage of the term in time series analysis would suggest, namely a device that allows some ranges of frequencies through and not others. Conventionally, a filter is a univariate transformation of a series through a well-defined transfer function. The HP filter is a such a device, namely a low pass filter that allows low frequencies to pass and eliminates others. In the context of the output gap, the former are interpreted as potential output shocks (in the prevailing jargon: supply shocks) and the latter as transitory shocks (in the prevailing jargon: demand shocks).

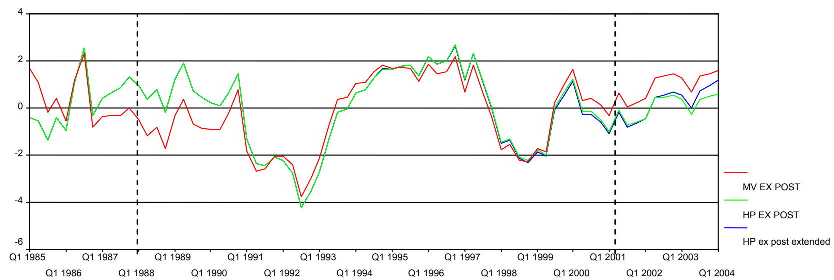
By contrast, the MV filter is a combination of the usual HP filter applied to the same GDP series and information from gaps signalled by additional indicators (labour market tightness, capacity utilisation and inflation, each expressed in deviations from their expected or equilibrium values) plus a stiffener that exerts a moderating influence on the last twelve quarters of GDP growth by incorporating the assumption that equilibrium growth equals an arbitrarily imposed value. This is a hybrid construction. The reason for insisting on this seemingly minor point of terminology is that the use of the term "filter" is disguising the fact that the MV filter is actually a multi-indicator model of inflation, so that the caveats about theoretical inference of causality from the functional approach apply. In other words, to the extent that the MV filter processes coincident information from indicators of inflationary pressure as well as inflation itself, the resulting potential output series is transformed into an endogenous variable that refers to inflation and inflationary pressure rather than being an independent estimate of productive capacity. To be sure, this procedure might be entirely sensible, but the more weight we put on the indicators, the less we can claim that we are really talking about an "output gap" (which

after all is a theoretically founded concept rather than a simple indicator approach to forecasting near-term inflation).[42]

Recall that the "true" value of potential output is inherently unobservable. Moreover, with the MV filter, we are producing a final potential output series that is not the same as trend output that emerges from a low pass filter.

Let us now look at how exactly the MV filter modifies the HP filtered output gap. Figure 16 plots the March 2004 MV filtered output gap along with the HP filtered output gap resulting from the same GDP data series, with filtering starting in 1985 and a range to 2004q1 and an extended filtering over twelve additional quarters of GDP forecasts from the Bank's March 2004 model run.

Figure 16:
MV versus HP filtered ex-post output gaps since 1985



Looking at figure 16, recall that the end point problem inherent to symmetric filters is a two-sided phenomenon and affects the left as well as the right margins of the series (though for policy purposes, the right boundary is what matters). Hence, the divergence between the output gaps for roughly the last *and* for the first three years must be considered preliminary in the sense that the series might undergo

42 Introducing the Bank of Canada's closely related EMV filter, Butler (1996, p. 53) is careful to remark that the resulting "output gap is best thought of not as an absolute concept, but as an indicator of excess demand or supply pressures … that is adapted to a particular model."

considerable revisions as one adds data referring to the future or to years before 1985. On the other hand, the central domain of figure 16, ranging from about 1988 to 2000 (delimited by the two dotted lines), can be considered as final, or at least very close to it.

Apparently, the difference between the naïve, univariate HP filtered and the more sophisticated series is negligible during most of the period. There are two exceptions to this general picture. Firstly, for 1985 the MV filter pushes the output gap up, but this is quickly reversed, and by 1987 we have entered a period of about four years during which the MV filter suggests considerably less strain on resources than does the HP filter. Since a good part of this divergence lies in the central part of the period, it is not due to a (left) end point problem, but rather to the information which is added by the MV filter. Secondly, beginning in the end of 1999, the MV filter causes an upward-shift of the output gap, which attains about one percentage point of GDP in 2001 and has stayed there ever since, thereby suggesting more inflationary pressure than the HP filter. In general terms, this confirms that the MV filter not only affects the last few observations. For certain periods, it produces *persistently* different output gaps.

Why is this? Since the HP GDP filter can be thought of as a special case of the MV filter with the weights put on the additional information set to zero, a divergence of the MV filtered output gap must result from a situation where the additional information is giving a different signal than the HP filter GDP decomposition. On the other hand, with positive weights on the additional information, similarity between the MV and HP filtered output gaps means that the additional information is in line with the GDP decomposition.

We can give a substantive ex-post explanation in growth terms for the first of the two persistent deviations between the two filters: As shown in figure 17, in the aftermath of the structural reforms, the late 1980s witnessed a period of decelerating GDP growth, which then recovered massively around 1993. This bulk of growth pushes the HP filtered trend up. With high unemployment and low capacity utilisation however, the MV filter signals less stain on resources and hence a lower output gap.

The present deviation of the MV filtered output gap from its HP filtered counterpart is not a growth story, it is a level story. As can be seen from figure 17, the growth rates of HP and MV filtered potential output have recently converged, but the difference between the output gaps persists. Figure 18 shows that this is a level phenomenon. Both HP and MV filters have potential output growing smoothly and with a similar rate, but the MV filter pulls the level of potential output down and prevents it from converging to the observed ("demand") output series, resulting in a persistently positive output gap.

Figure 17:
Growth rates of GDP and HP vs. MV filter potential

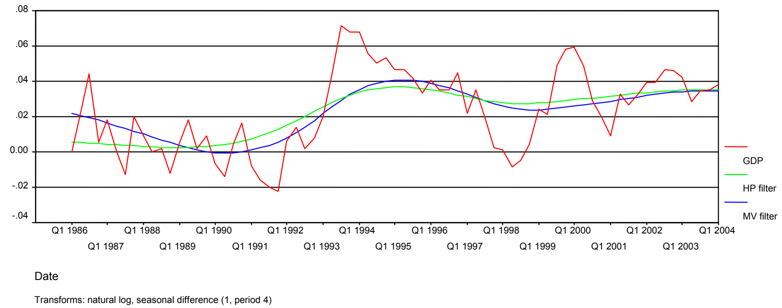
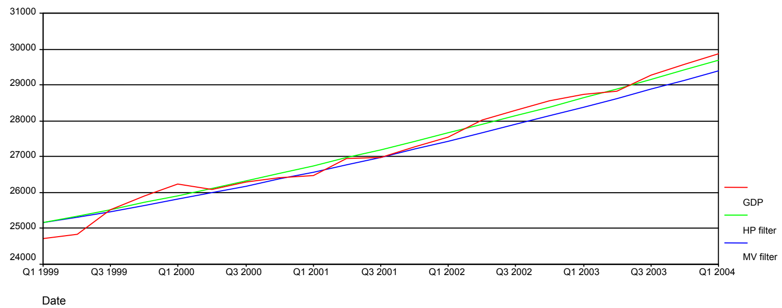


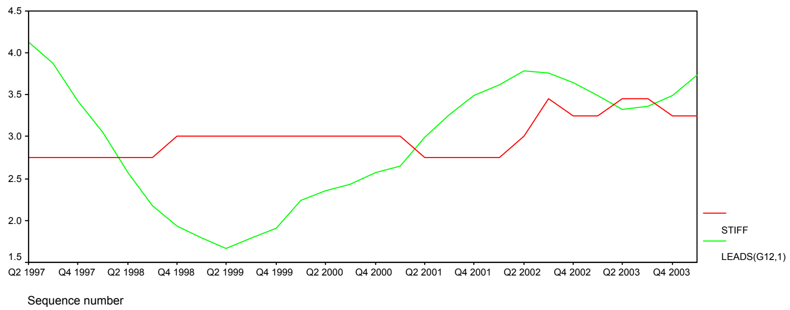
Figure 18:
GDP and HP vs. MV filter potential output



A minor part of this is presently done by the stiffener that imposes an exogenously chosen equilibrium (potential) GDP growth rate for the last three years. In 1997q2, the stiffener was originally set to 2.75 per cent. Since then, it has been changed on several occasions and is now at 3.25 per cent.[43]

Figure 19 plots the values of the stiffener from its implementation to the present, along with the twelve months prior moving average of annual real GDP growth. In 2001, prior average growth started to exceed the stiffener constraint, which – other things equal – pushes the output gap up, but in the middle of 2002, prior growth went back in line with how the stiffener would have it, and only recently do we see a new, albeit small, gap opening up.

Figure 19:
Stiffener values and the last 12 quarters GDP growth



As a result, with the present calibration of the MV filter, the stiffener now adds 0.17 percentage points to the HP filtered output gap, whereas capacity utilisation, unemployment and inflation, ie the indicator based additions to the HP filter, account for 0.91 percentage points.[44] Accordingly, the persistence of the present upward shift of the MV filtered output gap is mainly due to the additional indicators, which are signalling a continued strain on resources.

43 Moreover, the weight on the "stiffener has been increased twice, and since 1999q4 it is 35 times higher than in 1997q2.

44 These values are reported by Gray (2004).

Returning to figure 16, note that in 2002, the modified HP filtered gap that results from a GDP series that incorporates the latest GDP forecasts from the Bank's macro model starts to deviate from the simple HP filtered output gap and now occupies an intermediate position between the HP filtered and the MV filtered output gaps. How can we explain this pattern? Firstly, recall that by construction the additional data points only affect the right boundary of the HP filter, since before, there is no need to resort to forecasts. Secondly, the extended HP filtered series output gap is pushed up due to a future GDP track that starts with GDP presently well above potential (the MV filter provides the starting points) and therefore makes sure that near and medium term growth are coming down. This is *consistent*, though we cannot say that this adds further evidence that the MV filtered output gap is closer to the "true" value than the HP filtered counterpart.

A fundamental difference between the HP filter and the MV filter is that the first is a mechanical tool, and with wide acceptance of the smoothing parameter for quarterly data to be set at 1600, the output gap is determined. With the same GDP series to be filtered, everyone would come up with the same figures. However, this is not true for the MV filter, where judgement enters at various stages, so that with the same indicator set, one can compute any number of MV filtered output gaps. In particular, apart from the first decision to augment the HP filter with a specific vector of additional information, the following elements of the MV filter are based on judgement:

- The specification of expected inflation for the Phillips curve;
- the specification of the NAIRU for the employment gap;
- the specification of "equilibrium" capacity utilisation for capacity utilisation gap;
- the value of the "equilibrium" growth rate for the stiffener and the specification of what time window of observed (or forecasted) growth to compare this with;
- the lag operator $F(L)$, as well as the functions G and H to map the additional gaps into the output gap space;
- the weights θ, μ and ρ placed on the additional information.

Each of these points can at times have significant effects on the resulting output gap estimate, leaving some room for "tinkering around". The Reserve Bank has been careful to avoid this, and apart from a sequence of adjustments of "equilibrium" growth in the stiffener – which is a sensible adaptation to a changing environment – , there have been very few changes. After the MV filter was successfully installed, the only major revision was undertaken for the December 2002 model run.[45] This is reflected in figures 3.1 to 3.28, where we can observe a major level shift for most of the past output gaps occurring at 2002q4. (The other observable general level shift in 2001q1 is due to a major revision of the historical GDP data by Statistics New Zealand, and not by changes to the MV filter). Another way to visualise this is to compute the averages of the absolute impact of a given model run on all previous output gaps. The result of this exercise is plotted in figure 20, referring to output gaps from 2000q1 to 2003q4 and revisions from 1997q3 to 2004q1.

Figure 20:
Mean absolute impact on previous output gaps

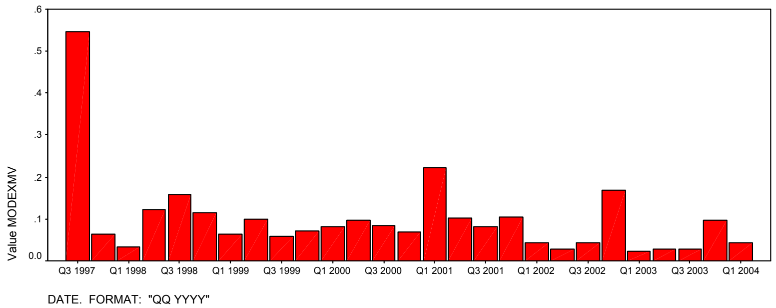


Figure 20 confirms what has been said so far. Apart from the large impact of the second model run in 1997q3, where major adaptations and learning are to be expected, there are only two model runs with exceptionally large impacts on previous output gaps: 2001q4 where the vintage of GDP available from Statistics New Zealand differed considerably from previous vintages due to methodical innovations and 2002q4 where the impact is due to changes to the MV filter.

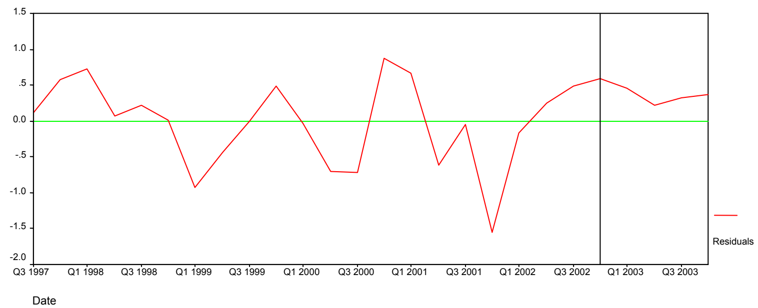
[45] The revision comprised new specifications of the NAIRU and π, alterations to the mapping functions as well as augmenting of selected weights.

Before we resume the discussion about the ontological status of the MV filtered output gap, let us see whether we can infer something like a "modelling reaction function" from the published record. The starting point is that in this case the "discretion" (m) to modify the MV filter is exerted in order to "optimise" its performance. Since the output gap itself is unobservable, in a central banking context, a modelling reaction function would relate to perceivable real-time mal-performance (here: failure of the unadjusted output gap g to reflect the output gap $g*$ that gives the "true" indication of inflationary pressure) and a recognition/action lag λ. Hence, in general terms, such a reaction function would be written as

$$m_{t+\lambda} = f(g - g^*)_t \,. \tag{6}$$

The implementation of this approach to the history of the factual MV filter derived output gap is straightforward. The event m that would have to be "explained" by the reaction function became effective in 2002q4. For the right hand side, let (g – g*) be represented by the output gap based inflation forecast error. Practically, starting with 1997q2 and moving step by step to the present, we consecutively regress the MV filtered real-time output gap series (advanced two quarters) on median non-tradables CPI inflation and save the residuals, which are identical to the state space forecast errors. In real time, there is a one quarter publication lag for inflation so that the forecast errors are shifted on the time axis to represent the real-time surprise. The result is plotted in figure 21.

Figure 21:
Real-time non-tradables π forecast residuals



Negative residuals imply that the real-time output gap signalled too much inflationary pressure, and positive residuals imply underestimation. The vertical line marks the quarter in which *m* became effective. Figure 21 supports the claim that the MV filter has not been hastily adapted to observable real-time mal-performance: before its major recalibration in 2002q4, there have been three episodes of considerable real-time underestimation of inflationary pressure and three episodes of overestimation (which might have been perceived as least severe). In other words, until 2002q4, the MV filter was operated with a "wait and see" philosophy. However, the 2002q4 recalibration coincides with another peak of real-time inflation underestimation, and though this is not the highest peak, it is the one that took the longest time to build up so that a careful observer would have had a chance to understand that something was going wrong. In other words, while there is evidence of caution not to change the basic calibration of the MV filter in every period of perceivably poor inflation forecasting ability (ie of an *ad hoc* fitting of the output gap to inflation), the one major recalibration occurred in such a situation.

Is this a possible source for concern? After all, in a central banking context, learning is highly desirable. Yet, in so far as the MV filter is adapted to maintain its correlation with median non-tradables inflation, this might impair its link to other variables in the system. If we want to see the output gap as a variable which is more than merely an indicator of inflationary pressure, we face problems in having its quantification moving around to maintain its indicator properties. The output gap is not identical with itself through time.

On the other hand, we could leave these concerns aside and take the MV filtered output gap as a peculiar but ultimately quite successful leading indicator of domestic inflationary pressure. Yet, in this case we should be sure that it is a superior indicator compared with other indicators that do not stick to the notion of "equilibrium" and "gaps". However, since this is a question that merits to be addressed in depth, we shall deal with it elsewhere.[46]

[46] Using a routine selection procedure to derive combined indicator series of economic processes (described in Graff/Etter 2004), a pre-study to this paper demonstrated that it is not difficult to construct a leading indicator of median non-tradables CPI inflation for a sample period covering the last ten years that is superior to the MV filtered output gap (ie having a longer lead and/or smaller forecast errors). However, this indicator was constructed ex post and had the advantage of hindsight, so that simulated real-time model building as well as out-of-sample tests have to be conducted to arrive at a more conclusive answer. For this, see Graff and Matheson (2004).

The final step of this paper will be to assess the reliability of our output gap *forecasts*. Though these are not produced by the MV filter, the real time MV filtered output gap provides an important input by providing the starting value for potential output. To the extent that this is flawed, the forecast accuracy will be impaired from the very beginning.

4.4 Has FPS produced reasonably reliable forecasts of the output gap in the near future?

For the purpose of this final step of the analyses reported in this paper, let us put aside the doubts concerning the empirical content of the MV filtered output gap and assume that – whatever it is that it exactly measures – it quantifies information that is crucial for the conduct of monetary policy. In other words, we take its validity as given. Its virtues in a forward looking monetary policy setup would then be founded on the ability of the macro model to produce reasonably reliable forecasts of the output gap in the future.

As above, the natural reference series for this exercise is the March 2004 ex post MV filtered output gap series itself, which we compare to the forecasts of this series for a successively higher number of quarters ahead. The data are from the factual projections that were documented following the MPS rounds since 1997q2. Table 2 shows four summary statistics for the goodness-of-fit for a forecast horizon of up to three quarters ahead.

For four quarters ahead the correlation between forecast and final series gets negative, which obviously does not make sense so that we have to disregard the possibility that FPS has a forecast horizon with respect to its own output gap of more than three quarters. Moreover, a meaningful and significant positive correlation between the real estimate and the final MV filter estimate has not been produced for more than one quarter ahead (though $R^2 = 0.41$ is not what one would call a close fit). For two quarters ahead, the coefficient of determination falls to 0.22 and fails to pass any conventional test of significance. For three quarters ahead, the correlation vanishes completely. In other words, judged on FPS's historical record of forecasting its own final output gap in real time, a forecast horizon of one quarter appears acceptable if one is inclined to base monetary policy recommendations on an ex-ante explanation of the final output gap in the range of 40 per cent (and a further 70 per cent correlation of the latter with non-tradables CPI inflation two more quarters ahead). For a forecast horizon of two quarters, there are still slight hints in real time of how the MV filter will eventually quantify the output gap, but after that the window into the future is obscured.

Table 2: Output gap forecasts vs March 2004 (1997q2–2000q4)

Projection	real-time	1 quarter ahead	2 quarters ahead	3 quarters ahead
R^2	0.54	0.41	0.22	0.001
p	0.002	0.014	0.11	0.91
MAF	0.97	0.99	1.10	1.28
max. AF	2.21	2.53	2.58	3.04
n	15	14	13	12

To be sure, FPS shares its inability to produce reasonably accurate forecasts for more than a very limited number of quarters with other medium-sized structural macro-models. Nevertheless, we should be aware of the fact that this also holds for our output gap forecasts.

5 Conclusion

This paper has taken a new look at the properties of our real-time estimates of the output gap in comparison to the final numbers. Though this question has been addressed before, our approach is markedly different. Instead of computing extended time series of *simulated* real-time data, which are always open to doubt, we have relied on the *factual* output of the MV filter since it was implemented as a regular tool in June 1997.

An elaboration of some methodological issues led us to the conclusion that the MV filter is a hybrid construct. The empirical reference to indicators of inflationary pressure distances it from the original concept of the output gap, where a deviation of observed from potential output is taken as a *cause* of inflationary pressure. To the extent that the MV filter is referring to indicators of inflationary pressure, the cause – in particular: the magnitude of the output gap – is inferred from the presumed results. With factual output given, this procedure endogenises potential output which thereby ceases to carry much substantive meaning in itself. Moreover, we found some indication that the major recalibration of the MV filter in 2002 helped to maintain the correlation with a target variable that it is supposed to "explain". This potentially amplifies the problems in attaching a substantive meaning to potential output.

There are three sources of revisions of the output gap: revisions of real GDP data, the end point problem of symmetric filters and changes to the calibration of the MV filter. During the recorded history of the MV filter, the revisions to real-time output gap have been no less than they would have been, had an HP filter been used (preferably applied to a GDP series that is extended by a number of forecasts). Moreover, the MV filter leads to permanently different levels of the output gap estimates if compared to a purely statistical trend.

The performance of the output gap with respect to signalling inflationary pressure, as measured by future median non-tradables (but not headline) CPI inflation has been reasonably good. Interestingly, this is true for both the real-time and the ex-post output gap, which means that we have not only found evidence for "suggested" but for "operational usefulness" of the MV filtered output gap. In particular, the real-time output gap has been better at predicting inflation than it has at predicting its own *ex post* realisations, which gives empirical support to our view that the MV filtered output gap can be thought of as a multi-variable indicator of inflationary pressure rather than as a true measure of potential output.

Finally, our window into the future with respect to the output gap is obscured after the second quarter. For three quarters ahead, the correlation of the real-time forecast with the values which by now can pragmatically be regarded as final is zero.

Given these findings, the doubts expressed in the emerging literature on the practical shortcomings the output gap to some extent also apply to our MV filtered output gap.

The analyses presented in this paper refer to data for New Zealand. Yet, the problems addressed are not specific to New Zealand, and the results fit well into the recent strand of empirical studies reporting problems in pinning down the output gap in real time. However, the finding that our MV filter produces a real-time output gap which is better at predicting inflationary pressure than at predicting its own final estimates, is not commonly reported in the literature. This is on the one hand reassuring with respect to the operational usefulness of the MV filtered output gap; on the other hand it is also disturbing with respect to its ontological status. Since the use of hybrid output filters (multivariate, Phillip curve relationship augmented, NAILO focussed etc.) is widely spread among central banks and other policy institutions, future research might usefully submit comparable data sets from other countries to the same type of analysis.

References

Benes, J and P N'Diaye (2004), "A Multivariate Filter for Measuring Potential Output and the NAIRU: Application to the Czech Republic", *IMF Working Paper* WP/04/45.

Billmeier, A (2004), "Measuring a Roller Coaster: Evidence on the Finnish Output Gap", *IMF Working Paper* WP/04/57.

Bruchez, P-A (2003), "A Modification of the HP Filter Aiming at Reducing the End-Point Bias", *Swiss Federal Finance Administration Working Paper* ÖT/2003/3.

Butler, L (1996), "A Semi-Structural Method to Estimate Potential Output: Combining Economic Theory with a Time-Series Filter", *Bank of Canada Technical Report* No 77.

Cayen, J-P and S van Norden (2002), "La fiabilité des estimations de l'écart de production au Canada", *Bank of Canada Working Paper* 2002-10.

Citu, F and J Twaddle (2003), "The output gap and its role in monetary policy decision-making", *Reserve Bank of New Zealand Bulletin,* 66 (1), 5–14.

Claus, I, P Conway and A Scott (2000), "The Output Gap: Measurement, Comparisons and Assessment", *Reserve Bank of New Zealand Research Paper* No 44, Wellington, June 2000.

Conway, P and B Hunt (1997), "Estimating Potential Output: a Semi-structural Approach", *Reserve Bank of New Zealand Discussion Paper* G97/9.

De Brouwer, G (1998), "Estimating Output Gaps", *Reserve Bank of Australia Research Discussion Paper* 9809.

Gaiduch, V and B Hunt (2002), "Inflation targeting under potential output uncertainty", *Reserve Bank of New Zealand Discussion Paper* DP2000/08.

Graff, M and R Etter (2004), "Coincident and Leading Indicators of Manufacturing Industry: Sales, Production, Orders and Inventories in Switzerland", *Journal of Business Cycle Measurement and Analysis*, 1, 109–131.

Graff, M and T Matheson (2004), "Leading indicators of inflation in New Zealand", Reserve Bank of New Zealand memo, forthcoming.

Gray, R (2004), "Conditioning Information and the MV Filtered Output Gap", Reserve Bank of New Zealand filenote.

Gruen, D, T Robinson and A Stone (2002), "Output Gaps in Real Time: Are They Reliable Enough to Use for Monetary Policy?" *Reserve Bank of Australia Research Discussion Paper* 2002-06.

Hagger, A J and N Groenewold (2003), "Time to Ditch the Natural Rate?" *Economic Record*, 79, 324–335.

Hirose, Y and K Kamada (2003), "A New Technique for Simultaneous Estimation of Potential Output and the Phillips Curve", *Bank of Japan Monetary and Economic Studies*, August 2003.

Kahn, J A and R Rich (2003), "Tracking the New Economy: Using Growth Theory to Detect Changes in Trend Productivity", *Federal Reserve Bank of New York Staff Reports* No 159.

Kaiser R and A Maravall (2001), Measuring Business Cycles in Economic Time Series, Springer, *New York*.

Lansing, K (2002), "Can the Phillip Curve Help Forecast Inflation?" *Federal Reserve Bank of San Francisco Economic Letter No* 2002-29.

Laxton, D and R Tetlow (1992), "A Simple Multivariate Filter for the Measurement of Potential Output", *Bank of Canada Technical Report* No 59.

McCaw, S and S Ranchhod (2002), "The Reserve Bank of New Zealand's forecasting performance", *Reserve Bank of New Zealand Bulletin,* 65 (4), 5–23.

Nelson, E and K Nikolov (2003), "UK Inflation in the 1970s and 1980s: the Role of Output Gap Mismeasurement", *Journal of Economics and Business*, 55 (4), 353–370.

Orphanides, A (2003), "Historical Monetary Policy Analysis and the Taylor Rule", *Journal of Monetary Economics*, 50 (5), 983–1022.

Orphanides, A and S van Norden (2002), "The Unreliability of Output-gap Estimates in Real Time", *Review of Economics and Statistics*, 84, 569–583.

Orphanides, A and S van Norden (2003), "The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time",www.hec.ca/pages/simon.vannorden/wps/RT2JMCB2.pdf.

Razzak, W (1997), "The Hodrick-Prescott technique: A Smoother versus a filter. An Application to New Zealand GDP", *Economics Letters*, 75, 163–168.

Razzak, W (2002), "Monetary policy and forecasting inflation with and without the output gap", Reserve Bank of New Zealand Discussion Paper 2002/03.

Robinson, T, A Stone and M van Zyl (2003), "The Real-Time Forecasting Performance of Phillips Curves" *Reserve Bank of Australia Research Discussion Paper* 2003-12.

Rünstler, G (2002), "The Information Content of Real-Time Output Gap Estimates: An Application to the Euro Area", *ECB Working Paper* No 182.

St-Amant, P and S van Norden (1997), "Measurement of the Output Gap: A Discussion of Recent Research at the Bank of Canada", *Bank of Canada Technical Report* No 79.

Twaddle, James (2002): "The MV filter and real-time data", *Reserve Bank of New Zealand memo.*

Twaddle, James (2003): "Evaluating starting point output gap estimate errors", www.rbnz.govt.nz/research/forecastper/0133055.pdf.

Van Norden, S (2002), "Filtering for Current Analysis", Bank *of Canada Working Paper* 2002-28.