

Measurement of the output gap: a discussion of recent research at the Bank of Canada

Pierre St-Amant and Simon van Norden¹

Introduction

Most macroeconomic models that are used for forecasting and policy analysis require an estimate of potential output. For example, at the Bank of Canada, estimates of potential output are important inputs in different "Phillips curve" models and in the staff's Quarterly Projection Model, where the gap between actual and potential output is a key variable determining the evolution of prices and wages. A level of real GDP above potential (a positive output gap) will often be seen as a source of inflationary pressures and a signal that monetary authorities interested in avoiding an acceleration of inflation should tighten monetary conditions. A level of real GDP below potential (a negative output gap) will have the opposite implication.

The output gap can thus be defined as the component of real output that is associated with changes in inflation.² Note that gaps could be calculated in markets other than that for goods and services. For example, gaps in the labour market have frequently been calculated and authors such as Hendry (1995) present "money gaps."

Unfortunately, measuring the output gap is not an easy task. Different sets of assumptions can be used together with different econometric techniques to provide different measures of the output gap. One common assumption is that the output gap is some part of the transitory (cyclical) component of real output. The methods discussed in this paper make that assumption.

The first group of methods we consider are those which simply use some (implicit or explicit) assumptions about the dynamics of real output to identify the output gap. For example, if one believed that real output was composed of a stationary component and a simple log-linear trend, the output gap could be measured as the residuals of a regression of log output on a linear time trend. Unfortunately, such a simple model does not adequately describe the behaviour of output, and measuring the temporary component in more complex models is problematic.

In this paper, we will assume that real output is $I(1)$; that is, that the level of output is subject to permanent shocks so there is no deterministic trend towards which output tends to revert.³

¹ The authors wish to thank Chantal Dupasquier, Paul Fenton, Gabriele Galati, Alain Guay, Seamus Hogan, Irene Ip, Robert LaFrance, René Lalonde, David Longworth, Tiff Macklem, John Murray and Brian O'Reilly for useful comments and discussions. They also thank Jennifer Page and Rebecca Szetto for their excellent research assistance. Of course, since the authors are solely responsible for the paper's content, none of the aforementioned are responsible for any remaining errors. The views expressed in this study are those of the authors and do not necessarily represent those of the Bank of Canada.

² To be more precise, we should take into account expected inflation, and therefore define the output gap with respect to changes in unexpected inflation. Some models also imply a relationship between the change in the gap and inflation.

³ This is the most common assumption in modern applied macroeconomics and is consistent with the view that real output can be permanently affected by shocks, such as technological innovations. An alternative view is that output is stationary around a time trend, but that this time trend is subject to occasional random changes in its slope and intercept. Evidence for such a view is discussed by Perron (1989) and Weber (1995). As detecting changes in the slope or intercept near the end of a sample is quite difficult, such models imply that one cannot reliably measure the current

Many approaches have been proposed to identify the permanent and cyclical components of real output in such models, such as those proposed by Hodrick and Prescott (1997), Watson (1986) or by Beveridge-Nelson (1981). The problem is that the measured cyclical component may differ considerably from method to method. Quah (1992) argues that this is an intrinsic problem and that "...without additional ad hoc restrictions those (univariate) characterizations are completely uninformative..."

These problems have not prevented the widespread use of Hodrick and Prescott's filter to identify the cyclical component of output.⁴ Arguments commonly made to justify its use are that:

- it extracts the relevant business-cycle frequencies of output;
- it closely approximates the cyclical component implied by reasonable time-series models of output.

We examine these arguments in Sections 1.2 and 1.3. We also note that, unlike much of the literature on "detrending", the focus of the problem confronting policymakers is to estimate the deviation from trend at the end rather than the middle of a data sample.⁵ We conclude that such methods are unlikely to be suitable for use in a policy context, and we discuss economic factors which limit our ability to estimate the current output gap.

An important class of alternatives to these univariate dynamic methods are those which combine their assumptions with information from assumed or "structural" relationships between the output gap and other economic variables, such as a Phillips Curve or Okun's Law. We examine some of these in Section 2. Among them are the multivariate HP filters (MHPF) proposed by Laxton and Tetlow (1992) and Butler (1996), which is the general approach currently used in the staff economic projection of the Canadian economy at the Bank of Canada. In Section 2, we note that calibration of the MHPF methods has been problematic and that despite the inclusion of structural information, their estimates of the output gap have wide confidence intervals. Spectral analysis of the Canadian output gap resulting from the application of the MHPH method also gives the "disturbing" result that it includes a very large proportion of cycles much longer than what is usually defined as being business cycles. A reaction to these methods is the "Trivial Optimal Filter that may be Useful" (TOFU) approach suggested by van Norden (1995), which replaces the HP smoothing problem with the simpler restriction of a constant linear filter. The TOFU approach has yet to be shown to be workable.

The third and final class of methods we consider uses multivariate rather than univariate dynamic relationships, often in combination with structural relationships from economic theory, to estimate output gaps as a particular transitory component of real output. Some of these are examined in Section 3. One example is the decomposition method suggested by Cochrane (CO, 1994). This method is based on the permanent income hypothesis and uses consumption to define the permanent component of output which can then be used as a measure of potential output. Multivariate extensions of the Beveridge-Nelson decomposition method (MBN) have also been proposed to identify the

deviations from trend. As we argue below that this is what policymakers wish to measure, adoption of the "breaking-trend" model per se is not a solution to the problems of measuring output gaps which we discuss below.

⁴ We will henceforth refer to this method as the HP filter, although Hodrick and Prescott note that their method is due to Whittaker (1923) and Henderson (1924). Also, although the Hodrick and Prescott article is to be published in 1997, their working paper dates from 1981.

⁵ This is an oversimplification. More accurately, policymakers will usually be most interested in expected *future* values of the output gap, particularly when these expectations are conditioned on specific policy actions. This is more demanding than simply estimating the output gap at the end of sample, so our discussion of the additional difficulties introduced by end-of-sample problems underestimates the true difficulty of the policy problem. For that reason, we think good end-of-sample performance is a necessary rather than a sufficient condition for reliable estimation of the deviation from trend.

permanent component of output (Evans and Reichlin, 1994). A major restriction, used by both the CO and the MBN methods, is that the permanent component of real output is a random walk.

Section 3.1 of this paper, which draws heavily from Dupasquier, Guay and St-Amant (1996), discusses the CO and MBN methodologies and compares them with a structural vector autoregression methodology based on long-run restrictions imposed on output (LRRO.) This method was proposed by Blanchard and Quah (1989), Shapiro and Watson (1988), and King et al. (1991). One characteristic of the LRRO approach is that it does not impose restrictions on the dynamics of the permanent component of output. Instead, it allows for a permanent component comprising an estimated diffusion process for permanent shocks that can differ from a random walk. The output gap then corresponds to the cyclical component of output excluding the diffusion process of permanent shocks which is instead assigned to potential output. Instead, it allows for a permanent component comprising an estimated diffusion process of permanent shocks which is instead assigned to potential output. Section 3.2 presents an application of the LRRO method to Canadian data.

In Section 3.3 (which draws from Lalonde, Page and St-Amant (forthcoming)) we present another methodology based on long-run restrictions imposed on a VAR that associates restrictions imposed to real output *and* inflation. The output gap is then a part of the cyclical component of real output that is consistent with changes in the trend of inflation.⁶

The final section concludes with some directions for future research.

1. The HP filter

In recent years, mechanical filters have frequently been used to identify permanent and cyclical components of time series. The most popular of these mechanical filters is that proposed by Hodrick and Prescott (1997). This section evaluates the basic Hodrick-Prescott (HP) filter's ability to provide a useful estimate of the output gap. Section 2 then discusses some extensions and alternatives to the basic HP filter that have recently been proposed.

Guay and St-Amant (1996) show that the HP filter does a poor job in terms of extracting business cycle frequencies from macroeconomic time series. As a consequence, it is not an adequate approach to estimating an output gap constrained to correspond to the business-cycle frequencies of real GDP. This is discussed in Section 1.2, where we further argue that constraining the output gap in that way is not very attractive in any case. Guay and St-Amant also show that the HP filter is likely to do a poor job in terms of extracting an output gap assumed to correspond to the unobserved cyclical component of real GDP. This is discussed in Section 1.3. In Section 1.4, we focus explicitly on the HP filter's end-of-sample problems and conclude that these raise further doubts about the appropriateness of using the HP filter to estimate the output gap. Finally, Section 1.5 investigates what economic theory has to say about the possible usefulness of filters for estimating output gaps at the end of sample.

Most of the arguments in this section of the paper are drawn from Guay and St-Amant (1996) and van Norden (1995). Note that Guay and St-Amant show that the main conclusions that they reach concerning the HP filter also apply to the band-pass filter proposed by Baxter and King (1995).

1.1 The optimization problem

The HP filter decomposes a time series y_t into additive components, a cyclical

⁶ Lalonde, Page and St-Amant also present a method associating the output gap with changes in the trend of inflation but which does not impose that the output gap is stationary.

component y_t^c and a growth component y_t^g ,

$$y_t = y_t^g + y_t^c \quad (1)$$

Applying the HP filter involves minimizing the variance of y_t^c subject to a penalty for the variation in the second difference of y_t^g . This is expressed in the following equation:

$$\left\{ y_t^g \right\}_{t=0}^{T+1} = \arg \min \sum_{t=1}^T \left[\left(y_t - y_t^g \right)^2 + \lambda \left[\left(y_{t+1}^g - y_t^g \right) - \left(y_t^g - y_{t-1}^g \right) \right]^2 \right] \quad (2)$$

where λ , the smoothness parameter, penalizes the variability in the growth component. The larger the value of λ , the smoother the growth component. As λ approaches infinity, the growth component corresponds to a linear time trend. For quarterly data, Hodrick and Prescott propose setting λ equal to 1,600. King and Rebelo (1993) show that the HP filter can render stationary any integrated process of up to the fourth order.

1.2 How well does the HP filter extract business cycle frequencies?

Authors such as Singleton (1988) have shown that the HP filter can provide an adequate approximation of a high-pass filter when it is applied to stationary time series. Here we need to introduce some elements of spectral analysis. A zero-mean stationary process has a Cramer representation such as:

$$y_t = \int_{-\pi}^{\pi} \varepsilon e^{i\omega t} dz(\omega) \quad (3)$$

where $dz(\omega)$ is a complex value of orthogonal increments, i is the imaginary number $(-1)^{1/2}$ and ω is frequency measured in radians, i.e. $-\pi \leq \omega \leq \pi$ (see Priestley (1981), chapter 4). In turn, filtered time series can be expressed as:

$$y_t^f = \int_{-\pi}^{\pi} \alpha(\omega) e^{i\omega t} dz(\omega), \text{ with} \quad (4)$$

$$\alpha(\omega) = \sum_{h=-k}^k a_h e^{-i\omega h} \quad (5)$$

Equation (5) is the frequency response (Fourier transform) of the filter. That is, $\alpha(\omega)$ indicates the extent to which y_t^f responds to y_t at frequency ω and can be seen as the weight attached to the periodic component $e^{i\omega t} dz(\omega)$. In the case of symmetric filters, the Fourier transform is also called the gain of the filter.

An ideal high-pass filter would remove low-frequency, or long-cycle, components and allow high-frequency, or short-cycle, components to pass through so that $\alpha(\omega)=0$ for $|\omega| \leq \omega^p$, where ω^p has some predetermined value and $\alpha(\omega)=1$ for $|\omega| > \omega^p$. Figure 1 shows the squared gain of the HP filter. Very high frequencies are left aside because we want to focus on business-cycle frequencies as defined by NBER researchers since Burns and Mitchell (1946); i.e. cycles lasting no less than 6 and no more than 32 quarters. We see that the squared gain is 0 at zero frequency and is close to 1 from around frequency $\pi/10$ (6 quarters) and up. On the basis of Figure 1, the HP filter would appear to be an adequate approximation of a high-pass filter in that it removes most low frequencies and passes through most higher frequencies including business-cycle frequencies.

Figure 1
Squared gain of the HP filter

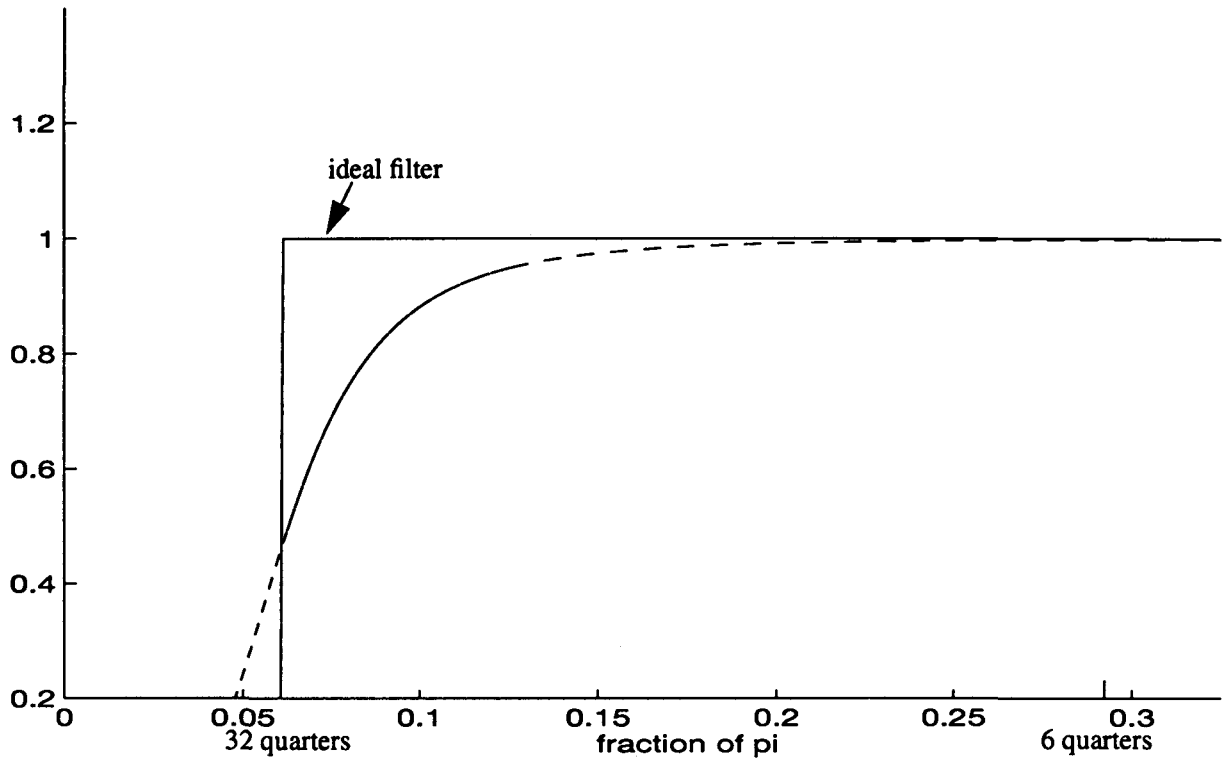
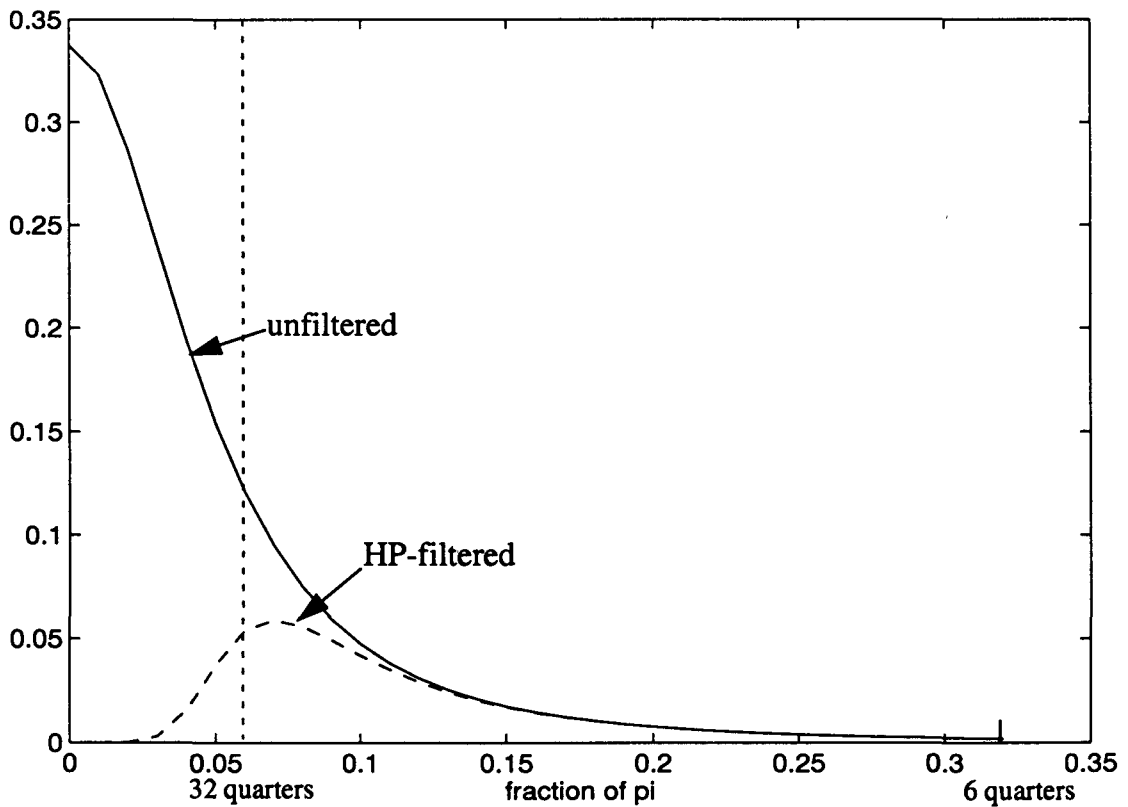


Figure 2
Series with the typical Granger shape - (AR(3) coefficients: 1.2, -0.11, -0.16)



One could associate the output gap with business cycle frequency plus higher frequency volatility in the data. Figure 1 would then suggest that the HP filter is an adequate measure of the output gap. One problem with this is that most macroeconomic time series are either integrated or highly persistent processes. In their study, Guay and St-Amant (1996) conduct a systematic investigation of the HP filter's ability to capture business-cycle frequencies; i.e. the area delimited by the spectrum of an original series at frequencies between 6 and 32 quarters. Their main finding is that, when the peak of a series is at zero frequency and the bulk of the variance is located in low frequencies, which is the shape described by Granger as typical for macroeconomic time series, the HP filter cannot capture business cycle frequencies adequately. This is illustrated by Figure 2, which shows the spectrum of an autoregressive process having its peak at zero frequency and that of the cyclical component resulting from the application of the HP filter.

In Figure 2, the spectrum of the cyclical component resulting from the application of the HP filter is very different from that of the original series. This comes as no surprise since the filter is designed to extract low frequencies from the data. However, we can see that business cycle frequencies are not left intact. In particular, the HP filter induces a peak inside business-cycle frequencies even though it is absent from the original series. Moreover, it fails to capture a significant fraction of the variance contained in business-cycle frequencies but captures some variance originating outside these frequencies. Guay and St-Amant (1996) show that this is typical of time series having the typical Granger shape; i.e. most macroeconomic series. Indeed, the unfiltered spectrum shown in Figure 2 is a parametric estimate of the spectrum of US real GDP.

The intuition behind this result is simple. Figure 1 shows that the gain of the HP filter at low business-cycle frequencies is smaller than that of the ideal filter. Indeed, the squared gain of the HP filter is 0.49 at frequencies corresponding to 32-quarter cycles and does not reach 0.95 before frequency $\pi/8$ (cycles of 16 quarters). Note also that the squared gain does not fall immediately to zero at lower frequencies. The problem is that a large fraction of the power of typical macroeconomic time series is concentrated in the band where the squared gain of the HP filter differs from that of an ideal filter. Also, the shape of the squared gain of the HP filter is such that when it is applied to typical macroeconomic time series a peak in the spectrum of the cyclical component is induced. In short, applying the HP filter to series dominated by low frequencies results in the extraction of a cyclical component that does not capture an important fraction of the variance contained in business-cycle frequencies of the original series captures an important part of the variance situated at lower frequencies than business-cycle frequencies but induces spurious dynamic properties.

An additional problem is that associating the output gap with the business-cycle frequencies in the data might not be a good idea in the first place. Note in particular that part of the variance associated with business-cycle frequencies could reflect the dynamics of shocks to potential output. As noted by King et al. (1991), "productivity shocks set off transitional dynamics, as capital is accumulated and the economy moves towards a new steady-state". To the extent that such dynamics reflect the evolution of potential output itself, one might prefer to use a different approach to identify potential output and the output gap. Section 3 of the paper provides a more detailed discussion of this point.

1.3 How well does the HP filter extract the cyclical component?

In the previous section, we have seen that the HP filter does not have spectral properties good enough to be able to isolate accurately the component of a series due to fluctuations at business cycle frequencies. As discussed by King and Rebelo (1993), another justification for the use of the HP filter is that in some cases it will be the optimal filter for identifying the cyclical component of a series. However, (King and Rebelo op. cit., and Harvey and Jaeger (1993)) these are cases when, in particular, the series is I(2), there are identical propagation mechanisms for innovations in the growth rate and in the cycle (or the transitory component is white noise), and the smoothing parameter λ is known. These conditions are rarely met in practice.

Of course, the fact that the HP filter is not an optimal filter does not necessarily mean that it will not be a good approximation of the optimal filter. We therefore turn to consider whether the HP filter can reliably isolate the cyclical component of a variety of time series.

It is often argued that macroeconomic time series are really comprised of a permanent component and a cyclical component. The permanent component could be driven by an I(1) technological process with drift, while monetary shocks, among others, could generate the cyclical component. In order to assess the HP filter's ability to extract such a cyclical component, consider the following DGP:

$$y_t = \mu_t + c_t, \text{ where} \quad (6)$$

$$\mu_t = \mu_{t-1} + \varepsilon_t \quad (7)$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \eta_t \text{ and} \quad (8)$$

$$\varepsilon_t \sim NID(0, \sigma_\varepsilon^2), \quad \eta_t \sim NID(0, \sigma_\eta^2) \quad (9)$$

Equation (6) defines y_t as the sum of a permanent component, μ_t , which in this case corresponds to a random walk, and a cyclical component, c_t .⁷ The dynamics of the cyclical component are specified as a second order autoregressive process so that the peak of the spectrum could be at zero frequency or at business-cycle frequencies. We assume that ε_t and η_t are uncorrelated.

Data are generated from equation (6) with ϕ_1 set at 1.2 and different values for ϕ_2 to control the location of the peak in the spectrum of the cyclical component. We also vary the standard-error ratio for the disturbances $\sigma_\varepsilon/\sigma_\eta$ to change the relative importance of each component. We follow the standard practice of giving the value 1,600 to λ , the HP filter smoothness parameter. We also follow Baxter and King's (1995) suggestion of dropping 12 observations at the beginning and at the end of the sample which should favour the filter considerably by abstracting partly from its end-of-sample problems (see Section 1.4). The resulting series contains 150 observations, a standard size for quarterly macroeconomic data. The number of replications is 500.

The performance of the HP filter is assessed by comparing the autocorrelation function of the cyclical component of the true process with that obtained from the filtered data. We also calculate the correlation between the true cyclical component and the filtered cyclical component and report their relative standard deviations ($\hat{\sigma}_c / \sigma_c$). Table 1 presents the results and illustrates that the HP filter performs particularly poorly when there is an important permanent component. Indeed, in most cases, for high $\sigma_\varepsilon/\sigma_\eta$ ratios, the correlation between the true and the filtered components is not significantly different from zero. The estimated autocorrelation function is invariant to the change in the cyclical component in these cases (the values of the true autocorrelation functions are given in parentheses.) When the ratio $\sigma_\varepsilon/\sigma_\eta$ is equal to 0.5 or 1 and the peak of the cyclical component is located at zero frequency ($\phi_2 < -0.43$), the dynamic properties of the true and the filtered cyclical components are significantly different, as indicated by the estimated parameter values. In general, the HP filter adequately characterizes the series' dynamics when the peak of the spectrum is at business-cycle frequencies and the ratio $\sigma_\varepsilon/\sigma_\eta$ is small. However, even when the ratio of standard deviations is equal to 0.01 (i.e. the permanent component is almost absent), the filter performs poorly when the peak of the spectrum of the cyclical component is at zero frequency. Indeed, for $\phi_2 = -0.25$, the dynamic properties of the filtered component differ significantly from those of the true cyclical component, the correlation is only equal to 0.66, and the standard deviation of the filtered cyclical component is half that of the true cyclical component.

⁷ This is Watson's (1986) specification for real GDP in the United States.

Table 1
Simulation results for the HP filter

GDP			Estimated values				
$\sigma_\varepsilon/\sigma_\eta$	ϕ_1	ϕ_2	Autocorrelations			correlation	$(\hat{\sigma}_c / \sigma_c)$
			1	2	3		
10	0	0	0.71[0] (0.59, 0.80)	0.46[0] (0.30, 0.60)	0.26[0] (0.08, 0.43)	0.08 (-0.07, 0.21)	12.96 (10.57, 15.90)
10	1.2	-0.25	0.71[0.96] (0.61, 0.80)	0.47[0.90] (0.31, 0.61)	0.27[0.84] (0.08, 0.44)	0.08 (-0.11, 0.28)	4.19 (2.77, 6.01)
10	1.2	-0.40	0.71[0.86] (0.60, 0.80)	0.46[0.63] (0.30, 0.60)	0.26[0.41] (0.08, 0.44)	0.13 (-0.12, 0.36)	6.34 (4.82, 8.07)
10	1.2	-0.55	0.71[0.77] (0.60, 0.80)	0.46[0.38] (0.29, 0.60)	0.26[0.03] (0.06, 0.43)	0.14 (-0.08, 0.33)	6.93 (5.36, 8.70)
10	1.2	-0.75	0.71 [0.69] (0.60, 0.78)	0.46[0.27] (0.30, 0.59)	0.25[-0.19] (0.07, 0.41)	0.15 (-0.01, 0.31)	6.37 (4.79, 7.95)
5	0	0	0.69[0] (0.58, 0.78)	0.45[0] (0.30, 0.58)	0.26[0] (0.09, 0.41)	0.15 (0.02, 0.27)	6.5 (5.28, 7.85)
5	1.2	-0.25	0.71[0.96] (0.61, 0.80)	0.46[0.90] (0.32, 0.61)	0.26[0.84] (0.08, 0.43)	0.16 (-0.01, 0.36)	2.11 (1.43, 3.04)
5	1.2	-0.40	0.72[0.86] (0.61, 0.80)	0.46[0.63] (0.31, 0.60)	0.25[0.41] (0.08, 0.42)	0.23 (-0.01, 0.45)	3.26 (2.47, 4.15)
5	1.2	-0.55	0.71[0.77] (0.61, 0.80)	0.46[0.38] (0.30, 0.59)	0.24[0.03] (0.06, 0.41)	0.24 (0.01, 0.44)	3.60 (2.83, 4.52)
5	1.2	-0.75	0.70[0.69] (0.61, 0.79)	0.43[0.27] (0.26, 0.57)	0.20[-0.19] (0.00, 0.38)	0.29 (0.11, 0.44)	3.3 (2.53, 4.17)
1	0	0	0.43[0] (0.27, 0.57)	0.28[0] (0.11, 0.42)	0.20[0] (-0.02, 0.31)	0.59 (0.49, 0.70)	1.61 (1.41, 1.85)
1	1.2	-0.25	0.76[0.96] (0.67, 0.83)	0.51[0.90] (0.37, 0.62)	0.29[0.84] (0.11, 0.44)	0.51 (0.33, 0.68)	0.66 (0.44, 0.91)
1	1.2	-0.40	0.75[0.86] (0.67, 0.81)	0.44[0.63] (0.28, 0.55)	0.16[0.41] (-0.03, 0.33)	0.71 (0.56, 0.82)	1.02 (0.83, 1.22)
1	1.2	-0.55	0.72[0.77] (0.66, 0.78)	0.34[0.38] (0.21, 0.47)	0.01[0.03] (-0.17, 0.19)	0.76 (0.56, 0.82)	1.15 (0.83, 1.22)
1	1.2	-0.75	0.68[0.69] (0.63, 0.72)	0.15[0.27] (0.04, 0.27)	-0.27[-0.19] (-0.44, 0.10)	0.83 (0.75, 0.89)	1.16 (1.04, 1.29)
0.5	0	0	0.16[0] (0.01, 0.32)	0.10[0] (-0.04, 0.24)	0.04[0] (-0.10, 0.18)	0.82 (0.75, 0.88)	1.16 (1.07, 1.27)
0.5	1.2	-0.25	0.79[0.96] (0.71, 0.85)	0.53[0.90] (0.38, 0.65)	0.30[0.84] (0.11, 0.46)	0.61 (0.41, 0.79)	0.55 (0.37, 0.76)
0.5	1.2	-0.40	0.77[0.86] (0.69, 0.81)	0.43[0.63] (0.29, 0.54)	0.13[0.41] (-0.05, 0.29)	0.84 (0.73, 0.92)	0.87 (0.74, 0.99)
0.5	1.2	-0.55	0.72[0.77] (0.67, 0.78)	0.28[0.38] (0.17, 0.39)	-0.10[0.03] (-0.25, 0.06)	0.89 (0.83, 0.94)	0.98 (0.89, 1.07)
0.5	1.2	-0.75	0.67[0.69] (0.63, 0.71)	0.07[0.27] (-0.03, 0.18)	-0.42[-0.19] (-0.57, -0.27)	0.94 (0.90, 0.96)	1.02 (0.97, 1.08)
0.01	0	0	-0.08[0] (-0.21, 0.06)	-0.06[0] (-0.21, 0.06)	-0.06[0] (-0.19, 0.06)	0.98 (0.96, 0.99)	0.97 (0.94, 0.99)
0.01	1.2	-0.25	0.80[0.96] (0.72, 0.86)	0.54[0.90] (0.38, 0.67)	0.30[0.84] (0.11, 0.48)	0.66 (0.45, 0.83)	0.51 (0.34, 0.69)
0.01	1.2	-0.40	0.78[0.86] (0.72, 0.83)	0.43[0.63] (0.30, 0.55)	12[0.41] (-0.05, 0.28)	0.90 (0.82, 0.96)	0.81 (0.71, 0.90)
0.01	1.2	-0.55	0.73[0.77] (0.67, 0.77)	0.26[0.38] (0.15, 0.37)	-0.14[0.03] (-0.30, 0.01)	0.96 (0.91, 0.99)	0.92 (0.86, 0.96)
0.01	1.2	-0.75	0.67[0.69] (0.62, 0.71)	0.02[0.27] (-0.08, 0.13)	-0.50[-0.19] (-0.61, -0.35)	0.99 (0.97, 1.0)	0.97 (0.95, 0.99)

It is interesting to note that the HP filter does relatively well when the ratio $\sigma_\varepsilon/\sigma_\eta$ is equal to 1, 0.5, or 0.01 and the spectrum of the original series has a peak at zero frequency and at business-cycle frequencies (i.e. the latter frequencies contain a significant part of the variance of the series). Consequently, the conditions required to adequately identify the cyclical component with the HP filter can be expressed in the following way: the spectrum of the original series must have a peak located at business-cycle frequencies, which must account for an important part of the variance of the series. If the variance of the series is dominated by low frequencies, which is the case for most macroeconomic series in levels, including real output, the HP filter does a poor job of extracting an output gap associated with the cyclical component of real output.

1.4 The HP filter at the end of samples

In examining the performance of the HP filter in the last two sections, we have looked at how well it isolates particular business cycle frequencies or the cyclical component of the series. Both cases implicitly looked at the performance of the HP filter over the available sample of data as a whole. However, it is useful to remember that the focus for policy advice is on estimating the *current* output gap. This is a more difficult task since future information will presumably be useful in determining whether recent changes in output are persistent or transitory. We should therefore consider how the conclusions from the two previous sections might be altered by this added complication.⁸

To understand how the HP filter behaves at the end of sample, recall that the optimization problem it solves trades off the size of deviations from trend with the smoothness of that trend. In the face of a transitory shock, the filter is therefore reluctant to change the trend very much since this implies raising the trend before the shock and lowering it afterwards. However, the latter penalty is absent, implying that the optimal trend will be more responsive to transitory shocks than in mid-sample.

We can show this difference in several ways. Figure 3 shows the HP filter trend expressed as a moving average of the unfiltered data. The weights in this moving average change as we move from the mid-sample towards the end of sample. The former gives us a smooth 2-sided average in which no observation receives more than 6% of the weight. However, the latter gives a 1-sided average where the last observation alone accounts for 20% of the weight. Not surprisingly, this makes the HP trend more variable at the end of sample. Figure 4 shows that the deviations from the HP trend result in different frequency responses. In particular, the 1-sided, or end-of-sample, filtered deviations from trend capture less of the variation at business cycle frequencies (indicated by the dotted vertical lines).⁹

Figure 5 and Figure 6 show how the deviations from the HP trend differ depending on whether we are at the end-of-sample or mid-sample. The solid line in Figure 5 shows the usual deviation from the HP trend for Canadian GDP. The dashed line then shows the estimate we get from the same if we only use data available up to that point in time (i.e. the corresponding end-of-sample estimate). Although the two series tend to move together, there are some important differences in size and timing. Comparing Figure 5 with Figure 6, we see that while deviations from trend are usually less than 3% of GDP, the difference between its mid-sample and end-of-sample estimates is often as large as 2% of GDP. If we accept that the difference between these two measures is just one

⁸ This problem has been mentioned in other studies as well. Much of the analysis we present can also be found in Butler (1996).

⁹ The squared-gains of the two HP trends also look quite different. At the frequency corresponding to cycles of 32 quarters, the end-of-sample filter has a squared-gain of about 1 while the mid-sample filter has a squared-gain of about 0.1. This is precisely the frequency at which an optimal filter would have a gain of zero.

Figure 3

MA representation of the HP filter as a function of sample position (128 observations, $\lambda = 1,600$)

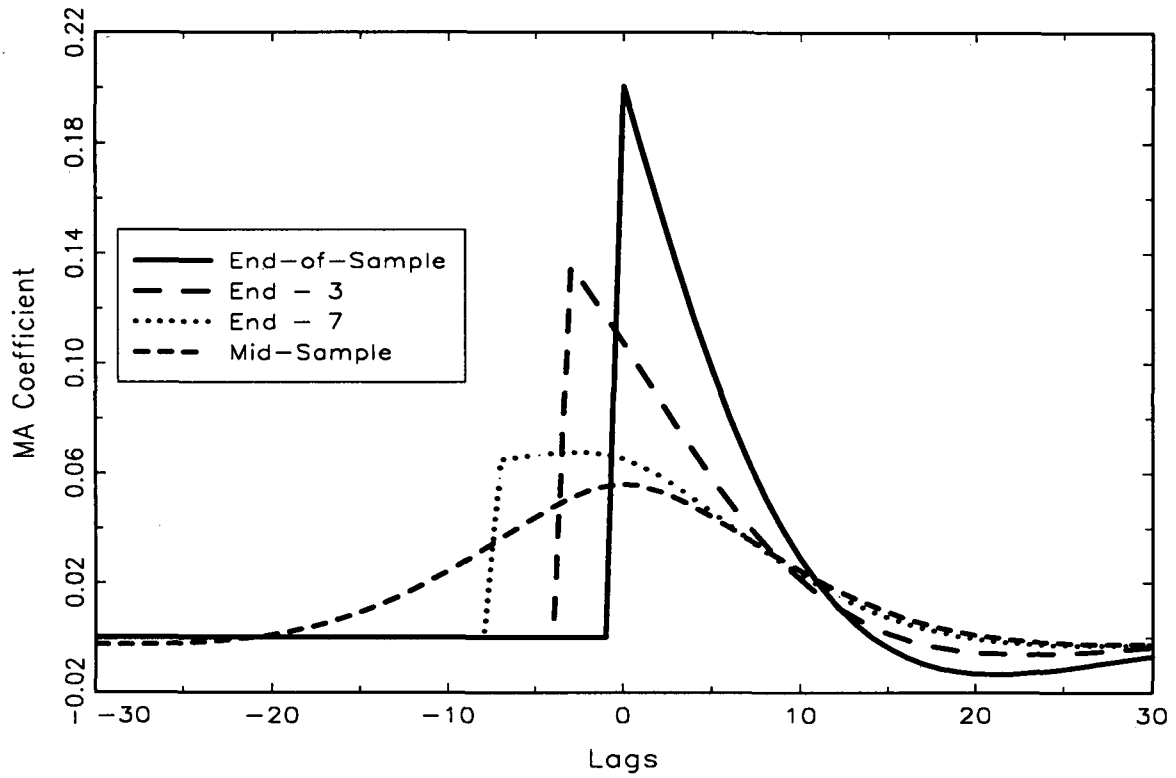


Figure 4

Squared gain of the HP filter (128 observations, $\lambda = 1,600$)

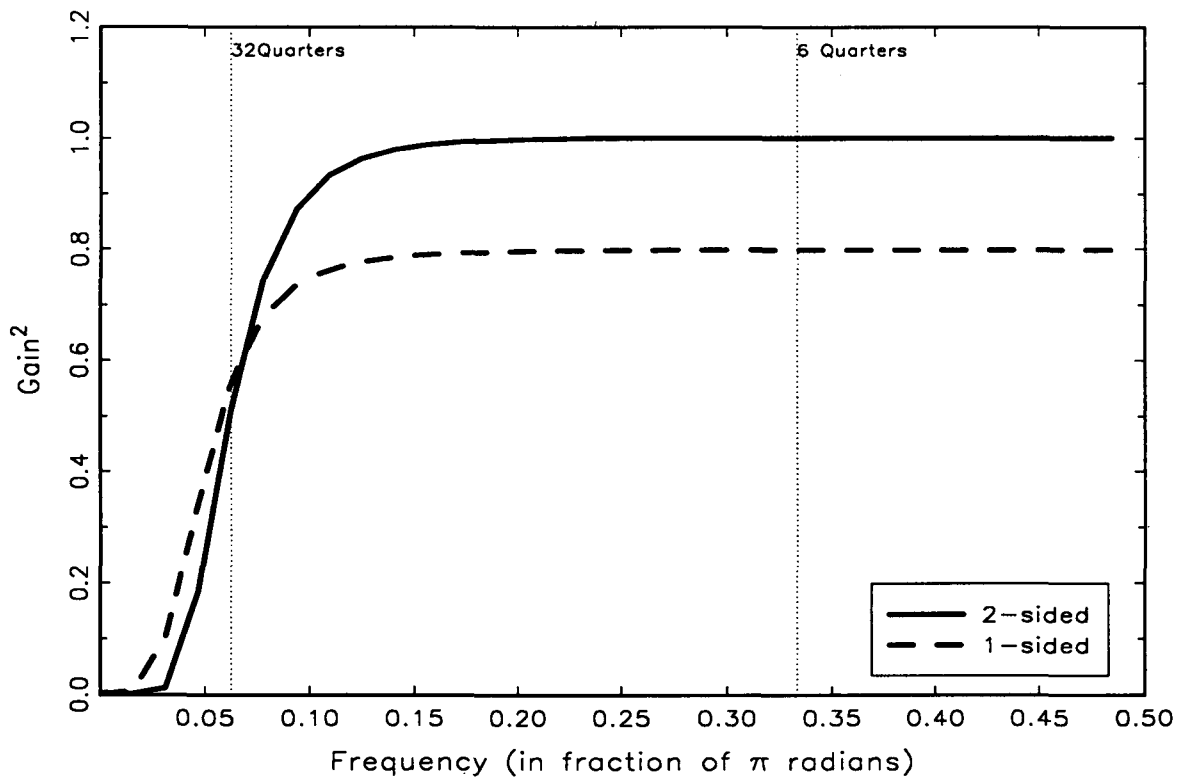


Figure 5

HP detrended real GDP (Canada, $\lambda = 1,600$)

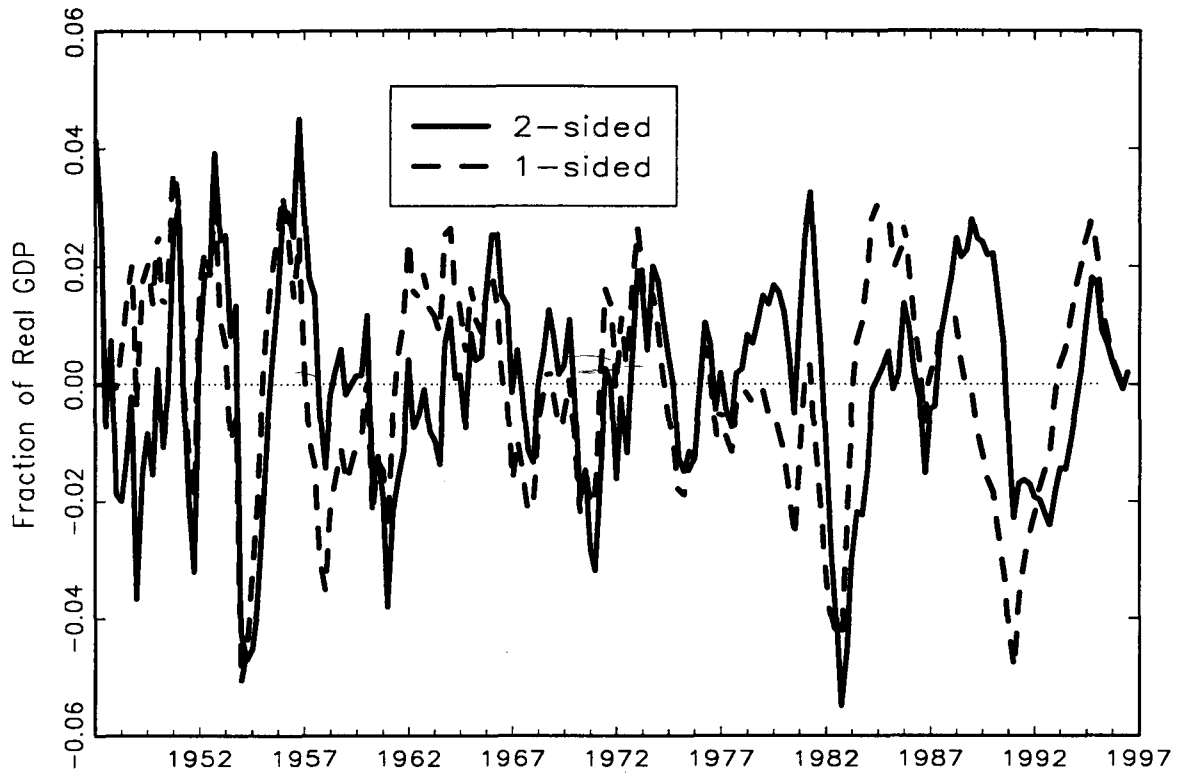


Figure 6

HP detrended real GDP mid-sample – end-of-sample (Canada, $\lambda = 1,600$)

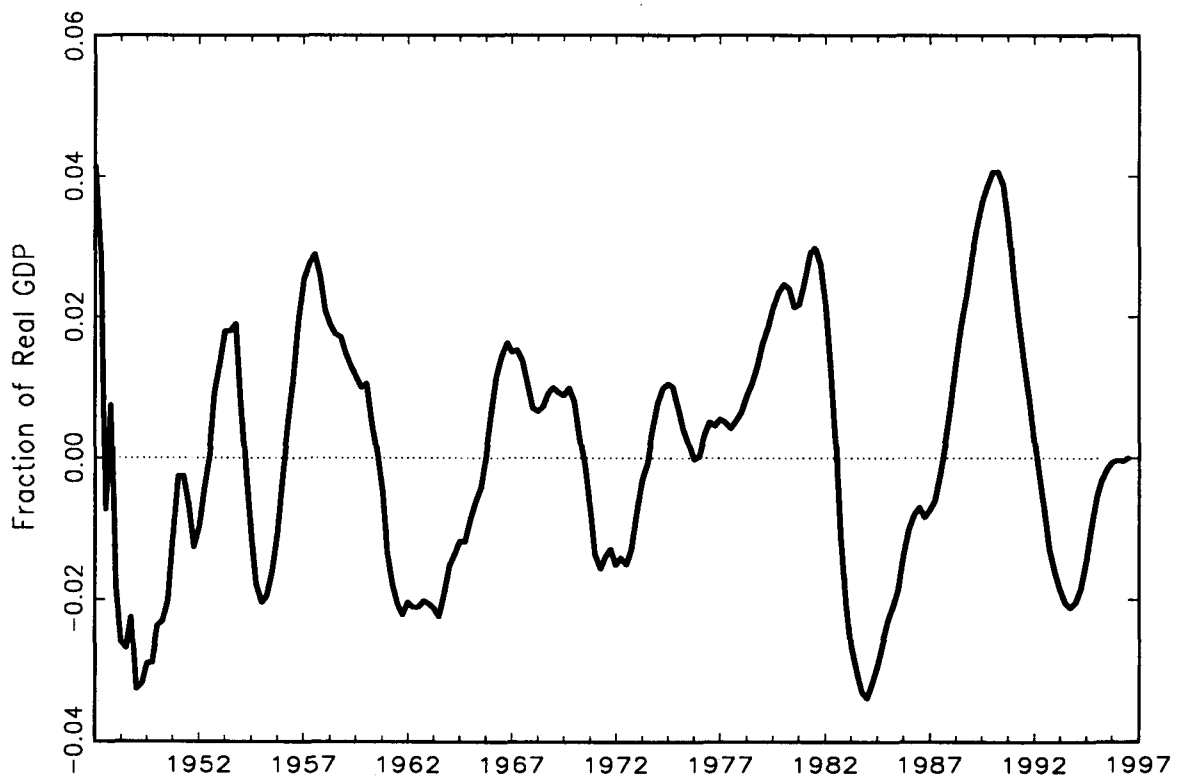
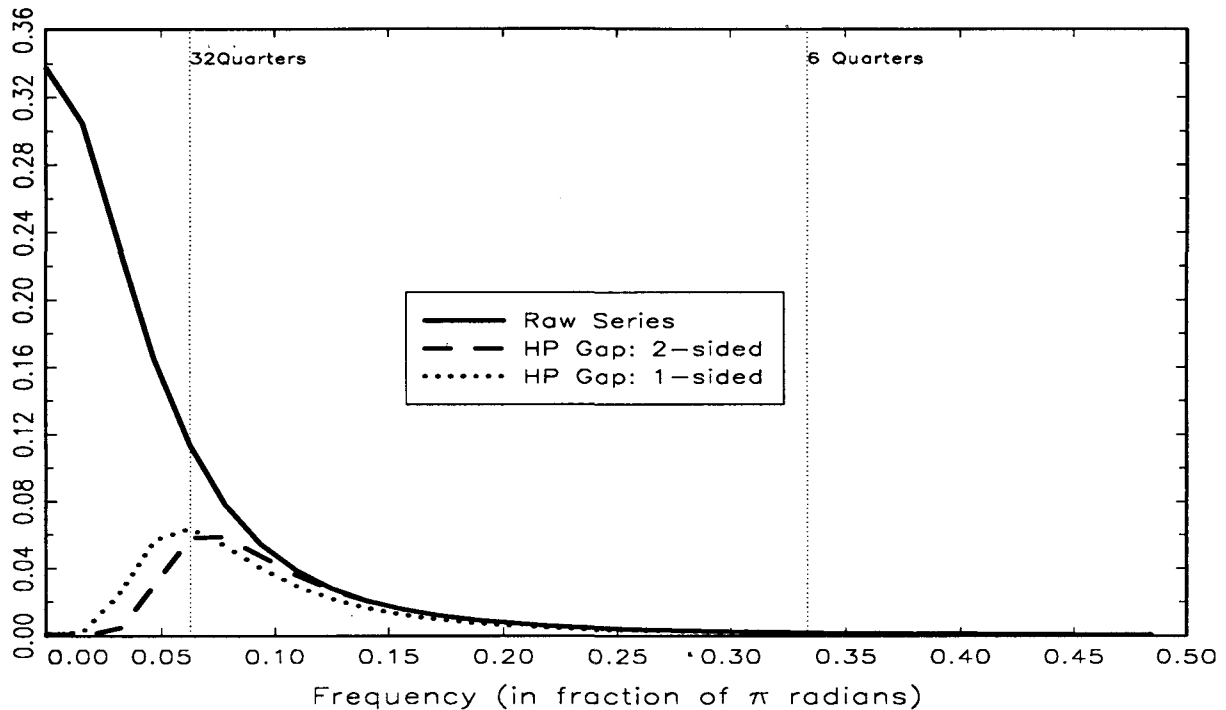


Figure 7

Spectrum of series with typical Granger shape (128 observations, $\lambda = 1,600$)



component of the measurement error of end-of-sample estimates, then the measurement errors of the latter must be roughly as large as the estimates themselves.¹⁰ Hence, end-of-sample estimates cannot be very reliable estimates of deviations from trend.

Figure 7 applies the HP filter to the "typical Granger Shape" series we considered previously. At the end-of-sample, even less of the variance of the deviations from the HP trend is due to variations at business cycle frequencies and more is due to "leakage" from lower frequencies. This suggests that the results we obtained in Section 1.2 probably overstate the reliability of the HP filter for identifying an output gap associated with business cycle frequencies. This is consistent with the results of Laxton and Tetlow (1992) and Butler (1996), who note that related filters also seem to perform worse at the end of samples. We turn to these related filters in Section 2.

1.5 Limits to 1-sided filtering

Part of the end-of-sample problem discussed in Section 1.4 reflects the fact that the HP filter behaves differently at the end-of-sample and at mid-sample, as shown in Figure 3. This suggests that other univariate filters might be able to measure output gaps more reliably. In this section, we consider one intrinsic limit to the ability of univariate filters to measure the current output gap, and show how this in turn will relate to beliefs about the economic relationships between actual and potential output. We show that models in which potential output is exogenous with respect to actual

¹⁰ We reach the same conclusion if we look at the range of the series, or at their standard deviations. The range (maximum - minimum) of the 1-sided estimate is 8.7% of GDP while the range of the difference between the 1 and 2-sided estimates is 7.5%; the comparable standard errors are 1.8% and 1.8%. These comparisons are only approximate; small sample problems in the 1-sided estimate at the beginning of the sample may make their difference appear excessively volatile, while constraining the two estimates to be identical at the end of the sample will tend to understate the volatility of their difference.

output and the output gap imply that univariate filters will never be able to give much information about contemporaneous output gaps.¹¹

Suppose that potential output can be expressed as a linear filter of actual output, so that:

$$q_t = A(L).y_t + \varepsilon_t \quad (10)$$

where q_t is (the log of) potential output, y_t is (the log of) actual output, ε_t is an innovations process that is uncorrelated with y_t at all leads and lags, and $A(L)$ is a two-sided polynomial in the lag operator (i.e. it takes a weighted sum of leads, lags and contemporaneous values of y_t). A sufficient but not necessary condition for such a representation to exist is that output y_t has a unit root and that the output gap $q_t - y_t$ is stationary.

We typically think of y_t as being non-stationary in mean, since it tends to drift upwards over time. To ensure that q_t and y_t move together in the long run (so that the gap is stationary), we will further assume that:

$$A(1) = 1 \quad (11)$$

which simply means that the weights in $A(L)$ must sum to one. This in turn implies that:

$$A(L) - 1 = (1 - L). \tilde{A}(L), \text{ and therefore that:} \quad (12)$$

$$q_t - y_t = \tilde{A}(L). \Delta y_t + \varepsilon_t \quad (13)$$

Therefore, we should be able to express the output gap as the weighted sum of past, present and future output growth. The difference between equation (13) and the HP filter representation of the output gap is that the HP filter implies a particular set of restrictions on $\tilde{A}(L)$ that vary with the position in the sample. Let:

$$E(q_t - y_t | H_y) = \tilde{A}(L). \Delta y_t \text{ and} \quad (14)$$

$$V((q_t - y_t) - E(q_t - y_t | H_y)) = \sigma^2 \quad (15)$$

where H_y is the set of all past, present and future values of y_t .

So far, we have assumed that $\tilde{A}(L)$ is two-sided, whereas its use for policy purposes requires that it be one-sided. To understand how such a restriction on $\tilde{A}(L)$ will affect the accuracy of our estimate, note that the law of iterated expectations and equation (14) imply:

$$E(q_t - y_t | H_y^-) = E(E(q_t - y_t | H_y) | H_y^-) = E(\tilde{A}(L). \Delta y_t | H_y^-) \quad (16)$$

where H_y^- is the set of all past values of y_t . If we define:

$$\tilde{A}(L) = \tilde{A}^-(L) + \tilde{A}^+(L) \quad (17)$$

where $\tilde{A}^-(L)$ has only positive powers of L and $\tilde{A}^+(L)$ only non-positive powers, then equation (16) implies:

$$E(q_t - y_t | H_y^-) = E(\tilde{A}^-(L). \Delta y_t | H_y^-) + E(\tilde{A}^+(L). \Delta y_t | H_y^-) \quad (18)$$

$$= \tilde{A}^-(L). \Delta y_t + \sum_{j=1}^m a_j^+ . E(\Delta y_{t+j} | H_y^-) = \bar{A}(L). \Delta y_t$$

¹¹ This section draws heavily on van Norden (1995).

where a_j^+ is simply the coefficient on L^{-j} in $\tilde{A}^+(L)$. Similarly, we can show that

$$V(q_t - y_t | H_y^-) = V(q_t - y_t | H_y) + V(\tilde{A}(L) \cdot \Delta y_t | H_y^-) = \sigma^2 + V(\tilde{A}^+(L) \cdot \Delta y_t | H_y^-) \quad (19)$$

where $V(X|\Omega)$ is the variance of the error in forecasting X given the information set Ω .

Equations (18) and (19) have an intuitive interpretation. The extent to which the 1-sided filter $\bar{A}(L)$ is less informative than $\tilde{A}(L)$ will depend on the weight which $\tilde{A}(L)$ puts on current and future values of Δy_t and the extent to which those future values can be predicted from current and past values. The former will in turn depend on the Granger causal relationship between $q_t - y_t$ and Δy_t , while the latter will depend on the degree to which output growth is serially correlated.

For most industrialized nations, 12 lags of quarterly output growth predicts only 20-40% of the variance of current output growth and much of this explanatory power seems to come from the first few lags. This suggests that since predictability can be low, the extent of Granger-causality will play an important role in determining how accurately the one-sided univariate filter can estimate the current output gap. For simplicity, we will discuss the role Granger-causality plays under the assumption that the past history of output growth is of no use in predicting present and future output growth.

From equation (18) we can see that $\bar{A}(L)$ will tell us as much about the output gap as $\tilde{A}(L)$ when $\tilde{A}^+(L) = 0$, which in turn implies that $q_t - y_t$ does not Granger-cause Δy_t . Van Norden (1995) shows that the latter condition in turn implies that q_t does not Granger-cause y_t . If y_t and q_t are cointegrated, this would imply that there is unidirectional causality from y_t to q_t . In other words, exogenous shocks to potential output would have no subsequent effect on actual output, but persistent shocks to actual output would eventually be followed by a similar change in potential. Such behaviour could describe a particularly severe form of hysteresis; one where output has no tendency to return to potential and, instead, potential output is driven in the long run only by previous variations in actual output. In this kind of world, univariate filters can hope to be as effective in estimating the current output gap as they are in estimating past output gaps.

The latter will not be the case when $q_t - y_t$ Granger-causes Δy_t . Therefore, so long as output growth appears to respond to some degree to past changes in the output gap, then univariate methods will estimate the current output gap less accurately than past output gaps. The intuition behind this result should be clear. Since future output growth will reflect the influence of the current output gap, one can gain information about the current gap by observing future growth.

Univariate methods will be of no use in estimating the current output gap when Δy_t not Granger cause $q_t - y_t$.¹² In other words, if faster, or slower, than normal output growth tends to have no subsequent effect on the size of the output gap, then time series estimates of the current gap will be as uninformative as possible. The intuition is similar to that above. Past output growth is the only information about the gap that we have; if it tells us nothing about the current gap, then our estimates will be unilluminating.

It is more difficult to say what kind of economic model will generate this kind of result since Granger-causality from Δy_t to $q_t - y_t$ does not directly correspond to any statement about

¹² Again, this conclusion assumes that past output growth is of no use in predicting future variations in output growth. As was mentioned earlier, the data show that there is some serial correlation in output growth, so time series methods would still have some explanatory power even in this case.

Granger-causality between y_t and q_t .¹³ However, it is possible to give examples in which this result would hold. One simple case would be where:

$$\Delta y_t = \alpha(y_{t-1} - q_{t-1}) + u_t \quad \text{and} \quad q_t = q_{t-1} + v_t \quad (20)$$

Potential output follows a random walk that is independent of the behaviour of output. Actual output in turn is generated by a simple error-correction model, which ensures that actual and potential output move together in the long run. Such a model precisely satisfies the condition for no Granger causality from Δy_t to $q_t - y_t$.

Clearly, there is a range of models in which univariate time-series methods will be of little use at the end of sample. Furthermore, it is the short-run dynamics of potential and actual output which are critical in determining whether models belong to this class. This is not an empirically testable question since we cannot directly observe potential. However, we can try to ensure that our views on the determination of potential are consistent with the methods we use to measure it.

2. Extensions of the HP filter

The Bank of Canada has used various extensions of the HP filter to obtain measures of the output gap and help guide policy. These "hybrid" methods were developed in the 1990s to try to balance strengths and weakness of "structural" and "astructural" approaches to measuring the output gap for policy makers. The key papers explaining the justification and implementation of this approach are Laxton and Tetlow (1992) and Butler (1996). Work in a similar spirit has been pursued both at some of the Federal Reserve Banks (such as Kuttner (1994)) and at the OECD (see Giorno et al. (1995)).

To understand the contribution of these methods, one needs to appreciate the problems that these authors were trying to avoid. Laxton and Tetlow argue that there is insufficient knowledge about the true structural determinants of the supply side of the economy to make the purely structural approach practicable. At the same time, for policy purposes we need to distinguish between movements in output caused by supply shocks and those caused by demand shocks, whereas most astructural (time-series) models attempt to distinguish permanent and transitory components of output. As an alternative, they suggest a way of combining the two approaches which we refer to as the multivariate HP filter.¹⁴

As we explain in Section 2.1, this methodology consists of adding the residuals of a structural economic relationship to the minimization problem that the HP filter is seeking to solve. Section 2.2 discusses the production function variant of this methodology. In Section 2.3, we examine additional modifications introduced to the filter to improve its performance at the end of the sample. Section 2.4 looks at these approaches from a different perspective and relates them to both the methods of Section 1 and other methods that use additional structural relationships.

¹³ See van Norden (1995).

¹⁴ Laxton and Tetlow refer to their specific filter as "The Multivariate Filter (MVF)" and Butler refers to his as "The Extended Multivariate Filter (EMVF)". In this paper, we broadly refer to all multivariate extensions of the univariate Hodrick-Prescott filter as Multivariate HP Filters (MHPF), which include the MVF and EMVF as special cases. The method currently used to estimate Canadian potential output for the Bank's staff projection will also be referred to as the EMVF. The latter differs somewhat from the implementation described in Butler (1996), but is conceptually the same.

2.1 A multivariate HP filter

As noted previously, the original HP filter chooses the trend as the solution to:

$$\{y_t^g\}_{t=0}^{T+1} = \arg \min \sum_{t=1}^T (y_t - y_t^g)^2 + \lambda (\Delta^2 y_{t+1}^g)^2 \quad (21)$$

where $\Delta^2 y_{t+1}^g = \Delta \cdot \Delta \cdot y_{t+1}^g$ and $\Delta y_t = y_t - y_{t-1}$. The HP filter adds a term,

$$\{y_t^g\}_{t=0}^{T+1} = \arg \min \sum_{t=1}^T (y_t - y_t^g)^2 + \lambda_g (\Delta^2 y_{t+1}^g)^2 + \lambda_\varepsilon \varepsilon_t^2 \quad (22)$$

where $\varepsilon_t = z_t - f(y_t^g, x_t)$. z_t is some other economic variable of interest, and $f(\cdot)$ models z_t as a function of both some explanatory variables x_t and the unobserved trend y_t^g . The new term in ε_t^2 has the effect of choosing the trend to simultaneously minimize deviations of output from trend, minimize changes in the trend's growth rate, *and* maximize the ability of the trend to fit some structural economic relationship $f(\cdot)$. λ_g and λ_ε reflect the relative weights of these different objectives.

The key to implementing the multivariate HP filter for the purpose of estimating potential output (or an output gap) is to specify $(z_t, f(y_t^g, x_t))$ in such a way as to capture some structural relationship that depends on either potential output or the output gap. For example, one could specify a Phillips curve equation that relates observed inflation to a measure of inflation expectations, the output gap, and perhaps additional explanatory variables (such as oil prices). ε_t would then be the residual from this Phillips curve equation and trend output would be chosen in part to improve the explanatory power of the output gap for inflation. Alternatively, one could use an Okun's Law relationship to link the rate of unemployment to the output gap and various structural variables determining the natural rate of unemployment. The trend of output would then be influenced by the evolution of the unemployment rate and its structural determinants.

Of course, there is no reason for restricting ourselves to a single structural relationship. Equation (22) can be generalized to include an arbitrary number n of structural relationships with a common trend y_t^g , giving

$$\{y_t^g\}_{t=0}^{T+1} = \arg \min \sum_{t=1}^T (y_t - y_t^g)^2 + \lambda_g (\Delta^2 y_{t+1}^g)^2 + \left(\sum_{i=1}^n \lambda_{\varepsilon_i} \varepsilon_{it}^2 \right) \quad (23)$$

The original Laxton and Tetlow (1992) paper used information from both a Phillips curve and an Okun's Law relationship and Butler (1996) also uses multiple structural relationships simultaneously.

The usefulness of the multivariate HP filter depends on several factors. Obviously, the extent to which it improves upon the original HP filter will depend on the reliability and information content of the structural relationship(s) with which it is combined. These potentially offer a way of mitigating the problems of HP filters noted in Section 1. However, given the importance of obtaining good end-of-sample estimates of output gaps, we require structural relationships that can give good contemporaneous information.¹⁵

For the particular DGP they examine, Laxton and Tetlow find that the degree to which their filter does better than the univariate HP filter at estimating the output gap increases with the relative importance of demand shocks to supply shocks. While the improvement produced by the

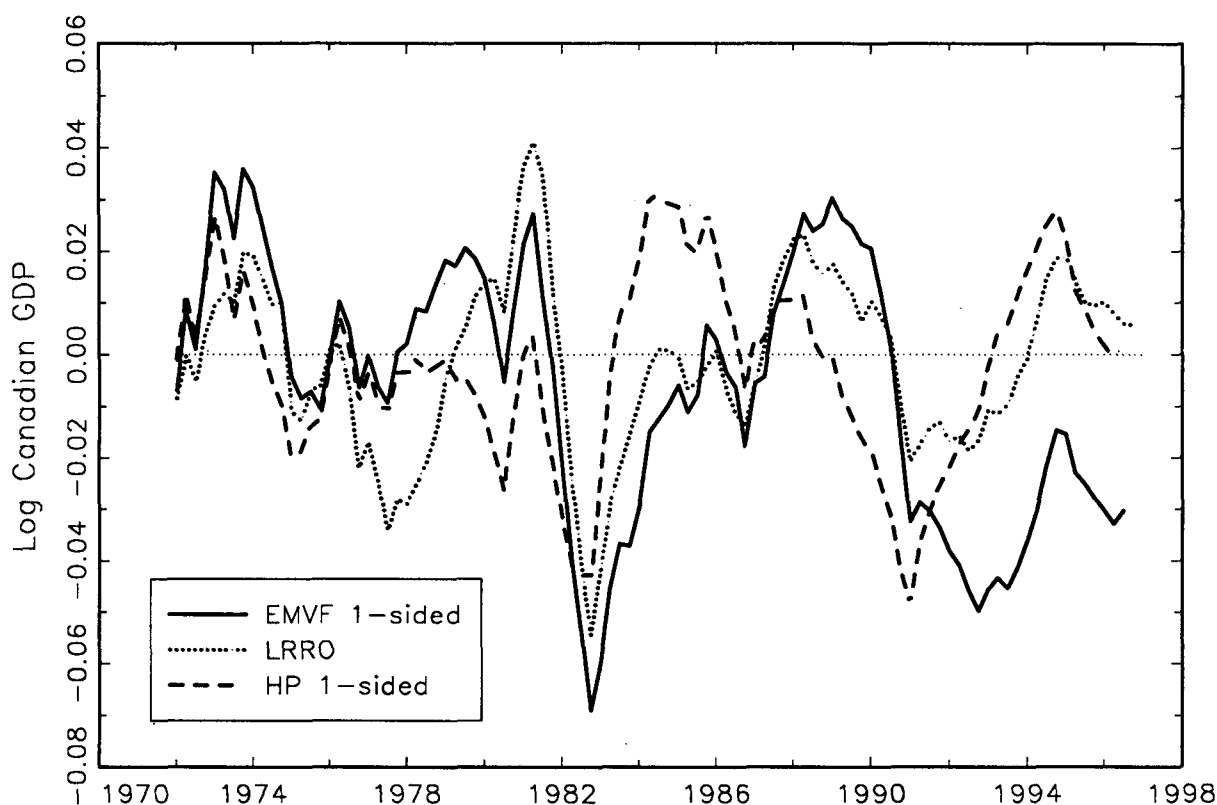
¹⁵ In that respect, there may be limitations to the information we can expect to gain from Phillips curve relationships if we believe that inflation responds to output gaps with a lag.

MHPF can be large, they find that there is still substantial uncertainty in their point estimates of the output gap and that this uncertainty is larger at the end of sample. In their base case, they find that the 95% confidence interval for the output gap at the end of sample is about 4%, which implies that policymakers would rarely observe statistically significant output gaps.

Another factor of key importance to the success of the multivariate HP filter is calibration. Instead of having a single λ parameter with a consensus value of 1,600, we now have vectors of parameters $\bar{\lambda}_{g,\epsilon}$ without a clear guide as to their appropriate values. In addition, we also need to estimate the form of the structural relationships involving potential output or the output gap. If we attempt to do this before calculating $\{y_t^g\}$, then we will be estimating a structural relationship which may be inconsistent with the values of $\{y_t^g\}$ the MHPF produces. Furthermore, theory will often not be a sufficient guide to allow us to tightly calibrate such a relationship. The approach used by Laxton and Tetlow (1992) and Butler (1996) is to experiment with alternative weightings to see which produce reasonable results and how sensitive the outcomes are to these choices.

Figure 8

Comparison of 3 different measures of the Canadian output gap



An alternative explored by Harvey and Jaeger (1993) and Côté and Hostland (1994) is to estimate the structural relationship simultaneously with $\{y_t^g\}$ and $\{\lambda_g, \bar{\lambda}_\epsilon\}$ via maximum likelihood methods.¹⁶ Côté and Hostland found that the results can be sensitive to the specification of the

¹⁶ Butler (1996) mentions that a direct maximum likelihood estimation was attempted, but that this did not produce reasonable results for the λ 's.

structural relationships,¹⁷ that the usefulness of the structural information vanishes when one considers only end-of-sample performance,¹⁸ that the structural parameters cannot be estimated with much accuracy, and that maximization of the likelihood function was problematic.

To give some idea of how such filters perform in practice, Figure 8 compares three different estimates of the output (GDP) gap. The first is that produced by the Butler (1996) filter (labelled EMVF).¹⁹ The others are those produced by a 1-sided HP (1600) filter and by the LRRO filter (which we introduce below, in Section 3). We can see from that figure that all three methods produce gaps of roughly the same amplitude, and that there is a tendency for the three series to rise and fall at similar times. While all three series show negative output gaps (i.e. excess supply) in the early 1990s, the LRRO and the HP show the economy returning to potential after a few years, while the EMVF shows large output gaps remaining through the end of the sample (1996Q3.) However, as we see in the next section, this last difference is more a reflection of the differences in the structural information used.

2.2 The production function approach

Another important feature of the EMVF filter is that rather than directly filtering output, output is decomposed into a number of components which are then individually filtered. This allows for a more direct link to sources of structural information as well as for an easier interpretation of the source of changes in the gap or potential.

The decomposition is based on an aggregate Cobb-Douglas constant-returns-to-scale production function:

$$Y = QN^{\alpha}K^{1-\alpha} \quad (24)$$

where Q is total factor productivity, N is labour, K the capital stock, and α is the labour-output elasticity (as well as labour's share of income). With some algebra, we can show that:

$$\mu \equiv \partial Y / \partial N = \alpha Y / N \Rightarrow y = n + \mu - \alpha \quad (25)$$

where lower-cases letter are the logs of the upper-case counterparts. This means that to estimating the trend in output, we instead estimate the trends in employment, the marginal product of labour and the labour-output elasticity and then sum them. One nice feature of the decomposition in equation (25) rather than equation (24) is that it allows us to avoid the problem of trying to estimate reliably the capital stock. We can then use the further decomposition that the log of total employment n is given by:

$$n = Pop + p + \log(1 - u) \quad (26)$$

where Pop is the log of the working-age population, p is the log of the participation rate and u is the rate of unemployment.

¹⁷ They find that specifying the dynamic relationship in levels or first differences has a large effect on the estimated values of λ_{ϵ} .

¹⁸ Côté and Hostland approximate the behaviour of the 1-sided filter by using only the lags from the mid-sample representation of the HP filter. They also obtain much more useful results when they apply the 2-sided filter at the end-of-sample using forecast values for the required leads in the filter.

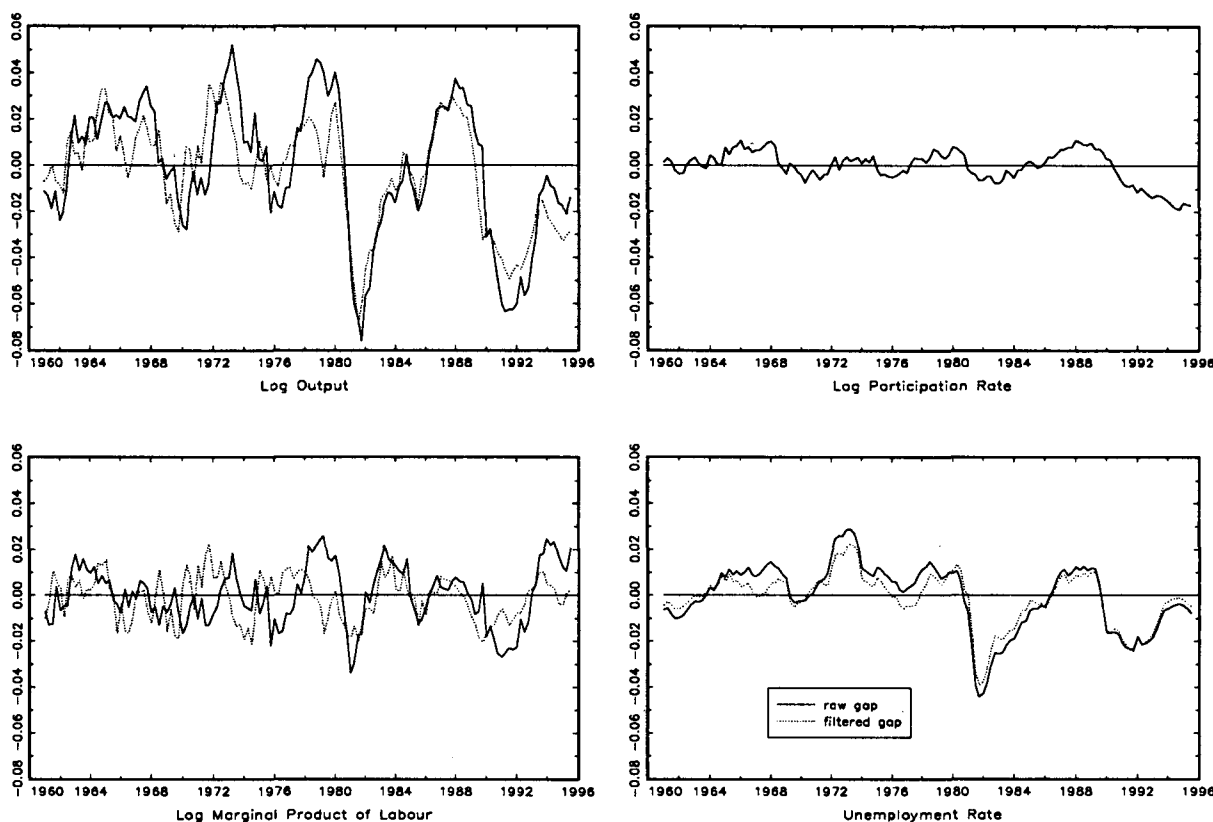
¹⁹ The details of the filter used for the Bank's staff projection of the Canadian economy evolve over time in response to ongoing research and may therefore, as noted in footnote 14, differ slightly from the exposition in Butler (1996). The EMVF gaps shown in the figure reflect the specification used in the staff's December 1996 Economic Projection.

Within this framework, the level of potential output is defined to be the level of output consistent with existing population, trend rates of unemployment and participation and trend levels of the marginal product of labour and the output share of labour. In practice, the trend levels of the participation rate and the output share of labour are determined by a combination of judgement, demographic factors and univariate HP smoothing. Separate MHPF systems are then used to identify trend levels of the marginal product of labour and unemployment. The trend rate for unemployment is estimated with the help of a long-run structural equation due to Côté and Hostland (1996) as well as a price-unemployment Phillips Curve described in Laxton, Rose and Tetlow (1993). Perhaps more novel is the use of a long-run relationship between the marginal product of labour and producer wages as well as an Okun's Law relationship to identify the trend income share of labour.

Analysis of the performance of the EMVF in Butler (1996) shows that this method has its own strengths and weaknesses. On the one hand, he notes that the rolling and full-sample estimates of the NAIRU and the equilibrium marginal product of labour are quite similar, and that the labour market gaps are highly correlated with inflation. On the other hand, he also notes that there is significant correlation in the errors across structural equations, suggesting that there may be further efficiency gains to be had.

Figure 9

Decomposition of the EMVF output gap



To understand the source of the persistent output gap the EMVF produces in the 1990s, we can decompose the output gap into its three components, as shown by the dotted lines in Figure 9. This shows that the filter gives an unemployment rate gap and a marginal product of labour gap that are roughly zero at the start of 1996. Instead, the aggregate output gap largely reflects a deviation of the participation rate from its structural trend level. (Again, note that the participation rate gap is not filtered before it is added to the other components.) However, it would be wrong to attribute the aggregate gap entirely to structural information, as we can see by comparing the filtered output gap (top-right corner, dotted line) with its unfiltered, or "raw", counterpart (same graph, solid line). This

shows that the effects of filtering over the most recent period has tended to increase the estimated size of the output gap by 1-2% of GDP. Since the filtered unemployment rate gap is very close to its unfiltered counterpart, this implies that most of the difference between the filtered and unfiltered output gap is due to the effects of the filter on the marginal product of labour gap.

2.3 The end-of-sample problem

Having seen that both the filter and the structural information play important roles in the EMVF's estimate of recent output gaps, nothing discussed to this point address the end-of-sample problems of HP filters noted in Section 1.4. However, the EMVF contains two novel features intended to modify its end-of-sample behaviour.

First, the EMVF contains an additional growth-rate restriction. If we temporarily ignore the structural information for expositional simplicity, the modified filter solves the problem:

$$\{y_t^g\}_{t=0}^{T+1} = \arg \min \left(\left(\sum_{t=1}^T (y_t - y_t^g)^2 + \lambda_g (\Delta^2 y_{t+1}^g)^2 \right) + \sum_{t=T-j}^T \lambda_{ss} (\Delta \tau_t - \mu_{ss}) \right) \quad (27)$$

where μ_{ss} is a constant equal to the steady-state growth rate of potential output and λ_{ss} is the weight put on the growth rate restriction. The key feature is that λ_{ss} only penalizes deviations from the steady-state growth rate in the last j periods of the sample, effectively "stiffening" the filter. This restriction assumes that the growth rate of potential reverts toward a constant, whereas the theoretical justification for HP filters as optimal filters (noted in Section 1) assumes that this growth rate contains a stochastic trend and therefore will not show any such reversion. Whether such a restriction leads to more accurate estimates of the output gap depends on the accuracy with which one can determine the appropriate value of μ_{ss} .

The second novel feature of the EMVF's treatment of end-of-sample problems is the introduction of a recursive updating restriction. This simply adds an additional term to equation (27), giving:

$$\{y_t^g\}_{t=0}^{T+1} = \arg \min \left(\left(\sum_{t=1}^T (y_t - y_t^g)^2 + \lambda_g (\Delta^2 y_{t+1}^g)^2 \right) + \lambda_{pr} (y_t^g - {}_{pr}y_t^g)^2 + \sum_{t=T-j}^T \lambda_{ss} (\Delta \tau_t - \mu_{ss}) \right) \quad (28)$$

where ${}_{pr}y_t^g$ is the t th element of $\{y_t^g\}_{t=0}^T$. This means that filter is restricted to choose $\{y_t^g\}$ while minimizing the degree to which new observations modify estimates of y^g based on shorter spans of observations. Perhaps not surprisingly, Butler (1996) shows that this gives a 1-sided estimate of the output gap that behaves more like the subsequent 2-sided estimate. While this makes estimates of the output gap behave in a more "orderly" fashion at the end of sample, the net effect on the accuracy of the estimated output gap is unclear.

One way to better understand the effects of these two changes at the end of samples is to compare the resulting 1-sided filter to the 1- and 2-sided HP filters examined in section 1.²⁰ As shown in Figure 10, these modifications cause the EMVF to put much less weight on the last few observations of the sample than the 1-sided HP filter, and overall make its weights more closely resemble those of the 2-sided HP filter. If we look in the frequency domain, however, we see that this change causes the 1-sided EMVF to pass more of the undesired low-frequency or "trend" components than either of the HP filters. In fact, Figure 11 shows that the squared-gain of the filter is greater than 0.2 for all frequencies.

²⁰ The remainder of this section expands the analysis of the EMVF filter properties presented in Butler (1996) to include the effects of the recursive updating restriction.

Figure 10

MA representation of the EMV & HP ($\lambda = 1,600$) filters (as a function of sample position)

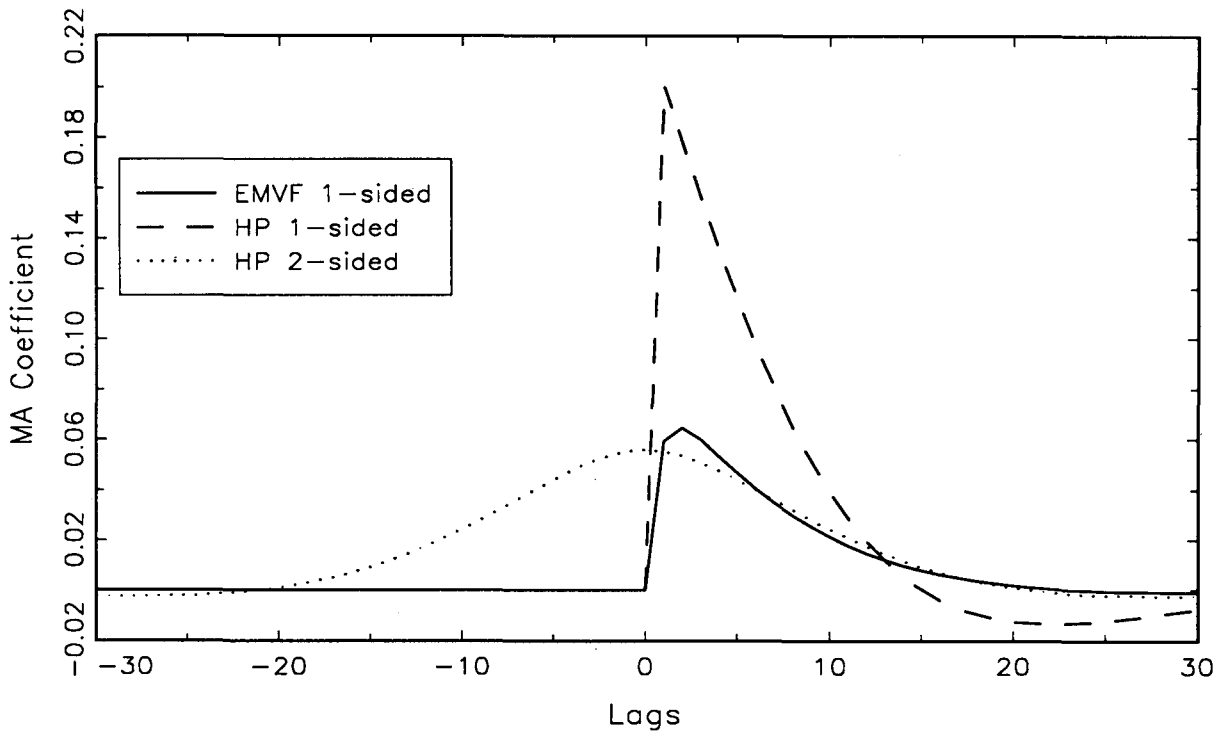


Figure 11

Squared gain of the EMV & HP ($\lambda = 1,600$) filters

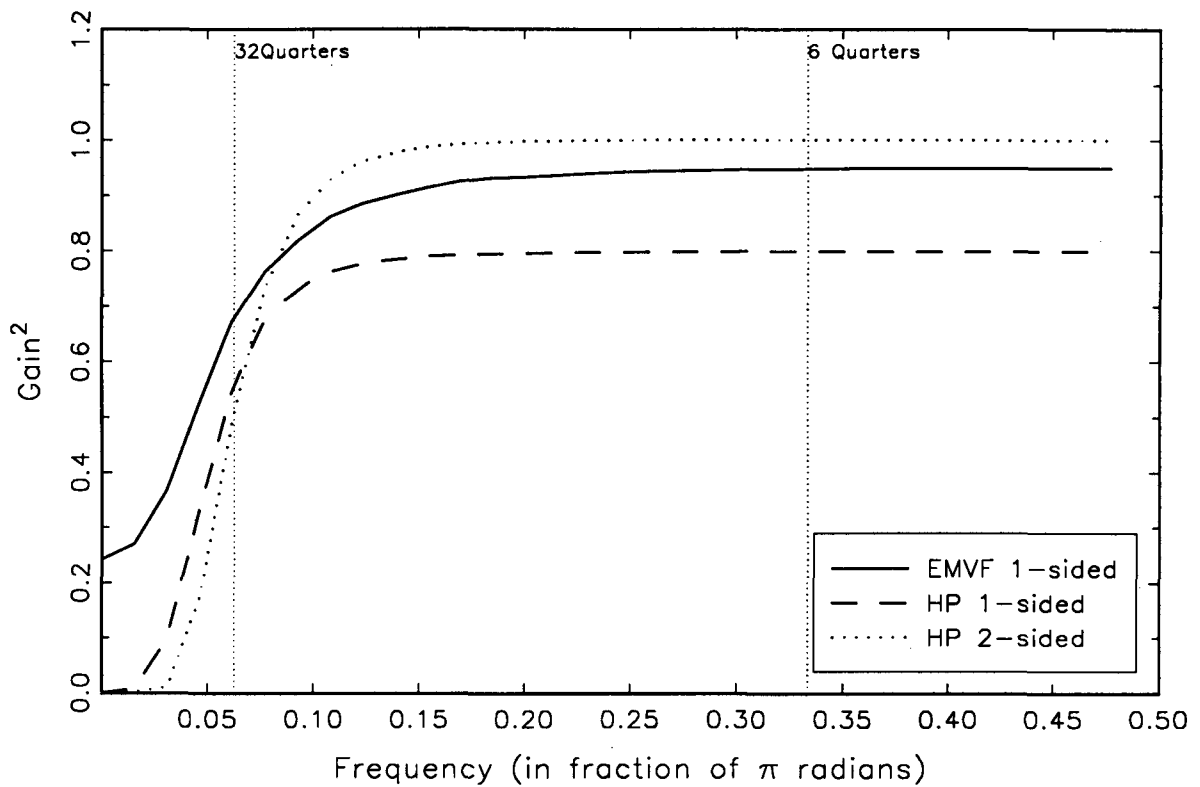


Figure 12

Spectrum of the EMVF gap and its components (Ar (3) fit for 1954Q4 - 1996Q4)

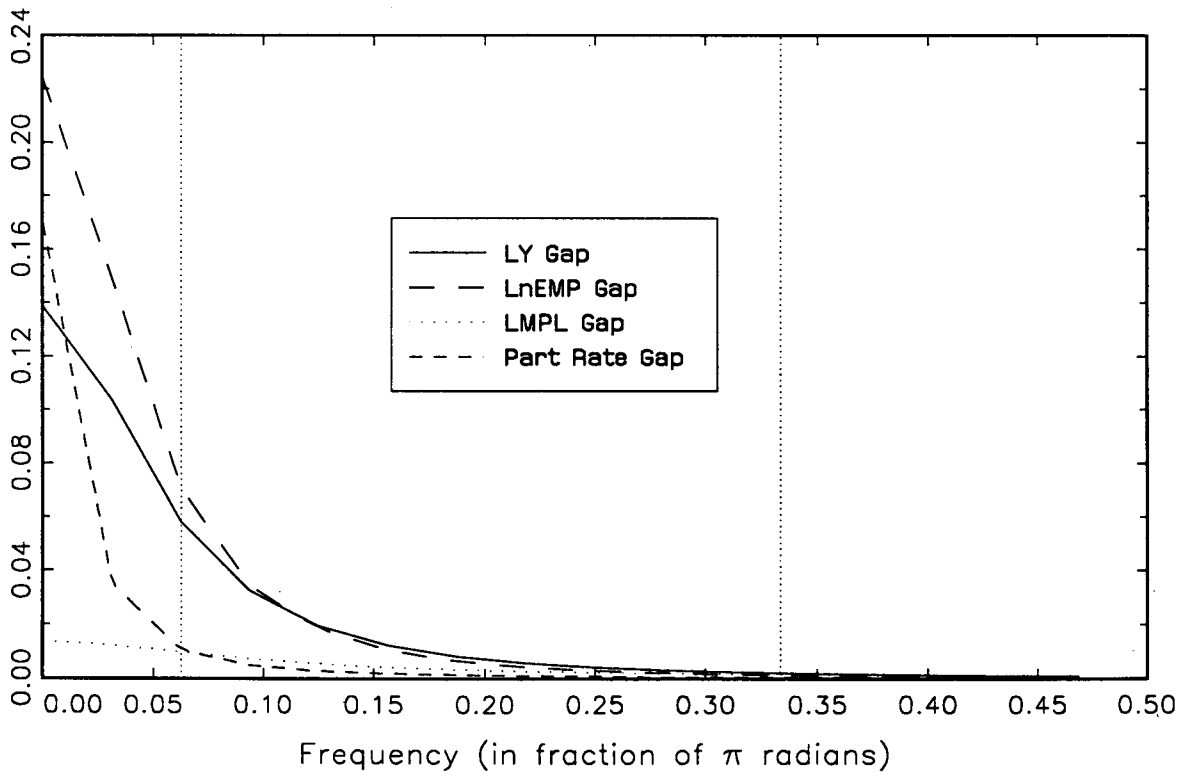
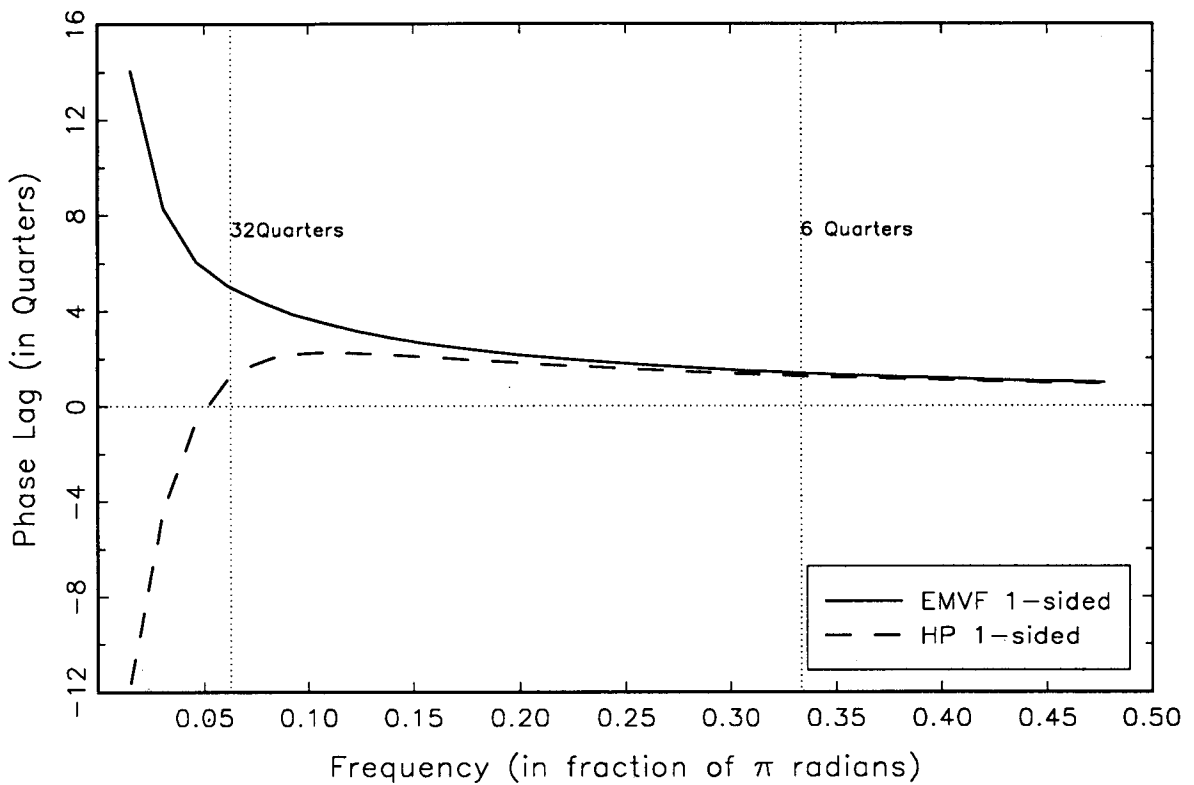


Figure 13

Phase shift of the EMV & HP ($\lambda = 1,600$) filters



The result of this is clearly shown in Figure 12, where we see that both the EMVF output gap and each of its three components appear to be dominated by low-frequency movements not normally associated with business cycles. Compared with Figure 7, it appears that the end-of-sample modifications of the EMVF make the filter much less able to isolate fluctuations at business-cycle frequencies than its simple HP filter counterpart. One way to quantify this effect is to use the estimated spectrum to calculate the correlation of the "measured" EMVF gap with that of an "ideally filtered" gap which perfectly isolated the business cycle frequencies. The result is that the measured EMVF output gap has a correlation with the "ideally filtered" gap of 31.4%, while its two filtered components have correlations of 24.9% (employment gap) and 44.0% (labour productivity gap).

Another implication of the differences in weights between the 1-sided HP and the 1-sided EMVF filter is that, since relatively more of its weight comes from observations with greater lags, the EMVF must have a greater phase shift than the HP filter at the end of sample. This in turn implies that the measured EMVF output gap will tend to lag the true output gap by more than the measured HP output gap. The extent of this difference depends on the frequency of the data series as shown in Figure 13. For all but the lowest of the business-cycle and the sub-business-cycle frequencies, the difference between the two is small and is roughly constant at a lag of about 2 quarters. For lower frequencies, however, the phase shift of HP falls to zero and then becomes negative, while that of the EMVF reaches 5 quarters by the lower bound of the business cycle frequencies and increases rapidly thereafter. If we weight these different phase shift by the relative importance of the different frequencies in measured output gaps, we obtain a weighted average measure of the overall phase lag for these different measures. This gives an overall phase lag of roughly 0 for the 1-sided HP filter, compared to a lag of 3.3 quarters for the EMVF output gap, 3.8 for the EMVF employment rate gap and 2.1 for the EMVF labour productivity gap.

2.4 TOFU

In Section 1, we looked at how well a time-series method (in this case, the HP filter) measured output gaps, and we concluded that by itself it did not produce very reliable estimates for policy makers. In Section 2 so far, we have looked at how adding sources of structural information to HP and related filters could improve the situation. One of the strengths of the MHPF approach is that it clearly states the problem which the resulting estimate of the output gap solves. However, some components of that optimization problem may be easier to accept than others. It seems reasonable to us that the estimated output gap be as consistent as possible with one or more structural economic relationships. The justification for the smoothing portion of the filter is more difficult, as was noted in Section 1, and it is unclear how helpful the particular assumptions of the HP filter are in identifying output gaps in the presence of structural information. Furthermore, as the complexity of the filter increases, the question of how to choose the parameters controlling the filter's behaviour becomes more difficult. While it is conceptually straightforward to estimate the filter parameters jointly with the structural relationships (as in Côté and Hostland (1994)) this can be quite difficult to implement. Accordingly, it would be useful to restate the optimization problem in a way that allows for easier estimation and has a filtering interpretation that is easier to justify.²¹ This leads to an alternative to the HP filter which we refer to as TOFU.

From equation (13) we see that we can estimate the output gap in terms of the observable variable Δy_t if we can identify $\tilde{A}(L)$. Presumably, if we know of an economic relationship that involves the output gap, we could use this to define an optimal estimate of $\tilde{A}(L)$, call it $\hat{A}(L)$. For example, we might consider a Phillips curve of the form:

²¹ This section draws heavily on van Norden (1995).

$$\pi_t = \alpha_0 + \gamma(q_t - y_t) + B(L) \cdot \pi_{t-1} + C(L) \cdot z_t + e_t \quad (29)$$

where z_t is a vector of additional observable variables, e_t is an i.i.d. mean zero error term and $B(L)$, $C(L)$ are one-sided polynomials in non-negative powers of L . We could substitute equation (13) into equation (29) to get:

$$\pi_t = \alpha_0 + \gamma(\tilde{A}(L) \cdot \Delta y_t) + B(L) \cdot \pi_{t-1} + C(L) \cdot z_t + \tilde{e}_t \quad (30)$$

where $\tilde{e}_t = e_t + \gamma \epsilon_t$. Equation (30) can now be estimated by conventional methods to obtain optimal estimates of $\tilde{A}(L)$, since it is now specified entirely in terms of observable variables. This would allow us to estimate $\hat{A}(L) \cdot \Delta y_t$ and use equation (13) to estimate the output gap.²²

We refer to this estimator of the output gap as TOFU; a Trivial Optimal Filter that may be Useful. It is optimal in the sense that estimation by maximum likelihood is straightforward, so our estimates of $\tilde{A}(L)$ will be efficient. The estimator imposes quite general assumptions on the time series properties of the series involved, so the restrictions are hopefully reasonable. It incorporates a simple structural relationship in order to identify the output gap. Furthermore, if we wished to estimate the output gap at the end-of-sample, we can simply replace $\tilde{A}(L)$ with $\bar{A}(L)$ (i.e. use only lagged values of Δy_t). This estimator therefore potentially avoids some of the problem mentioned at the outset of this section.

Of course, the TOFU estimate of the output gap is related to the estimate which would be obtained by "inverting" the structural relationships. The difference is simply that inversion calculates the implicit value of the output gap which would exactly fit the structural relationship. TOFU is therefore a half-way house between such methods and the MHPF methods of Section 2.1. The HP methods are optimal filters only for quite special cases whereas the TOFU methods gives us the optimal linear filter estimate of the output. TOFU offers less smoothing than MHPF methods but more than simple inversion of the structural equations.

Unfortunately, the information gained from structural relationships is contaminated by considerable "noise". Inversion of structural equations is therefore rarely used as a guide to policy, since the resulting estimates of the output gap are usually considered to be too volatile to be of practical use. Whether estimated TOFU filters can reduce this noise enough to be a useful tool for policymakers remains to be seen. If so, they may offer a tractable alternative to the MHPF methods. If not, it suggests that MHPF estimates may be dominated by the arbitrary assumptions they impose on the dynamics of the output gap rather than on the information coming from structural relationships. This suggests looking for other sources of information on these dynamics, which we turn to in the next section.

²² We should note two minor caveats. First, equation (30) identifies $\tilde{A}(L) \cdot \Delta y_t$ only up to the scaling factor γ . Strictly speaking, therefore, we only recover an index of the output gap. This should be reasonable for the purpose of, say, deciding whether interest rates should be higher or lower to achieve a given target, since the current value of the index can be readily compared to its historical values. Second, consistent estimation of this relationship requires an implicit assumption. OLS estimation requires $\text{cov}(e_t, \Delta y_{t-j}) = 0 \quad \forall j$ for consistency. If this condition is not satisfied, then instruments for Δy will be required for estimation. Consistent IV estimation in turn will depend on the assumption that the chosen instruments are valid. This estimator can be extended in a number of ways. If the output gap were to enter the structural equation in a non-linear fashion, we could estimate the system via GMM rather than least-squares techniques. If we had a series of structural equations involving the output gap, we could estimate them simultaneously subject to cross-equation restriction on the coefficients of Δy_t .

3. Using long-run restrictions to estimate the output gap

In this section, we discuss approaches based on long-run restrictions imposed on a VAR. These approaches allow the identification of structural shocks and structural components on the basis of a limited number of economic restrictions imposed on an estimated VAR. The chosen restrictions can be ones that are widely agreed upon in the literature. No arbitrary mechanical filter has to be imposed on the data. Other characteristics distinguishing the methods discussed in this section from those based on mechanical filters are that they do not suffer from obvious end-of-sample problems and that they provide forecasted values of the output gap.

In Section 3.1, we discuss the method based on long-run restrictions imposed on output (LRRO) put forward by Blanchard and Quah (1989), Shapiro and Watson (1988), and King et al. (1991). This method is compared with two VAR-based alternatives: the multivariate Beveridge-Nelson method (MBN) and Cochrane's (1994) method (CO). We argue that one important advantage of the LRRO approach over the MBN and CO approaches is that it allows for the diffusion process of shocks to potential output to be estimated. Section 3.2 considers an application of the LRRO methodology to Canadian data. Many of the arguments used in Sections 3.1 and 3.2 are drawn from Dupasquier, Guay and St-Amant (1997).

In Section 3.3, we present a method involving restrictions imposed on real output and inflation (LRROI) giving an output gap corresponding to that part of the cyclical component of real output which is associated with the trend of inflation. This should be of interest for policymakers concerned about an output gap associated with movements in that trend as opposed to cyclical movements of output unrelated with that trend. This method is discussed in more detail by Lalonde, Page and St-Amant (forthcoming).

3.1 The LRRO, CO and MBN methodologies²³

Let Z_t be a $n \times 1$ stationary vector including a n_1 -vector of $I(1)$ variables and a n_2 -vector of $I(0)$ variables such that $Z_t = \left(\Delta X_{1t}', X_{2t}' \right)'$. By the Wold decomposition theorem, Z_t can be expressed as the following reduced form:

$$Z_t = \delta(t) + C(L)\varepsilon_t \quad (31)$$

where $\delta(t)$ is deterministic, $C(L) = \sum_{i=0}^{\infty} C_i L^i$ is a matrix of polynomial lags, $C_0 = I_n$ is the identity matrix, the vector ε_t is the one-step-ahead forecast errors in Z_t given information on lagged values of Z_t , $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_t') = \Omega$ with Ω positive definite. We suppose that the determinantal polynomial $|C(L)|$ has all its roots on or outside the unit circle, which rules out the non-fundamental representations emphasized by Lippi and Reichlin (1993).

Beveridge and Nelson (1981) show that equation (31) can be decomposed into a long-run component and a transitory component:

$$Z_t = \delta(t) + C(1)\varepsilon_t + C^*(L)\varepsilon_t \quad (32)$$

with $C(1) = \sum_{i=0}^{\infty} C_i$ and $C^*(L) = C(L) - C(1)$. We define $C_1(1)$ as the long-run multiplier of the

²³ See Cogley (1996) for another comparison of the MBN and CO methodologies.

vector X_{1t} . If the rank of $C_1(1)$ is less than n_1 , there exists at least one linear combination of the elements in X_{1t} that is I(0). In other words, there exists at least one cointegration relationship between these variables.

The LRRO approach assumes that Z_t has the following structural representation:

$$Z_t = \delta(t) + \Gamma(L)\eta_t \quad (33)$$

with η_t an n -vector of structural shocks, $E(\eta_t) = 0$, and $E(\eta_t \eta_t') = I_n$ (a simple normalization). We can retrieve the structural form (equation (33)) from the estimated reduced form by using the following relationships: $\Gamma_0 \Gamma_0' = \Omega$, $\varepsilon_t = \Gamma_0 \eta_t$, and $C(L) = \Gamma(L) \Gamma_0^{-1}$.

The long-run covariance matrix of the reduced form is equal to $C(1)\Omega C(1)'$. From equations (31) and (33) we have:

$$C(1)\Omega C(1)' = \Gamma(1)\Gamma(1)' \quad (34)$$

This relation suggests that we can identify the matrix Γ_0 with an appropriate number of restrictions on the long-run covariance matrix of the structural form. Blanchard and Quah (1989) and Shapiro and Watson (1988) use long-run restrictions to identify shocks with $C(1)$ having full rank. King et al. (1991) work in a context where the rank of $C(1)$ is less than n_1 and use cointegration restrictions.

Let us assume that the log of real output is the first variable in the vector Z_t . It is then equal to

$$\Delta y_t = \mu_y + \Gamma_1^p(L)\eta_t^p + \Gamma_1^c(L)\eta_t^c \quad (35)$$

where η_t^p is the vector of permanent shocks affecting output, η_t^c is the vector of shocks having only transitory effects on output, and $\{\Gamma_1^p(L), \Gamma_1^c(L)\}$ reflect the dynamic effects of these shocks. Potential output growth based on the LRRO method can then be defined as:

$$\Delta y_t^p = \mu_y + \Gamma_1^p(L)\eta_t^p \quad (36)$$

Thus, "potential output" corresponds to the permanent component of output. The part of output due to purely transitory shocks is defined as the "output gap."

The MBN decomposition defines potential output as the level of real output that is reached after all transitory dynamics have worked themselves out. With reference to equation (32), where real output is the first element of Z_t , we write the following decomposition: $\Delta y_t = \mu_y + C_1(1)\varepsilon_t + C_1^*(L)\varepsilon_t$ (37)

Potential output can be defined as the first two terms on the right-hand side of equation (37):

$$\Delta y_t^p = \mu_y + C_1(1)\varepsilon_t \quad (38)$$

It is thus simply a random walk with drift.

Note that the MBN approach gives an output gap that is sensitive to the choice of variables included in the VAR. In general, the more information that is brought into the VAR, the more important the transitory component will be. This is not the case with the LRRO approach. Adding additional information may or may not add to the importance of the cyclical component.

Cochrane (1994) uses a two-variable VAR including GNP and consumption to identify the permanent and transitory components of GNP. The bivariate representation is augmented with lags of the ratio of consumption to GNP. The permanent income theory implies that consumption is a random walk (for a constant real interest rate). In addition, if we assume that GNP and consumption are cointegrated, then fluctuations in GNP with consumption unchanged must be perceived as transitory. It is on that basis that Cochrane decomposes real GNP into permanent and transitory components. To extract potential output, the errors of the VAR are orthogonalized so that consumption does not respond contemporaneously to GNP shocks.

Cochrane shows that, if GNP and consumption are cointegrated and consumption is a random walk, identification based on the LRRO method and conventional orthogonalization (i.e. a Choleski decomposition) amount to the same thing. Moreover, if consumption is a pure random walk, Cochrane's decomposition corresponds exactly to the Beveridge-Nelson decomposition based on output and consumption.

In order to better compare the LRRO approach with the CO and MBN approaches, let us first write the structural form (equation (33)) in terms of the log of real GDP (y_t) and the log of real consumption (c_t) decomposed between permanent and transitory shocks (we assume that y_t and c_t are cointegrated):

$$\Delta y_t = \mu_y + \Gamma_y^p(1)\eta_t^p + \Gamma_y^{p*}(L)\eta_t^p + \Gamma_y^c(L)\eta_t^c \quad (39)$$

$$\Delta c_t = \mu_c + \Gamma_c^p(1)\eta_t^p + \Gamma_c^{p*}(L)\eta_t^p + \Gamma_c^c(L)\eta_t^c \quad (40)$$

where $\Gamma^p(1)$ is the long-run multiplier of permanent shocks and $\Gamma_c^{p*}(L) = \Gamma^p(L) - \Gamma^p(1)$ is their transitory component. The MBN method considers only the first component of the permanent shocks plus the drift term, i.e. $\mu + \Gamma_y^p(1)\eta_t^p$. The LRRO approach is different in that it also includes the dynamics of permanent shocks to real output ($\Gamma_y^{p*}(L)$) in potential output.

With the CO approach, potential output is constrained to be a random walk to the extent that consumption is a random walk. Indeed, the validity of the permanent income hypothesis would imply that the last two terms of equation (39) are equal to zero and that $\Gamma_y^p(1) = \Gamma_c^p(1)$. It is not clear what the CO decomposition would correspond to if consumption is not a random walk.²⁴

As pointed out by Lippi and Reichlin (1994), modelling the trend in real output as a random walk is inconsistent with most economists' interpretation of productivity growth. Indeed, it is

²⁴ Stochastic growth models – such as in King et al. (1988) or King et al. (1991) – imply that the ratio of the log of GNP to the log of consumption is stationary but that consumption is not a random walk because the real interest rate is not constant. In these models, the transitory component of permanent shocks to consumption is not equal to zero. The LRRO decomposition is compatible with the predictions of these models.

generally believed that technology shocks are absorbed gradually by the economy. Adjustment costs for capital and labour, learning and diffusion processes, habit formation, and time to build imply richer dynamics than a random walk for these shocks. Again, a decisive advantage of the LRRO approach is that it lets the data determine the shape of the diffusion process of permanent shocks.²⁵

One implication of defining potential output as a random walk with drift is that, when the contemporary effect of a positive permanent shock is smaller (greater) than its long-run effect, the output gap, defined as observed output minus potential, is negative (positive). For example, a positive technological shock whose short-term impact is smaller than its long-term impact will cause a transitory negative output gap. Many researchers and/or policymakers will find that this feature of the MBN and CO approaches (in the latter case under the assumption that consumption is a random walk) reduces their attractiveness. It will often appear preferable to include the diffusion process associated with permanent shocks in potential output since the economy is likely to remain on its production possibility frontier as adjustments unfold. There should be no reason for the trend of inflation to change during that adjustment process.

3.2 An application of the LRRO approach to Canadian data

For our applications of the LRRO methodology to Canadian data, we assume that the growth rate of real output (Δy_t) follows a stationary stochastic process responding to two types of structural shocks: permanent (ε_p) and transitory (ε_T). Also included in the estimated VARs are the first differences of, respectively, inflation ($\Delta\pi$), the unemployment rate (Δu) and the real interest rate (Δr). We assume that these series are $I(0)$ and that there is no cointegration involved.²⁶

Adding money or the exchange rate to the estimated VARs would have little impact on the results. In selecting the variables, one hopes to include in the VAR the information necessary to identify the structural components of interest. Of course, there is a cost in adding information in terms of lost degrees of freedom and less precise estimates.

It is important to note that the assumptions made on the level of integration of the series could be changed and that they are not part of the LRRO methodology per se (except for real output which has to be $I(1)$).²⁷ Changing these assumptions could have a significant impact on the results.

Our objective is to illustrate the methodology. A practitioner interested in using the approach to estimate the output gap might well choose to use additional information and be able to obtain a better estimate of the output gap.

The structural shocks and the variables used in the VAR can be expressed in the following vector form:

$$\eta_t = \begin{bmatrix} \varepsilon_p \\ \varepsilon_{T_1} \\ \varepsilon_{T_2} \\ \varepsilon_{T_3} \end{bmatrix} \text{ and } Z_t = \begin{bmatrix} \Delta y \\ \Delta\pi \\ \Delta u \\ \Delta r \end{bmatrix} \quad (41)$$

²⁵ Kuttner (1994) proposes a method based on the univariate unobserved stochastic trend decomposition of Watson (1986) augmented with a Phillips curve equation. As with the Beveridge-Nelson decomposition, Kuttner's approach constrains potential output to follow a random walk process.

²⁶ Unit root tests support these assumptions. Results are available on request.

²⁷ Indeed, DeSerres, Guay and St-Amant (1995) assume that money growth is $I(0)$.

We use quarterly data on real GDP. Our measure of inflation is the total CPI. The real interest rate is proxied by the overnight rate (see Armour et al. (1996) for a discussion of that series) minus inflation (quarterly growth rate). Our sample extends from the first quarter of 1970 to the fourth quarter of 1996 so that we focus on the flexible exchange rate period. The estimation of the output gap based on the LRRO methodology is fairly robust to the choice of the sample period, however.

The autoregressive reduced-form of the model is first estimated:

$$Z_t = \sum_{i=1}^q \Pi_i Z_{t-i} + e_t \quad (42)$$

with q the number of lags and e_t a vector of estimated residuals with $E(e_t e_t') = \Sigma$.

It is crucial that the estimated VARs include enough lags. Indeed, Monte Carlo simulations by DeSerres and Guay (1995) show that using a too parsimonious lag structure can significantly bias the estimation of the structural components. We decided to use 8 lags. However, we verified that using 6 or 10 lags had little impact on the results.

The LRRO approach involves the identification of structural shocks (ϵ_t) from reduced-form shocks (e_t) and their variance. For this, we need to provide enough identifying restrictions to evaluate the 16 elements in Γ_0 . Given that Σ is symmetric, we need to impose 6 additional restrictions. The matrix of long-run effects of reduced-form shocks, $C(1)$, is related to the equivalent matrix of structural shocks, $\Gamma(1)$, as follows:

$$\Gamma(1) = C(1)\Gamma_0 \quad (43)$$

where the matrix $C(1)$ is calculated from the estimated VAR. To identify the system we simply impose the condition that $\Gamma(1)$ is triangular, i.e. three shocks have no long-run effect on real output and two have no long-run effect on inflation. There are then three transitory components of real output which we do not need to identify separately.

Figure 14 presents the impulse response of real output to a one standard deviation permanent output shock. (Blanchard and Quah suggest interpreting this as an aggregate supply shock.) The horizontal axis represents the number of years. Confidence intervals were generated using Monte Carlo simulations in RATS with 1,000 replications.

The important message of Figure 14 is that permanent shocks are characterized by statistically significant dynamics, i.e., potential output has richer dynamics than a simple random walk.²⁸ As mentioned above, this could be due to factors such as adjustment costs on capital and labour, learning, habit formation, or time to build. One implication of the rejection of the random walk assumption is that methods which do not take the diffusion process of permanent shocks into account could miss an important part of potential output. Indeed, Dupasquier, Guay and St-Amant (1997) show that the correlation between the output gaps calculated on the basis of the LRRO, CO and MBN approaches applied to US data is relatively small. Part of this can be attributed to the different treatments of the diffusion process of permanent shocks

Figure 15 shows the output gaps calculated on the basis of the LRRO methodology as it is applied in this paper, together with 90 and 67% confidence intervals.

²⁸ Dupasquier, Guay and St-Amant (1997), Blanchard and Quah (1989) and Gali (1993), among others, report similar results for the US economy.

Figure 14

Response of real GDP to a permanent output shock

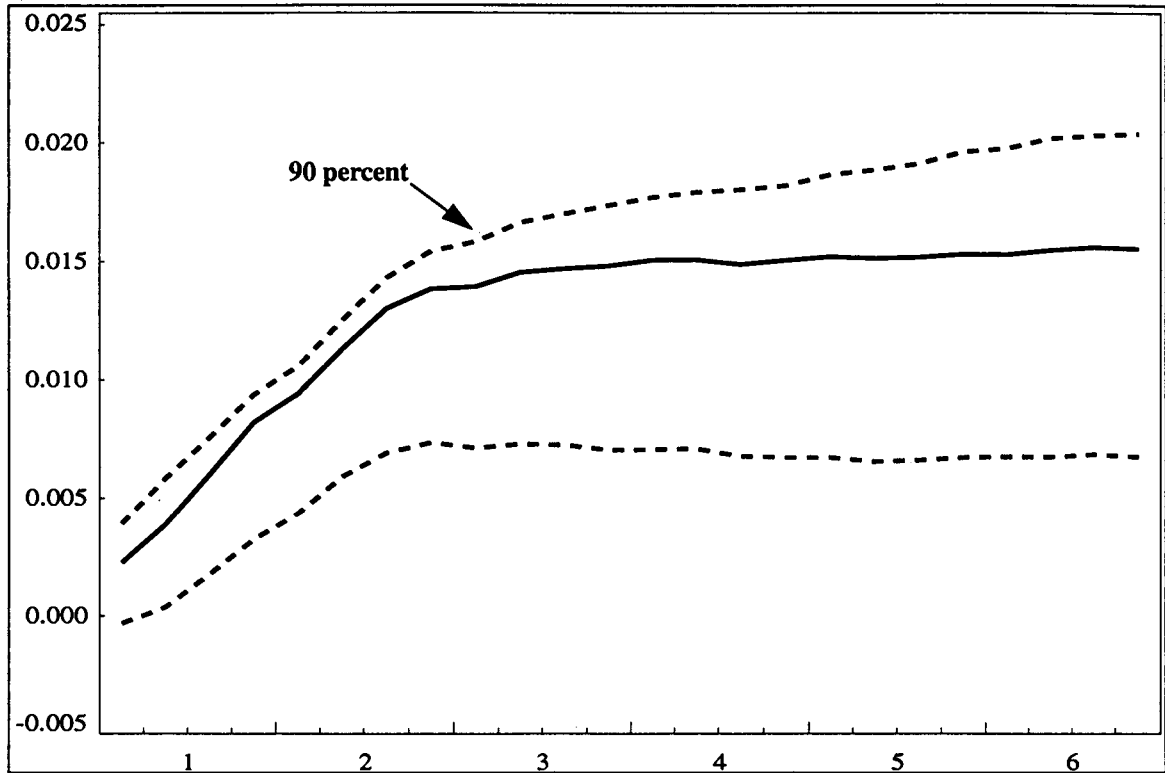
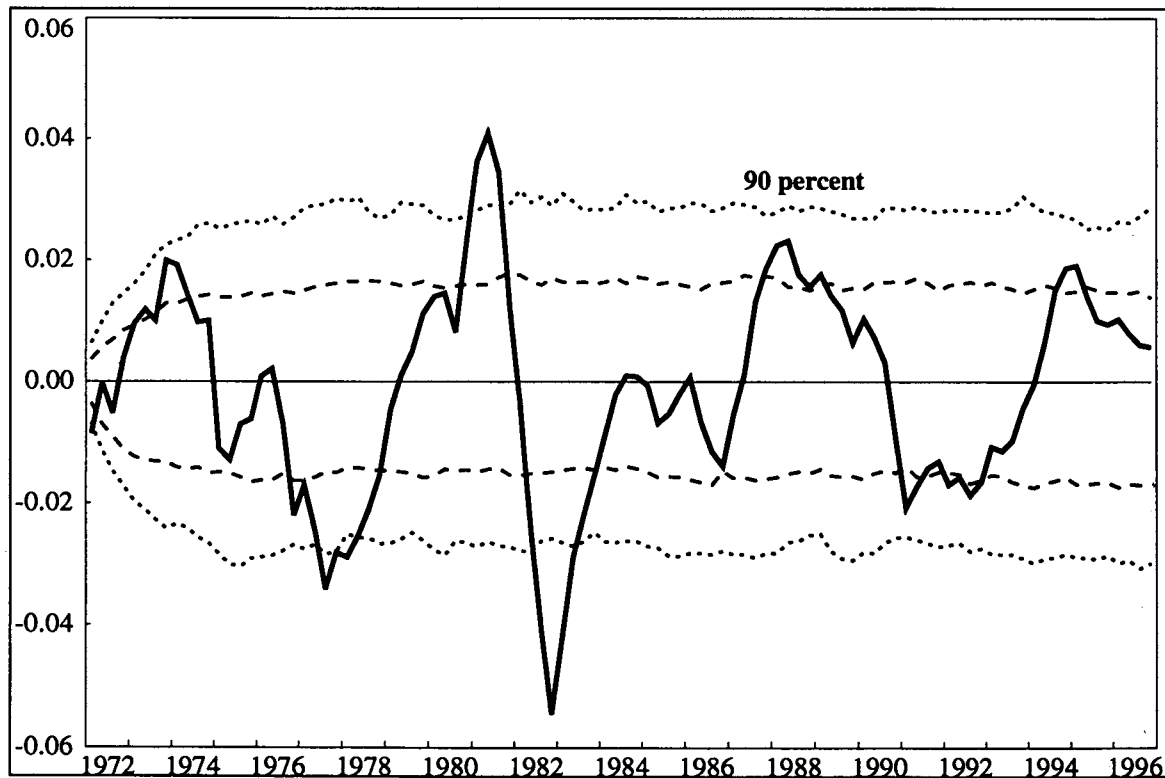


Figure 15

LRRO-based Canadian output gap



One important message of Figure 15 is that there is a high degree of uncertainty surrounding the estimation of the output gap. Dupasquier, Guay and St-Amant (1997) reach the same conclusion using US data. Staiger, Stock and Watson (1996), using a different methodology, arrive at a similar conclusion concerning the estimation of the NAIRU for the United States. Some of that uncertainty is attributable to the large number of lags that have to be included in the estimated VARs. As mentioned above, DeSerres and Guay (1995) show that many lags have to be used to provide an unbiased decomposition into permanent and transitory components with structural VARs. To a large extent, the purpose of these lags is to approximate the moving-average part of the underlying DGP. Preliminary results obtained at the Bank of Canada suggest that the estimation of VARMA's instead of VARs could reduce parameter uncertainty by allowing the use of more parsimonious models.

Still, there are episodes of significant output gaps at either the 90% level or the 67% level. These output gaps appear reasonable in that positive output gaps are associated with episodes of accelerating inflation while negative output gaps correspond to episodes of decelerating inflation. Interestingly, we find that the Canadian economy was either at potential or in slight excess demand in the mid-1990s.

It should be of interest to see whether the structural shocks used to calculate the output gap account for an important part of the variance of inflation at different horizons. That information is presented in Table 2.

Table 2
Variance decomposition of Canada's inflation (LRRO method)
 Relative contribution of the different types of shocks, in percentage points

Horizon (quarters)	Permanent output shock	Transitory output shock
1	11 (0-43)*	89 (57-100)
4	8 (2-36)	92 (64-98)
8	6 (4-30)	94 (69-97)
16	13 (5-49)	87 (51-95)
32	23 (5-66)	77 (34-95)
long-term	37 (3-84)	63 (16-97)

* 90% confidence interval.

We can see that transitory shocks affecting real output account for a large part of the variance of Canadian inflation. This result suggests that the component of output that we identify includes much of the information one would want to include in the output gap. Note however that there appears to be some additional information related to permanent shocks, especially at longer horizons. Lalonde, Page and St-Amant (forthcoming) discuss this and propose another approach which does not impose that the output gap is part of the cyclical component of output.

Instead, in this paper we keep our focus on methods imposing that the output gap is stationary. Although Table 2 suggests that the output gap based on the LRRO approach accounts for a large fraction of fluctuations in the trend of inflation, it may also be that a part of that gap is unrelated to the trend. For example, it might include very-high-frequency cycles having little to do with the trend. This leads us to another method which involves imposing restrictions on both real output *and* inflation in order to produce an output gap that is constrained to be associated with movements in the trend of inflation.

3.3 The LRROI approach

It is possible, with the four-variables system presented in Section 3.2, to identify shocks that have no long-run effects on real output but affect the trend of inflation. Real output was indeed decomposed in the following way:

$$\Delta y_t = \mu_y + \Gamma_y^p(1)\eta_t^p + \Gamma_y^{p*}(L)\eta_t^p + \Gamma_y^{cp}(L)\eta_t^{cp} + \Gamma_y^{cc}(L)\eta_t^{cc} \quad (44)$$

with certain shocks having no long-run effect on real output but changing the trend of inflation (η_t^{cp}), certain shocks having no long-run effect on real output *and* inflation (η_t^{cc}),²⁹ and certain shocks having long-run effects on real output but whose effect on inflation is left unconstrained

(η_t^p). The component $\Gamma_y^{cp}(L)\eta_t^{cp}$ could be used as a measure of the output gap. This is what we call the LRROI method. In Section 3.2 we were not interested in distinguishing $\Gamma_y^{cp}(L)\eta_t^{cp}$ from $\Gamma_y^{cc}(L)\eta_t^{cc}$ so that we simply added these components to form one component that was the LRRO output gap.

The LRROI method gives a measure of the output gap that is more constrained than the LRRO method in that it combines restrictions on real output *and* inflation. While the LRRO approach only requires that real output is I(1), the additional assumption that inflation is better characterized as being I(1) over the sample is a necessary feature of the LRROI approach. Of course, that assumption is not uncontroversial. However, we think it is reasonable. Assuming that inflation is I(0) is in some sense equivalent to assuming that, whatever the actions of the monetary authorities, it has to go back to a constant mean. We do not think this is realistic. At least, inflation has to be modelled as a process, the mean of which changes through time. One way to do it is to assume that it is I(1).

The LRROI method proposes a measure of the output gap that should be attractive for policymakers who focus on that part of the cyclical component of real output that is associated with movements in the trend of inflation as opposed to short-run fluctuations of that series. Instead, the LRRO method provides a measure of the output gap including cyclical movements that may not be of interest to such policymakers. In some sense, the LRROI method could provide policymakers with a less "noisy" indicator of changes in the trend of inflation. Short-run fluctuations of inflation could be caused by factors such as transitory fluctuations in the exchange rate or changes of indirect taxes which may not require any change in the policy stance. A caveat to this is of course that there might be instances when policymakers will be interested in reacting to fluctuations of output associated with transitory changes of inflation. Unlike the LRRO gap, the LRROI gap would not include the effect of such shocks.

Figure 16 presents the LRROI output gap calculated on the basis of the Canadian data used in Section 3.3. Figure 17 compares that output gap with the one resulting from the LRRO approach to the gap from the one-sided HP filter. Table 3 presents the variance decomposition of inflation corresponding to the LRROI decomposition.

We can see in Figure 16 that there is as much uncertainty surrounding the estimation of the LRROI output gap as there is with the LRRO gap. We also see that the LRROI gaps are generally smaller than those produced by the LRRO method. This indicates that some cycles included in the LRRO gap are not related to the trend of inflation. Finally, the variance decomposition of inflation suggests that the LRROI approach includes most of the information related to the cyclical component of output that is relevant for monitoring or forecasting medium-and-long-term inflation. However, the

²⁹ Since there are two shocks which have no long-run impact on the trend of output or inflation (the last two of equation (44)), we simply consider their sum.

Figure 16
LRROI-based Canadian output gap

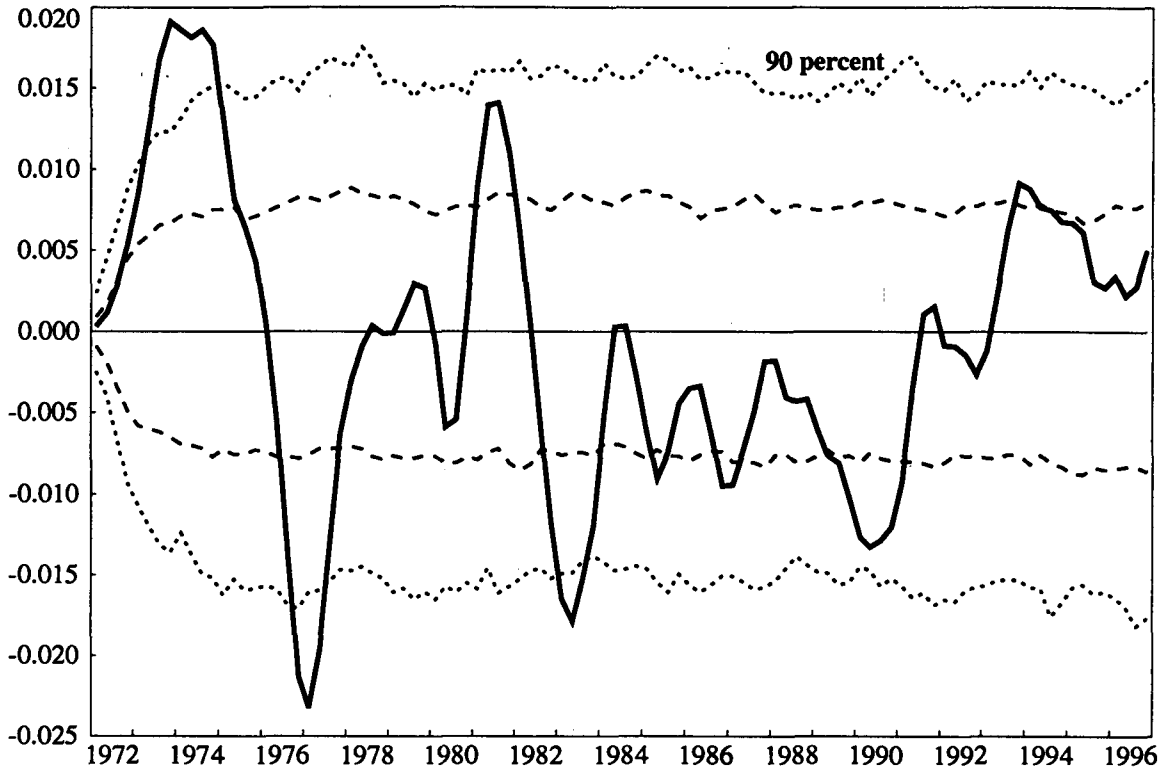


Figure 17
LRRO, LRROI and one-sided HP-based Canadian output gaps

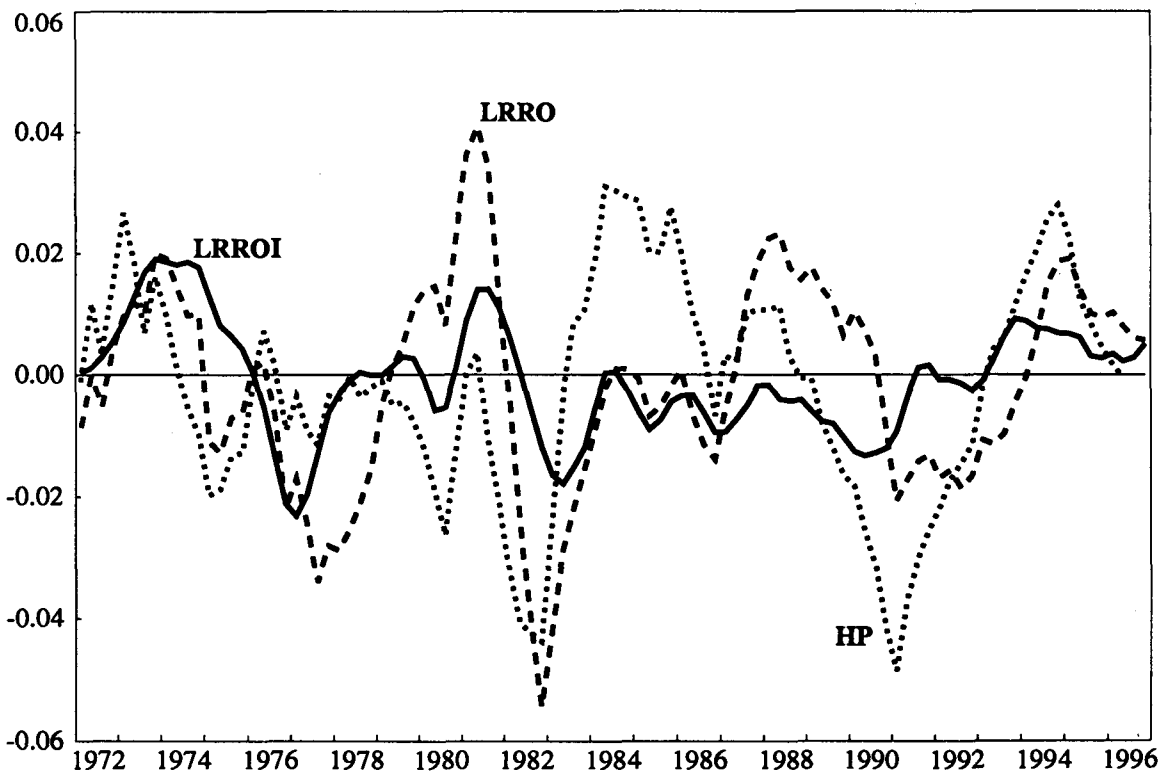


Table 3

Variance decomposition of Canada's inflation (LRROI method)
Relative contribution of the different types of shocks, in percentage points

Horizon (quarters)	Permanent output shock	Transitory output shocks affecting the trend of inflation	Transitory output shocks that do not affect the trend of inflation
1	11 (0-43)*	54 (6-88)	35 (3-81)
4	8 (2-36)	63 (18-85)	29 (7-68)
8	6 (4-30)	70 (33-80)	24 (11-53)
16	13 (5-49)	72 (34-80)	15 (8-35)
32	23 (5-66)	68 (27-83)	9 (4-23)
long-term	37 (3-84)	61 (16-95)	1 (0-4)

* 90% confidence interval.

LRRO gap appears to include non-negligible information about movements of inflation lasting as much as a couple of years.

To conclude this section we would like to emphasize the fact that the LRRO and LRROI methods are two variants of a more general VAR-based approach. Other variants could be considered. For example, one might want to take into account possible non-linearities in the relationship between output and inflation (e.g. Fillion and Léonard (1997)). Evans (1992) proposes a time-varying method allowing for such non-linearities. One might also want to include different sets of restrictions, including cointegration and bayesian types of restrictions, to ensure that the estimated gap is compatible with some specific macroeconomic model.

Summary

Much research has gone into measuring output gaps and much still remains to be done. On the basis of the results discussed, we think that there are three main lessons to be learned.

1. Univariate time-series methods, such as the HP filter, are not a reliable way to measure the output gap

Univariate methods rely on an arbitrary decomposition of a series into a trend and a cyclical component. However, changing the decomposition method can make large differences to the measured output gap, and economic theory usually has little or nothing to say about which method should be favoured. In addition, the causal relationship between potential and the gap can limit the information we can hope to gain about the current gap, and some popular economic models imply that univariate filters will never be able to give much information.

In the particular case of the HP filter, although this filter is thought to be close to an ideal high-pass filter, it does not accurately measure the components from business-cycle frequencies when series have the typical Granger shape. Although the HP filter can also be justified as an optimal filter

for particular cases, these cases do not appear to be realistic approximations of output and the filter is generally not a reliable way to estimate the "cyclical" component. HP filters also behave very differently at the end of the sample, which is what policymakers care most about, and little is known about the trade-off between phase shift and smoothness at the end of sample.

2. Existing hybrid methods that combine univariate dynamic methods and structural relationships are not a panacea

In practical terms, existing hybrid methods have proved to be hard to estimate. Moreover, they may not be robust to alternative reasonable calibrations and it is difficult to calculate their appropriate confidence intervals. In the case of the EMVF, which is the filter used to estimate the output gap in the Bank of Canada's staff projections, its ability to isolate business cycle frequencies is worse than that of the HP filter and its estimates lag the true output gap by just under a year. These problems seem to be the result of features which were introduced to improve the filter's estimate of current and recent output gaps.

More generally, the "hybrid" approach is not only driven by a desire to include structural relationships, but also by a pragmatic desire for "smooth" estimates of the output gap. However, if the structural relationships are very informative then the "smoothing" assumptions may be unnecessary. Otherwise, it is hard to argue that these assumptions will be innocuous, for the reasons we mentioned above in the context of univariate methods. For that reason, and because they can incorporate the same sources of structural information, further investigation of TOFU methods may provide a good benchmark for hybrid methods. If TOFU estimates of the output gap are useful, then the strict (or ad hoc) filters used in existing hybrid methods are not required. If TOFU methods are not useful, then structural information alone is not sufficient to identify the output gap. Relying instead on ad hoc dynamic assumptions for identification raises questions about the reliability and significance of the resulting estimated output gaps.

3. Methods combining estimated dynamics with structural information offer an interesting alternative which needs to be further explored

Examples of these methods explored in this paper are VARs with long-run restrictions suggested by economic theory. Their advantages include an absence of arbitrary dynamic assumptions, straightforward estimation, and an ability to estimate both current and expected future output gaps. On the other hand, it is not always clear which variables have to be included in the VAR and work to date suggests that the estimated output gaps have wide confidence intervals comparable to those of other methods. More work is needed to evaluate the extension of these methods to VARMA models. It might also be of interest to explore VARs with time-varying parameters (as is proposed by Evans (1992)) in order to deal with possible non-linearities in structural or dynamic relationships.

The VAR-based methods considered in this paper impose relatively little economic structure on the data and allow the dynamics properties of the estimated output gaps to be data-determined. This is an advantage insofar as economic theory provides little guidance on what the dynamics should be. Of course, one has to assume that the estimated VAR includes the information relevant for the identification of the output gap.

On the other hand, there might be instances when practitioners would want to impose more economic structure on their estimation of the output gap by incorporating more economic relationships. To some extent, this could be accommodated in the VAR (or VARMA or VECM) framework. Many types of economic or statistical restrictions have been discussed in the literature including bayesian priors and cointegration relationships. Such restrictions could be used to ensure that the estimated output gap is broadly compatible with some structural model of interest.

However, the TOFU approach might provide a better framework for those interested in imposing detailed economic relationships on the data (such as a specific Phillips curve or a NAIRU equation). One advantage of using detailed economic structure is that the derived measures of potential output can then be embedded in a model that is consistent with that structure. As noted briefly in van Norden (1995), extensions of the TOFU approach to multiple detailed structural relationships appears to be straightforward, although this remains to be explored in detail. It should be kept in mind, however, that the usefulness of TOFU output gaps would depend on the validity of the economic structure imposed on the data and that economics provides very few non controversial structural relationships.

In our view, the VAR-based and TOFU approaches deserve further research, while our findings suggest that one should avoid univariate methods and move away from arbitrary smoothing methods wherever practicable. Variants of the VAR-based and TOFU approaches could be considered and their usefulness in terms of monitoring and projecting inflation evaluated.

Bibliography

- Baxter, M. and R. G. King (1995): "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series". *Working Paper*, No. 5022. National Bureau of Economic Research.
- Beveridge, S. and C. R. Nelson (1981): "A New Approach to the Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle". *Journal of Monetary Economics*, No. 7, pp. 151-74.
- Blanchard, O. J. and S. Fischer (1989): *Lectures on Macroeconomics*, MIT Press, Cambridge (Massachusetts).
- Blanchard, O. J. and D. Quah (1989): "The Dynamic Effect of Aggregate Demand and Supply Disturbances". *American Economic Review*, No. 79, pp. 655-73.
- Boschen, J. and L. Mills (1990): "Monetary Policy with a New View of Potential GNP". *Federal Reserve Bank of Philadelphia Business Review*, pp. 3-10, July-August.
- Burns, A. M. and W. C. Mitchell (1946): *Measuring Business-Cycles*. New York: National Bureau of Economic Research.
- Butler, L. (1996): "A Semi-Structural Method to Estimate Potential Output: Combining Economic Theory with a Time-Series Filter". *Technical Report*, No. 77, Bank of Canada.
- Cochrane, J. H. (1994): "Permanent and Transitory Components of GNP and Stock Prices". *Quarterly Journal of Economics*, No. 61, pp. 241-65.
- Cogley, T. (1996): "Evaluating Non-Structural Measures of the Business Cycle". Draft Paper, Federal Reserve Bank of San Francisco.
- Cogley, T. and J. Nason (1995): "Effects of the Hodrick-Prescott filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research". *Journal of Economic Dynamics and Control*, No. 19, pp. 253-78.
- Côté, D. and D. Hostland (1994): "Measuring potential output and the NAIRU as unobserved variables in systems framework", in *Economic Behaviour and Policy Choice Under Price Stability*, Bank of Canada, pp. 357-418.
- DeSerres, A. and A. Guay (1995): "The Selection of the Truncation Lag in Structural VARs (or VECMs) with Long-Run Restrictions". *Working Paper*, No. 95-3, Bank of Canada.

- DeSerres, A., A. Guay and P. St-Amant (1995): "Estimating and Projecting Potential Output using Structural VAR Methodology: the Case of the Mexican Economy". *Working Paper*, No. 95-2, Bank of Canada.
- Dupasquier C., A. Guay and P. St-Amant (1997): "A Comparison of Alternative Methodologies for Estimating Potential Output and the Output Gap". *Working Paper*, No. 97-5, Bank of Canada.
- Dupasquier. C., R. Lalonde and P. St-Amant (1997): "Les zones monétaires optimales: une application au Canada et aux États-Unis". Forthcoming.
- Engle, R. F. and C. W. J. Granger (1987): "Cointegration and Error Correction: Representation, Estimation and Testing". *Econometrica*, No. 55 (March), pp. 251-76.
- Evans, G. and L. Reichlin (1994): "Information, Forecasts, and Measurement of the Business Cycle". *Journal of Monetary Economics*, No. 33, pp. 233-54.
- Evans, M. D. D. (1992): "The Changing Nature of the Output-Inflation Trade-off". NYU Stern School of Business, *Working Paper*, No. 92-17.
- Gali, J. (1992): "How well does the IS-LM Model fit Postwar US Data?" *Quarterly Journal of Economics*, No. 107, pp. 709-38.
- Giorno, C., P. Richardson, D. Reveare and Paul van den Noord (1995): "Estimating Potential Output, Output Gaps and Structural Budget Balances". OCED Economic Department, *Working Paper*, No. 152.
- Gouriéroux, C. and M. Monfort (1995): "Séries Temporelles et Modèles Dynamiques". *Economica*.
- Guay, A. and P. St-Amant (1996): "Do Mechanical Filters Provide a Good Approximation of Business Cycles?" Bank of Canada, *Technical Report*, No. 78.
- Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press.
- Harvey, A. C. and A. Jaeger (1993): "Detrending, Stylized Facts and the Business Cycle". *Journal of Applied Econometrics*, No. 8, pp. 231-47.
- Henderson, R. (1924): "A New Method of Graduation". *Actuarial Society of America Transactions*, No. 25, pp. 29-40.
- Hendry, S. (1995): "Long-Run Demand for M1". *Working Paper*, No. 95-11, Bank of Canada.
- Hodrick, R. J. and E. Prescott (1997): "Post-war U. S. Business-Cycles: An Empirical Investigation". *Journal of Money Credit and Banking*, forthcoming.
- Jeong, J. and G. S. Maddala (1993): "A Perspective on Application of Bootstrap Methods in Econometrics", in G. S. Maddala, C. R. Rao and H. D. Vinod (eds.), *Handbook of Statistics*, Vol. 11, pp. 573-610, New York: North-Holland.
- King, R. G., C. I. Plosser and S. T. Rebelo (1988): "Production Growth and Business Cycles: II. New Directions". *Journal of Monetary Economics*, No. 21(2-3), pp. 309-41.
- King, R. G., G. H. Plosser, J. H. Stock and M. W. Watson (1991): "Stochastic Trends and Economic Fluctuations". *American Economic Review*, No. 81 (September), pp. 819-40.
- King, R. G. and S. Rebelo (1993): "Low Frequency Filtering and Real Business-Cycles". *Journal of Economic Dynamics and Control*, No. 17, pp. 207-31.
- Kuttner, K. N. (1994): "Estimating Potential Output as a Latent Variable". *Journal of Business and Economic Statistics*, No.12, pp. 361-68.
- Lalonde, R., J. Page and P. St-Amant (1997): "Trois méthodes d'estimation de l'écart de production reposant sur des restrictions de long terme". Forthcoming.
- Laxton, D. and R. Tetlow (1992): "A Simple Multivariate Filter for the Measurement of Potential Output". *Technical Report*, No. 59. Ottawa: Bank of Canada.

- Laxton, D., D. Rose and R. Tetlow (1993): "Is the Canadian Phillips Curve non-linear?" Bank of Canada, *Working Paper*, No. 93-7.
- Lippi, M. and L. Reichlin (1993): "The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment". *American Economic Review*, No. 83, pp. 644-52.
- Lippi, M. and Reichlin (1994): "Diffusion of Technical Change and the Decomposition of Output into Trend and Cycle". *Review of Economic Studies*, No. 61, pp. 19-30.
- Nelson, C. R. and C. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series". *Journal of Monetary Economics*, No. 10, pp. 139-67.
- Priestly, M. (1981): *Spectral Analysis and Time Series*. Academic Press.
- Quah, D. (1992): "The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds". *Econometrica*, No. 60(1), pp. 107-18.
- Reinsel, G. C. and S. H. Ahn (1988): "Asymptotic Distribution of the Likelihood Ratio Test for Cointegration in the Nonstationary VAR Model". *Technical Report*, University of Wisconsin, Madison.
- Runkle, D. E. (1987): "Vector Auto-regression and Reality". *Journal of Business and Economic Statistics*, No. 5, pp. 437-54.
- Sargent, T. J. (1987): *Macroeconomic Theory*. Academic Press.
- Schwert, G. W. (1987): "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data". *Journal of Monetary Economics*, No. 20, pp. 1037-87.
- Shapiro, M. D. and M. W. Watson (1988): "Sources of Business Cycle Fluctuations". *NBER Working Paper*, No. 2589.
- Singleton, K. (1988): "Econometric Issues in the Analysis of Equilibrium Business-Cycle Models". *Journal of Monetary Economics*, No. 21, pp. 361-86.
- Staiger, D., J. Stock and M. W. Watson (1996): "How Precise are Estimates of the Natural Rate of Unemployment?" Presented at the *Monetary Policy and Low Inflation Conference*, National Bureau of Economic Research.
- van Norden, S. (1995): "Why is it so Hard to Measure the Current Output Gap?" *Mimeo*, Bank of Canada.
- Watson, M. W. (1986): "Univariate Detrending Methods with Stochastic Trends". *Journal of Monetary Economics*, No. 18, pp. 49-75.
- Watson, M. W. (1993): "Vector Autoregressions and Cointegration". *Handbook of Econometrics*, Vol. 4.
- Whittaker, E. T. (1923): "On a New Method of Graduation". *Proceedings of the Edinburgh Mathematical Society*, No. 41, pp. 63-75.

**Comments on: "Measurement of the output gap: a discussion of
recent research done at the Bank of Canada"
by Pierre St-Amant and Simon van Norden**

by Florence Béranger and Gabriele Galati

This is a very interesting and well-written survey paper on different methods to estimate the output gap. The authors compare three approaches to measuring potential output and the output gap – a mechanical approach based on univariate dynamics, a pragmatic approach that augments univariate dynamics with structural relationships, and an approach that combines multivariate dynamics with long-run restrictions based on economic theory.

The discussion of the mechanical approach in the first section of the paper focuses on the popular univariate Hodrick-Prescott (HP) filter and its shortcomings. The HP filter seems ill suited as a tool for policy makers because it suffers from end-of-sample problems that make measures of the current output gap very unreliable and, moreover, it cannot be used to provide forecasts of future output gaps, which are of great interest for policy makers. Another limitation of the HP is that it associates the output gap with business cycle frequencies in the data. Instead, these may reflect the (transitional) dynamics of shocks to potential output. Also, the HP filter method produces spurious cyclicity when it is applied to integrated or near-integrated data. Finally, it performs poorly in identifying the cyclical component of time series – like those of most macroeconomic variables – that have a spectrum or pseudo-spectrum with typical Granger shape.

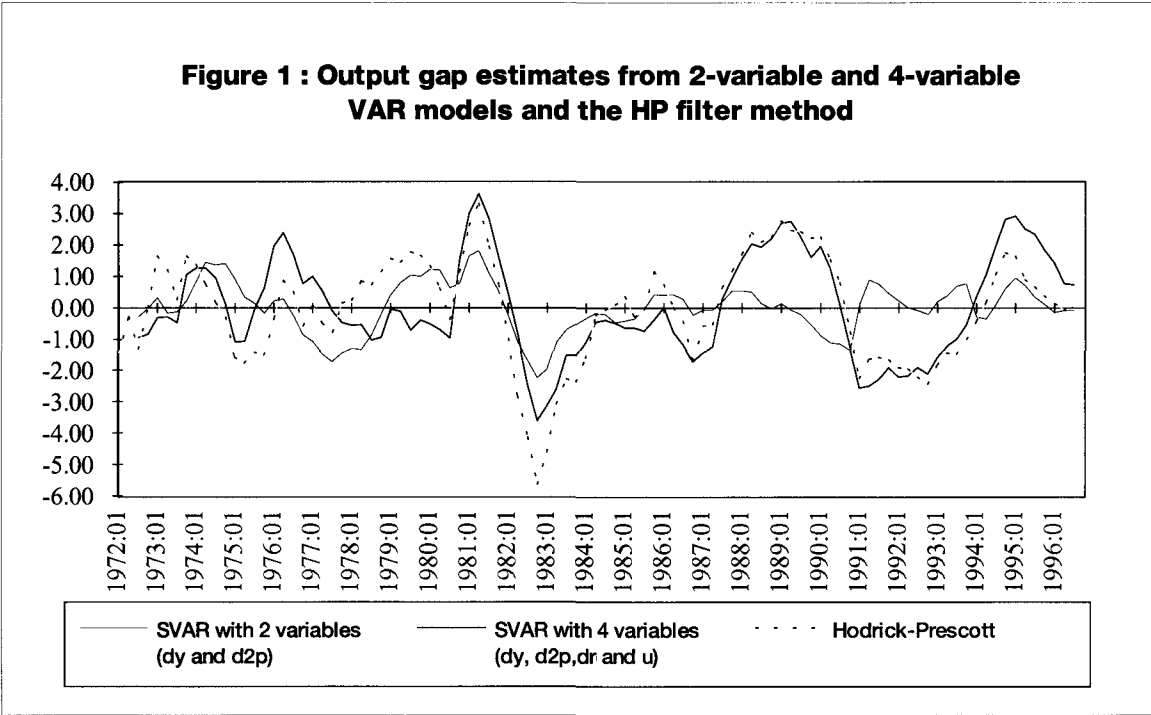
The paper then describes methods that augment univariate dynamics with structural relationships by adding the residuals of a structural economic relationship to the minimisation problem of the HP filter. These methods are more pragmatic but, at the same time, there are no guidelines on how to calibrate the structural relationships. They also fail to solve the end-of-sample problems and usually provide only very imprecise estimates of potential output and the output gap. The authors further describe a method (the TOFU) based on Maximum Likelihood estimates of the inverse of a simple structural relationship, but it is hard to get a clear sense of how well this method can perform in practice.

The most innovative and, in our view, interesting part of the paper describes methods of multivariate dynamics based on structural relationships. These methods use VAR models that incorporate long-run restrictions based on economic theory following the methodology of Blanchard and Quah (1989). Within this category the authors distinguish three methods: the long-run restrictions on output method (LRRO) which defines potential output as the permanent component of output and the output gap as that part of output which is due to purely transitory shocks (it should be noted that the authors are careful in avoiding the tricky issue of how to interpret permanent and transitory shocks in terms of demand and supply shocks); the Multivariate Beveridge-Nelson (MBN) method which defines potential output as the level of output that is reached after all transitory dynamics have worked themselves out; and the method proposed by Cochrane (CO) which derives identifying restrictions for a bivariate VAR model from the permanent income hypothesis. Most of the third section is devoted to a comparison of these three methods. The authors show that the LRRO outperforms the MBN and the CO methods because it lets the data determine the shape of the diffusion process of permanent shocks; i.e. it does not impose restrictions on the short-run dynamics of the permanent component of output or assume that output follows a random walk with drift.¹ In our discussion we will, therefore, focus on the LRRO method, noting some of the main advantages of this method and the problems and issues

¹ For the CO method this is true when consumption follows a random walk.

that it entails. It should be stressed, though, that the authors are aware of all these problems and issues, which makes their paper an excellent guide to empirical work on estimating the output gap.

A main advantage of the LRRO method is that it uses economic theory to derive restrictions on the VAR models that pin down the estimates of the output gap. An important issue concerns the choice of variables that are included in the VAR model. In order to capture the main features of real output in Canada, the authors include four variables in their VAR model – the growth rate of real output, the first difference of inflation, the unemployment rate and the first difference of the real interest rate. We can, however, think of alternative specifications, for example Blanchard and Quah's bivariate VAR model with output and unemployment, or a model that includes only output and inflation. It is important, therefore, to motivate the choice of variables. Figure 1 shows how the estimates of the output gap can be sensitive to the specification of the underlying model. It compares an estimate of the output gap for Canada derived from the four-variable model used in the paper with an estimate derived from a bivariate model with the growth rate of output and the first difference of inflation. The two measures of the output gap look quite different and have a correlation coefficient of only 0.34.²

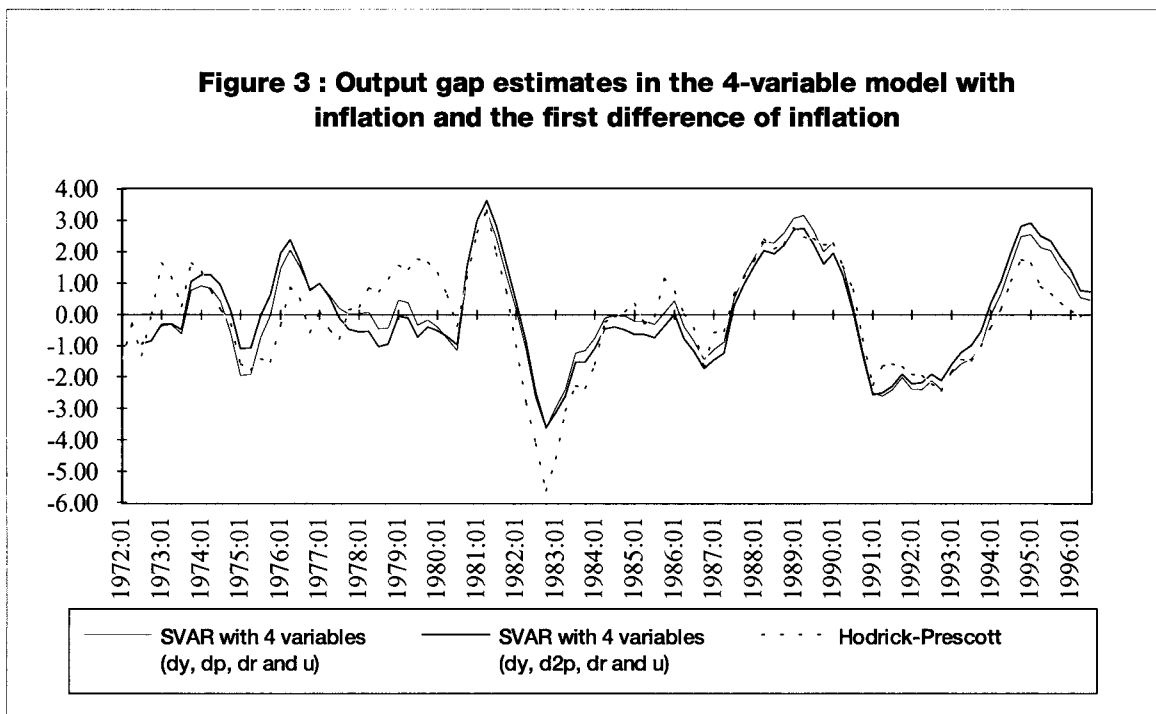
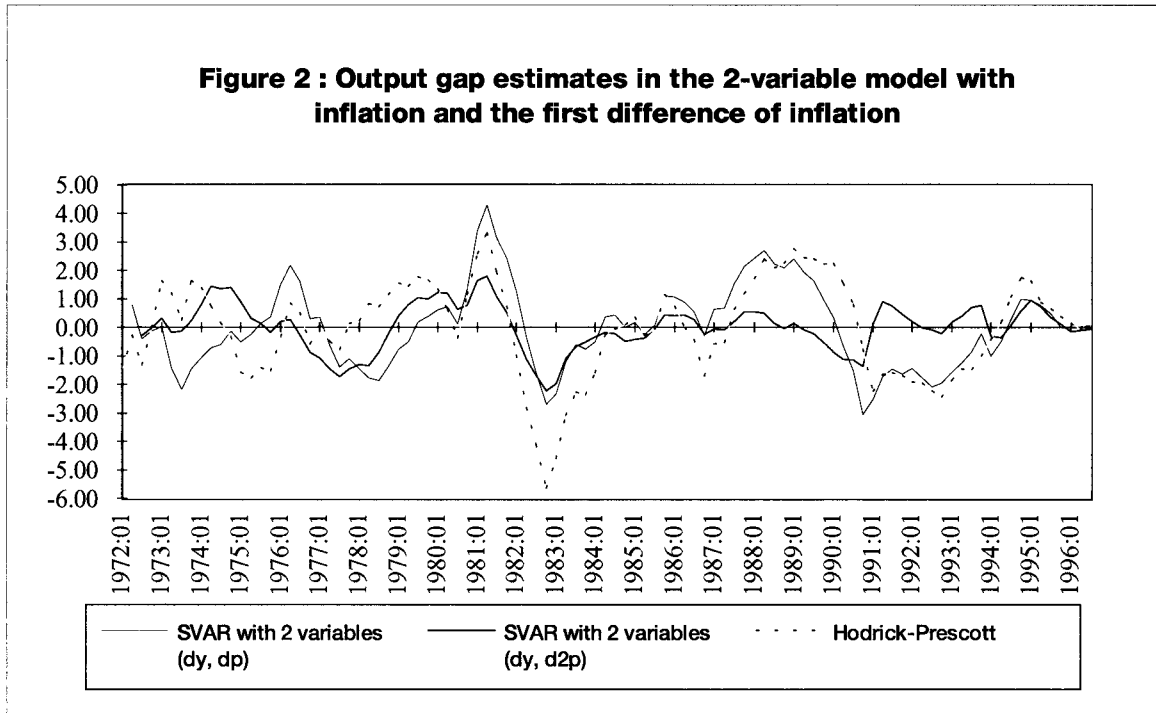


Another advantage of the LRRO method is that it is flexible with respect to the properties of the variables used in the model. As the authors state, "the assumptions on the level of integration are not part of the methodology per se" and "changing these assumptions could have a significant impact on the results". The degree of differencing to be used for variables other than output (which should be $I(1)$) is, however, a tricky issue. The authors assume that both the real interest rate and the inflation rate have unit roots, but if these series have almost but not exactly unit roots, important information on the short-run dynamics of output may be lost by differencing them³. We can get a sense of whether we over-difference and lose important information by comparing the impulse

² The correlation coefficient rises however to 0.75 when both models are estimated using the level of inflation. Figure 1 reports also the HP filter estimate of the output gap.

³ The estimates would still be consistent if the series are not differenced.

response functions of the bivariate and the four-variable models. Figure 2 shows that the output gap derived from the bivariate model looks very different depending on whether we difference inflation. The two series have a correlation coefficient of only 0.47. In the authors' four-variable model, however, it does not matter that much whether inflation and the interest rate are differenced or not (Figure 3).



Finally, an advantage of the LRRO method over the mechanical approaches is that it provides estimates of the current as well as of the expected future output gap. However, these estimates are generally very uncertain. There is a trade-off in choosing the number of lags in a VAR model between keeping it low to reduce the uncertainty of the estimates and keeping it high to avoid the bias in the estimation of the structural components. Still, Figure 15 in the paper shows several episodes where the LRRO output gap estimate for Canada is significantly different from zero in several instances. The authors also find that their results are robust to using 6 or 10 lags instead of 8.

References

Blanchard, Olivier J. and Danny Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances". *American Economic Review*, No. 79, pp. 655-673.