DEUTSCHE
BUNDESBANK

The reliability of Canadian output gap estimates

Jean-Philippe Cayen
(HEC Montréal and Bank of Canada)
Simon van Norden
(HEC Montréal, CIRANO and CIREQ)

Discussion Paper
Series 1: Studies of the Economic Research Centre
No 29/2004

Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-1
Telex within Germany 41227, telex from abroad 414431, fax +49 69 5601071

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax No +49 69 9566-3077

ISBN 3–86558–018–1

Abstract:

Recent work on policy rules under uncertainty have highlighted the impact of output gap measurement errors on economic outcomes and their importance in the formulation of appropriate policy rules. This paper investigates the reliability of current estimates of the output gap in Canada. We begin by assembling a new data base of quarterly real-time output estimates which spans the post-WWII period and contains data vintages dating back to 1972. We use this with a broad range of univariate and multivariate output gap models to recreate "contemporary" estimates of the output gap and then study how these estimates are revised over time. The nature and sources of these revisions are used to draw conclusions about the overall measurement errors associated with current estimates of the output gap. Relative to similar recent work with US real-time data, we find that revisions in Canadian output gaps are more important and that the role of data revision is less innocuous than previously indicated. We also show that using the change rather than the level of the output gap may only modestly reduce the measurement problem, and we investigate the relative importance of model uncertainty to overall measurement uncertainty.

Keywords: output gap, business cycle, real-time data, policy rules, monetary policy, Canada

JEL-Classification: C32, E32

Non Technical Summary

Recent work on policy rules under uncertainty have highlighted the impact of output gap measurement errors on economic outcomes and their importance in the formulation of appropriate policy rules. Four distinct issues comlicate measurement of the output gap in real time. First, output data may be revised, implying that output gaps estimated from real-time data may differ from those estimated from data for the same period published later. Second, as data on output in subsequent quarters become available, hindsight may clarify our position in the business cycle even in the absence of data revision. Third, the arrival of new data may instead made us revise our model of the economy. Fourth, different available models may give quite different estimates of the output gap. This paper investigates the relevance of these issues for the measurement of the output gap in Canada.

Nicht technische Zusammenfassung

Neuere Untersuchungen über die Anwendung geldpolitischer Regeln in einem mit Unsicherheit behafteten Umfeld haben die Auswirkung von Messfehlern bei der Bestimmung der Produktionslücke auf die wirtschaftlichen Ergebnisse sowie die Bedeutung solcher Messfehler bei der Formulierung geeigneter geldpolitischer Regeln verdeutlicht. Die Messung der Produktionslücke in Echtzeit wird durch vier unterschiedliche Faktoren erschwert. Erstens können die Produktionsdaten Revisionen unterliegen, so dass sich anhand von Echtzeitdaten geschätzte Produktionslücken von Schätzungen unterscheiden können, die auf später veröffentlichten Daten zum selben Zeitraum basieren. Zweitens kann die Veröffentlichung von Produktionsdaten in nachfolgenden Quartalen auch ohne Datenrevisionen zu einer nachträglich veränderten Einschätzung der Konjunkturlage führen. Drittens können neue Datenveröffentlichungen auch zu einer Änderung des verwendeten Modells führen. Viertens können verschiedene Modelle recht unterschiedliche Schätzungen der Produktionslücke ergeben. In diesem Diskussionspapier wird die Relevanz dieser Faktoren bei der Messung der Produktionslücke in Kanada untersucht.

Contents

Lists of Tables and Figures

The Reliability of Canadian Output Gap Estimates

1 Introduction

Recent research on monetary policy rules has rekindled interest in the measurement of the output gap. In his seminal paper, Taylor (1993) suggests that the central bank set interest rates on the basis of only the rate of inflation and the output gap. In discussing Taylor, McCallum (1993) pointed out that such rules are not operational because the output gap is not directly observed. More recent analysis (for example, see Orphanides et al. (2002), Orphanides (2003a), McCallum and Nelson (2004) or the survey by Walsh (2003c)) shows that the weights which policy makers should attach to indicator variables in simple policy rules generally depend on the precision with which they measure the underlying state of the economy. This has in turn focussed attention on the precision and accuracy with which one can measure the current output gap.

There are conflicting schools of thought on the potential accuracy and precision of current estimates of the output gap. Orphanides and van Norden (2002) showed that a broad range of univariate methods for estimating current output gaps must lack precision because estimates are considerably revised over time. The wide divergence of estimates they find across different models also suggests that many models produce inaccurate estimates, which further contributes to the uncertainty facing decision makers. Kuttner (1994), Gerlach and Smets (1997), and Orphanides and van Norden (2002) all find relatively large confidence intervals surrounding estimated output gaps using bivariate models models of output and inflation for the US and Canada. Related work using unemployment and inflation has reached similar conclusions for NAIRU gaps; for example, see Staiger, Stock and Watson (1997) or Laubach (2001). However, a few studies have suggested that the uncertainty may be much less. For example, Gruen, Robinson and Stone (2002) use Australian data to argue that careful specification of the Phillips Curve may greatly reduce output gap uncertainty.

Another way to incorporate information from additional variables is to assume that several different variables share a common (but unobserved) cyclic component. Early examples of this approach include Clark (1989), who examines a bivariate system of output and unemployment, and Apel, Jansson and Lindberg (1995) who also add inflation. These papers found output gap estimates to be relatively imprecise. Recent papers using euro-zone data, such as Fabiani and Mestre (2003), Camba-Mendez and Rodriguez-Palenzuela (2003) and Rnstler (2002) conclude that these approaches lead to much more reliable estimates. However Azevedo, Koopman and Rua (2003) use a larger but less structural model and find broad confidence intervals. Previous attempts to incorporate structural economic information using different techniques (Structural VARs) and other data sets (the US, Germany, Canada and New Zealand) have found that estimates of the gap remain relatively imprecise. (For example, see Dupasquier, Guay and St-Amant (1999), Lalonde, Page and St-Amant (1998), Lalonde (1998), Kichian (1999), Claus (2003) or Mitchell (2003).) Camba-Mendez and Rodriguez-Palenzuela (2003) examine SVAR models for the eurozone and argue that they give precise estimates. Mitchell (2003 p. 6 and Table 1) reports quite different results using the same model and data. Rnstler (2002) finds that the addition of capacity utilisation reduces the uncertainty further still for the euro-zone, a claim also made by Longworth (2003) for Canada.

Four distinct issues complicate measurement of the output gap in real time. First, output data may be revised, implying that output gaps estimated from real-time data may differ from those estimated from data for the same period published later. Second, as data on output in subsequent quarters become available, hindsight may clarify our position in the business cycle even in the absence of data revision. Third, the arrival of new data may instead make us revise our model of the economy which in turn revises our estimated output gaps. Fourth, different available models may give quite different estimates of the output gap given the same available data. Most of the above papers focus on the third of these four issues. Orphanides and van Norden (2002) also address issues one and two, but find that the third dominates.

One unrealistic feature of this analysis is that policy makers consider the behaviour of many macroeconomic variables in arriving at their estimates of the current output gap. It is therefore possible that policymakers are able to estimate the current gap much more precisely by incorporating additional sources of information. Orphanides (2001, 2003b) analyses the US Federal Reserve Boards own estimates of the output gap and finds that the size and persistence of the revisions in that series are similar to those found in Orphanides and van Norden (2002). Nelson and Nikolov (2003) attempt to reconstruct official UK Treasury estimates of the output gap and find that revisions in that series appear to be larger still. These results appear to support the conclusion that policymakers estimates of the current output gap are subject to considerable uncertainty. Another shortcoming of this literature lies in the fact that forward-looking policy implies that the accuracy of *forecasts* of the output gap are the key factor, but this distinction has been largely overlooked in the literature thus far. As a result, the studies discussed above likely understate the importance of measurement uncertainty facing policy makers.

This paper investigates the relevance of these issues for the measurement of the output gap in Canada since the 1970s using several well-known detrending methods. We begin by creating a new "real-time" database for Canadian output along the lines of that produced for the US by Croushore and Stark (2001) Then, following earlier work by Orphanides and van Norden (2002), we examine the behavior of end-of-sample output gap estimates and of the revision of these estimates over time. Presuming that revisions "improve" our estimates, the total amount of revision gives us a lower bound on the measurement error thought to be associated with real-time output gaps. This is informative when and if we find that revision errors are relatively large since we can conclude that the total error of these estimators must be larger still. Furthermore, such results are quite general; they apply regardless of whether output gaps are used to cyclically-adjust budget balances, to forecast inflation or for other purposes, and do not require a priori assumptions on the true structure of the economy or on the true time-series model generating observed output. We also decompose the revisions into their various sources, including that due to revisions of the underlying output data and that due to re-estimation of the process generating potential output. We explore whether the rate of change in the gap may be estimated more precisely than its level and the extent of revisions associated with output gap forecasts. Finally, we also examine the extent to which uncertainty about the correct model may contribute to estimation uncertainty.

The next section documents the data used in the paper and describes of some of the features of Canadian output revisions. Section three introduces the various models and techniques used to estimate the output gap. (Readers familiar with such techniques may wish to skip this section at first and refer back to it as needed.) Section Four reviews the construction of the revisions and the interpretation of their decomposition. It then presents the main results on their properties and sources. Section Five considers whether measurement errors are less serious for the change in the rather than its level. Section Six documents the extent to which output gap forecasts are revised, and Section Seven explores the relative importance of model uncertainty. The final section concludes.

2 Real Time Data

2.1 Data Sources and Definitions

The raw data for real GDP were taken from quarterly series published in various issues of *National Revenue and Expenditure Accounts* by Statistics Canada and occasional special Statistics Canada publications which documented major historical revisions. [1] Although Statistics Canada originally started publishing quarterly national accounts in 1961, the series were extensively revised in the late 1960s and we were unable to obtain complete and reliable data vintages prior to 1972Q1. These GDP series were originally analysed in an

[1]What we refer to for simplicity as real GDP is more accurately referred to as real output. Vintages up to 1986Q1 are in fact real GNE, with Statistics Canada thereafter switching to the GDP concept. The series were reported quarterly in Statistics Canada's *National Income and Expenditure Accounts*, catalogue #13-001. We use deseasonalized figures throughout.

earlier version of this paper (Cayen and van Norden (2002).) The range of available data vintages used here has been extended so that the "final" data vintage now corresponds to 2003Q4 (i.e. data available as of mid-February 2004, so data series end in 2003Q4) rather than 1999Q4 as in the earlier paper.

Statistics Canada's national accounts were originally published with historical quarterly estimates reaching back to 1947Q1. Starting with the National Accounts revision of 1997, however, regularly published vintages included estimates back to only 1961Q1. The earlier version of this paper used no historical data for periods prior to 1961Q1. This in turn limited the range over which revisions of some of the output gap estimates could be calculated. We have therefore collected additional data allowing us to extend most of our data vintages back to 1947Q1. When the most recently published vintage does not start in 1947Q1, we follow the procedure of Orphanides and van Norden (2002) and splice the most recent vintage to the most recently published estimates for the missing periods. [2]

Some of the output gap estimates also use data on the Consumer Price Index (CPI) and some also use interest rates. Interest rate data are never revised, so we used final data throughout, using Bank of Canada quarterly data on auction yields on 3-month Treasury Bills. [3] Revisions in the CPI occur but are minor; we therefore follow the practise of Cayen and van Norden (2002), Orphanides and van Norden (2002) and Orphanides and van Norden (2003) and use final estimates of the seasonally adjusted All-Items Canadian CPI throughout.[4] This may therefore slightly understate the extent to which related output gaps are revised *ex post*; if so, the bias should be quite small.

[2] To extend a vintage *V1* which starts at period n using an earlier vintage *V0*, the missing entries *1* through *n-1* of *V1* are replaced by those of *V0* and adjusted by a scaling constant of $(V1_n/V0_n)$ to adjust for possible changes in the base year.

[3] Quarterly observations are the last auction of each quarter, recorded in Statistics Canada series V122541. Cayen and van Norden (2002) used overnight call money rates. The relevant differences between these series are trivial at quarterly frequencies.

[4] Data from 1992 onwards are CANSIM series V18702611; data prior to 1992 were provided by the Bank of Canada.

2.2 Revisions in Canadian GDP

To understand the effects of data revision on the measurement of GDP, we construct what Croushore and Stark (2001) call a "real-time" GDP series. This series simply collects the first published estimates of GDP; each observation in the series therefore comes from a different vintage. (The term realtime is a misnomer because these first estimates are published with a lag. For example, the "realtime" estimate for 2003Q4 was published in February 2004.) We similarly construct realtime estimates of output growth by collecting the first estimates of the 1, 2, 4 or 8 quarter change in log GDP. Again, each growth rate observation in the realtime series comes from a single data vintage.[5] These realtime series should therefore reflect the most timely information available to analysts and policymakers.

Realtime estimates of GDP are then compared to "final" estimates to measure the extent to which data are eventually revised. Throughout this study, we treat the last available data vintage (2003Q4) as our "Final" estimate; to the extent that the data may yet be further revised, we will tend to underestimate the extent of data revision. Revisions are defined simply as the difference between our realtime and final estimates. Table 1 provides some descriptive statistics for the Realtime series and their revisions. Additional detail is provided in Figures 1 through 3.

We make no attempt here to quantify the dynamics of the revision process. However, they appear to be qualitatively very similar to those previously described in the US data. Statistics Canada revises GDP figures up to four years after their original publication, with most of the revision happening in the first year.[6] Revisions thereafter are very infrequent but frequently large, usually reflecting either a rebasing of the data (e.g. from 1992 to

[5] Real GDP figures are periodically "rebased." This rebasing has no appreciable effect on our realtime growth rate series because all growth rates are constructed using observations measured in the same units. However, our GDP level series has been adjusted to remove discontinuities due to rebasing. This is done by restricting the 1-quarter change in realtime log GDP in a rebasing quarter to equal that in the 1-quarter realtime log GDP change series.

[6] Seasonal factors for the last four years are usually revised with the release of data for the first quarter of the following year.

1997 constant dollars) or a change in methodology (e.g. change from GNE to GDP, or fixed-weights to chain-weights.)

As Table 1 and Figure 1 make clear, revisions are small relative to GDP. Final and realtime estimates of log GDP look very similar and have a correlation coefficient ¿ 0.998. Revisions have a standard error of 0.015 and revisions larger than ± 0.03 are occasionally observed. Although small relative to total ouput, GDP data revisions may therefore still be important relative to the size of the business cycle. Mean revisions are positive, implying that on average GDP figures tend to be revised upwards. However, the mean revision is small relative to both the standard error of the revisions and the mean of the realtime series.

Data revision becomes more important if we focus on GDP growth rather than its level. Figure 2 shows scatterplots of final and realtime growth estimates. While there is a clear positive relationship, there is considerable scatter about the 45-degree line that diminishes as we move from 1Q to 8Q growth measures. This is confirmed by the correlations shown in Table 1, which rise from 75% for 1Q changes to almost 94% for 8Q changes. The table also shows that the revision standard errors fall relative to those of the realtime series, moving from roughly a 2/3 Noise/Signal ratio for 1-quarter changes to a 1/3 ratio for 8-quarter changes. The declining relative importance of the revision is also apparent from Figure 3.

Finally, revisions appear to have complex dynamics. Revisions to the level of GDP are highly persistent as shown by a first-order autocorrelation coefficient of almost 90%. In contrast, revisions to 1-quarter changes are negatively autocorrelated, although this quickly becomes positive for changes over longer horizons.

The impact of these data revisions on the accuracy of realtime output gap measurement ultimately depends on the extent to which they cause estimates of the gap to be revised rather than estimates of potential output. Certainly the size of the revisions in real GDP are important relative to most estimates of the size of the business cycle. However, to the extent that output gap estimators can be thought of as filters of output growth rates, the degree to which revisions affect realtime estimates of output gaps will depend in large measure on the persistence of the revisions to growth rates.[7] The fact that data revisions become relatively less important as we move from 1-quarter to 8-quarter growth rates is therefore something which might limit their impact on output gap estimation. However, quantifying their impact requires that we construct realtime output gap estimates, a subject to which we turn in the next section.

3 Estimating Output Gaps

The estimation of the output gap requires a detrending method which decomposes the log of real output, q_t, into a trend component, μ_t, and a cycle component, y_t.

$$q_t = \mu_t + y_t \tag{1}$$

Some methods use the data to estimate the trend, μ_t, and define the cyclical component as the residual. Others specify a dynamic structure for both the trend and cycle components and estimate them jointly. We examine detrending methods that fall into both categories. The output gap estimates examined here include all those examined in Cayen and van Norden (2002) plus the Breaking-Trend and Band-Pass methods described below.

3.1 Deterministic Trends

The first set of detrending methods we consider assume that the trend in (the logarithm of) output is well approximated as a simple deterministic function of time. The linear trend (LT) is the oldest and simplest of these models. The quadratic trend (QT) is a popular alternative.

Because of the noticeable downturn in GDP growth after 1973, another simple deterministic technique is a breaking linear trend, one that allows for a slowdown beginning in the 1970s. Our implementation of the breaking trend method (BT) incorporates the

[7] See St-Amant and van Norden (1998) and van Norden (2002) for a discussion of this point.

assumption that the location of the break is fixed and known.[8] In their analysis of US data, Orphanides and van Norden (1999) found that a break at the end of 1973 would have been detected by statistical tests in 1977, which conformed with the actual practice of the Council of Economic Advisors. Canadian evidence is mixed however, with breakpoint tests on realtime data giving only intermittent evidence of a break in trend in the mid 1970s, and detecting it later than in the US data. Based on Andrews-Quandt and Andrews-Ploberger tests for a break in trend, we assume that a break in the trend at the start of 1974 would have been incorporated starting with the 1979Q2 vintage.

3.2 Mechanical Filters

Dissatisfaction with the assumption of deterministic trends (see Nelson and Plosser (1982)) has led many to prefer business cycle models which allow for stochastic trends. Harvey (1989) popularized the structural time series model, which attempts to separately characterize the dynamics of the trend and the cycle. A simpler (but related) approach prefers to remain agnostic about their dynamics, or to simply model their joint dynamics as some ARIMA process. We examine two examples of this approach; the Hodrick-Prescott (HP) and the Band-Pass (BP) filter.[9]

The earliest approach modelled potential output with a smoothing spline, as in the popular filter proposed by Hodrick and Prescott (1997) (the HP filter).[10] We apply the HP filter to quarterly data with their recommended smoothing parameter of 1600.

Another more recent approach to cycle-trend decomposition aims to isolate specific frequencies in the data via the use of band-pass filters. The clearest exponent of this approach is Baxter and King (1999), who suggest the use of truncated versions of the ideal

[8]To the extent that the existence and location of the break is unknown, our assumption to the contrary may tend to understate the degree of uncertainty surrounding the output gap.

[9]The distinction between these approaches and the structural time series approach is somewhat arbitrary as both of these models can be expressed in the form of an Unobserved Components model – see Gomez(2001) and Harvey and Trimbur (2003). However, this is rarely done in practice.

[10]The development of smoothing splines dates back to the work of Whittaker (1923) and Henderson (1924) and discussion of its use for measuring business cycles may be found in Orphanides and van Norden (1999). St-Amant and van Norden (1998) discuss the properties of the HP filter at the end of sample.

(and therefore infinitely long) filter passing fluctuations with durations between 6 and 32 quarters in length.[11] Stock and Watson (1998) adapt this for use at the end of data samples by padding the available observations with forecasts from a low-order AR model fit to the data series. Following their application to US data, we use a filter 25 observations in width and pad using an AR(4) forecast.[12]

3.3 The Beveridge-Nelson Decomposition

A very different approach is that of Beveridge and Nelson (1981), who consider the case of an ARIMA(p,1,Q) series q, which is to be decomposed into a trend and a cyclical component. For simplicity, we can assume that all deterministic components belong to the trend component and have already been removed from the series. Since the first-difference of the series is stationary, it has an infinite-order MA representation of the form

$$\Delta q_t = \varepsilon_t + \beta_1 \cdot \varepsilon_{t-1} + \beta_2 \cdot \varepsilon_{t-2} + \cdots = e_t \tag{2}$$

where ε is assumed to be an innovations sequence. The change in the series over the next s periods is simply

$$q_{t+s} - q_t = \sum_{j=1}^{s} \Delta q_{t+j} = \sum_{j=1}^{s} e_{t+j} \tag{3}$$

The trend is defined to be

$$\lim_{s\to\infty} E_t(q_{t+s}) = q_t - \lim_{s\to\infty} E_t(\sum_{j=1}^{s} e_{t+j}) \tag{4}$$

Since changes in the trend are therefore unforecastable, this has the effect of decomposing the series into a random walk and a cyclical component.[13] Using the Beveridge-Nelson decomposition requires us to select an appropriate ARIMA model: based on results for the

[11]One justification frequently given for the use of the HP filter is that it approximates such a band-pass filter. See Baxter and King 1999 for a discussion.

[12]Christiano and Fitzgerald (2003) and van Norden (2002) discuss the construction and properties of optimal band-pass filters which avoid the need to pad the series.

[13]We follow Morley (2002) in defining the trend to be *difference* between actual and cyclic output. A previous draft of this paper followed Beveridge and Nelson (1981) in defining it to be the *sum* of actual and cyclic output. Morley, Nelson and Zivot (2003) discuss the relationship between the Beveridge-Nelson decomposition and the unobserved components models discussed next.

full sample, we assume an ARIMA(2,1,1) throughout. We then use the results of Morley (2002) to calculate the Beveridge-Nelson (BN) decomposition in state-space form.

3.4 Unobserved Component Models

Unobserved component (UC) models offer a general framework for decomposing output into an unobserved trend and a cycle, allowing for explicit dynamic structures for these components. We examine three such alternatives, by Watson (1986), by Harvey (1985) and Clark (1987), and by Harvey and Jaeger (1993). All are estimated by maximum likelihood.

The Watson model (WT) modifies Harvey's Linear Level model to allow for greater business cycle persistence. Specifically, it models the trend as a random walk with drift and the cycle as an AR(2) process:

$$\mu_t = \delta + \mu_{t-1} + \eta_t \tag{5}$$

$$y_t = \rho_1 \cdot y_{t-1} + \rho_2 \cdot y_{t-2} + \varepsilon_t \tag{6}$$

Here ε_t and η_t are assumed to be i.i.d mean-zero, Gaussian and mutually uncorrelated and δ, ρ_1 and ρ_2, and the variances of the two shocks are parameters to be estimated (5 in total).

The Harvey-Clark model (CL) similarly modifies Harvey's Local Linear Trend model:

$$\mu_t = g_{t-1} + \mu_{t-1} + \eta_t \tag{7}$$

$$g_t = g_{t-1} + \nu_t \tag{8}$$

$$y_t = \rho_1 \cdot y_{t-1} + \rho_2 \cdot y_{t-2} + \varepsilon_t \tag{9}$$

Here η_t, ν_t, and ε_t are assumed to be i.i.d mean-zero, Gaussian, mutually uncorrelated processes and ρ_1 and ρ_2 and the variances of the three shocks are parameters to be estimated (5 in total).

The Harvey-Jaeger model (HJ) differs from the Harvey-Clark model only in the dynamics specified for the cyclic component, which are now

$$\begin{bmatrix} y_t \\ y_t^* \end{bmatrix} = \rho \cdot \begin{bmatrix} cos\lambda_c & sin\lambda_c \\ -sin\lambda_c & cos\lambda_c \end{bmatrix} \cdot \begin{bmatrix} y_{t-1} \\ y_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} \tag{10}$$

where κ_t and κ_t^* are assumed to be i.i.d mean-zero, Gaussian, mutually uncorrelated processes with variance σ_κ. Nowρ, λ_c and the three shock variances are parameters to be estimated (again, 5 in total).

3.5 Unobserved Component Models with a Phillips Curve

Multivariate formulations of UC models attempt to refine estimates of the output gap by incorporating information from other variables linked to the gap. Examples of this approach include Kuttner (1994), Gerlach and Smets (1997) and Kichian (1999); the former adds a Phillips Curve to the Watson model, the latter two add the curve to the Harvey-Clark model. We follow the latter approach, modifying the Harvey-Clark model by adding the following equation

$$\Delta\pi_t = \phi_1 + \phi_2 \cdot y_t + e_t + \phi_3 \cdot e_{t-1} + \phi_4 \cdot e_{t-2} + \phi_5 \cdot e_{t-3} \tag{11}$$

In each case the shock e_t is assumed i.i.d. mean zero Gaussian and uncorrelated with shocks driving the dynamics of the trend and cycle components of output in the model. Thus, by adding the Phillips curve, this Gerlach-Smets (GS) model introduces an additional six parameters that require estimation ($\{\phi_1, ..., \phi_5\}$ and the variance of e_t).[14]

3.6 The Structural VAR Approach

The Structural VAR measure of the output gap is based on a VAR identified via restrictions on the long-run effects of the structural shocks, as proposed by Blanchard and Quah (1989). Our implementation is identical to that of Cayen and van Norden (2002), who use a trivariate

[14]Cayen and van Norden (2002) examined the performance of Kichian's model, which uses a more elaborate Phillips Curve containing additional exogenous variables. Difficulty in extending the real-time data set for these exogenous regressors led us to examine instead this simpler specification.

system including output, CPI and yields on 3-month treasury bills. Earlier work by Lalonde, Page and St-Amant (1998) found that the inclusion of interest rates improved the model's ability to distinguish real and nominal shocks. Lag lengths for the VAR are selected using finite-sample corrected LR tests and a general-to-specific testing approach; the lag length is re-optimized every time the VAR parameters are re-estimated. We examine the same two variants of the SVAR output gap analyzed in these earlier studies; RLTP (BQ), which is the transitory component of output, and RLTP1 (SV), which excludes those transitory output shocks which have permanent effects on inflation.

4 Evaluating Output Gap Revisions

4.1 The Components of Output Gap Revisions

We use our data with each of the detrending methods described above to produce estimated output gap series. We apply each detrending method in a number of different ways in order to estimate and decompose the extent of the revisions in the estimated gap series.

The first of these estimates for each method simply takes the last available vintage of data (2003:4) and detrends it. The resulting series of deviations from trend constitutes our Final (FL) estimate of the output gap corresponding to that method.

The Real-Time (RT) estimate of the output gap is constructed in two stages. First, we detrend each and every vintage of data available to construct an ensemble of output gap series. Next, we use these different output gap vintages to construct a new series which solely consists of output gap estimates for the last observation in each vintage. The resulting series represents the most timely estimates of the output gap which could be constructed using the method employed.[15]

The difference between the Real-Time and the Final estimate gives us the total revision in the estimated output gap at each point in time. This revision may have several sources,

[15] Note that in any quarter T, the available output series ends at T-1. The real-time gap estimates are therefore estimates for the *last quarter's* output gap. This usage of the term is consistent with that of Orphanides and van Norden (2002), among many others.

one of which is the ongoing revision of published data. To isolate the importance of this factor, we define a third output gap measure, the Quasi-Real estimate. The Quasi-Real estimate of the output gap is simply the rolling estimate based on the Final data series. That is, the gap at period t is calculated using only observations 1 through t to estimate the long-run trend and the deviations around it. The difference between the Real-Time and the Quasi-Real series is entirely due to the effects of data revision, since estimates in the two series at any particular point in time are based on data samples covering exactly the same period.

For unobserved component (UC) models, we further decompose the revision in the estimated gap by defining a Quasi-Final (QF) estimate. UC models use the data in two distinct phases. First, they use the available data sample to estimate the parameters of a time-series model of output. Next, they use these estimated parameters to construct filtered and smoothed estimates of the output gap. For this class of models, smoothed estimates of the output gap are used to construct the Final series, while filtered estimates are used for the Quasi-Final series. In both cases, the UC model's parameters are estimated using the full sample of the same data which is then used for filtering and smoothing. We construct QF estimates for the WT, CL, HJ and GS models.[16]

The difference between the Quasi-Final and the Quasi-Real series reflects the use of different parameter estimates (i.e. full-sample ones versus partial-sample ones) to filter the data. The extent of the difference will reflect the importance of parameter instability in the underlying UC model. The difference between the Quasi-Final and the Final series reflects the importance of ex post information in estimating the output gap given the parameter values of the process generating output.[17]

[16] While they can be put in state-space form, we make no attempt to construct QF estimates for the HP, BQ, SV or BN models. Note that for the latter model, the QF estimates are identically equal to the FL estimates.

[17] St-Amant and van Norden (1998) argue that the degree to which the subsequent behavior of output is informative about the output gap is linked to presence or absence of hysteresis in output.

4.2 The Size and Sources of Canadian Output Gap Revisions

Descriptive statistics for all the output gap estimates are shown in Table 2, while Figure 4 provides a visual comparison. From the figure, it is apparent that while some detrending methods produce similar output gap estimates (e.g. compare the Breaking Trend and the Gerlach-Smets models,) other produce quite different estimates (e.g. compare the Quadratic Trend, Blanchard-Quah and Beveridge-Nelson models.) We will defer discussion of these cross-model comparisons for now and focus here on how the gap estimates are revised over time.

Figure 4 roughly groups the different detrending methods based on some characteristics of their results. In the first row are three models (LT, WT and CL) for which the real-time gap estimates are almost always negative while the final gap estimates have a mildly positive mean. In the case of the LT model, Table 2 shows the differences in mean to be over 0.12, while differences for the other two are less than 0.04, but still larger than for most other models. In the case of the LT model, this reflects the steady downward revision in estimated trend growth rates following the growth slowdown of the early 1970s. The WT model, which also assumes a constant trend growth rate, suffers from the same problem to a lesser extent. (The QT model, shown in the last row of Figure 4, also shares this problem.) The CL model allows for a time-varying trend growth rate and has the smallest difference in mean of the three. Comparison of the means of its QF and QR estimates shows that not much of the overall difference in mean is attributable to changing model parameter estimates; rather, as difference in the means of the QF and FL estimates show, most of the difference is due to *downward* revisions in estimated trend output as new observations arrive.

The second row of Figure 4 groups together three models of the gap (BQ, SV and BN) whose estimates are relatively small and have little persistence. Their standard deviations and extreme values include the smallest of all the models examined. One noteworthy feature in these models is the usually large revision caused by the revision of output data in the 1980s. As shown by the differences in the RT and QR estimates, the change is most dramatic for the BQ model from 1982-84, and for the SV model from 1984-1986, as large negative real-time gap estimates were revised to roughly zero. The WT and CL models show somewhat similar behaviour in the latter period. In contrast, the BN model shows unusually large downward revisions for the 1981-83 period.

The last two rows of Figure 4 show the remaining six models (BT, HJ, GS, BP, HP and QT). Most display a visually similar pattern, with persistent output gaps that are usually positive in the early 70s, late 80s and around 2000, with mostly negative output gaps around the recessions of 1982 and 1991.[18]

To better understand the relative importance of *ex post* revisions of output gaps, the last two columns of Table 2 provide two simple measures of the similarity between the Final and other estimates of the gap. The first of these is simply the correlation with the Final estimate; a low correlation with the Final estimate necessarily implies that output gap revisions are relatively important in size. The correlations shown in Table 2 show that all estimates (except for the Harvey-Clark model) are positively correlated with the Final estimate and the correlation increases as we move from the RT to the QR to the QF estimate (except for the Beveridge-Nelson and Watson models, where data revision *lowers* the correlation.)

Just over half the models (7/12) have a correlation of less than 50% between the Real-Time and Final estimates. Put another way, this means that Real-Time estimates can account for less than one quarter of the *variance* of the Final estimates in most of the models we examine. This is considerably lower than in Cayen and van Norden (2002), where 5/10 models had a correlation of over 80%, and in Orphanides and van Norden (2002), where 4/8 exceeded 80%. The former study used Canadian output vintages from 1982Q2 to 1999Q4 while the latter used US data vintages from 1966Q1 to 1997Q4. This suggests that

[18]See Cross(1996, 2001) for a chronology of Canadian business cycles.

neither the nature of Canadian output data revisions nor the sample period examined can, by themselves, account for these low correlations. If we look instead at the correlation of Quasi-Real and Final estimates, Table 2 shows that 10/12 models have correlations below 70% and 7/12 are below 60%. This compares with 5/10 having correlations above 80% in Cayen and van Norden (2002) and 5/8 in Orphanides and van Norden (2002). This is further evidence that the revision of the Canadian data is not responsible for these results; rather, it appears to result from poor stability of many of these models' estimates when applied to the postwar Canadian data.

Note that even such low correlations may understate the relative importance of the revisions. This is because a high correlation is not necessarily evidence of small revisions. For example, while the Linear Trend model appears remarkable for the size of its revisions in Figure 4, Table 2 shows that it has by far the highest correlation between Final and Real-Time estimates.[19] The last column of Table 2 therefore produces an alternative measure of association; the fraction of observations in which the Final estimate has the same sign as another estimate. Such a directional estimate may be of particular interest if the output gap is used primarily to determine whether policy is too loose or too tight. The absence of any revisions would produce a value of 0, while replacing Real-Time gap estimates with purely random noise should give a value close to 0.50. The results in Table 2 from this measure tend to confirm the disappointing correlation results discussed above. Real-Time estimates from all models give the "wrong" sign at least 25% of the time; 5/12 exceed 40% and 3/12 models exceed 50%. This is somewhat worse than the results in Cayen and van Norden (2002), where 3/10 models gave the "wrong" sign less than 25% of the time, but 4/10 exceeded 40% of the time. The differences are largest for the Harvey-Clark (where the rate rises from 28% in the earlier study to 62.5% here) and the BQ, SV and BN models, which produce some of the best results in either paper but whose misclassification rates rise

[19] This is due in part to the fact that correlations ignore differences in the means of the two series. Orphanides and van Norden (2002) similarly found the linear trend model to produce large revisions, but high correlations between RT and FL estimates.

roughly 15%. Comparing results for the RT and QR estimates, we can see that data revisions cause important deteriorations in the reliability of a few models' estimates (particularly the BT and SV models), but that they also appear to have no negative effect on 4/12 models (LT, HJ, WT and GS). Instead, other factors again appear to be the dominant source of output gap revisions.

More evidence on the relative importance and sources of output gap revisions is given in Figure 5 and Table 3. The Figure compares the Real-Time gap estimates for each model (shaded area) with the total revision of the estimated gap (green line, calculated as FL-RT) and the revision resulting from data revision (blue line, calculated as QR-RT.) It is apparent from the figure that the size and variability of the total revisions are not trivial relative to the size and variability of the Real-Time gap estimates. It also appears that for several models (e.g. LT, HJ, GS) the revision due to data revision is small relative to the total revision. This does not appear to the be the case for models where the output gap has relatively little persistence (BQ, SV and BN.) Other models seem to occupy a middle-ground, where the contribution of data revision appears to be neither trivial nor dominant.

The statistics presented in Table 3 allow us to make more precise statements about the relative importance of output gap revisions and their sources. In addition to the total revision (FL-RT), the table presents descriptive statistics for each of the revision components discussed in the previous section. While the average revision is less than 1% for 6/12 models, it is larger for 3% for the WT and QT models, and over 12% for the LT model. All these models assume a deterministic trend growth rate, which causes an important bias in their Real-Time gap estimates. The revisions are highly persistent. The first-order autocorrelation coefficient (ρ) is greater than 0.95 for half of the models and greater than 0.80 for all but two models, although the persistence of the revisions appears to be lower for those models with the least persistent gap estimates (BN, BQ and SV.)[20] These statistics

[20]This does not imply that revisions are forecastable, nor does it imply that gaps are revised slowly.

are potentially serious, since they suggests that output gap revisions are sometimes of an economically important size and could lead to persistent misperceptions about the state of the business cycle.

A better sense of the relative importance of these revisions is given by the noise-signal ratio, shown in the last column of the table. This number is the ratio of the root mean squared revision (noise) to the standard deviation of the Real-Time output gap (signal). Since the numerator includes the squared bias but the denominator does not, this ratio can easily exceed 1. In fact, Table 3 shows that it is less than 1 for only 1 of the 12 models examined (0.924 for BQ) and it exceeds 2 for the LT and WT models. This again appears to be more pessimistic than the US results in Orphanides and van Norden (2002), who find ratios ranging between 1.32 and 0.69 for the eight models they examine.

The latter study also emphasized that data revisions usually played a minor role in accounting for the total revisions of output gap estimates. The evidence for this in the Canadian data is less clear cut. For the three models which produce low-persistence output gaps (BQ, SV and BN), the standard deviation of the QR-RT revisions show that data revision is at least as important a source of overall output gap revisions as all other sources combined. On the other hand, the importance of observations on the future (captured by the FL-QF revision for the UC models and the FL-QR revision for the other models) continues to be the dominant source of revisions for several models, particularly for the deterministic trend models (LT, QT and BT), the mechanical filters (HP and BP) and some of the UC models (GS and HJ). The Watson and the Harvey-Clark models show large contributions to revisions from several sources. Both of these models appeared to suffer from parameter instability, with estimated parameters sometime switching between distinct local peaks in the likelihood function as observations were added or data were revised. This had important effects on the estimated gaps, as can be seen in Figure 4 by the abrupt changes in the Real-Time gap estimates in both models around the 1972-75 period and again in the mid-1980s.

Rather, it implies that future information will affect gap estimates in consecutive periods similarly.

Depending on how this parameter instability interacted with data revision, these changes were captured to varying degrees by either the QR-RT revision or the QF-QR revision.

While evidence on the sources of the revisions varies across models, the results on the correlations and sign disagreements between the RT and FL estimates, as well as the results on the N/S ratios show that gap revisions in the Canadian data have been important across all models and larger than earlier work would have suggested. In the next section, we therefore examine an alternative approach which has been proposed in part to circumvent some of the measurement problems associated with the output gap.

5 Speed Limit Rules and Measurement Precision

The absence of sufficient historical dependence or monetary policy inertia in forward-looking models of monetary policy can lead to inefficient outcomes (Woodford 1999, 2003); McCallum and Nelson (2004) find that the resulting inefficiency may be quantitatively important. In response, some advocate targeting the change in the output gap, steering policy by what Walsh (2003a) refers to as "Speed Limit" rules. On a theoretical level, he notes that such rules implicitly introduce lagged output as a state variable, thereby making policy history-dependent and hence improving the trade-off between inflation and output stability. Longworth (2003) and Lam and Pelgrin (2004) discuss the robustness of these results in the face of other sources of inertia.

Other papers have advocated Speed Limit rules for a different reason. In particular, Orphanides et al. (2000), Orphanides (2003a) and Walsh (2003b,c) argue that targeting the change in the output gap avoids much of the measurement problem associated with measuring the level of the output gap.[21] This is based in large measure on the finding in Orphanides (2003b) that US output gap mismeasurement from the mid-sixties to the mid-nineties has been characterized by a level shift, causing the revision in the estimated gaps

[21] Orphanides refers to such rules as *natural growth rate* rules. See Orphanides and Williams (2002) for related analysis of NAIRU-based rules.

to have a mean quite different from zero. In contrast, revisions to the *change* in US output gaps have historically been much more modest.[22] The question of whether such a change in the defined policy target would similarly reduce measurement errors in other countries or other historical episodes has not been systematically addressed to our knowledge. In this section, we therefore examine the conditions under which differencing should reduce measurement error and investigate whether Canadian output gap changes appear to have less important revisions than their corresponding levels.

5.1 Persistence and Precision

Suppose $x_t^0 = x_t + \theta_t$ where x_t^0 is the measured gap and θ_t is the measurement error. Suppose that $\theta_t = \rho_\theta \theta_{t-1} + v_t$ where ρ_θ is close to one. Walsh (2003c) then points out that

> The variance in the measurement error for the level of the gap is $\sigma_v^2/(1-\rho_\theta^2)$; the variance of the error in the measured change in the gap is $2\sigma_v^2/(1+\rho_\theta)$. Thus, as long as $\rho_\theta > 0.5$, the measurement error in the change is smaller than that in the level.[23]

He notes that for $\rho_\theta \approx 0.9$, measurement error variance in changes will be only 20% as large as that in levels. (However, the estimates of ρ presented in Table 3 are ¿ 0.95 for most models, suggesting a reduction in error variance of 90% or more.) Both Walsh (2003b) and Orphanides (2003b) presents further evidence of such improvements by comparing the size of revisions in the estimated US output gap with those in its changes, finding that the latter are much smaller just as the above argument predicts.[24] Walsh concludes

> This suggests that, rather than ignoring the output gap altogether, policymakers might be better off focusing on the gap between the *growth rate* of output and the *growth rate* of trend output, essentially the change in the output gap.[25]

[22]For example, see Figure 2 in Orphanides (2003b).
[23]Walsh(2003c), p. 14.
[24]Walsh (2003b), p. 2, especially Figure 2, and Figure 2 in Orphanides (2003b).
[25]Walsh 2003b, p. 2, emphasis in original.

However, this argument probably exaggerates the *general* benefits from using changes in the gap. To understand why, recall that the absolute size of measurement errors is less important than their size *relative* to what we seek to measure (i.e. the gap or its change.) The true output gap is also believed to be persistent; suppose it has dynamics similar to those of the measurement error, so that $x_t = \lambda x_{t-1} + u_t$ where the error term u_t has variance σ_u^2. The variance of the true output gap is therefore $\sigma_u^2/(1-\lambda^2)$ and that of its change is $2\sigma_u^2/(1-\lambda^2)$. The ratio of the measurement error's variance to that of the gap (i.e. the noise-to-signal ratio) is therefore

$$Q = \frac{\sigma_v^2}{\sigma_u^2} \cdot \frac{1-\lambda^2}{1-\rho_\theta^2} \tag{12}$$

while the ratio for changes in the gap is

$$\frac{2\sigma_v^2}{2\sigma_u^2} \cdot \frac{1-\lambda}{1-\rho_\theta} = Q \cdot \frac{1+\rho_\theta}{1+\lambda} \tag{13}$$

As the above formula makes clear, the relative usefulness of the level and change of the gap depends not on the persistence of the measurement errors but on their persistence *relative* to that of the gap itself. The improvement in N/S due to differencing depends only on $(1+\rho_\theta)/(1+\lambda)$. When the gap is the more persistent of the two ($\lambda > \rho_\theta$), its changes will be less accurately measured than its levels.

This is consistent with the above-mentioned evidence that Orphanides and Walsh present on the revision in *changes* in US output gaps. One striking feature of these revisions is that their mean is large relative to their variability; put another way, US potential output was almost always overestimated during the period these authors examine.[26] This implies that measurement errors over this period were persistent in the extreme, with $\rho_\theta \to 1$. This is therefore the kind of circumstance in which we might expect the improvements from first-differencing to be large. However, evidence presented in Table 3 on the properties of

[26] These authors use "official" output gap estimates and focus on the period immediately following the slowdown in US productivity growth in the early 1970s. Orphanides and van Norden (2002) examine revisions of US output gaps over a longer period and constructed with a variety of detrending methods; they find that some methods give less biased estimates and that the bias may be smaller in 1990s.

Canadian output gap revisions suggest that the mean and the persistence of their measurement error may vary considerably across methods. It is therefore of interest to understand the extent to which Speed Limit rules may avoid measurement error problems in this setting.

5.2 Evidence from the Revision of Gap Estimates

We cannot directly test the relative persistent of the measurement errors and the unobserved "true" output gaps to see whether this last condition is satisfied. However, following Orphanides (2003b) and Walsh (2003b,c), we can compare the properties of the revisions (FL-RT) for the level and the change in our estimated output gaps. Results are shown in Table 4; figures shown for the level of the gap repeat those previously shown in Table 3. The figures for the mean and standard deviations of the revisions echo the US findings for all twelve models; revisions are smaller for the change in the output gap than for its level. We usually (but not always) also find that the maximum and minimum revisions are smaller in size for the change in the output gap than for its level. The last column presents estimated noise-signal ratios. In some respects, these seem to confirm the US findings; changes are subject to relatively less revisions than levels for all but one model (BN). The improvement is particularly large for models with deterministic trends (LT, QT, BT) and those UC models which suffered from severe parameter instability (WT, CL). For these models, the N/S ratio is always more than 50% higher in levels than in changes; in two cases it is more than double. These are also models which which Real-Time estimates made persistent mistakes in identifying the *level* of potential output. For the remaining models, the gains are more modest, with typical improvements ranging around 20%.

Overall, these result suggest that Speed Limit rules will suffer less from measurement error than standard Taylor-type rules in Canadian data as well as in US data. However, in some cases these gains appear to be much more modest than the US historical experience would suggest. It should also be noted that even in changes, measurement error for the Canadian output gap remains quantitatively important, with noise-signal ratios ranging from 0.678 to 1.557.

6 Output Gap Forecast Uncertainty

Up to this point, we have tried to gauge the precision of estimated output gaps using the behaviour of their ex post revisions; that is, the degree to which Real-Time estimates are subsequently revised. One shortcoming of this approach is that a decision maker may not be interested in the Real-Time estimate. As mentioned above,[27] our *Real-Time* estimate for the gap in period t is not available until *t+1* due to a one quarter lag in data reporting. This means, for example, that a decision maker in May 2004 has data series which end in 2004Q1, and so the last available estimated gap is that for 2004Q1. However, the decision maker cares about the state of the economy in the second quarter or, if policy needs to be forward-looking, the future state(s) of the economy. Since no data is available for any of these periods, they must instead rely on *forecasts* of the output gap. The precision of these forecasts will generally be different from that of our Real-Time estimates. However, relatively little is known about how the reliability of forecast gaps or how their reliability varies with the forecast horizon. The remainder of this section considers both these questions.

6.1 Measuring the Reliability of Output Gap Forecasts

Suppose we have quarterly macroeconomic data series x_t available over the period $\{0, \ldots, t, \ldots, T\}$. Up to this point we have used various detrending methods to produce estimates of the output gap at various points in the interval $[0, T]$. Now we wish to produce estimates for some period $T + h$ where $h > 0$. This can be done in two logically distinct steps;

1. **Forecast Step:** create forecasts of the missing data $\{\hat{x}_{T+1}, \ldots, \hat{x}_{T+h}\}$
2. **Detrending Step:**run the usual detrending procedure on the "padded" data set $\{x_0, \ldots, x_T, \hat{x}_{T+1}, \ldots, \hat{x}_{T+h}\}$

[27] See section 4.1.

Obviously the first step requires a forecasting model of the data (or at least of those elements which are relevant for the detrending step.) The choice of this forecasting model is not innocuous; it will generally affect the properties of the resulting gap forecasts. However, our deterministic trend and mechanical filter methods of trend estimation are silent on the question of how to forecast the data. The performance of their forecast gaps is therefore as much a reflection of the forecasting model selected as it is a property of the detrending method. For that reason, we do not investigate the performance of these methods and instead focus on the five models we estimate in a UC framework: BN, WT, CL, HJ and GS.[28] These models specify the dynamics of all the data series used to estimate the output gap; this is all that is required to create model-consistent forecasts of the data series and of the output gaps.[29]

Once gap forecasts have been constructed, we can treat them analogously to Real-Time estimates and examine the properties of their revisions.[30] That is, for a given h -horizon forecast FC(h), we calculate the revision (FL - FC(h)) and compare its properties to those of (FL - RT). (Note that FC(0) = RT, so for $h = 0$ we obtain the familiar FL-RT revision.) For brevity, we restrict our attention to our three preferred measures of precision: ρ, the correlation between FL and FC(h), f, the frequency with which FL and FC(h) have the opposite sign, and noise-to-signal ratio, calculated as the root mean squared revision divided by the standard deviation of FC(h). We are interested in how the properties of these revisions evolve as h increases from 0. To put these into perspective, we also calculate

[28]Recall that our application of the Band-Pass filter pads the available data with forecasts from an AR model of output to produce BP estimates of the gap at the end of sample. It is mechanically straightforward to pad further still with this same forecasting model and examine the resulting gap forecasts. Our decision not to examine such forecasts is therefore somewhat arbitrary; however we judged that forecasts would place too much weight on the auxiliary forecasting model rather than on the Band-Pass filter itself.

[29]For ease of calculation, the Forecasting and Detrending steps mentioned above were replaced by standard formulas for the minimum MSE forecasts of the state vector from a conventional UC model. Note that although they are not included here, the two SVAR gap models (BQ and SV) contain all the information required to produce model-consistent output gap forecasts, and therefore could be treated similarly.

[30]Given that we have a complete model of the data generating process, we could calculate the standard errors of the forecasts directly and examine these instead. However, incorporation of parameter uncertainty into the standard errors of the estimated state vector can be problematic and there is not straightforward way to incorporate the effects of data revision. We therefore leave this for future research.

revisions for $h < 0$. In the latter case, the revision is again (FL - FC(h)) where FC(h) is now the smoothed estimate of the output gap at $T + h < T$ (i.e. an h - period smooth.) This type of revision provides a measure of how quickly gap estimates converge to their Final values.

6.2 Reliability and Forecast Horizon

Table 5 and Figure 6 present evidence on the properties of Canadian output gap forecasts; note that the bottom panel of Figure 6 has a logarithmic vertical scale. A horizon of 0 gives us the results for Real-Time estimates previously presented in Tables 2 and 3. As expected, all five models show a smooth increase in the importance of the ex post revisions as the forecast horizon increases from -16 to 16 quarters, regardless of the three statistical measures we examine.

In all cases, we see important changes in the properties of the gap estimates as the horizon changes from -16 to 0, implying that a substantial portion of the total revisions take place in the first four years. These changes appear to be the smallest for the BN model and largest for CL and WT. However, even after four years, important revisions still occur for some model's estimates. At $h = 16$, BN and WT still have noise-to-signal ratios above 0.8, have signs opposite to those of the Final estimates more the 20% of the time, and have the lowest correlations with the Final estimates. The best-performing at this point are the GS and HJ models, which have N/S ratios below 50% and correlations over 90%.

As the forecast horizon changes from 0 to 16 quarters, however, the performance of some models appears to deteriorate continuously with the increase in the horizon while that of others appears to be unchanged. The Harvey-Jaeger model shows relatively steady deterioration across all three measures, with its correlation falling from roughly 44% to less than zero, the frequency of opposite signs rising from 35% to roughly 50% and the noise-to-signal ratio rising from under 1.3 to over 10. In contrast, the Beveridge-Nelson estimator quickly loses all forecast power; for horizons from 1 to 16 its correlation with the final estimate is roughly constant at 0 and the frequency with which the two have opposite signs hovers just below 50%.[31] The Watson and Harvey-Clark models also show little deterioration, with roughly constant N/S ratios and frequencies of opposite signs. The latter may simply reflect the fact it would be hard for these models to do worse; the sign of their forecast disagrees with that of the Final estimate *more than 50% of the time* for all these horizons, and seems to improve slightly for the longest horizons. Judged by it correlation with the final estimate and the opposite sign frequency, the Gerlach-Smets model is the only one of the five that appears to do better than a random guess at horizons of a few years.[32]

Given the differences across models, it is difficult to draw firm conclusions for policy. The most pessimistic results suggest that output gap forecasts are no better than random noise and that substantial measurement errors may persist even after several years. However, the most optimistic results imply that forecasts of output gaps two to three quarters ahead perform about as well as estimates of the current gap; this suggests that there is no reason to make forward-looking policy more cautious due to measurement problems. The differences in these conclusions highlights the potential importance of uncertainty about the correction specification of the output gap, a question we explore more fully in the next section.

7 Model Misspecification and Output Gap Uncertainty

To this point, we have focussed exclusively on the properties of output gap revisions, using these as an indicator of the reliability of output gap estimates. It is well understood that these revisions should *understate* the measurement errors associated with real-time output gap estimates, since even Final estimates contain often important estimation uncertainty. For that reason, some authors attempt to estimate the overall statistical uncertainty associ-

[31] The explosive behaviour of its noise-to-signal ratio reflects the fact that the variance of the forecast is collapsing towards zero as the variability of the revisions stays roughly constant.

[32] Again, the explosive behaviour of its noise-to-signal ratio reflects the fact that the variance of the forecast is collapsing towards zero as the variability of the revisions stays roughly constant.

ated with end-of-sample output gap estimates, such as Orphanides and van Norden (2002), or Rnstler (2002). This approach also produces underestimates of the total uncertainty owing to the difficulty of incorporating the effects of data revision into a formal statistical framework. However, both this statistical approach and the revision-based approach we have used to this point overlook another potentially important source of output gap estimation uncertainty. Our modest goal in this section is to describe the nature of this omission and to provide some qualitative indications of the severity of this problem.

An assumption common to both the statistical approach and the revision approach to output gap assessment is that the underlying model of the output gap is correctly specified. In the statistical approach, this model commonly takes the form of a particular state-space representation. This representation is then combined with the data to produce confidence intervals for one-sided (i.e. filtered) estimates of the output gap. These intervals may or may not account for estimation uncertainty of some of the model parameters, and they may be produced by analytic or by simulation methods. In all cases, however, the confidence intervals are estimated under the maintained assumption that the state-space model is correctly specified. Similarly, the revision approach restricts its attention to the consistency over time of different estimates of the gap from a given model. Neither approach provides advice about which of the many available models is the best one.

However, it has long been understood that the choice of model can make an important difference to the estimated output gap. Comparison of the different Real-Time estimates shown in Figure 4 show that the available models usually produce a broad dispersion of estimates at any point in time over the past thirty years. Policy makers therefore face model risk; the possibility that the model they use contains errors serious enough to affect their gap estimates in a substantial way. Furthermore, this risk exists over and beyond any estimation risk which the methods discussed above attempt to quantify, since their estimates are made conditional on a particular model of the gap. The pertinent question is whether this additional source of uncertainty is large or small relative to the size of the estimation uncertainty we have documented so far.[33]

Properly quantifying the extent of model risk requires some way of assessing the appropriate weights to assign to a number of different specifications. If these specifications nested, for example, one could use likelihood ratio statistics to assess their relative goodness of fit. Another approach has been to use inflation-forecasting performance as a guide to model selection. Unfortunately, most studies find that output gaps have little ability to improve inflation forecasts out of sample (see Clark and McCracken (2003), Orphanides and van Norden (2003) or Runstler (2002).) Cayen and van Norden (2002) examine this question for Canada and conclude that, with the exception of the BN model, Canadian gap estimates do not significantly improve inflation forecasts, making it therefore difficult to choose among them. In light of these problems, it appears that there is genuine uncertainty about the correct model to use, and a rigorous quantification of the degree of model risk is beyond the scope of this paper.

Instead, we attempt to provide a qualitative measure of model risk, based on the implicit assumption that there is little reliable information available to guide our model selection. To do so, we simply extend our revision methodology to take account of the possibility of a particular kind of misspecification. Specifically, instead of using the difference between the Real-Time and Final estimates from Model A as a measure of estimation uncertainty, we use the difference between the Real-Time estimate from Model A and the Final estimate of Model B as a measure of overall uncertainty. By allowing for the possibility that the RT estimates are formed using the "wrong" model, this approach therefore includes model risk as a factor which may cause revisions. The difficulty of this approach is that the results depend on the pair of models selected in constructing the revision and may therefore vary widely. Due to the absence of a simple method for weighting these results, we restrict

[33] McCallum and Nelson (2004) provide an quantitative example of the potential efficiency loss due to output gap misspecification.

ourselves to two relatively agnostic approaches. First, we present results for all (12 x 12 = 144) possible pairs of models, allowing the reader to apply their own judgement in weighing the various outcomes. Second, we calculate medians over these outcomes. The median has the desirable property that it is robust to outliers; this is important in the absence of a formal model weighting scheme since model selection could be used to bias the mean outcome.[34]

To measure the impact of model risk, we compare the "cross-model" revisions described in this section with the total revisions examined in previous sections. We focus on three measures of the importance of these revisions; the correlation between the Real-Time and the Final estimates, the frequency with which these two estimates have opposite signs, and the Noise-to-Signal ratio. Estimates of these three statistics are presented in Tables 5A through 5C. Figures at the bottom of each table give the median value over all 144 model pairs. Note that the values along the main diagonals of the Tables are simply the figures for the model-consistent revisions previously presented in Tables 2 and 3. This provides a natural benchmark against which to judge the incremental contribution of model risk. Table 7 uses these to express the conditional median (i.e. the median for a given row in Table 6) as an increase or decrease relative to this benchmark.

Looking first at the correlations of the RT and FL gap estimates, we see that model risk appears to be important in the sense that it decreases the correlations substantially. Table 6A shows that the median correlation across all model pairs is less than 15%, and many of the figures in the table are negative. Table 7 shows that the impact of model risk on correlation appears to vary greatly across models. In many cases, incorporating model risk causes a fall in the correlation between Real-Time and Final estimates of 50% or more. For the SV and BN models, the median correlation becomes zero or negative, suggesting

[34] For example, to reduce estimates of model uncertainty, one could simply include many models whose differences are trivial and which produce very similar estimates. Similarly, to increase estimates of model uncertainty, one could simply include highly improbable models which produce extreme output gap estimates. Given the similarities of some of the UC models we consider on the one hand and the obvious instability of some of the deterministic trend models on the other, we take both possibilities seriously.

very important model risk. At the other extreme, the BT model sees a median correlation only 10% less than without model risk, and incorporating model risk seems to *improve* the correlation for the WT model. Looking at Table 6A, we can see that the latter is because the Real-Time gap estimate from the Watson model is often more highly correlated with Final gaps from other models than with its own Final estimate.

Turning next to the frequency with which the Real-Time estimates appear to have the wrong sign, Table 6B shows that regardless of which model pairs we consider, this frequency is never below 30%, frequently exceeds 50% and is sometimes reaches 70%. The median frequency over all model pairs is just over 50%, suggesting that model risk reduces the accuracy with which one can identify the sign of the output gap to roughly that of a random guess. Table 7 shows that conditioning on the estimation model does not have a major impact on the results, with median frequencies lying in a narrow range from 38-58%. However, the *relative* importance of model risk varies considerably, however, with some models showing little or no negative impact (LT, BT, BN and WT) while others (particularly those models with not very persistent gap estimates) show important increases in the frequency of sign errors.

Finally, evidence on the noise-to-signal ratios shown in Table 6C show a wide range of estimates. The median across all model pairs is 1.8 but the distribution is skewed, with very few less than 1 but a few exceeding 10. Conditional on the estimating model, the median N/S ratios are all above 1.1 and most are below 2.0. Again, the relative contribution of model risk varies from model to model. Once more, its addition has little or no negative effect on the LT or BT models and its relative importance is greatest for those models with not very persistent gap estimates. Surprisingly, it also tends to *improve* the N/S ratios for the UC models.

In summary, there is evidence to suggest that model risk may be an important omitted component of the overall uncertainty facing policy makers who use estimates of the output gap. Based on medians calculated across a broad range of models, the accuracy of real-time output gap estimates appears to be still worse than the relatively pessimistic estimates presented earlier suggest. However, there are indications that the degree of model risk may vary with the model used and that in some cases, incorporation of model risk may reduce estimates of the overall risk associated with a particular model.

8 Conclusions

This study has assembled and analyzed a new database of real-time estimates of Canadian output. It shows that revisions are small relative to GDP but are potentially important relative to the size of the business cycle. Data revision also becomes more important if we focus on GDP growth rather than its level. Revisions to the level of GDP are highly persistent; incontrast, revisions to 1-quarter changes are negatively autocorrelated. Mean revisions are positive, implying that GDP gures have historically tended to be revised upwards.

Results from a variety of measures and a broad range of output gap estimates suggest that measurement error in Canadian data may be more severe than previously thought. Most Real-Time estimates have a less than 50% correlation with their corresponding Final estimates, for most models these two gap estimates have opposite signs more than 40% of the time, and the noise-to-signal ratios for the Real-Time gaps are generally above 1. This is more pessimistic than results reported by Orphanides and van Norden (2002) using US data and by Cayen and van Norden (2002) using a shorter sample of Canadian data. Although we find evidence to support Walsh (2003b,c)'s contention that changes in the output gap suffer less from measurement error, the noise-to-signal ratios indicate that measurement error remains substantial in first differences. Further analysis of output gap forecasts and of model risk is not conclusive and results vary considerably from model to model. However, there is evidence to suggest that model risk may be an important omitted component of the overall uncertainty facing policy makers, and that the accuracy of real-time output gap estimates may be still worse than the relatively pessimistic revision-based estimates suggest.

References

Apel, Mikeal, Per Jansson and Hans Lindberg (1995) Estimates of poential output using a multiple indicator unobserved components approach: Evidence from quarterly Swedish time-series data 1970-1995. mimeo, 42 p. Azevedo, Koopman and Rua (2003)

Baxter, Marianne; King, Robert G. (1999) "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series" *Review of Ecnomics and Statistics*, 81, 575-93.

Beveridge, S and C. R. Nelson (1981) "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'," *Journal of Monetary Economics,* 7, 151-174.

Camba-Mendez, G. and D. Rodriguez-Palenzuela (2003) Assessment Criteria for Output Gap Estimates Economic Modelling 20(3) 529-562.

Cayen, Jean-Philippe and Simon van Norden (2002) "La fiabilité des estimations de lécart de production au Canada. *Bank of Canada working paper* 2002-10, 73 p.

Christiano, Lawrence J., and Terry J. Fitzgerald (2003) "The Band Pass Filter," *International Economic Review*, 44(2), 435-465.

Clark, Peter K. (1987) "The Cyclical Component of U.S. Economic Activity," *Quarterly Journal of Economics* 102(4), 797-814.

Clark, Peter K. (1989) Trend Reversion in Real Output and Unemployment Journal of Econometrics, 40, 15-32

Clark, Todd E. and Michael W. McCracken (2003) "The predictive content of the output gap for inflation: resolving the in-sample and out-of-sample evidence." mimeo.

Claus, Iris (2003) Estimating Potential Output for New Zealand Applied Economics, 35, 751-760.

Cross, Philip (1996) Alternative Measures of Business Cycles in Canada: 1947-1992 *Canadian Economic Observer*, Statistics Canada, No. 11-010-XPB.

Cross, Philip (2001) Expenditure on GDP and Business Cycles *Canadian Economic Observer*, Statistics Canada, No. 11-010-XPB.

Croushore, Dean and Tom Stark (2001) "A Real-Time Data Set for Macroeconomists," *Journal of Econometrics*, 105, 111-130.

Dupasquier, Chantal, Alain Guay and Pierre St-Amant (1999) A Survey of Alternative Methodologies for Estimating Potential Output and the Output Gap. Journal of Macroeconomics, 21(3), 577-95.

Fabiani, Silvia and Ricardo Mestre (2001) A System Approach for Measuring the Euro Area NAIRU European Central Bank Working Paper, No. 65. (forthcoming Empirical Economics 2003)

Gerlach, Stefan and Frank Smets (1997) “Output Gaps and Inflation: Unobserable-Components Estimates for the G-7 Countries.” Bank for International Settlements mimeo, Basel.

Gomez, Victor (2001) ”The use of Butterworth filters for trend and cycle estimation in economic time series.” *Journal of Business and Economic Statistics*, 19, 365-373.

Gruen, David, Time Robinson and Andrew Stone (2002) Output Gaps in Real Time: Are They Reliable Enough to Use for Monetary Policy? Reserve Bank of Australia Research Discussion Paper, 2002-06, 49 p.

Harvey, Andrew C. (1989) *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, 554 p.

Harvey, Andrew C. (1985) “Trends and Cycles in Macroeconomic Time Series,” *Journal of Business and Economic Statistics*, 3 216-227.

Harvey, Andrew C. and A. Jaeger (1993) “Detrending, Stylized Facts, and the Business Cycle,” *Journal of Applied Econometrics*, 8, 231-247.

Harvey, Andrew C. and Thomas M. Trimbur (2003) “General Model-based Filters for Extracting Cycles and Trends in Economic Time Series” *The Review of Economics and Statistics*, 85(2), 244-255.

Henderson, R. (1924). ”A New Method of Graduation.” *Actuarial Society of America Transactions*, 25, 29-40.

Hodrick, R, and E. Prescott (1997) “Post-war Business Cycles: An Empirical Investigation,” *Journal of Money, Credit, and Banking,* 29, 1-16.

Kichian, Maral (1999) ”Measuring Potential Output within a State-Space Framework”, *Bank of Canada working paper* 99-9.

Kuttner, Kenneth N. (1994) “Estimating Potential Output as a Latent Variable,” *Journal of Business and Economic Statistics*, 12(3), 361-68.

Lalonde, René (1998) ”Le PIB potentiel des états-unis et ses déterminants: la productivité de la main d’oeuvre et le taux d’activité” *Bank of Canada working paper* 98-13.

Lalonde, René, Jennifer Page and Pierre St-Amant (1998) “Une nouvelle méthode d’estimation de l’écart de production et son application aux États-Unis, au Canada et à l’Allemagne, *Bank of Canada Working Paper*98-21, 38 p.
Lam and Pelgrin (2004)

Laubach, Thomas (2001) Measuring the NAIRU: Evidence from Seven Economies The Review of Economics and Statistics, 83(2), 218-231.

Longworth, David (2003) Comments on Implications of a Changing Economic Structure for the Strategy of Monetary Policy mimeo, presented at the Federal Reserve Bank of Kansas Citys Jackson Hole Conference August 29-30 2003, 11 p.

McCallum, Bennett T. (1993) Discretion versus Policy Rules in Practice: A Comment Carnegie-Rochester Conference Series on Public Policy, 39, 215-220.

McCallum, Bennett T. and Edward Nelson (2004) "Timeless Perspective vs. Discretionary Monetary Policy in Forward-Looking Models" *Federal Reserve Bank of St. Louis Review*, 86(2), 43-56.

Mitchell, James (2003) Should we surprised by the unreliability of real-time output gap estimates? Density estimates for the Eurozone. National Institute of Economic and Social Research mimeo, 47 p.

Morley, James C. (2002) "A state-space approach to calculated the Beveridge-Nelson decomposition" *Economics Letters*, 75, 123-127.

Morley, James C., Charles R. Nelson and Eric Zivot (2003) Why are the Beveridge-Nelson and Unobserved Components decompositions of GDP so different? *Review of Economics and Statistics*, LXXXV(2), 235-243.

Nelson, Edward and Kalin Nikolov (2003) "UK Inflation in the 1970s and 1980s: The Role of Output Gap Mismeasurement" *Journal of Economics and Business* 55(4), 353-70.

Nelson, Charles R. and Charles I. Plosser (1982), "Trends and Random Walks in macroeconomic time series: some evidence and implications." *Journal of Monetary Economics*, 10(2), 139-62.

Orphanides, Athanasios (2001) Monetary Policy Rules Based on Real-Time Data. *American Economic Review*, 91(4), 964-85.

Orphanides, Athanasios, Richard D. Porter, David Reifschneider, Robert Tetlow and Frederico Finan (2000) "Errors in the Measurement of the Output Gap and the Design of Monetary Policy" *Journal of Economics and Business*, 52:117-141.

Orphanides, Athanasios (2003a) "Monetary Policy Evaluation with Noisy Information" *Journal of Monetary Economics*, 50, 605-631.

Orphanides, Athanasios (2003b) "The quest for prosperity without inflation" *Journal of Monetary Economics*, 50, 633-663.

Orphanides, Athanasios and Simon van Norden (1999) "The Reliability of Output Gap Estimates in Real Time," *Finance and Economics Discussion Series* 1999-38.

Orphanides, Athanasios and Simon van Norden (2002) "The Unreliability of Output Gap Estimates in Real Time," *Review of Economics and Statistics*, 84(4) 569-83.

Orphanides, Athanasios and Simon van Norden (2003) “The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time, *CIRANO working paper* 2003s-01, 49 p.

Orphanides, Athanasios and John C. Williams (2002) ”Robust Monetary Policy Rules with Unknown Natural Rates” Brookings Papers on Economic Activity, 2, 63- 118.

Rudebusch, Glenn D. (2001) Is the Fed Too Timid? Monetary Policy in an Uncertain World. The Review of Economic and Statistics, LXXXIII(2), 203-217.

Rnstler, Gerhard (2002) The information content of real-time output gap estimates: an application to the euro area European Central Bank Working Paper, No. 182.

Staiger, Douglas, James H. Stock, and Mark W. Watson (1997) “How Precise are Estimates of the Natural Rate of Unemployment?” in Romer, Christina and David Romer, eds. *Reducing Inflation: Motivation and Strategy*, Chicago: University of Chicago Press.

St-Amant, Pierre and Simon van Norden, “Measurement of the Output Gap: A discussion of recent research at the Bank of Canada,” Bank of Canada Technical Report No. 79, 1998.

Stock, James and Mark Watson (1998) Business Cycle Fluctuations in U.S. Macroeconomic Time Series. *NBER Working Paper* No. 6528, 1998, 83 p., prepared for *The Handbook of Macroeconomics*, edited by John B. Taylor and Michael Woodford. Taylor (1993)

van Norden, Simon, “Filtering for Current Analysis” Bank of Canada Working Paper, 2002.

Walsh, Carl (2003a) ”Speed Limit Policies: The Output Gap and Optimal Monetary Policy” *American Economic Review*, 93(1), 265-278.

Walsh, Carl (2003b) ”Minding the Speed Limit” *FRBSF Economic Letter.* 2003-14, 3 p.

Walsh, Carl (2003c) ”Implications of a Changing Economic Structure for the Strategy of Monetary Policy,” in *Monetary Policy and Uncertainty: Adapting to a Changing Economy*, Federal Reserve Bank of Kansas City, 2003.

Watson, Mark W. (1986) “Univariate Detrending Methods with Stochastic Trends,” *Journal of Monetary Economics*, 18, 49-75.

Whittaker, E. T. (1923) “On a New Method of Graduation.” *Proceedings of the Edinburgh Mathematical Society*, 41, 63-75.

Woodford, Michael (1999) ”Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability?” in *New Challenges for Monetary Policy: A Symposium Sponsored by the Federal Reserve Bank of Kansas City*, Federal Reserve Bank of Kansas City, 277-316.

Woodford, Michael (2003) *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, 800 p.

Table 1: Revisions in Real GDP 1972Q1-2003Q4

RealTime	Level	1Q Change	2Q Change	4Q Change	8Q Change
Mean	13.453205	0.006253	0.012878	0.026419	0.054596
Std. Error	0.256788	0.008877	0.014867	0.024403	0.036285
Min.	12.961364	−0.020835	−0.044006	−0.064129	−0.050265
Max.	13.912242	0.028367	0.056836	0.077548	0.140081
$\rho(RT, FL)$	0.998423	0.750134	0.868334	0.924300	0.937800
Revisions					
Mean	0.002006	0.001315	0.002278	0.004547	0.008056
Std. Error	0.015412	0.006045	0.007393	0.009315	0.012659
Min.	−0.031025	−0.015131	−0.013328	−0.013638	−0.023424
Max.	0.038499	0.020421	0.024344	0.032228	0.037645
$\rho(r_t, r_{t-1})$	0.883964	−0.207029	0.252752	0.597021	0.550104

Notes:

RealTime refers to the real-time estimate; that is, for the last period for which published data are available.

Revisions = final estimate - realtime estimate.

Level refers to the estimate of log real GDP, after adjustments for rebasing in 1975Q1, 1986Q2, 1990Q1, 1997Q3 and 2001Q1.

***n*Q Change** refers to the realtime estimate of the change in real GDP over the most recent n quarters. Because this does not require an adjustment for rebasing, this differs slightly from the n-quarter change in *Level*.

Std. Error is the standard error of the data.

$\rho(RT, FL)$ is the correlation between the realtime and final estimates.

$\rho(r_t, r_{t-1})$ is the correlation between the current and previous period's revision. *Ceteris paribus*, this will increase with n due the increasing overlap of the differencing periods.

Table 2: Output Gap Summary Statistics 1972Q1-2003Q4

Gap	Ver.	Mean	Std. Dev.	Min.	Max.	Corr(X,FL)	Opp. Sign
LT	RT	−0.1114	0.0641	−0.2012	0.0351	0.8147	0.507
LT	QR	−0.0973	0.0569	−0.1896	0.0156	0.8607	0.531
LT	FL	0.0137	0.0921	−0.1231	0.1409	1.0000	0.000
QT	RT	−0.0292	0.0461	−0.1478	0.0623	0.5965	0.398
QT	QR	−0.0089	0.0397	−0.1030	0.0781	0.6110	0.375
QT	FL	0.0041	0.0394	−0.0726	0.0565	1.0000	0.000
BT	RT	−0.0236	0.0319	−0.0950	0.0351	0.7005	0.390
BT	QR	−0.0113	0.0286	−0.0724	0.0443	0.7427	0.242
BT	FL	0.0019	0.0292	−0.0515	0.0551	1.0000	0.000
BP	RT	−0.0104	0.0107	−0.0436	0.0123	0.6370	0.437
BP	QR	−0.0074	0.0101	−0.0378	0.0117	0.6735	0.367
BP	FL	0.0008	0.0142	−0.0489	0.0324	1.0000	0.000
HP	RT	−0.0030	0.0175	−0.0495	0.0320	0.3842	0.445
HP	QR	−0.0008	0.0164	−0.0465	0.0299	0.4468	0.398
HP	FL	0.0003	0.0150	−0.0512	0.0313	1.0000	0.000
BQ	RT	−0.0042	0.0136	−0.0656	0.0195	0.4881	0.273
BQ	QR	−0.0020	0.0087	−0.0285	0.0233	0.5720	0.210
BQ	FL	−0.0007	0.0090	−0.0189	0.0215	1.0000	0.000
SV	RT	−0.0015	0.0064	−0.0288	0.0103	0.2224	0.343
SV	QR	−0.0017	0.0056	−0.0258	0.0154	0.6132	0.210
SV	FL	−0.0004	0.0025	−0.0097	0.0071	1.0000	0.000
BN	RT	0.0005	0.0015	−0.0030	0.0050	0.5242	0.328
BN	QR	0.0002	0.0012	−0.0018	0.0046	0.4953	0.273
BN	FL	0.0004	0.0019	−0.0046	0.0056	1.0000	0.000
WT	RT	−0.0237	0.0147	−0.0635	−0.0017	0.3017	0.578
WT	QR	−0.0175	0.0163	−0.0713	−0.0014	0.1122	0.578
WT	QF	−0.0050	0.0155	−0.0292	0.0212	0.9606	0.226
WT	FL	0.0076	0.0216	−0.0264	0.0364	1.0000	0.000
CL	RT	−0.0139	0.0159	−0.0732	−0.0009	−0.1888	0.625
CL	QR	−0.0073	0.0067	−0.0434	0.0004	−0.0477	0.601
CL	QF	−0.0068	0.0051	−0.0171	0.0011	0.7529	0.554
CL	FL	0.0023	0.0139	−0.0237	0.0198	1.0000	0.000
HJ	RT	−0.0036	0.0190	−0.0450	0.0256	0.4352	0.351
HJ	QR	−0.0037	0.0196	−0.0363	0.0329	0.5138	0.390
HJ	QF	−0.0031	0.0176	−0.0476	0.0335	0.7019	0.351
HJ	FL	0.0014	0.0237	−0.0676	0.0500	1.0000	0.000
GS	RT	−0.0064	0.0131	−0.0302	0.0238	0.4702	0.382
GS	QR	−0.0053	0.0138	−0.0328	0.0239	0.5345	0.414
GS	QF	−0.0061	0.0159	−0.0475	0.0261	0.7144	0.382
GS	FL	−0.0002	0.0262	−0.0669	0.0419	1.0000	0.000

Notes:
All statistics based on gap estimates from 1972Q1-2003Q4 (128 observations.) Univariate gaps are estimated using log output data starting in 1947Q1. Gaps requiring prices are estimated using price data starting in 1949Q1.
Gap refers to the model used to estimate the output gap
Version refers to the method used to estimate the output gap model (RT = Real Time, QR = Quasi-Real, QF = Quasi-Final, FL = Final.) Final estimates are those based on 2003Q4 vintage.
Mean is the average value of the estimated gap.
Std. Dev. is the standard deviation of the estimated gap.
Min. / Max. is the minimum /maximum value of the estimated gap.
Corr(X,FL) is the correlation of estimated gap with the Final estimate.
Opp. Sign is the fraction of the total number of observations in which the estimated gap has a sign different from that of the final estimate.

Table 3: Output Gap Revisions 1972Q1-2003Q4

Gap	Revision	Mean	Std. Dev.	Min.	Max.	ρ	N/S
LT	FL-RT	0.1251	0.0545	0.0000	0.2235	0.990	2.126
LT	QR-RT	0.0141	0.0169	−0.0218	0.0562	0.907	0.342
LT	FL-QR	0.1110	0.0520	0.0000	0.1696	0.999	1.909
QT	FL-RT	0.0333	0.0389	−0.0242	0.1138	0.985	1.107
QT	QR-RT	0.0203	0.0126	−0.0135	0.0537	0.872	0.516
QT	FL-QR	0.0130	0.0349	−0.0438	0.0632	0.998	0.804
BT	FL-RT	0.0255	0.0238	−0.0340	0.0703	0.956	1.091
BT	QR-RT	0.0123	0.0115	−0.0218	0.0319	0.837	0.527
BT	FL-QR	0.0132	0.0207	−0.0155	0.0668	0.981	0.769
BP	FL-RT	0.0111	0.0111	−0.0152	0.0443	0.885	1.465
BP	QR-RT	0.0030	0.0034	−0.0049	0.0115	0.724	0.424
BP	FL-QR	0.0081	0.0105	−0.0111	0.0376	0.921	1.240
HP	FL-RT	0.0033	0.0182	−0.0356	0.0391	0.933	1.048
HP	QR-RT	0.0023	0.0061	−0.0095	0.0212	0.601	0.367
HP	FL-QR	0.0011	0.0165	−0.0289	0.0407	0.977	0.939
BQ	FL-RT	0.0036	0.0121	−0.0222	0.0467	0.844	0.924
BQ	QR-RT	0.0023	0.0120	−0.0222	0.0634	0.827	0.893
BQ	FL-QR	0.0013	0.0082	−0.0167	0.0286	0.684	0.606
SV	FL-RT	0.0011	0.0063	−0.0129	0.0242	0.645	1.002
SV	QR-RT	−0.0002	0.0050	−0.0194	0.0144	0.561	0.783
SV	FL-QR	0.0013	0.0045	−0.0097	0.0216	0.675	0.734
BN	FL-RT	−0.0001	0.0017	−0.0061	0.0039	−0.350	1.108
BN	QR-RT	−0.0003	0.0014	−0.0051	0.0036	−0.317	0.955
BN	FL-QR	0.0002	0.0016	−0.0072	0.0055	0.274	1.087
WT	FL-RT	0.0313	0.0222	0.0000	0.0991	0.974	2.608
WT	QR-RT	0.0062	0.0098	−0.0186	0.0317	0.882	0.785
WT	QF-QR	0.0124	0.0231	−0.0016	0.0914	0.972	1.782
WT	FL-QF	0.0127	0.0080	−0.0001	0.0268	0.989	1.018
CL	FL-RT	0.0162	0.0231	−0.0116	0.0929	0.921	1.763
CL	QR-RT	0.0066	0.0105	−0.0062	0.0511	0.722	0.776
CL	QF-QR	0.0005	0.0084	−0.0101	0.0407	0.933	0.527
CL	FL-QF	0.0091	0.0106	−0.0101	0.0244	0.984	0.874
HJ	FL-RT	0.0050	0.0231	−0.0401	0.0643	0.957	1.237
HJ	QR-RT	−0.0001	0.0076	−0.0195	0.0160	0.805	0.400
HJ	QF-QR	0.0006	0.0078	−0.0156	0.0208	0.970	0.407
HJ	FL-QF	0.0045	0.0169	−0.0295	0.0361	0.967	0.918
GS	FL-RT	0.0062	0.0231	−0.0463	0.0537	0.954	1.818
GS	QR-RT	0.0011	0.0057	−0.0135	0.0174	0.715	0.440
GS	QF-QR	−0.0007	0.0066	−0.0189	0.0202	0.841	0.502
GS	FL-QF	0.0058	0.0185	−0.0321	0.0394	0.959	1.477

Notes:
All statistics based on gap estimates from 1972Q1-2003Q4 (128 observations.) Univariate gaps are estimated using log output data starting in 1947Q1. Gaps requiring prices are estimated using price data starting in 1949Q1.
Gap refers to the model used to estimate the output gap
Revision refers to the series used to construct the revisions. *FL-RT* is the total revision and may be decomposed into *QR-RT* (data revision) + *FL-QR*, or into *QR-RT* + *QF-QR* (parameter revision) + *FL-QF*.
Mean is the average value of the revision.
Std. Dev. is the standard deviation of the revision.
Min. / Max. is the minimum /maximum value of the revision.
ρ is the autocorrelation coefficient of the revision.
N/S is the noise/signal ratio, calculated as the root-mean-squared revision divided by the standard deviation of the Real-Time estimate of the output gap.

Table 4: Total Revisions: Levels versus Differences 1972Q1-2003Q4

Model	Form	Mean	Std. Dev.	Min.	Max.	ρ	N/S
LT	Level	0.1251	0.0545	0.0000	0.2235	0.990	2.126
LT	Diff.	−0.0008	0.0076	−0.0305	0.0226	−0.129	0.719
QT	Level	0.0333	0.0389	−0.0242	0.1138	0.985	1.107
QT	Diff.	−0.0001	0.0067	−0.0196	0.0211	−0.144	0.678
BT	Level	0.0255	0.0238	−0.0340	0.0703	0.956	1.091
BT	Diff.	0.0001	0.0070	−0.0170	0.0245	−0.059	0.689
BP	Level	0.0111	0.0111	−0.0152	0.0443	0.885	1.465
BP	Diff.	0.0000	0.0054	−0.0150	0.0168	−0.030	1.291
HP	Level	0.0033	0.0182	−0.0356	0.0391	0.933	1.048
HP	Diff.	0.0000	0.0068	−0.0169	0.0203	−0.012	0.842
BQ	Level	0.0036	0.0121	−0.0222	0.0467	0.844	0.924
BQ	Diff.	−0.0001	0.0068	−0.0249	0.0204	−0.123	0.790
SV	Level	0.0011	0.0063	−0.0129	0.0242	0.645	1.002
SV	Diff.	−0.0001	0.0053	−0.0190	0.0183	−0.185	0.973
BN	Level	−0.0001	0.0017	−0.0061	0.0039	−0.350	1.108
BN	Diff.	0.0000	0.0027	−0.0089	0.0093	−0.700	1.271
WT	Level	0.0313	0.0222	0.0000	0.0991	0.974	2.608
WT	Diff.	−0.0005	0.0050	−0.0363	0.0142	0.358	1.048
CL	Level	0.0162	0.0231	−0.0116	0.0929	0.921	1.763
CL	Diff.	−0.0003	0.0091	−0.0492	0.0552	−0.257	1.060
HJ	Level	0.0050	0.0231	−0.0401	0.0643	0.957	1.237
HJ	Diff.	0.0001	0.0069	−0.0337	0.0193	0.047	1.072
GS	Level	0.0062	0.0231	−0.0463	0.0537	0.954	1.818
GS	Diff.	0.0000	0.0071	−0.0193	0.0174	0.127	1.557

Notes:
Model refers to the model used to estimate the output gap
Form indicates whether the revisions (Final - Real-Time) are calculated based on the level or the first difference of the output gap.
Mean is the average value of the revision.
Std. Dev. is the standard deviation of the revision.
Min. / Max. is the minimum /maximum value of the revision.
ρ is the autocorrelation coefficient of the revision.
N/S is the noise/signal ratio, calculated as the root-mean-squared revision divided by the standard deviation of the Real-Time estimate of the output gap (Level) or its first difference (Diff.)

Table 5 - Revision of Gap Forecasts

Correlation(FL, FC_h)					
Horizon	BN	WT	CL	HJ	GS
−16	0.693	0.843	0.873	0.915	0.947
−12	0.682	0.822	0.798	0.913	0.933
−8	0.598	0.675	0.514	0.886	0.905
−4	0.504	0.487	0.154	0.753	0.795
−2	0.567	0.389	−0.036	0.622	0.677
−1	0.587	0.343	−0.123	0.529	0.594
0	0.524	0.302	−0.189	0.435	0.470
1	−0.105	0.266	−0.214	0.441	0.473
2	−0.013	0.226	−0.224	0.413	0.467
4	0.048	0.129	−0.231	0.329	0.423
8	0.038	−0.103	−0.270	0.093	0.372
12	0.008	−0.316	−0.364	−0.078	0.353
16	0.049	−0.479	−0.453	−0.163	0.294

Frequency of Opposite Signs					
Horizon	BN	WT	CL	HJ	GS
−16	0.205	0.223	0.152	0.188	0.063
−12	0.216	0.267	0.190	0.172	0.086
−8	0.300	0.383	0.308	0.183	0.117
−4	0.315	0.508	0.476	0.266	0.258
−2	0.325	0.571	0.556	0.317	0.333
−1	0.299	0.575	0.606	0.339	0.378
0	0.328	0.578	0.625	0.352	0.383
1	0.465	0.574	0.620	0.357	0.372
2	0.462	0.569	0.615	0.369	0.369
4	0.462	0.561	0.606	0.386	0.348
8	0.471	0.544	0.588	0.449	0.324
12	0.457	0.529	0.571	0.514	0.371
16	0.403	0.514	0.556	0.486	0.361

Noise-to-Signal Ratio					
Horizon	BN	WT	CL	HJ	GS
−16	0.852	0.810	0.550	0.428	0.378
−12	0.862	1.058	0.744	0.455	0.420
−8	0.926	1.441	1.081	0.574	0.533
−4	1.077	1.914	1.385	0.876	0.893
−2	1.036	2.288	1.597	1.097	1.301
−1	0.984	2.473	1.679	1.196	1.583
0	1.108	2.608	1.763	1.237	1.819
1	2.594	2.647	1.791	1.297	1.913
2	2.919	2.660	1.798	1.428	2.079
4	3.996	2.673	1.793	1.884	2.527
8	7.910	2.698	1.785	4.011	3.678
12	15.267	2.706	1.805	8.355	5.275
16	26.396	2.689	1.823	10.319	7.521

Notes:

Correlation is the correlation between the Final estimate and the h - period ahead forecast for the same period.

Frequency of Opposite Signs is the fraction of observations in which the above two estimates have opposite signs.

Noise-to-Signal Ratio $= \sqrt{T^{-1} \cdot \sum(FL - FC_h)^2}/\sigma_{FC_h}$

Table 6A
Cross-Model Correlations of RT and FL Estimates

Final:	LT	QT	BT	BP	HP	BQ	SV	BN	WT	CL	HJ	GS
Real-Time:												
LT	0.81	0.77	0.48	0.45	0.41	0.23	−0.18	−0.19	0.85	0.86	0.56	0.75
QT	−0.27	0.59	0.39	0.45	0.46	−0.03	−0.14	−0.24	−0.08	0.26	0.62	0.59
BT	0.18	0.78	0.70	0.66	0.63	0.12	−0.14	−0.29	0.36	0.62	0.78	0.80
BP	−0.01	0.39	0.34	0.63	0.63	0.13	−0.02	−0.48	0.08	0.27	0.56	0.52
HP	−0.13	−0.08	0.03	0.39	0.38	0.10	0.18	−0.56	−0.13	−0.10	0.13	0.05
BQ	−0.11	0.15	0.27	0.53	0.58	0.48	−0.09	−0.14	−0.01	0.13	0.47	0.36
SV	−0.40	−0.06	0.03	0.00	0.06	0.14	0.22	−0.06	−0.32	−0.19	0.10	−0.00
BN	−0.10	−0.15	−0.08	−0.14	−0.13	−0.07	−0.16	0.52	−0.11	−0.14	−0.12	−0.15
WT	0.14	0.37	0.63	0.22	0.25	0.39	0.03	0.17	0.30	0.42	0.47	0.41
CL	−0.40	−0.11	0.15	−0.17	−0.12	0.12	0.11	0.28	−0.30	−0.18	−0.00	−0.12
HJ	−0.22	0.23	0.33	0.44	0.45	−0.07	−0.03	−0.28	−0.10	0.09	0.43	0.34
GS	−0.24	0.40	0.30	0.47	0.48	−0.11	−0.13	−0.21	−0.10	0.17	0.52	0.47
Median = 0.1428												

Table 6B
Cross-Model Frequency of Opposite Signs (RT and FL)

Final:	LT	QT	BT	BP	HP	BQ	SV	BN	WT	CL	HJ	GS
Real-Time:												
LT	0.50	0.52	0.60	0.50	0.46	0.46	0.51	0.58	0.50	0.54	0.50	0.53
QT	0.69	0.39	0.46	0.44	0.42	0.55	0.51	0.60	0.68	0.51	0.39	0.42
BT	0.55	0.30	0.39	0.36	0.35	0.57	0.53	0.55	0.54	0.37	0.32	0.32
BP	0.54	0.51	0.52	0.43	0.39	0.48	0.49	0.64	0.53	0.50	0.46	0.50
HP	0.58	0.63	0.59	0.46	0.44	0.52	0.53	0.72	0.57	0.60	0.57	0.62
BQ	0.57	0.50	0.46	0.41	0.36	0.27	0.40	0.55	0.56	0.53	0.44	0.48
SV	0.64	0.52	0.53	0.52	0.53	0.42	0.34	0.52	0.64	0.54	0.52	0.50
BN	0.48	0.48	0.46	0.51	0.53	0.56	0.55	0.32	0.47	0.47	0.51	0.49
WT	0.58	0.60	0.57	0.57	0.52	0.39	0.45	0.58	0.57	0.62	0.57	0.59
CL	0.58	0.60	0.57	0.57	0.52	0.39	0.45	0.58	0.57	0.62	0.57	0.59
HJ	0.52	0.41	0.39	0.36	0.38	0.57	0.50	0.63	0.51	0.39	0.35	0.40
GS	0.62	0.35	0.47	0.40	0.39	0.54	0.53	0.57	0.61	0.46	0.37	0.38
Median = 0.5195												

Table 6C
Cross-Model Noise/Signal Ratios

Final:	LT	QT	BT	BP	HP	BQ	SV	BN	WT	CL	HJ	GS
Real-Time:												
LT	2.12	1.91	1.97	1.97	1.97	1.98	2.00	2.01	1.99	1.95	1.95	1.88
QT	2.62	1.10	1.16	1.10	1.09	1.19	1.18	1.19	1.38	1.17	1.02	1.01
BT	3.10	1.15	1.09	1.08	1.08	1.23	1.24	1.26	1.38	1.13	0.99	0.93
BP	8.94	3.65	2.80	1.46	1.47	1.51	1.38	1.48	2.74	1.83	2.14	2.30
HP	5.53	2.55	1.92	1.02	1.04	1.07	0.99	1.07	1.78	1.36	1.58	1.74
BQ	7.06	2.96	2.14	1.04	1.01	0.92	1.06	1.08	2.08	1.41	1.59	1.83
SV	15.0	6.37	4.68	2.46	2.50	1.60	1.00	1.10	4.09	2.63	3.77	4.23
BN	61.3	26.2	19.3	9.55	10.0	6.11	2.16	1.10	15.1	9.43	15.8	17.4
WT	6.68	3.11	2.33	2.06	2.04	1.82	1.87	1.91	2.60	2.04	2.23	2.28
CL	6.46	2.97	2.17	1.71	1.70	1.36	1.30	1.32	2.33	1.76	2.03	2.19
HJ	5.22	2.11	1.54	0.96	0.97	1.14	1.02	1.04	1.69	1.21	1.23	1.40
GS	7.46	2.85	2.23	1.20	1.20	1.34	1.13	1.15	2.27	1.47	1.64	1.81
Median = 1.829												

Table 7 - Model Risk Conditional on Estimation Model

RT Model:	Correlation	Decrease	Opp. Sign	Increase	N/S	Increase
LT	0.522	0.359	0.511	0.007	1.973	−0.071
QT	0.330	0.446	0.492	0.235	1.170	0.056
BT	0.632	0.096	0.382	−0.020	1.147	0.051
BP	0.310	0.512	0.507	0.160	1.990	0.357
HP	0.045	0.882	0.582	0.307	1.475	0.407
BQ	0.217	0.553	0.476	0.742	1.503	0.626
SV	−0.000	1.002	0.523	0.522	3.204	2.198
BN	−0.133	1.255	0.488	0.488	12.58	10.356
WT	0.336	−0.115	0.578	0.000	2.148	−0.176
CL	−0.119	0.365	0.578	−0.075	1.898	0.076
HJ	0.165	0.620	0.410	0.166	1.226	−0.009
GS	0.237	0.495	0.468	0.224	1.561	−0.141

Notes:
RT Model refers to the model used to estimate the Real-Time output gap
Correlation refers to the correlation between the Real-Time estimate from the model indicated the Final estimate from other models. The figure given is the median value across all models for the Final estimates. See Table 6A for details.
Decrease is the above correlation, expressed as a fractional decrease of the own-model correlation reported in Table 2. (E.g. 1.0 implies that the median correlation is zero.)
Opp. Sign is the fraction of the total number of observations in which the Real-Time estimate of the gap has a sign different from that of the final estimate. The figure given is the median value across all models for the Final estimates. See Table 6B for details.
Increase is the above frequency, expressed as a fractional increase over the own-model frequency reported in Table 2. (E.g. 1.0 implies that the median frequency across all models is double that of the own-model frequency.)
N/S is the noise/signal ratio, calculated as the root-mean-squared revision divided by the standard deviation of the Real-Time estimate of the output gap. The figure given is the median value across all models for the Final estimates. See Table 6C for details.
Increase is the above ratio, expressed as a fractional increase over the own-model ratio reported in Table 3. (E.g. 1.0 implies that the median N/S ratio across all models is double that of the own-model ratio.)

Figure 1: Log Canadian Real GDP

Estimates and Revisions 1972Q1-2003Q4

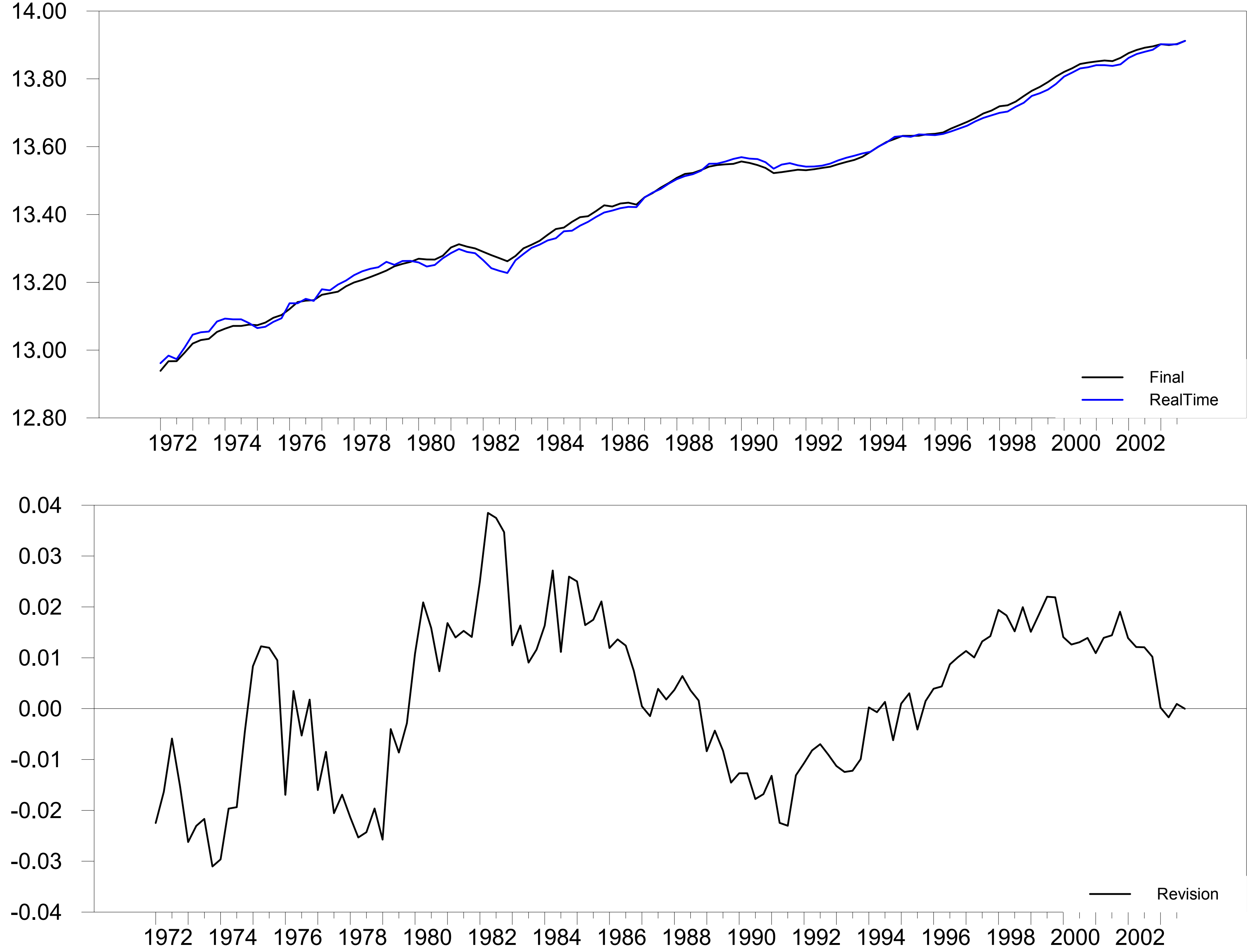


Figure 2: Change in Log Canadian GDP

1972Q1-2003Q4

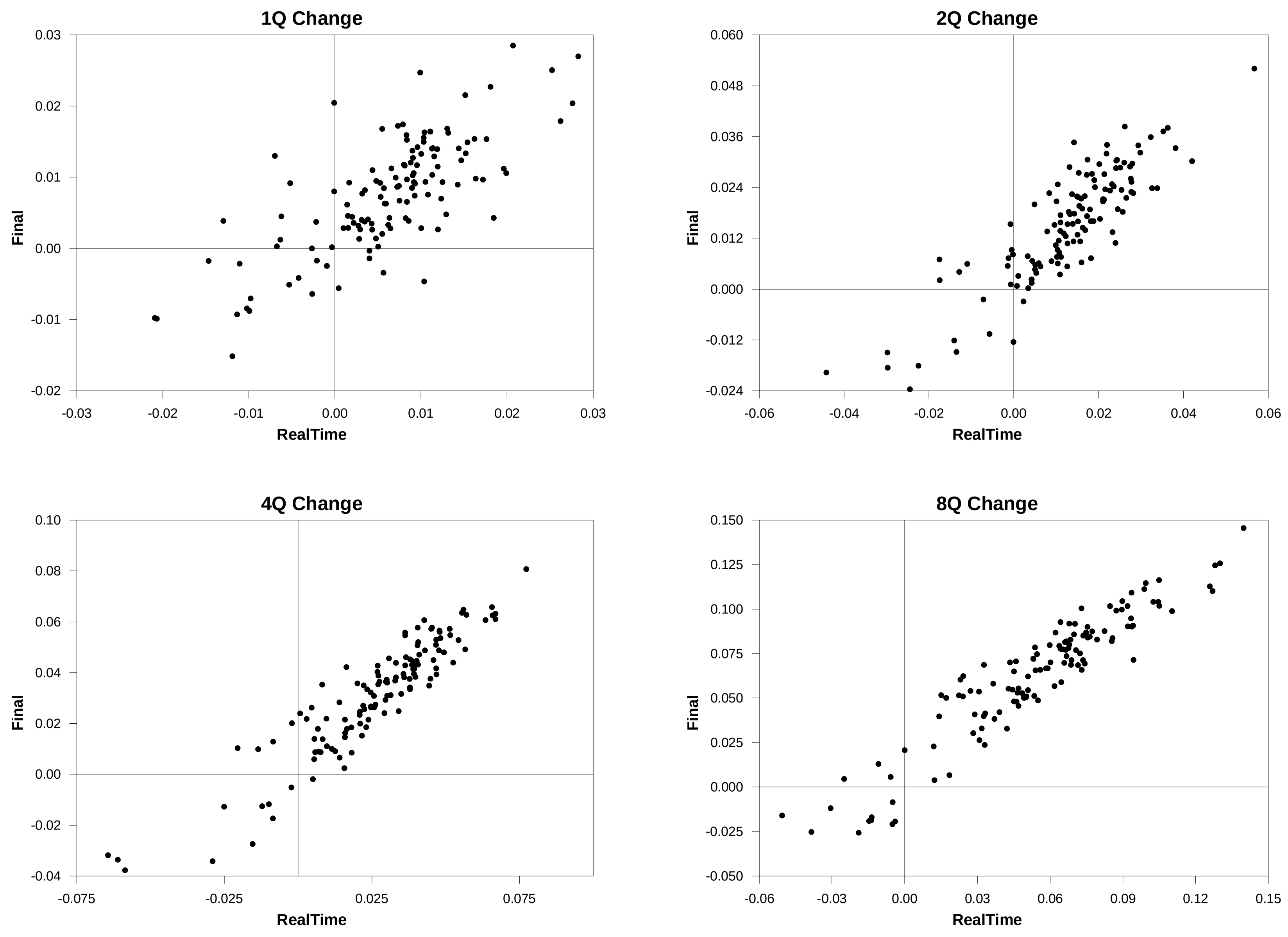


Figure 3: Change in Log Canadian GDP

Real-Time and Revisions 1972Q1-2003Q4

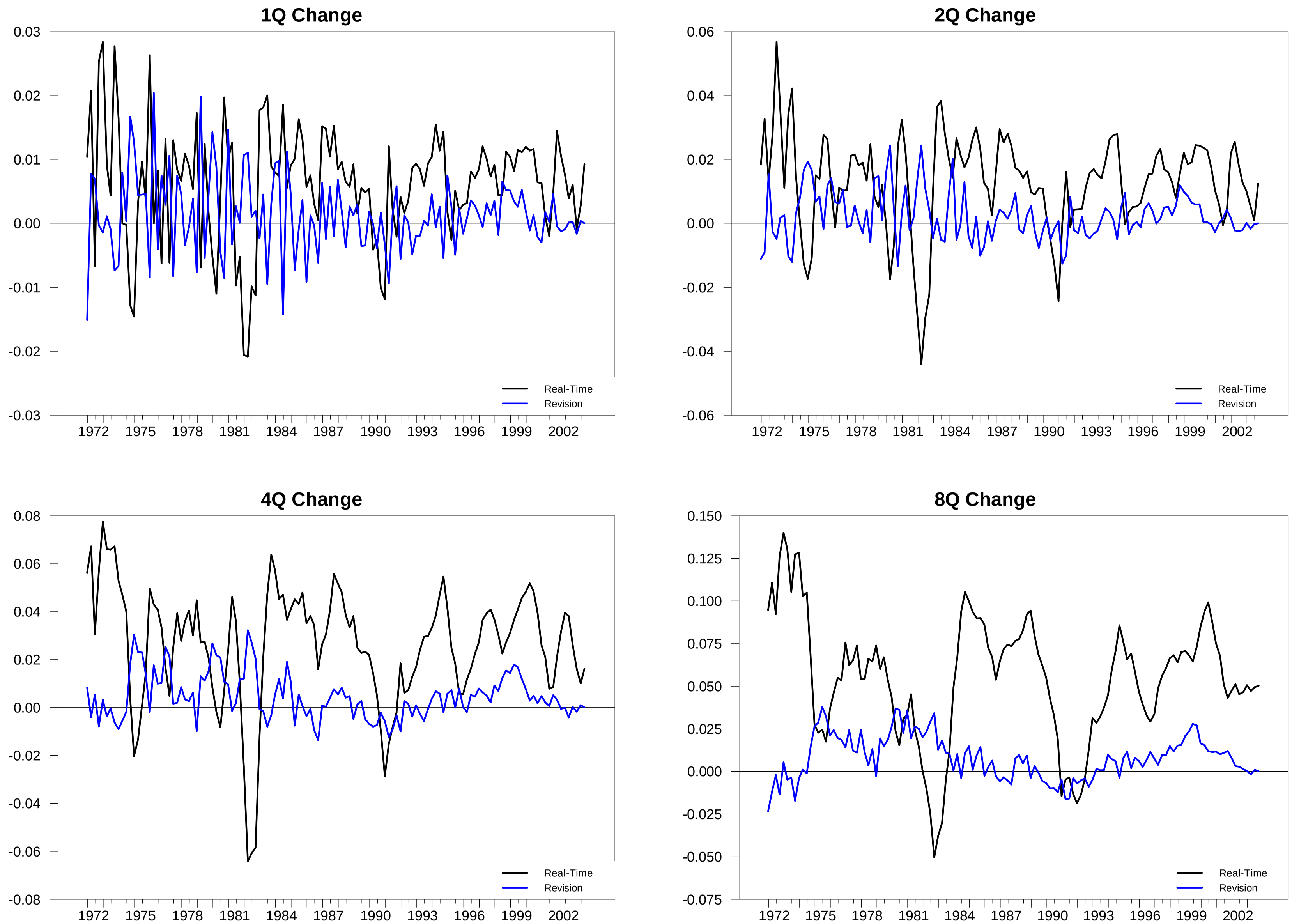


Figure 4: Output Gap Estimates

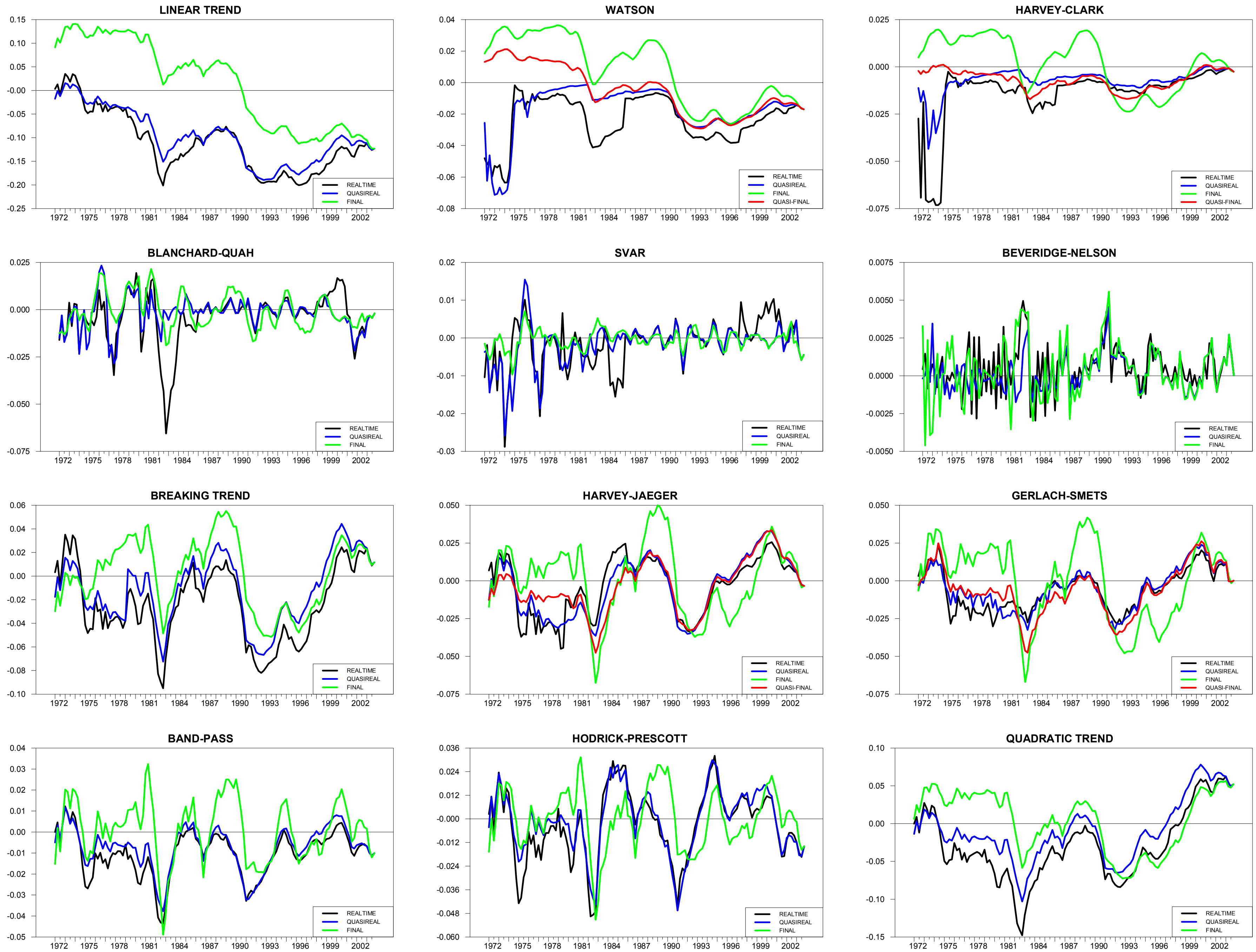


Figure 5: Real-Time Gaps and Revisions

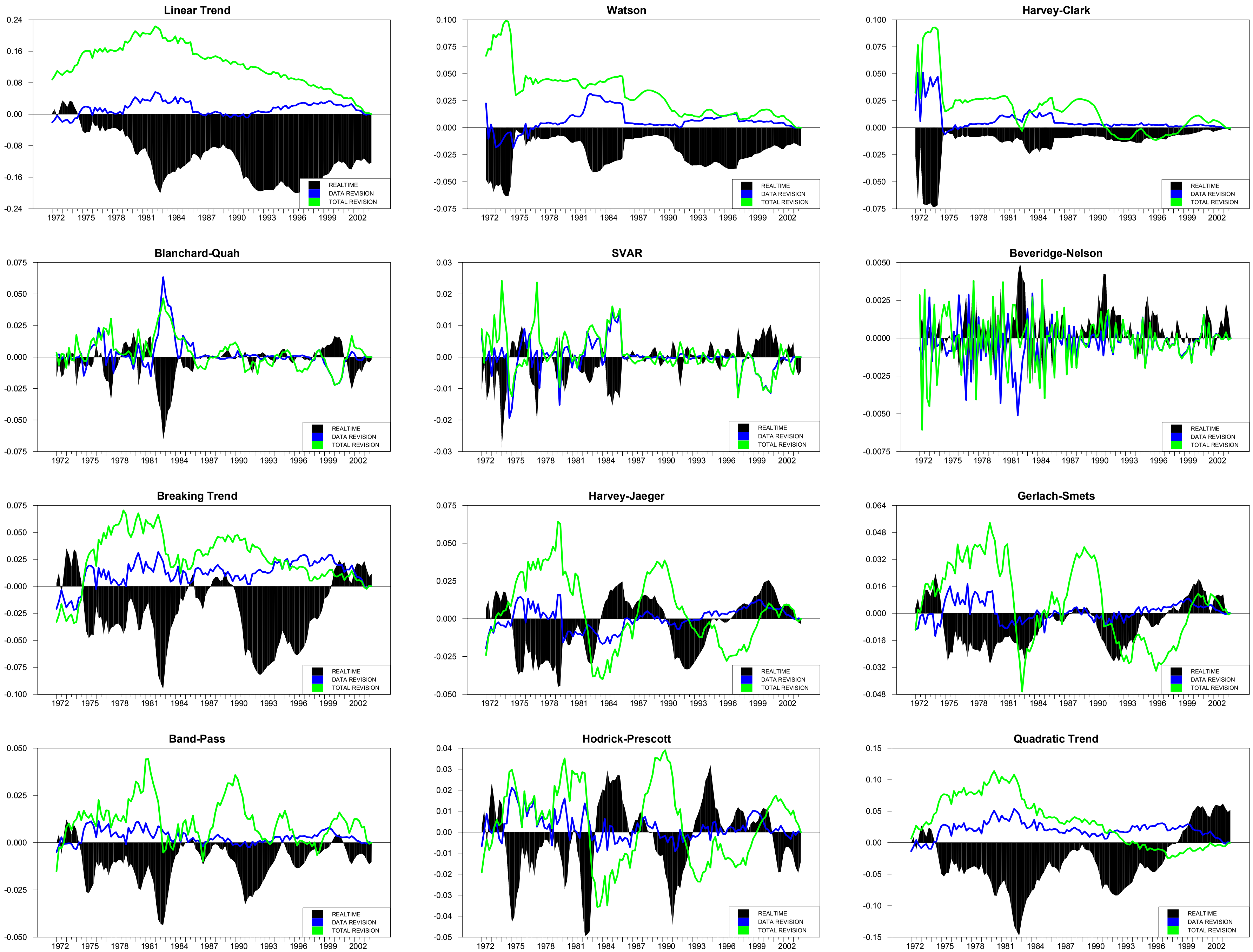
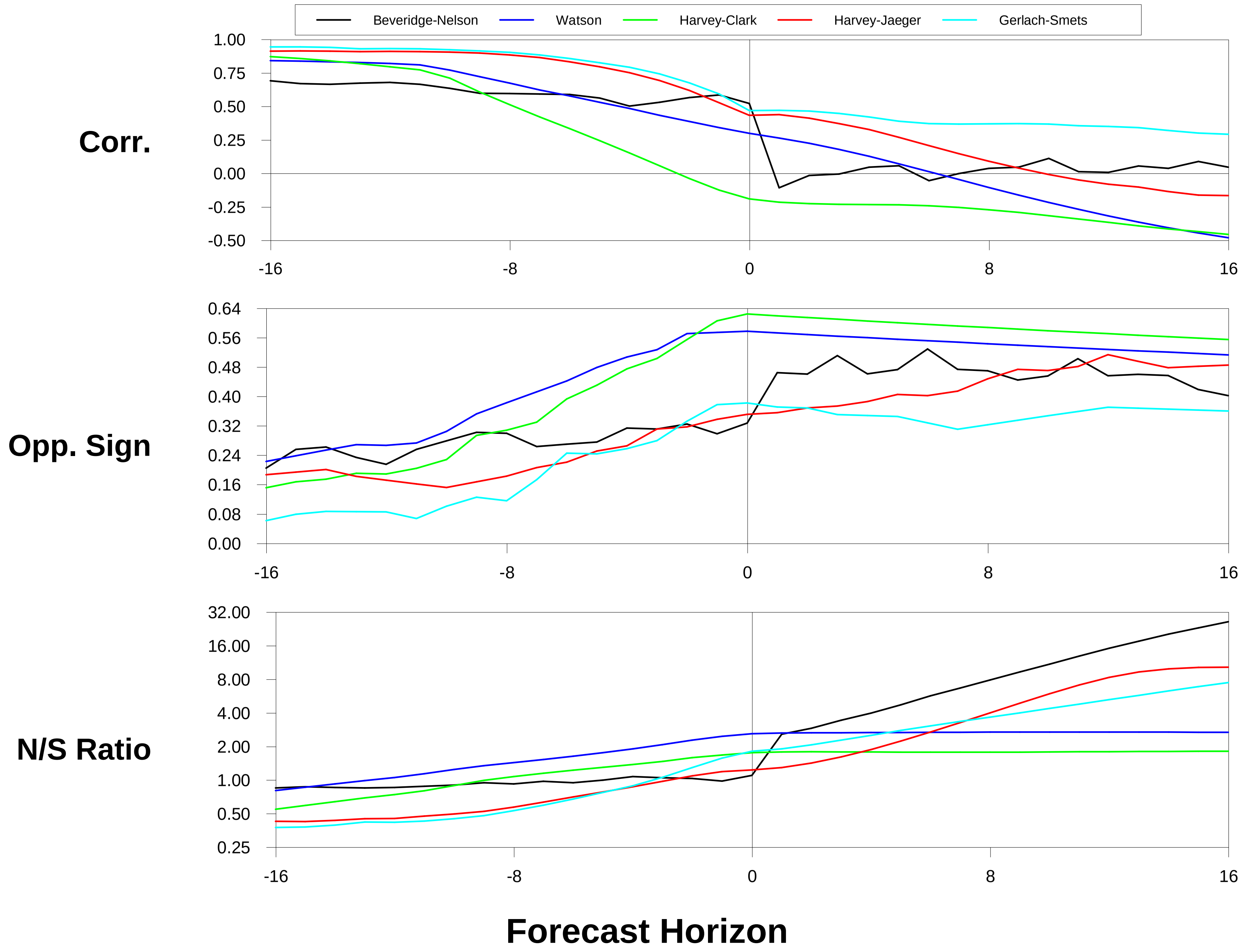


Figure 6: Revisions and Forecast Horizons



The following Discussion Papers have been published since 2003:

Series 1: Studies of the Economic Research Centre

January	2003	Testing mean-variance efficiency in CAPM with possibly non-gaussian errors: an exact simulation-based approach	Marie-Claude Beaul Jean-Marie Dufour Lynda Khalaf
January	2003	Finite-sample distributions of self-normalized sums	Jeong-Ryeol Kim
January	2003	The stock return-inflation puzzle and the asymmetric causality in stock returns, inflation and real activity	Jeong-Ryeol Kim
February	2003	Multiple equilibrium overnight rates in a dynamic interbank market game	Jens Tapking
February	2003	A comparison of dynamic panel data estimators: Monte Carlo evidence and an application to the investment function	Andreas Behr
March	2003	A Vectorautoregressive Investment Model (VIM) And Monetary Policy Transmission: Panel Evidence From German Firms	Joerg Breitung Robert S. Chirinko Ulf von Kalckreuth
March	2003	The international integration of money markets in the central and east European accession countries: deviations from covered interest parity, capital controls and inefficiencies in the financial sector	Sabine Herrmann Axel Jochem
March	2003	The international integration of foreign exchange markets in the central and east European accession countries: speculative efficiency, transaction costs and exchange rate premiums	Sabine Herrmann Axel Jochem

March	2003	Determinants of German FDI: New Evidence from Micro-Data	Claudia Buch Jörn Kleinert Farid Toubal
March	2003	On the Stability of Different Financial Systems	Falko Fecht
April	2003	Determinants of German Foreign Direct Investment in Latin American and Asian Emerging Markets in the 1990s	Torsten Wezel
June	2003	Active monetary policy, passive fiscal policy and the value of public debt: some further monetarist arithmetic	Leopold von Thadden
June	2003	Bidder Behavior in Repo Auctions without Minimum Bid Rate: Evidence from the Bundesbank	Tobias Linzert Dieter Nautz Jörg Breitung
June	2003	Did the Bundesbank React to Stock Price Movements?	Martin T. Bohl Pierre L. Siklos Thomas Werner
15	2003	Money in a New-Keynesian model estimated with German data	Jana Kremer Giovanni Lombardo Thomas Werner
16	2003	Exact tests and confidence sets for the tail coefficient of α-stable distributions	Jean-Marie Dufour Jeong-Ryeol Kurz-Kim
17	2003	The Forecasting Performance of German Stock Option Densities	B R Craig, E Glatzer, J Keller, M Scheicher
18	2003	How wacky is the DAX? The changing structure of German stock market volatility	Jelena Stapf Thomas Werner

1	2004	Foreign Bank Entry into Emerging Economies: An Empirical Assessment of the Determinants and Risks Predicated on German FDI Data	Torsten Wezel
2	2004	Does Co-Financing by Multilateral Development Banks Increase “Risky” Direct Investment in Emerging Markets? – Evidence for German Banking FDI	Torsten Wezel
3	2004	Policy Instrument Choice and Non-Coordinated Monetary Policy in Interdependent Economies	Giovanni Lombardo Alan Sutherland
4	2004	Inflation Targeting Rules and Welfare in an Asymmetric Currency Area	Giovanni Lombardo
5	2004	FDI versus cross-border financial services: The globalisation of German banks	Claudia M. Buch Alexander Lipponer
6	2004	Clustering or competition? The foreign investment behaviour of German banks	Claudia M. Buch Alexander Lipponer
7	2004	PPP: a Disaggregated View	Christoph Fischer
8	2004	A rental-equivalence index for owner-occupied housing in West Germany 1985 to 1998	Claudia Kurz Johannes Hoffmann
9	2004	The Inventory Cycle of the German Economy	Thomas A. Knetsch
10	2004	Evaluating the German Inventory Cycle Using Data from the Ifo Business Survey	Thomas A. Knetsch
11	2004	Real-time data and business cycle analysis in Germany	Jörg Döpke
12	2004	Business Cycle Transmission from the US to Germany – a Structural Factor Approach	Sandra Eickmeier

13	2004	Consumption Smoothing Across States and Time: International Insurance vs. Foreign Loans	George M. von Furstenberg
14	2004	Real-Time Estimation of the Output Gap in Japan and its Usefulness for Inflation Forecasting and Policymaking	Koichiro Kamada
15	2004	Welfare Implications of the Design of a Currency Union in Case of Member Countries of Different Sizes and Output Persistence	Rainer Frey
16	2004	On the decision to go public: Evidence from privately-held firms	Ekkehart Boehmer Alexander Ljungqvist
17	2004	Who do you trust while bubbles grow and blow? A comparative analysis of the explanatory power of accounting and patent information for the market values of German firms	Fred Ramb Markus Reitzig
18	2004	The Economic Impact of Venture Capital	Astrid Romain, Bruno van Pottelsberghe
19	2004	The Determinants of Venture Capital: Additional Evidence	Astrid Romain, Bruno van Pottelsberghe
20	2004	Financial constraints for investors and the speed of adaption: Are innovators special?	Ulf von Kalckreuth
21	2004	How effective are automatic stabilisers? Theory and results for Germany and other OECD countries	Michael Scharnagl Karl-Heinz Tödter
22	2004	Asset Prices in Taylor Rules: Specification, Estimation, and Policy Implications for the ECB	Pierre L. Siklos Thomas Werner Martin T. Bohl

23	2004	Financial Liberalization and Business Cycles: The Experience of Countries in the Baltics and Central Eastern Europe	Lúcio Vinhas de Souza
24	2004	Towards a Joint Characterization of Monetary Policy and the Dynamics of the Term Structure of Interest Rates	Ralf Fendel
25	2004	How the Bundesbank really conducted monetary policy: An analysis based on real-time data	Christina Gerberding Andreas Worms Franz Seitz
26	2004	Real-time Data for Norway: Challenges for Monetary Policy	T. Bernhardsen, Ø. Eitrheim, A.S. Jore, Ø. Røisland
27	2004	Do Consumer Confidence Indexes Help Forecast Consumer Spending in Real Time?	Dean Croushore
28	2004	The use of real time information in Phillips curve relationships for the euro area	Maritta Paloviita David Mayes
29	2004	The reliability of Canadian output gap estimates	Jean-Philippe Cayen Simon van Norden

Series 2: Banking and Financial Supervision

1	2003	Measuring the Discriminative Power of Rating Systems	B. Engelmann, E. Hayden, D. Tasche
2	2003	Credit Risk Factor Modeling and the Basel II IRB Approach	A. Hamerle, T. Liebig, D. Rösch
1	2004	Forecasting Credit Portfolio Risk	A. Hamerle, T. Liebig, H. Scheule
2	2004	Systematic Risk in Recovery Rates – An Empirical Analysis of US Corporate Credit Exposures	Klaus Düllmann Monika Trapp
3	2004	Does capital regulation matter for bank behaviour? Evidence for German savings banks	Frank Heid Daniel Porath Stéphanie Stolz
4	2004	German bank lending during emerging market crises: A bank level analysis	F. Heid, T. Nestmann, B. Weder di Mauro, N. von Westernhagen

Visiting researcher at the Deutsche Bundesbank

The Deutsche Bundesbank in Frankfurt is looking for a visiting researcher. Visitors should prepare a research project during their stay at the Bundesbank. Candidates must hold a Ph D and be engaged in the field of either macroeconomics and monetary economics, financial markets or international economics. Proposed research projects should be from these fields. The visiting term will be from 3 to 6 months. Salary is commensurate with experience.

Applicants are requested to send a CV, copies of recent papers, letters of reference and a proposal for a research project to:

Deutsche Bundesbank
Personalabteilung
Wilhelm-Epstein-Str. 14

D - 60431 Frankfurt
GERMANY