THE REAL-TIME FORECASTING PERFORMANCE OF PHILLIPS CURVES

Tim Robinson, Andrew Stone and Marileze van Zyl

Research Discussion Paper
2003-12

December 2003

Economic Research Department
Reserve Bank of Australia

The authors are grateful to David Gruen, Simon van Norden, Athanasios Orphanides and colleagues at the Reserve Bank, especially Alex Heath and Anthony Richards, for helpful comments. Responsibility for any errors rests with the authors. The views expressed in this paper are those of the authors and should not be attributed to the Reserve Bank.

Abstract

Analysts typically use a variety of techniques to forecast inflation. These include both ‘bottom-up’ approaches, for near-term forecasting, as well as econometric methods (such as mark-up models of inflation, which have been found to perform quite well for Australia – see de Brouwer and Ericsson (1998)).

One of the econometric approaches to inflation forecasting which is sometimes considered is the use of Phillips curves based on estimates of the output gap. This paper suggests, however, that the real-time capacity of such Phillips curves to forecast inflation is limited, relative even to such simple benchmark forecasting approaches as an autoregressive (AR) model of inflation or a random walk assumption. It appears that the lack of precision with which output-gap-based Phillips curves can be estimated in real time limits their usefulness as a means of forecasting inflation in isolation.

Phillips curve-based forecasts may, however, perform a little better than AR model-based ones in at least predicting whether inflation will increase or decrease from its current level. Moreover, combining Phillips curve-based forecasts with those from simple, alternative approaches does seem to offer some scope for improving the real-time forecast accuracy of the latter. These observations suggest that, in spite of their generally disappointing performance as a means of forecasting inflation in isolation, output-gap-based Phillips curves may continue to be useful in real time – as a tool for conditioning gap estimates within a multivariate filtering framework, and as a possible complement to other, alternative inflation forecasting approaches.

JEL Classification Numbers: E32, E37, E52, E60

Keywords: monetary policy, forecasting inflation, output gaps, real-time data

Table of Contents

THE REAL-TIME FORECASTING PERFORMANCE OF PHILLIPS CURVES

Tim Robinson, Andrew Stone and Marileze van Zyl

1. Introduction

Recent years have seen growing interest in the implications of real-time data issues for the use of output-gap estimates, and of Phillips curves based on these estimates, in policy analysis and macroeconomic modelling.[1] This paper aims to contribute to that research, focusing on the real-time inflation forecasting performance of output-gap-based Phillips curves.

Analysts typically use a variety of techniques to forecast inflation, rather than rely upon a single method. One approach frequently used, especially over short horizons, is to adopt a 'bottom-up' perspective, using a combination of partial indicators, liaison with businesses and government, and judgment. Additionally, analysts may use a variety of econometric models of inflation, reflecting alternative methodologies or views about the structure of the economy.[2,3]

One of these econometric approaches to inflation forecasting sometimes considered, particularly by academics and macroeconomic modellers, is the use of Phillips curves based on estimates of the output gap. These posit a relationship between the inflation rate and the level of capacity utilisation in the economy, as measured by the deviation of actual output (GDP) from the economy's potential

---

1 For example, a number of overseas central banks have been conducting research into these issues, including the US Federal Reserve (Orphanides *et al* 1999) and the Bank of England (Nelson and Nikolov 2001), amongst others.

2 For example, mark-up models of inflation have been found to perform quite well for Australia – see de Brouwer and Ericsson (1998).

3 Stevens (2001) provides a description of the policy formulation process employed by the Reserve Bank of Australia (RBA), and the roles of 'bottom-up' analysis, econometric models and judgment in this process. For a corresponding description for the US, see Reifschneider, Stockton and Wilcox (1997).

level of output.[4] However, a problem with this approach is that it requires estimation of an unobservable quantity, the level of potential output, and this must be done in *real time*, rather than with the benefit of hindsight.

There are several different aspects to the problem of reliably estimating the output gap in real time, and these are discussed in detail in Gruen, Robinson and Stone (2002). One is simply that GDP is revised over time as more information becomes available to the statistician.[5] Recent work, however, has suggested that data revisions are not the primary source of problems in the real-time estimation of the gap (Orphanides and van Norden 2002). Rather, for Australia, the results of Gruen *et al* (2002) suggest that revisions to the output gap instead arise primarily from end-point problems associated with econometric methods used to estimate potential output. In this regard, Gruen *et al* find that the use of a Phillips curve relationship, to condition such estimates on observed inflation, can help to mitigate these end-point problems – allowing them to obtain real-time output-gap estimates which bear a fair resemblance to those obtained with the benefit of hindsight.

While the reliability of output-gap estimates is of interest to analysts in its own right, our focus in this paper is, however, on the accuracy of inflation forecasts from Phillips curves and output gaps estimated in real time. The accuracy of these forecasts provides an alternative metric by which the usefulness of output gaps and Phillips curves for policy analysis and forecasting can be assessed.

From this perspective, it is not enough to have a method for estimating potential output under which the *ex post* relationship between the output gap and inflation is strong. Rather, it must be possible to obtain *real-time* output-gap estimates and Phillips curves which provide a reliable guide to future inflation. Potential sources of error in such forecasts include the difficulties of determining both the correct specification of the Phillips curve in real time, and the correct coefficient estimates in this specification. They also include the direct impact of real-time errors which inevitably arise in estimating the output gap – notwithstanding the

---

4 Phillips curves based on other measures of aggregate activity, such as an unemployment gap, are also commonly used to model inflation. However, we do not consider these further here.

5 Analysis for Australia by Stone and Wardrop (2002) suggests that, while the typical size of revisions to quarterly GDP growth has decreased over recent decades, sizeable changes do still occur, sometimes well after the initial data release for the period in question.

strong correlation which Gruen *et al* (2002) were able to obtain between their final and real-time gap estimates.

In this companion piece to Gruen *et al* (2002), we use the more than 120 vintages of Phillips curves and output gaps from that paper to investigate whether these obstacles are sufficient to undermine the usefulness of Phillips curves for forecasting inflation. We find that they result in inflation forecasts which, while unbiased, are excessively volatile. This makes their performance disappointing, relative even to simple alternatives such as univariate time-series models of inflation, or a 'no change' (random walk) assumption. Our results also suggest that it is the imprecision with which Phillips curve relationships can be estimated, in real time, which primarily undermines their usefulness for forecasting inflation.

In spite of their disappointing performance in isolation, however, it appears that Phillips curve-based forecasts do still contain some useful information about inflation, distinct from that available from other simple forecasting approaches. This appears to be especially so at the one-year-ahead forecasting horizon. Moreover, this added information can be used to improve the inflation predictions from these other forecasting approaches in real time – although the extent of the resulting improvement is only modest.

2. Recent Literature

The ability of Phillips curves to forecast inflation has been the focus of much recent research. In this section, we briefly review the relevant literature.

Stock and Watson (1999) consider year-ahead forecasts for US inflation obtained from Phillips curves based on an unemployment gap – the difference between the unemployment rate and the estimated Non-Accelerating Inflation Rate of Unemployment (NAIRU). They then assess the accuracy of these, relative to predictions from both univariate autoregressive models and from a 'no change' assumption.[6] They interpret their results as being generally positive regarding the

[6] This comparison is made on a real-time basis: for each period, Stock and Watson estimate their various models based on data up until that period, then calculate forecasts for the year ahead, and then repeat this process for the next period. Since the US unemployment rate only gets revised though seasonal re-analysis, the effect of which is small, the real-time data for this series in each period is essentially equivalent to the final data up until that period.

ability of such Phillips curves to forecast inflation. However, the outperformance by their Phillips curve-based forecasts is sometimes only marginal, with the mean squared errors (MSEs) of these forecasts little different from those of their alternative forecasting methods, for certain periods or measures of inflation.

There are two characteristics of the Phillips curve specifications estimated by Stock and Watson (1999) which mean that their results may not be directly applicable to Australia. First, they assume that the NAIRU is constant through time, over their sample period from 1959 to 1997. While this may be an appropriate assumption for the US, the work of Gruen, Pagan and Thompson (1999) suggests that, for Australia, there are likely to have been sizeable movements in the NAIRU over recent decades.[7] Secondly, they omit supply-side variables, such as import and oil prices, from their Phillips curve specifications. Such variables were found by Gruen *et al* (2002) to be important for modelling inflation in Australia – a result which is unsurprising given that Australia is a small, open economy.

In contrast to Stock and Watson's broadly positive endorsement, several recent papers have questioned the usefulness of Phillips curves for forecasting inflation. For example, Atkeson and Ohanian (2001) estimate alternative versions of the Phillips curves in Stock and Watson (1999) and find their performance to be no better than a naive alternative.[8] However, Sims (2002) argues that Atkeson and

---

[7] Gruen *et al* (2002) find that it is important that the possibility of corresponding large shifts in the growth rate of potential output over time be allowed for, in any output-gap-based Phillips curve specifications for Australia.

[8] Atkeson and Ohanian's Phillips curves are more restricted than those of Stock and Watson, in the sense that they allow only one lag at a time of the monthly change in inflation, and of the level of the unemployment gap, in their specifications (although they do examine the effect of varying the lengths of these lags, up to one year, on the forecasts). They also do not allow a contemporaneous unemployment gap in their Phillips curve specifications, which the results in Gruen *et al* (2002) suggest would be an undesirable restriction for Australia.

Ohanian's results are sensitive to the sample they use (1984:Q1 to 1999:Q3), so that the poor Phillips curve performance they report may not be instructive.[9]

The paper with the most direct relevance to our own study is the recent work by Orphanides and van Norden (2003), who examine inflation forecasts derived from output-gap-based Phillips curves. They consider the wide array of different output-gap series estimated, using real-time US GDP data and both univariate and multivariate techniques, in Orphanides and van Norden (2002). To forecast inflation, they place each of these alternative output-gap series – a dozen in all – in a separate 'forecasting relationship', namely a simple Phillips curve with a constant, lags of inflation and lags of the output gap. In a procedure similar to Stock and Watson (1999), they re-estimate each forecasting relationship (including choice of lag lengths) for each forecast period, using only data available up to that date. They then compare the forecasts generated by these relationships with those from an alternative, autoregressive model of inflation.

Orphanides and van Norden (2003) find that, while several of their Phillips curves yield more accurate forecasts over their full sample, the improvement is typically not statistically significant, and is frequently reversed when the analysis is restricted to one or other half of their evaluation period. Moreover, this finding of little or no consistent improvement in forecast accuracy is found to be robust to a wide array of variations to their general framework for assessing the forecast performance of their Phillips curves and alternative output-gap estimates.[10] In all, they conclude that while 'a historical Phillips curve is suggested by the data, and *ex post* estimates of the output gap are useful for understanding historical movements in inflation ... our simulated real-time forecasting experiment suggests, instead, that [their] predictive ability is mostly illusory' (p 24).

One possible explanation for these disappointing results is that Orphanides and van Norden's use of such a simple Phillips curve framework, with no role for

---

9 When this sample is extended back only slightly, to 1979, so as to incorporate a period when US inflation was more volatile, Sims finds that Phillips curve-based forecasts do outperform the naive model over this longer sample.

10 Orphanides and van Norden's long evaluation period for forecasts is 1969:Q1 to 1998:Q4. Splitting this sample into two, their latter sub-sample is 1984:Q1 to 1998:Q4, comparable to that used in Atkeson and Ohanian (2001).

supply side influences, may be limiting the performance of their Phillips curve-based inflation forecasts.[11] Moreover, for their multivariate filter-based output-gap estimates, their use of separate 'forecasting relationship' Phillips curves to generate their inflation forecasts is somewhat counterintuitive, given that the derivation of these gap estimates used *different* Phillips curves, embedded in the multivariate filtering.[12]

Orphanides and van Norden's approach, however, simply reflects their somewhat different overall goal from our own. They use their framework to assess the value, for Phillips curve-based inflation forecasting, of a wide range of alternative output-gap estimates. Their use of a common, simple framework to do this allows them to readily test the sensitivity of their results to various plausible variations to this framework, such as to the method for selecting lag lengths in their forecasting Phillips curves in real time.

By contrast, our aim is to focus on only one suite of Phillips curves and associated output-gap data vintages – those assessed by Gruen *et al* (2002) to be the best performed from a range of alternatives considered there. We then assess the real-time forecasting performance of these Phillips curves when coupled with corresponding optimally-specified equations for forecasting the output gap. We thus aim to assess the real-time inflation forecasting performance of Phillips curves for Australia, when care has been taken to try to make these Phillips curves and output-gap equations as well and richly specified as possible.

---

11 Orphanides and van Norden note, in this regard, that 'the forecasting problem faced by policy-makers in practice is more complex than the one we consider. One obvious and important difference is that the information set available to policy-makers is much richer'. They argue, however, that 'it is therefore possible that output gaps might improve on simple univariate forecasts of inflation but not on forecasts using a broader range of inputs. For this reason, we feel that the experiment we perform may actually overstate the utility of empirical output gap models'. Certainly, the comparison exercise they conduct prevents *both* their Phillips curves and competing univariate models from using additional information which might naturally be incorporated into inflation forecasts – such as data on oil and import prices. In this sense, their comparison is designed to isolate what, if any, real-time benefit is gained solely through the addition of estimates of the output gap to a univariate inflation-forecasting framework.

12 The Phillips curves incorporated in these multivariate filters were, however, relatively simple – for example, they also did not include any supply-side variables. Orphanides and van Norden's 'forecasting relationship' Phillips curves may therefore not be a bad approximation to them.

3. Methodology

The starting point for our analysis is the paper by Gruen *et al* (2002), which provides a suite of 125 real-time Phillips curves and output-gap estimates, covering the data vintages from 1971:Q4 to 2002:Q4.[13] We wish to study the forecasting performance of these Phillips curves – focusing on the one- and four-quarter-ahead forecast horizons. However, we encounter an immediate methodological obstacle, namely: how should we handle the various explanators in our Phillips curves over the forecast period?

An important finding in Gruen *et al* (2002) was that, to generate reasonably reliable real-time output gaps using the filtering procedure adopted there, it is 'important that the Phillips curve[s] be information-rich and as well specified as possible' (p 32). Reflecting this, their preferred Phillips curves employ not just the output gap, but also a variety of other possible explanators of consumer price inflation. Often, these explanators enter their Phillips curves with a lag of less than four quarters, so that some assumption needs to be made about their future behaviour to enable out-of-sample inflation forecasts for the coming year to be generated in each period.[14] We now describe how we overcome this difficulty (by expanding our set of Phillips curves into a more general suite of *two-equation forecasting models* – consisting of a Phillips curve and an output-gap equation – one model for each data vintage). In Sections 3.2 and 3.3 we then describe a number of simple, alternative benchmark models for forecasting inflation.

---

13 The way in which these Phillips curves were selected, and corresponding vintages of output-gap data generated, was described in detail in Gruen *et al* (2002). It is summarised briefly in Appendix A, which also contains a complete listing of these specifications. Note that Gruen *et al* actually provides only 121 Phillips curves. We have increased this number to 125 by also applying the final Phillips curve from Gruen *et al* to the four data vintages from 2002. In arriving at this specification for these additional vintages we applied the same specification search procedure as that used to obtain the original 121 Phillips curves in Gruen *et al*.

14 This difficulty does not arise in Orphanides and van Norden (2003) because their forecasting Phillips curves do not include explanators other than the output gap and lags of inflation itself. Moreover, the output gap is not allowed to enter their Phillips curves with a lag less than the length of their inflation forecast horizon, so that they even avoid the need to extrapolate the output gap into the forecast period. While this simplifies the construction of their real-time inflation forecasts, it raises the question: do these constraints on the specifications of their Phillips curves bias their results against finding a role for these curves in forecasting inflation?

3.1 A Suite of Phillips Curve-based Models for Forecasting Inflation

The Phillips curves from Gruen *et al* (2002) typically contain a contemporaneous output-gap term, as well as terms involving import price inflation, oil price inflation and bond market inflation expectations. We now specify how each of these explanators is to be extrapolated over the forecast period.

Forecasting the output gap

The explanator in our Phillips curves which is most likely to be the source of any real-time problems with forecasting inflation is the output gap. For this variable, our approach is therefore to assemble separate, estimated gap equations for each data vintage, using an approach similar to that used in Gruen *et al* (2002) to select our real-time Phillips curves. We now describe this approach in greater detail.

To obtain simple, estimated gap equations, we start with a general specification which includes, as explanatory variables: the real cash rate (rcash); changes in the terms of trade (Δtot); changes in the real trade-weighted exchange rate (Δrtwi); and a measure of the disequilibrium between the levels of the real exchange rate and the terms of trade (all with lags of up to 8 quarters).[15] Following Beechey *et al* (2000), we also include a de-trended real share accumulation index (rshare), to capture the share market's influence on real activity (for example, through wealth and confidence effects). The Southern Oscillation Index (soi) is also included as a possible explanator, as a proxy for the effects of the weather on the agricultural sector.

Prior to the mid 1980s, the Australian financial system was heavily regulated. As a result, the overnight cash rate is a less reliable indicator of the overall tightness

---

[15] Note that, for each period, to generate our short-term real interest rate data we use the inflation expectations measure implicit in the Phillips curve specification for that vintage. This is done to ensure consistency in the measure of inflation expectations used within our model. This measure computes inflation expectations as a linear combination of lagged inflation and bond market inflation expectations, with the weights for each vintage being the coefficients estimated in the Phillips curve – see Gruen *et al* (2002) for further details. This adds a real-time aspect to our measure of inflation expectations, since these weights evolve slowly over time.

or looseness of credit conditions for this period than for the post-deregulation era (Bullock, Morris and Stevens 1989). We therefore also experimented with adding to the general specification various alternative measures of the state of credit conditions. However, none of these variables were found to be consistently significant in any of our regressions.[16]

To arrive at our final specifications for each data vintage, both recursive regressions (in which the start date was held fixed at 1961:Q2) and 14-year rolling regressions were estimated. These regressions were used to sequentially eliminate variables on the basis of both *t*- and *F*-tests. Also, rather than conduct a new specification search for each new data vintage, we adopted the approach of revisiting our output-gap equation specifications: whenever significant deterioration was observed in the performance either of the overall equation or of its components; or, in any event, roughly every 10–12 years. Overall, this procedure was very similar to that used in Gruen *et al* (2002) to select our real-time Phillips curve specifications for each data vintage.

Ultimately, the specifications of our output-gap equations change only slowly over time. Table 1 shows the estimation results for our final vintage output-gap equation, estimated on 2002:Q4 data. All terms appear with the expected signs, and the regression standard error suggests that the equation does a fair job of explaining movements in the output gap over the past four decades.

The real-time output-gap equation specifications for our complete set of 125 data vintages are summarised in Table 2, and are set out in full in Appendix B. Several characteristics of the specifications in Table 2 stand out. First of all, the first lag of the gap itself appears in every equation vintage (with the fifth lag also appearing in specifications for the late 1970s and most of the 1980s). Fairly consistent roles are

---

[16] The variables we experimented with included: the quarterly change in the Statutory Reserve Deposit (SRD) ratio, the proportion of their deposits which trading banks were, until August 1988, required to hold at the Reserve Bank as cash reserves; the 90-day bank bill rate; and a 'weighted interest rate' (constructed as an average of the prevailing interest rates for different debt maturities, weighted by the proportion of total credit outstanding at these maturities). Real-time data issues ruled out consideration of a number of other, otherwise plausible candidates, such as various monetary aggregates.

also found for: the de-trended real share accumulation index lagged one period; one or more lags of the quarterly change in the real trade-weighted exchange rate (except in early and late equation vintages); and the first lag of the Southern Oscillation Index (except in early vintages). Finally, an explicit role for lags of the real cash rate is only identified for equation vintages in the post-financial deregulation era, specifically from December quarter 1985 onwards.[17]

Table 1: Estimation Results for the Final-vintage Output-gap Equation

$$gap_t = \alpha + \beta_1 gap_{t-1} + \gamma_1 rshare_{t-1} + \phi_1 soi_{t-1} + \psi_2 (rcash_{t-2} + rcash_{t-3} + \ldots + rcash_{t-7}) + \varepsilon_t$$

Coefficient	Value	t-statistic
α	0.0016	1.1514
β_1	0.8705	27.1327
γ_1	0.0255	3.5064
ϕ_1	0.0002	2.1276
ψ_2	−0.0138	−2.1278
Summary statistics	**Value**	
R^2	0.8947	
Adjusted R^2	0.8921	
Standard error of the regression	0.0117	
Breusch-Godfrey LM test for autocorrelation (p-value):		
First order	0.2590	
First to fourth order	0.1845	

Note: The sample is 1961:Q2 to 2002:Q4 ($n = 167$).

17 Our failure to find an explicit role for some measure of conditions in credit markets, in the pre-deregulation era, may reflect our inability, discussed earlier, to identify a good, single summary indicator for this variable for this period.

Table 2: Summary of Real-time Output-gap Equation Specifications

Date of vintage	Broad equation specification
1971:Q4 to 1978:Q1	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \gamma_1 \text{rshare}_{t-1} + \varepsilon_t$
1978:Q2 to 1985:Q3	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \beta_5 \text{gap}_{t-5} + \gamma_1 \text{rshare}_{t-1} + \delta_3 \Delta\text{tot}_{t-3} + \lambda_2 \Delta\text{rtwi}_{t-2} + \lambda_5 \Delta\text{rtwi}_{t-5} + \phi_1 \text{soi}_{t-1} + \varepsilon_t$
1985:Q4 to 1989:Q1	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \beta_5 \text{gap}_{t-5} + \gamma_1 \text{rshare}_{t-1} + \delta_3 \Delta\text{tot}_{t-3} + \lambda_1 \Delta\text{rtwi}_{t-1} + \phi_1 \text{soi}_{t-1} + \psi_2 (\text{rcash}_{t-2} + \text{rcash}_{t-3}) + \varepsilon_t$
1989:Q2 to 1995:Q2	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \gamma_1 \text{rshare}_{t-1} + \lambda_1 \Delta\text{rtwi}_{t-1} + \phi_1 \text{soi}_{t-1} + \psi_2 (\text{rcash}_{t-2} + \text{rcash}_{t-3}) + \varepsilon_t$
1995:Q3 to 2002:Q4	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \gamma_1 \text{rshare}_{t-1} + \phi_1 \text{soi}_{t-1} + \psi_2 (\text{rcash}_{t-2} + \text{rcash}_{t-3} + \cdots + \text{rcash}_{t-7}) + \varepsilon_t$

Note: Start of sample for all regressions is 1961:Q2.

Extrapolating explanators other than the output gap

For all variables other than the output gap, our approach is simply to assume perfect foresight on the part of our forecasters, in each period. In other words, we assume that in each period our forecasters, although uncertain about the future trajectories of consumer price inflation and the output gap, know precisely how such variables as oil and import price inflation, bond market inflation expectations, and the Southern Oscillation Index, will develop over the coming year.

The obvious alternative to this approach would have been to construct real-time models in each quarter for the evolution of all these variables (as well as any other explanators which, in turn, arose in the equations for these variables). However, with 125 vintages of data, such an approach was beyond the scope of this project. Moreover, none of these variables is subject to serious real-time measurement problems, the implications of which are the focus of this paper.

Nevertheless, our assumption of perfect foresight, in relation to these variables, should be borne in mind in the discussion of our results in Section 4. It certainly imputes to our Phillips curve-based forecasters, in each period, unrealistic prescience about these inputs to their forecasts – and so presumably biases our results somewhat in favour of our Phillips curve-based forecasting approach.

Summary

In summary, we now have a small, estimated suite of models, one for each data vintage from 1971:Q4 to 2002:Q4, consisting of: a Phillips curve from Gruen *et al* (2002); an output-gap equation; and the assumption of perfect foresight on the part of forecasters with regard to oil and import price inflation, bond market inflation expectations, and the various explanators arising in that vintage's gap equation (such as the nominal cash rate, the Southern Oscillation Index, and so forth). For each period we use this model to generate real-time inflation forecasts at one- and four-quarter-ahead horizons. This constitutes our Phillips curve-based approach to forecasting Australian consumer price inflation.

3.2 Alternative Benchmark Models for Forecasting Inflation

To judge the real-time performance of our Phillips curve-based approach to forecasting inflation, we need some benchmarks against which we can compare its performance. We consider two alternative, univariate benchmarks: a 'no-change' assumption for inflation; and an autoregressive (AR) model of inflation.

The former treats inflation as a random walk.[18] The latter models inflation in terms of the first two lags of itself (together with a constant), with the equation re-estimated for each data vintage. In selecting this general specification we allowed, in principle, for longer lags of inflation. However, we found that the autocorrelation and partial autocorrelation functions for inflation regularly suggested focusing on the first three lags, and that the performance of an AR(2)

[18] Strictly speaking, we actually use two distinct random walk models, one for quarterly and one for year-ended inflation. In other words, in each period we assume: that quarterly inflation in the coming quarter will be the same as in the current quarter; and that year-ended inflation over the coming year will be the same as over the year to the current quarter.

model was typically closely comparable to that of an AR(3). Hence we preferred the slightly more parsimonious AR(2) specification.[19]

3.3 Two Further Benchmarks for Comparison

As a further reference point for assessing the performance of our suite of Phillips curve-based models, we also consider the forecasting performance of our final vintage (2002:Q4) Phillips curve, estimated on final data. By this we mean the results which would have been achieved by a forecaster who, in each quarter of the evaluation period, happened to know in advance: the final, 2002:Q4 vintage Phillips curve specification, including its 'true' coefficient values (as estimated on final data); the 'true' (final vintage) profile of the output gap up to that period; and, finally, the correct future paths of oil prices, import prices, bond market inflation expectations, and the output gap, over the year ahead. In Sections 4 and 5, we refer to these simply as the 'final vintage Phillips curve' forecasting results. Likewise, we also consider the corresponding forecasting performance of our 'final vintage (2002:Q4) AR model'.

4. Results

Table 3 summarises the performance of our three main alternative approaches to forecasting inflation in real time (Phillips curve-based; AR model-based; and random walk), as well as the performance of our final vintage Phillips curve and final vintage AR model. The sample used for the forecast comparison is 1976:Q1 to 2002:Q4. Here and henceforth, dates for forecasts refer to the period *for which* the forecast was being made, rather than *in which* the forecast was being made.

Note that, since we have data vintages (with corresponding real-time Phillips curves and output-gap equations) from 1971:Q4 onwards, we could in principle compare our various forecasting approaches over the longer sample 1972:Q4 to 2002:Q4 (or even 1972:Q1 to 2002:Q4, for one-quarter-ahead inflation forecasts).

---

[19] We also considered using a more general ARMA model, to allow for a possible indeterministic (moving average) component which might be useful in better capturing the persistence in inflation. However, while an ARMA(1,1) model also generally works well, it does not outperform our chosen AR(2) specification, which we therefore preferred as marginally simpler and more intuitive.

Table 3: **Root Mean Squared Error and Bias of Alternative Inflation Forecasting Methods**

Method	Final vintage PC	Final vintage AR	Real-time PC	Real-time AR	Random walk
Forecasts for quarterly inflation (one quarter ahead)					
RMSE	0.31	0.32	0.44	0.33	0.38
Bias	0.01	0.03	–0.01	0.05	0.03
Forecasts for year-ended inflation (one year ahead)					
RMSE	0.87	1.16	1.55	1.23	1.57
Bias	0.08	0.16	0.03	0.33*	0.46*

Notes: Results are calculated over the forecast sample period 1976:Q1 to 2002:Q4. 'RMSE' is the root mean squared error between actual inflation and the series of real-time inflation forecasts generated by the method shown, in percentage points; 'Bias' is the average percentage point error in the forecast series over the evaluation period. * indicates that the bias is significant at the 5 per cent level.

One reason we do not do this is that the early to mid 1970s were a period of atypically extreme inflation volatility in Australia, partly driven by the first OPEC oil price shock in March quarter 1974. Comparisons of forecast performance which include this period thus tend to be unduly dominated by it.

A second reason is that inflation outcomes for this period were strongly influenced by decisions of Australia's then heavily centralised wage-setting system.[20] Since none of our forecasting methods are directly able to take account of these wage developments, which Gruen *et al* (2002) characterise as being 'at least to some extent ... unrelated to the state of the economy and the size of the output gap' (p 16), we focus our forecast evaluation on the period from 1976:Q1 onwards.

---

[20] During this period Australia had an official body with legislated powers, the *Conciliation and Arbitration Commission*, which set wages for much of the workforce. In early 1973, following a change of government, the Commission awarded a 17.5 per cent increase in minimum wages, at a time when consumer price inflation, although rising, was running at an annual rate of less than 6 per cent. Further large award/minimum wage rises were mandated in May and December 1974. While wages in Australia in subsequent years were also affected by frequent Commission decisions – see Appendix A of Gruen *et al* (1999) – increases in average weekly earnings over these years did not again approach the extremes which occurred during 1973 and 1974.

4.1 Final Vintage Forecasting Performance

The first column of Table 3 shows that generally good inflation forecasting performance would have been achieved, over the past 27 years, by a forecaster with access, in each period, to the final vintage Phillips curve specification and historical output-gap profile, together with perfect foresight as to the future paths of the output gap, oil prices, import prices and bond market inflation expectations.[21] The root mean squared error (RMSE) achieved by such a forecaster, over the sample 1976:Q1 to 2002:Q4, would have been 0.31 percentage points for quarterly inflation one quarter ahead, and 0.87 percentage points for year-ended inflation over the coming year. This compares favourably with both our real-time AR model-based approach to forecasting inflation (RMSEs of 0.33 and 1.23 percentage points), and our random walk model (RMSEs of 0.38 and 1.57 percentage points). It also compares favourably with the performance which would have been achieved, at both the one- and four-quarter-ahead horizons, by our final vintage AR model (RMSEs of 0.32 and 1.16 percentage points).

These results for the predictive performance of our final vintage Phillips curve seem promising. The key question, however, is: does this good Phillips curve-based performance carry over from the unrealistic setting of forecasting based on our final vintage Phillips curve, to the situation we are actually interested in, out-of-sample forecasting in real time?

4.2 Real-time Forecasting Performance

Unfortunately, here the results for our Phillips curves are much less impressive. From the third column of Table 3 we see that, over the sample 1976:Q1 to 2002:Q4, our real-time Phillips curve-based inflation forecasts are unbiased, at

[21] There is, of course, some circularity in this finding since, as described in Gruen *et al* (2002), the final vintage output-gap profile is conditioned on information about inflation over the full sample from 1961:Q2 to 2002:Q4. Hence, 'forecasts' generated using this gap have some information about future inflation already built into them. This circularity, however, does not apply to our real-time Phillips curve-based forecasts, since the gap estimates used in generating these forecasts in each period are conditioned only on historical inflation data, not on information about future inflation.

both the one- and four-quarter-ahead horizons.[22] This accords with the findings of Orphanides and van Norden (2003) for the US, and contrasts with the bias apparent in the corresponding real-time AR model-based and random walk forecasts for inflation at the one-year-ahead horizon. Despite this bias, however, comparison of columns 3 to 5 of Table 3 shows that, in RMSE terms, both these simple, benchmark inflation forecasts perform essentially as well as, or better than, our Phillips curve-based forecasts in real time – despite the perfect foresight built into the latter in relation to variables such as oil and import prices. This is true both for one-quarter-ahead inflation forecasts, and for forecasts for inflation over the year ahead.

This disappointing forecasting performance in real time by our Phillips curve-based models is illustrated by Figures 1 and 2, which compare the real-time forecasts for year-ended inflation from our Phillips curve-based approach with those from our alternative benchmarks (and with actual inflation), over the sample 1976:Q1 to 2002:Q4.

These figures confirm that the large RMSEs for our Phillips curve-based inflation forecasts are mainly attributable to excessive volatility in these forecasts, rather than to forecast bias – although there are periods where our real-time Phillips curve-based forecasts for inflation persistently over- or under-shoot actual year-ended inflation, such as in the early 1990s. Interestingly, during this period of disinflation in Australia, while both of our real-time benchmark models repeatedly over-estimate inflation over the year ahead (as one would expect), our real-time Phillips curve-based forecasts are instead consistently too low for several years, starting in 1991. In part this reflects that the early 1990s is a period for which our generally reliable real-time output-gap estimates perform particularly poorly, underestimating the 'true' gap (as estimated on final vintage data) by around 3 percentage points for several years – see Figure 4 in Gruen *et al* (2002).

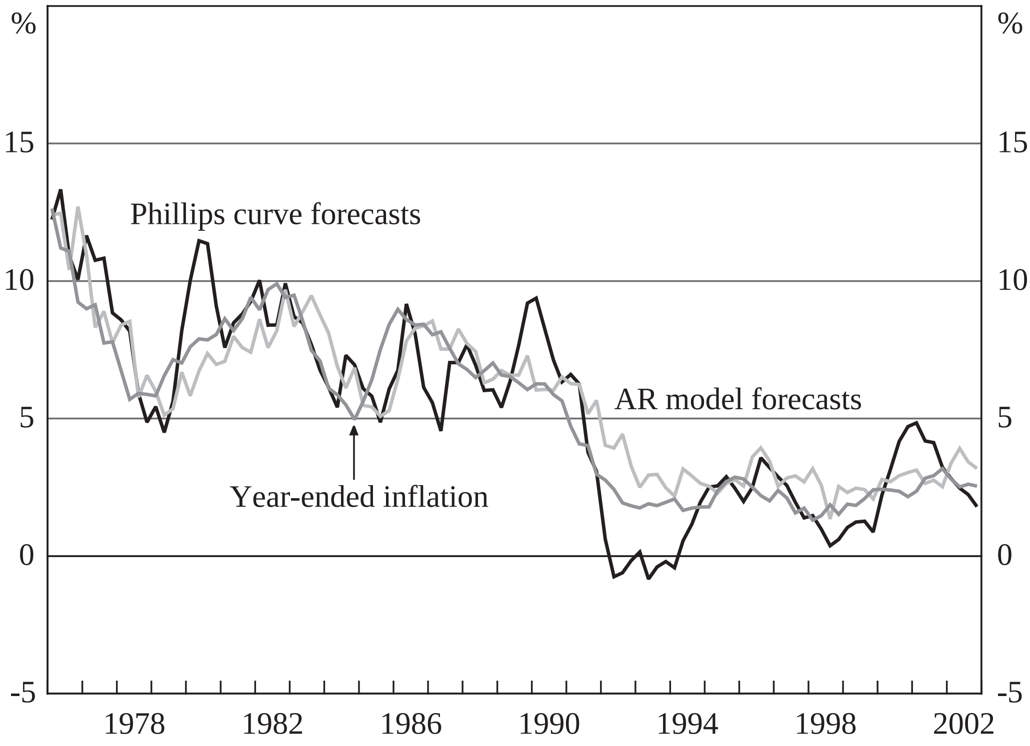
Figures 1 and 2 also highlight that the *relative* performance of our alternative, real-time approaches to forecasting inflation varies over the evaluation period. This

[22] To formally test that the one-quarter-ahead forecasts are, on average, correct, we simply regress the forecast error series against a constant (Holden and Peel 1990); while for the four-quarter-ahead forecasts, we also allow up to a third-order moving-average process in the corresponding regression (Diebold and Lopez 1996), to capture that the four-quarter-ahead forecasts actually reflect a path of forecasts for each quarter from the beginning of the forecast horizon.

point is further illustrated by Table 4, which breaks the RMSE results reported in Table 3 into results for five-year sub-periods – except the last, which covers only the two years from 2001:Q1 to 2002:Q4. For quarterly inflation one quarter ahead, our real-time Phillips curve-based approach outperforms our AR model-based and random walk approaches, in RMSE terms, only over the last few years (although it also achieves broadly comparable performance to our random walk forecasts in the late 1970s and early 1980s). It performs markedly worse than either real-time benchmark from the mid 1980s to the mid 1990s.

Figure 1: Real-time Forecasts for Year-ended Inflation

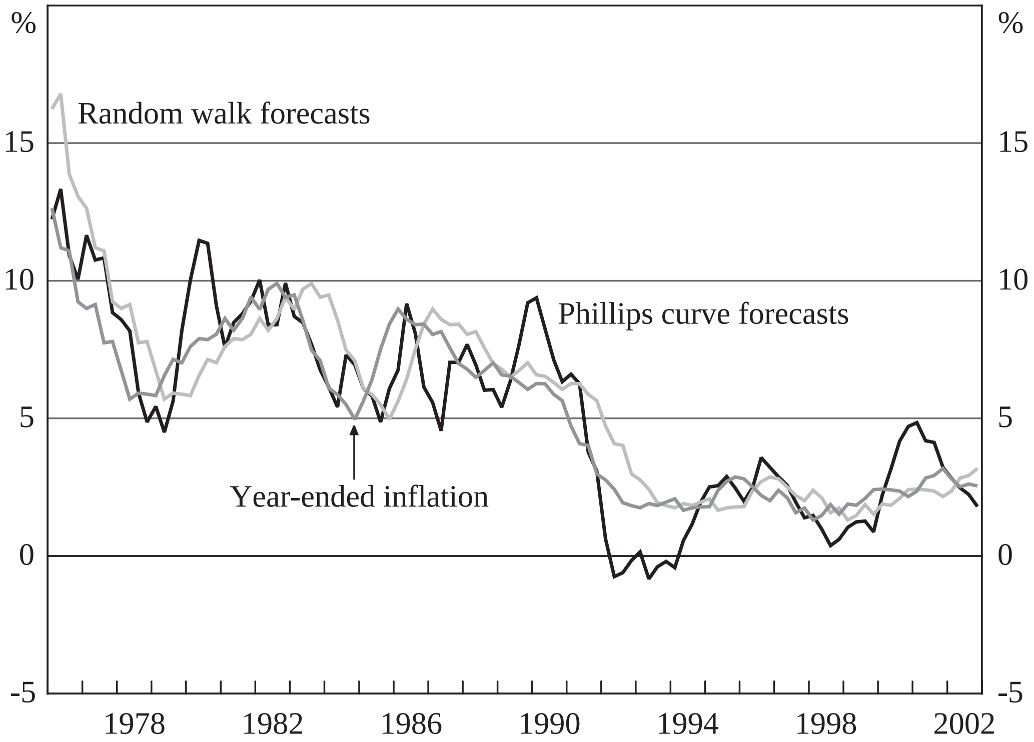
Phillips curve forecasts versus AR model forecasts



For year-ended inflation a similar story holds. Our real-time Phillips curve-based forecasts do better in RMSE terms than either real-time benchmark in the first half of the 1980s; and also do better than our random walk forecasts in the 1976 to 1980 sub-period. However, in all other sub-periods they do worse than either benchmark – sometimes dramatically so. For example, from 1996 onwards, a period of low and stable inflation, they perform around twice as badly in RMSE terms as our random walk forecasts (which in fact outperform all the other models considered in Table 4 over this period).

Figure 2: Real-time Forecasts for Year-ended Inflation

Phillips curve forecasts versus random walk forecasts



Given this generally poor performance of our Phillips curve-based forecasts in real time, it is interesting to ask: do they contain *any* useful additional information about future inflation, distinct from that provided by our simple alternative benchmarks? An indication that they may do comes from considering an alternative, weaker metric for measuring forecast performance, namely: how frequently does the forecast method at least predict correctly the *direction of change* of inflation, relative to its last value in history? In part, our interest in this metric is prompted by Fisher, Liu and Zhou (2002), who find that, for the US, Phillips curves do add some value, relative to autoregressive forecasts, by forecasting this direction of change in inflation more accurately. Table 5 shows the results for our alternative inflation forecasting models with respect to this metric.[23]

[23] We exclude our random walk model from this analysis, since by definition it never forecasts inflation to change direction over the forecast horizon.

Table 4: **Root Mean Squared Error of Alternative Inflation Forecasts Over Five-Year Sub-samples**

Method	Final vintage PC	Final vintage AR	Real-time PC	Real-time AR	Random walk
RMSE of forecasts for quarterly inflation (one quarter ahead)					
1976–1980	0.36	0.43	0.53	0.44	0.54
1981–1985	0.34	0.40	0.42	0.39	0.43
1986–1990	0.30	0.25	0.50	0.26	0.32
1991–1995	0.35	0.29	0.47	0.30	0.32
1996–2000	0.27	0.24	0.31	0.24	0.27
2001–2002	0.13	0.20	0.14	0.20	0.21
RMSE of forecasts for year-ended inflation (one year ahead)					
1976–1980	0.81	1.34	1.95	1.46	2.57
1981–1985	0.70	1.60	1.18	1.56	1.79
1986–1990	0.90	0.86	1.75	0.84	1.17
1991–1995	1.18	1.12	1.76	1.34	1.31
1996–2000	0.77	0.86	1.06	0.94	0.58
2001–2002	0.68	0.73	1.13	0.74	0.54

Note: ‘RMSE’ is the root mean squared error between actual inflation and the series of real-time inflation forecasts generated by the method shown, in percentage points.

We see from Table 5 that, at least for annual inflation over the year ahead, our real-time Phillips curve-based forecasts predict the direction of the change in inflation correctly more often than do our real-time AR model-based forecasts (assessed over the full evaluation period). Indeed, they even outperform the final vintage AR model forecasts under this metric – by around 7 percentage points.

Table 5: **Accuracy of Direction of Change Predictions for Alternative Inflation Forecasting Methods**

Method	Final vintage PC	Final vintage AR	Real-time PC	Real-time AR
Forecasts for quarterly inflation (one quarter ahead)				
Direction correctly predicted	67.6	65.7	62.0	63.9
Forecasts for year-ended inflation (one year ahead)				
Direction correctly predicted	76.9	63.0	70.4	62.0

Notes: Results are calculated over the sample 1976:Q1 to 2002:Q4. Results are reported in percentage terms – that is, the number of forecasts for which each method correctly predicts the direction of the change in inflation over the quarter or year ahead, as a percentage of the total number of forecasts made.

4.3 Combining Forecasts

These latter results suggest that, despite their poor performance in RMSE terms, the real-time forecasts from our Phillips curves may contain some useful information, additional to that from our benchmark models. If so, it might be possible to combine these forecasts with those from these benchmark models, to produce projections which are superior to either set in isolation. The intuition here is that, even where one set of forecasts performs much better than another, some value may be able to be gained by combining the two, if the errors in each are not perfectly correlated. We now test this possibility formally by estimating several sets of suitable 'forecast combination' equations.

Ex post forecast combination tests

We begin with an *ex post* forecast combination analysis of relative forecast performance. This involves regressing actual quarterly or year-ended inflation against two alternative forecasts for the same quantity, over our full evaluation period 1976:Q1 to 2002:Q4, and observing how much weight the regression chooses to place on each of the two competing forecasts. Specifically, for the case of quarterly inflation, let f^{pc} denote our real-time Phillips curve-based forecasts, and let f^{alt} represent an alternative set of forecasts, either those from our real-time AR models or from our random walk benchmark (denoted f^{ar} and f^{rw} respectively). Then we are interested in the weight, λ, which the regression chooses to place on f^{pc}, relative to that placed on f^{alt}, in the regression: [24]

$$\pi_t = \lambda\, f_t^{pc} + (1-\lambda)\, f_t^{alt} + \varepsilon_t \quad , \quad \varepsilon_t \sim N(0, \sigma^2)\,. \tag{1}$$

[24] Chong and Hendry (1986) refer to such a regression as an 'encompassing test'. If $\lambda = 1$ then the first set of forecasts are said to encompass the second, which add no additional information; if $\lambda = 0$ then the opposite conclusion holds. If $0 < \lambda < 1$ then neither model encompasses the other, and Chong and Hendry argue that this indicates that one should *adapt* one of the models, building into it elements from the other, until it does fully encompass the latter. However, there are obstacles to our proceeding in this way in the current setting, partly stemming from real-time considerations. See Appendix C for further details.

For year-ended inflation the corresponding regression is

$$\Delta_4 p_t = \lambda f_t^{pc} + (1-\lambda) f_t^{alt} + \varepsilon_t \quad , \quad \varepsilon_t \sim N(0,\sigma^2) \tag{2}$$

where f^{pc} and f^{alt} now denote forecasts for year-ended inflation.[25]

In all, we therefore conduct four separate *ex post* forecast combination tests: two for our one-quarter-ahead forecasts (real-time Phillips curve-based forecasts versus our real-time AR model-based and random walk benchmark forecasts, in turn); and another two for our one-year-ahead forecasts. The results of these regressions, estimated by OLS, are as follows:

$$\pi_t = 0.22 f_t^{pc} + 0.78 f_t^{ar} + \varepsilon_t \tag{3}$$

$$\Delta_4 p_t = 0.36 f_t^{pc} + 0.64 f_t^{ar} + \varepsilon_t \tag{4}$$

$$\pi_t = 0.36 f_t^{pc} + 0.64 f_t^{rw} + \varepsilon_t \tag{5}$$

$$\Delta_4 p_t = 0.51 f_t^{pc} + 0.49 f_t^{rw} + \varepsilon_t \, . \tag{6}$$

These *ex post* forecast combination tests suggest that our real-time AR model-based and random walk forecasts already contain much of the information incorporated in our real-time Phillips curve-based forecasts – although not all of it. For the year-ahead forecasts especially, the Phillips curves appear to contain some information on inflation distinct from that from our AR model-based and random walk forecasts, with weights of $\lambda = 0.36$ and $\lambda = 0.51$ assigned to the Phillips curve-based forecasts in these cases.[26]

[25] Regressions (1) and (2) incorporate a number of implicit restrictions, such as that the weights on the competing sets of forecasts in each case sum to one. See Footnote 27 for a discussion, in the setting of real-time forecast combination, of the impact of relaxing these restrictions.

[26] If we were to treat the forecasts f^{pc}, f^{ar} and f^{rw} in Regressions (3) to (6) as exogenous, then in each case the weight on the Phillips curve-based forecasts would, in fact, be significantly different from 0 at the 5 per cent level. However, these series represent constructed data which are subject to uncertainty – stemming, for example, from errors in coefficient estimation in the models used to produce them. Hence, OLS standard errors are not valid here: see West (2000).

Real-time forecast combination

In the same way, however, that final vintage RMSE results do not accurately reflect the *real-time, operational* usefulness of a given forecasting technique, so the above forecast combination tests do not properly capture the real-time usefulness or otherwise of our Phillips curve-based forecasts, as a supplement to those from our alternative benchmark models. This is because the weights in Equations (3) to (6) are those from *ex post* regressions, using our various alternative real-time forecasts over the full evaluation period, 1976:Q1 to 2002:Q4. A policy-maker in the middle of this period, trying to construct optimal combined inflation forecasts in real time, would not know these optimal full-sample weights.

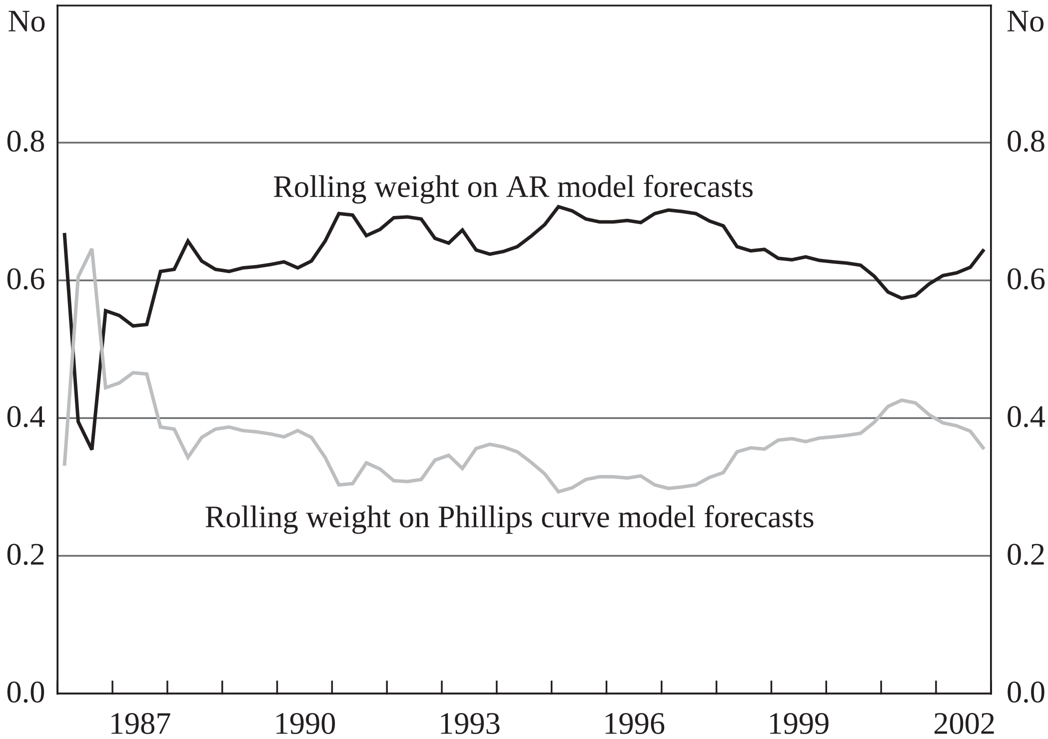
To more accurately represent the policy-maker's problem in real time, we repeat the estimation of the forecast combination equations described above (see Regressions (1) and (2)), but now on a rolling basis.[27] Specifically, we use a 10-year lagged window to determine, in each period, the weights which a policy-maker would, in real time, choose to place on each of the competing forecasts for inflation. For year-ended inflation, Figure 3 shows the profile over time of the estimated real-time weights placed on our Phillips curve-based forecasts, and on our AR model-based forecasts, under this real-time approach to combining these forecasts. Figure 4 does the same for the combination of our Phillips curve-based forecasts with those from our random walk benchmark.[28]

[27] We also considered a wide range of variations to Equations (1) and (2), for the general form of our combining equations. Alternatives considered included: relaxing the constraint that the weights on the two sets of competing forecasts sum to one; allowing a constant in the regressions; and allowing for the possibility in Regression (2) that the errors ε_t follow a moving average process (Diebold and Lopez 1996). These alternatives were considered both separately and jointly. None of these variations, however, yielded any significant improvement in out-of-sample combined forecast accuracy (judged in RMSE terms), either for the combination of our Phillips curve and AR model-based forecasts, or our Phillips curve and random walk forecasts.

[28] Note that Figures 3 and 4 cover the truncated evaluation period 1986:Q1 to 2002:Q4, reflecting that our use of rolling regressions to determine the combining weights in each period means that we lose 10 years from the sample for which our real-time combined forecasts are available.

Figure 3: Real-time Forecast Combination Weights

Based on 10-year rolling encompassing tests

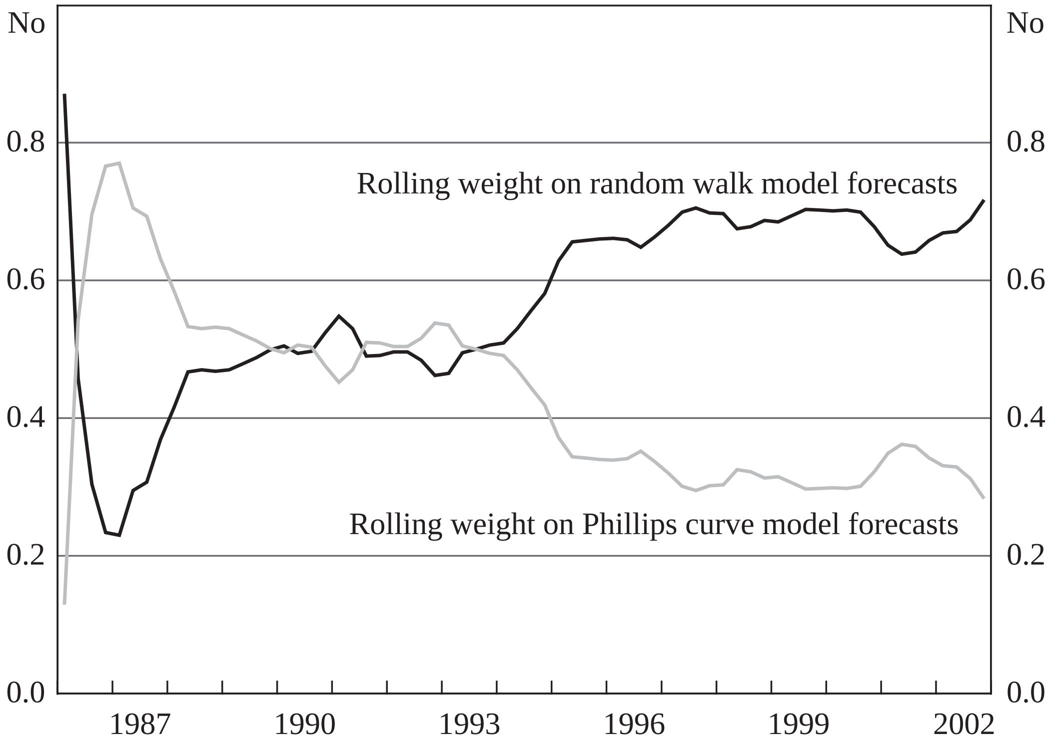


As is evident from Figure 3, the weight placed on our Phillips curve-based forecasts for year-ended inflation, when combined in real time with our AR model-based projections, is broadly stable over time at between 29 and 43 per cent (except in the first few years of the sample). By contrast, Figure 4 shows that, when combined in real time with our random walk projections for year-ended inflation, the weight placed on our Phillips curve-based forecasts exhibits two distinct phases (again leaving aside the period of extreme volatility from 1986:Q1 to 1988:Q1). Between 1988:Q2 and 1993:Q4 this weight fluctuates narrowly between 45 and 54 per cent, before undergoing a downward shift during 1994. Thereafter, it is again fairly stable, ranging now between 28 and 38 per cent.[29]

[29] The chief reason for the decline over 1994 is that 1983 was a year for which our real-time Phillips curve-based forecasts perform extremely well, whereas our random walk forecasts perform rather poorly (see Figure 2). Hence, as 1983 gradually drops out of the 10-year rolling window used to estimate this weight, which occurs over the course of 1994 in Figure 4, our Phillips curve-based forecasts are assessed to lose a good deal of their value as a guide to future inflation, relative to our random walk benchmark.

Figure 4: Real-time Forecast Combination Weights

Based on 10-year rolling encompassing tests



Using the weights shown in Figures 3 and 4 also yields new series of real-time, combined, out-of-sample forecasts for inflation, over the truncated evaluation period 1986:Q1 to 2002:Q4. The relative performance of these forecasts is summarised in Table 6, from which several interesting results emerge.

First, before even turning to the combined forecasts, it is notable that our real-time AR model-based forecasts for year-ended inflation are considerably more biased (58 basis points) over the sample considered in Table 6, than they are over the longer evaluation period considered in Table 3 (33 basis points). The reverse, however, is true for our random walk forecasts, whose bias declines from 46 to 26 basis points over the shorter sub-sample. Both methods display lower forecast RMSE over the shorter sub-sample, consistent with the period from 1976 to 1985 having been a more difficult period for which to forecast inflation than the post-1985 period.

Table 6: **Root Mean Squared Error and Bias of Alternative Real-time Inflation Forecasting Methods**

Method	Real-time PC	Real-time AR	Random walk	Combined PC and AR	Combined PC and RW
Forecasts for quarterly inflation (one quarter ahead)					
RMSE	0.41	0.26	0.30	0.26	0.29
Bias	–0.04	0.09*	0.02	0.04	0.07*
Forecasts for year-ended inflation (one year ahead)					
RMSE	1.51	1.03	1.02	0.90	0.97
Bias	–0.20	0.58*	0.26*	0.32*	–0.04

Notes: Results are calculated over the forecast sample period 1986:Q1 to 2002:Q4. 'RMSE' is the root mean squared error between actual inflation and the series of real-time inflation forecasts generated by the method shown, in percentage points; 'Bias' is the average percentage point error in the forecast series over the evaluation period. * indicates that the bias is significant at the 5 per cent level.

Secondly, combining our real-time Phillips curve-based forecasts with those from our real-time AR models does generate improved forecasts, measured relative to our real-time Phillips curve-based ones alone. This improvement is evident at both the one-quarter and one-year-ahead horizons. However, relative to our real-time AR model-based forecasts, the improvement in performance of these combined forecasts is less striking. No reduction in forecast RMSE is achieved at the one-quarter-ahead horizon; and, although there is some improvement for the year-ahead forecasts, it is still fairly small (around 13 basis points). The combining process does, however, reduce the forecast bias problem present in the AR model-based forecasts since the mid 1980s – from 58 to 32 basis points.

A similar picture emerges for the forecasts obtained by combining our real-time Phillips curve-based projections with those from our random walk benchmark. Overall, these results confirm our earlier intuition that, while our Phillips curves do seem to add some power to our capacity to forecast inflation, relative to simple alternative models, the extra information they provide is only modest.

5. Why Do Our Phillips Curves Forecast Poorly?

In Section 4 we saw that the capacity of our real-time Phillips curve-based models to forecast inflation is disappointing, relative to some simple alternatives. We now try to identify what factors are responsible for this poor performance.

Recall that our Phillips curve-based forecasts have perfect foresight built into them with regard to variables such as oil and import prices, and bond market inflation expectations. This leaves two main candidates for the source of errors in these forecasts.

One is that our real-time estimates of the output gap are subject to error, which may be *directly* contributing to problems with our Phillips curve-based forecasts of inflation. Errors in historical estimates of the gap in any given quarter feed through to corresponding errors over the forecast period. These, in turn, affect our Phillips curve-based inflation forecasts – even in the absence of any specification or coefficient estimation problems in the Phillips curve or gap equation for that quarter. The other possibility is that, in part because of real-time problems with estimating the gap, our Phillips curves and gap equations themselves may be flawed, either through mis-specification or mis-estimation of parameters.[30]

To assess which of these two factors may be the more important, we introduce two alternative suites of 'intermediate models', both of which lie between our real-time Phillips curve-based models and our final vintage Phillips curve. For the first of these sets of intermediate models, in each period we use the same Phillips curves and output-gap equations, estimated on that period's data, as in our real-time Phillips curve-based models. However, we use the final vintage data for the gap (and for inflation expectations) up to that period, in generating forecasts for inflation over the year ahead. In terms of bridging the gap between our real-time Phillips curve-based models (real-time equations and real-time data) and our final vintage Phillips curve (final equations and final data), this first set of intermediate

[30] It is also possible that a number of other factors may be contributing to the disappointing performance of our real-time Phillips curve-based forecasts. Such factors include our imposition of a long-run verticality condition on our Phillips curve specifications, and our choice of output-gap measure. However, for the US, Orphanides and van Norden (2003) find that the relative forecasting performance of output-gap-based Phillips curves with and without a long-run verticality condition imposed is about the same. Likewise, their results for the US suggest that our choice of output-gap measure is unlikely to be responsible for the unimpressive forecasting performance of our real-time Phillips curves: they examine a dozen alternative gap measures, and find all of them to be of generally low value in helping to predict inflation in real time. Moreover, it would be problematic, on other grounds, for us to switch to an alternative approach to estimating the gap, so as to try to improve our capacity to forecast inflation using Phillips curves, since the results of Gruen *et al* (2002) suggest that any gain on this front would be likely to come only at the cost of worse real-time reliability of the estimates of the gap itself.

models may therefore be thought of as representing 'real-time equations but final data'.

Our second set of intermediate models simply reverses the order in which the gap between our real-time Phillips curve-based models and final vintage Phillips curve is bridged, using instead 'final equations but real-time data'. Specifically, for this second set of intermediate models we therefore use, in each period, the final vintage Phillips curve and output-gap equation, as estimated on final vintage data. However, we use real-time data for the gap (and for inflation expectations) up to that period, in generating forecasts for inflation over the year ahead. A comparison of the forecasting performance of these two sets of intermediate models with that of our real-time and final vintage Phillips curves is provided in Table 7.

Table 7: Root Mean Squared Error and Bias of Alternative Inflation Forecasting Methods

Method	Final vintage PC	Intermediate models (I) Real-time eqns, final data	Intermediate models (II) Final eqns, real-time data	Real-time PC
Forecasts for quarterly inflation (one quarter ahead)				
RMSE	0.31	0.44	0.32	0.44
Bias	0.01	–0.02	0.03	–0.01
Forecasts for year-ended inflation (one year ahead)				
RMSE	0.87	1.59	1.03	1.55
Bias	0.08	–0.05	0.22*	0.03

Notes: Results are calculated over the forecast sample period 1976:Q1 to 2002:Q4. 'RMSE' is the root mean squared error between actual inflation and the series of real-time inflation forecasts generated by the method shown, in percentage points; 'Bias' is the average percentage point error in the forecast series over the evaluation period. * indicates that the bias is significant at the 5 per cent level.

Interestingly, we see that, on average over the whole sample from 1976:Q1 onwards, our first set of intermediate models performs no better (in RMSE terms) in forecasting inflation than our real-time Phillips curve-based models, at either the one- or four-quarter-ahead horizons. It appears that, if forced to work with Phillips curves estimated in real time, then one may as well continue to use the same flawed, real-time output-gap data used in the equation specification and estimation process, when forecasting inflation. Doing so produces forecasts which perform as well as predictions based on more accurate gap information, determined using revised output data and with the benefit of hindsight.

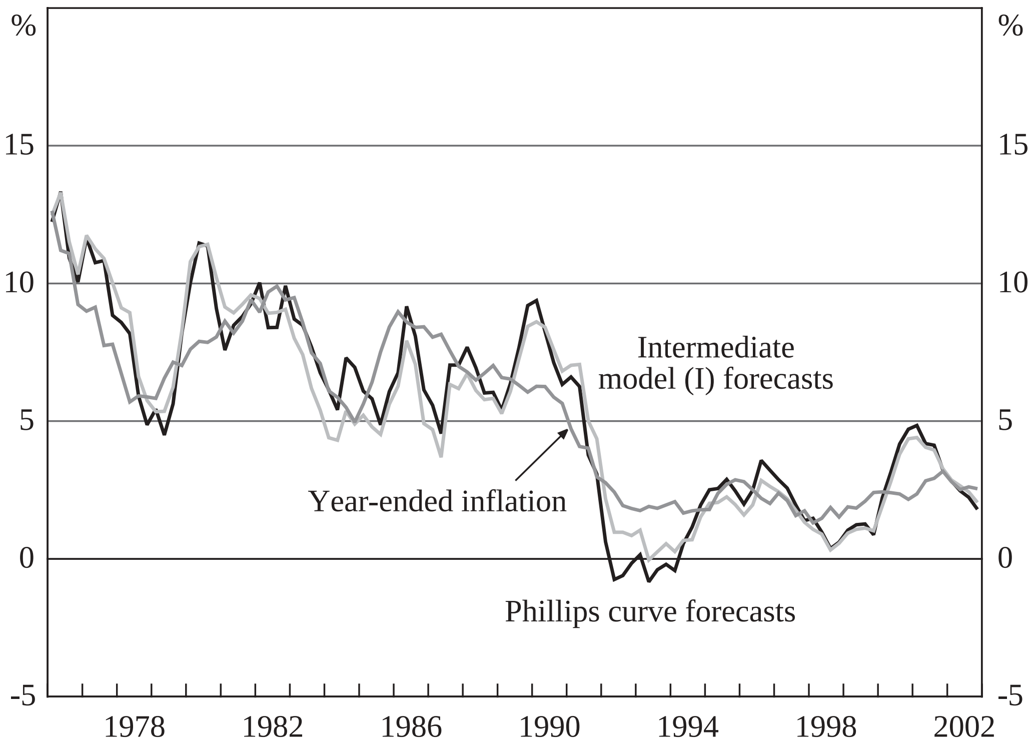
By contrast, while our second set of intermediate models performs worse in forecasting inflation than our final vintage Phillips curve, it does markedly better than our real-time Phillips curve-based models. This is true, in RMSE terms, at both the one- and four-quarter-ahead horizons – although, at the latter horizon, the forecasts from our second set of intermediate models do exhibit some bias.

In combination, these results strongly suggest that the principal problem with the use of Phillips curve-based models to forecast inflation in real time lies in identifying the correct specification and coefficient estimates for the model equations. By contrast, the *direct* effects of output-gap mis-estimation on the inflation forecasts seem relatively unimportant – a finding which aligns with that of Orphanides and van Norden (2003) for the US.[31] There have, however, been particular periods where the direct effects of gap mis-estimation have been important, such as the early 1990s recession (already discussed in Section 4.2). The persistent mis-estimation of the output gap during this period, by around 3 percentage points, is a major source of errors in our real-time Phillips curve-based forecasts for year-ended inflation for this period – as illustrated by Figure 5, which compares these forecasts (real-time equations, real-time data) with those from our first set of intermediate models (real-time equations, final data).

[31] Orphanides and van Norden observe that their results 'provide no evidence to suggest that the relative inaccuracy of [their various] real-time gaps is responsible for their poor forecasting performance', but rather suggest 'quantitatively important instability in the relationship between output gaps (however measured) and inflation' (p 16).

Figure 5: Real-time Forecasts for Year-ended Inflation

Phillips curve forecasts versus intermediate model (I) forecasts



6. Conclusions

Among the many techniques used to forecast inflation are Phillips curves based on estimates of the output gap. This paper suggests, however, that their real-time capacity to do so is limited, relative even to such simple alternative forecasting approaches as an AR(2) model or a random walk assumption.[32] It appears that, while the Phillips curve relationship is useful in real time as a source of information upon which to condition estimates of the output gap, the lack of precision with which the relationship can be estimated in real time limits its usefulness as a means of forecasting inflation.[33] This is so despite our having

[32] This conclusion accords with the recent results of Orphanides and van Norden (2003) for the US, regarding the real-time forecasting power of Phillips curves, notwithstanding the different frameworks used to examine the issue in the two studies.

[33] By comparison, the direct impact of real-time output-gap mis-estimation, on the performance of Phillips curve-based inflation forecasts, appears to be of secondary importance.

taken care to try to make our Phillips curve-based models as richly specified and realistic as possible.

Our Phillips curve-based forecasts may, however, perform a little better than AR model-based ones in at least predicting whether inflation will increase or decrease from its current level. Moreover, combining Phillips curve-based forecasts with those from our alternative, benchmark approaches, does seem to offer at least some scope for improving the real-time out-of-sample forecast accuracy of the latter.

Finally, an inflation-targeting central bank may, in any case, wish to react to anticipated spare capacity in the economy, beyond its expected implications for inflation.[34] Whether the output gap can be estimated sufficiently accurately in real time for such a purpose remains open, but at least the findings of Gruen *et al* (2002) on that score were more promising than those of this paper.

These latter observations point to a possible ongoing role for output-gap-based Phillips curves, beyond their value as an *ex post* tool for understanding historical movements in inflation. They suggest that, in spite of their generally disappointing performance as a means of forecasting inflation in isolation, such Phillips curves may continue to be useful in real time – as a tool for conditioning gap estimates within a multivariate filtering framework, and as a possible complement to other, alternative inflation forecasting approaches.

[34] It is increasingly accepted that flexible inflation-targeting central banks, including the Reserve Bank of Australia, focus not only on deviations of forecast inflation from target, but also of forecast output from potential – see, for example, Bean (2003).

Appendix A: Construction of Real-time Output Gaps and Phillips Curves

Gruen *et al* (2002) examine a range of univariate and multivariate methods for estimating the output gap in real time for Australia. They conclude that the best approach is to use an econometric technique based on inverting a Phillips curve, subject to a smoothness constraint. Since we use both the real-time output gaps and Phillips curves which they generate via this technique, we give a brief outline of it here. For a detailed description of the technique, see Gruen *et al* (2002).

Gruen *et al* start with an expectations-augmented Phillips curve of the form

$$\pi_t = \pi_t^e + \gamma(y_t - y_t^*) + \theta Z_t + \varepsilon_t \tag{A1}$$

where π_t denotes quarterly core consumer price inflation (which has the desirable property for real-time analysis that it is not revised); π_t^e denotes inflation expectations; y_t and y_t^* denote actual and potential output (in logs); Z_t represents a vector of other variables (which may include changes in the output gap); and ε_t denotes an error term.[35] Then, for each vintage of data, they seek the *smooth* path for potential output that gives the best fit to this Phillips curve. Formulated mathematically, this entails finding the values for the parameter γ, the parameter vector θ, and the potential output series $\{y_t^*\}$ which minimise the loss function

$$\mathscr{L} = \sum_{t=1}^{n} \varepsilon_t^2 + \lambda_{PC} \sum_{t=2}^{n-1} \left((y_{t+1}^* - y_t^*) - (y_t^* - y_{t-1}^*) \right)^2 \tag{A2}$$

where, as in the usual H-P filter, λ_{PC} is a smoothing parameter to be chosen.[36]

35 The Z_t variables are constructed to be zero in the long run, so that the Phillips curve defined by Equation (A1) is vertical in the long run, with output at potential when inflation is equal to expected inflation.

36 Details of the choice of λ_{PC} are set out in Gruen *et al* (2002). Note also that this loss function is similar to that underlying the conventional H-P filter, but with the sum of squared residuals from the Phillips curve in place of the sum of squared deviations of potential output from actual output. This has the desirable effect of building economic information into the filter: rather than potential output being penalised for differing from actual output in each period, such deviations are only penalised in Equation (A2) to the extent that they do not help to explain observed inflation. Incorporating economic information into the filter in this way also has the benefit that it reduces the end-point problems associated with H-P filters (since observed inflation then helps to determine the level of the output gap in each period more tightly).

For each data vintage, Gruen *et al* use a general-to-specific approach to choose an appropriate specification for the Phillips curve – as described in detail in Gruen *et al* (2002). Overall, the Phillips curve specifications selected vary only slowly over time. Table A1 lists these real-time specifications in full.

Table A1: **Complete List of Phillips Curve Specifications**

Date of vintage	Precise equation specification
1971:Q4 to 1972:Q4	$\pi_t = 0.5(\pi_{t-2}+\pi_{t-3})+\beta_2\text{bond}_{t-2}+\beta_3\text{bond}_{t-3}+\beta_4\text{bond}_{t-4}+ \gamma(y_t-y_t^*)+\delta_4\Delta(y_{t-4}-y_{t-4}^*)+ \xi_3\text{import}_{t-3}+\xi_4(\text{import}_{t-4}+\text{import}_{t-5}+\text{import}_{t-6})+\varepsilon_t$
1973:Q1 to 1973:Q2	$\pi_t = 0.5(\pi_{t-2}+\pi_{t-3})+\beta_2\text{bond}_{t-2}+\beta_3\text{bond}_{t-3}+\beta_4\text{bond}_{t-4}+ \gamma(y_t-y_t^*)+ \xi_3\text{import}_{t-3}+\xi_4(\text{import}_{t-4}+\text{import}_{t-5}+\text{import}_{t-6})+\varepsilon_t$
1973:Q3	$\pi_t = 0.5(\pi_{t-2}+\pi_{t-3})+\beta_2\text{bond}_{t-2}+\beta_3\text{bond}_{t-3}+\beta_4\text{bond}_{t-4}+ \gamma(y_t-y_t^*)+\xi_3\text{import}_{t-3}+\varepsilon_t$
1973:Q4 to 1974:Q2	$\pi_t = 0.5(\pi_{t-2}+\pi_{t-3})+\beta_2\text{bond}_{t-2}+\beta_3\text{bond}_{t-3} \gamma(y_t-y_t^*)+\eta_3\text{oil}_{t-3}+\varepsilon_t$
1974:Q3	$\pi_t = 0.25(\pi_{t-2}+\pi_{t-3}+\pi_{t-4}+\pi_{t-5})+\beta_1\text{bond}_{t-1}+ \gamma(y_t-y_t^*)+\eta_2\text{oil}_{t-2}+\varepsilon_t$
1974:Q4 to 1975:Q3	$\pi_t = 0.25(\pi_{t-2}+\pi_{t-3}+\pi_{t-4}+\pi_{t-5})+\beta_1\text{bond}_{t-1}+ \gamma(y_t-y_t^*)+\eta_2\text{oil}_{t-2}+\eta_3\text{oil}_{t-3}+\varepsilon_t$
1975:Q4 to 1986:Q2	$\pi_t = 0.25(\pi_{t-2}+\pi_{t-3}+\pi_{t-4}+\pi_{t-5})+\beta_1\text{bond}_{t-1}+ \gamma(y_t-y_t^*)+\eta_2\text{oil}_{t-2}+\eta_3\text{oil}_{t-3}+\eta_7\text{oil}_{t-7}+\varepsilon_t$
1986:Q3 to 1998:Q2	$\pi_t = 0.25(\pi_{t-2}+\pi_{t-3}+\pi_{t-4}+\pi_{t-5})+ \zeta_2(\pi_{t-2}-\pi_{t-6})+\zeta_3(\pi_{t-3}-\pi_{t-7})+ \beta_1\text{bond}_{t-1}+\beta_2\text{bond}_{t-2}+ \gamma(y_t-y_t^*)+\eta_2\text{oil}_{t-2}+\eta_3\text{oil}_{t-3}+\eta_7\text{oil}_{t-7}+\varepsilon_t$
1998:Q3 to 2002:Q4	$\pi_t = 0.25(\pi_{t-1}+\pi_{t-2}+\pi_{t-3}+\pi_{t-4})+ \zeta_2(\pi_{t-2}-\pi_{t-6})+\beta_1\text{bond}_{t-1}+ \gamma(y_t-y_t^*)+\eta_2\text{oil}_{t-2}+\eta_7\text{oil}_{t-7}+ \xi_0\text{import}_t+\xi_1\text{import}_{t-1}+\varepsilon_t$

Note: Start of sample for all regressions is 1961:Q2.

Appendix B: Selection of Real-time Output-gap Equations

The five broad specification types for our chosen output-gap equation specifications are summarised in Table 2 in Section 3. Table B1 provides a complete listing of the specifications used for each of our 125 data vintages.

Table B1: Complete List of Output-gap Equation Specifications

Date of vintage	Precise equation specification
1971:Q4 to 1978:Q1	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \gamma_1 \text{rshare}_{t-1} + \varepsilon_t$
1978:Q2 to 1981:Q3	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \beta_5 \text{gap}_{t-5} + \gamma_1 \text{rshare}_{t-1} + \delta_1 \Delta\text{tot}_{t-1} + \delta_3 \Delta\text{tot}_{t-3} + \lambda_5 \Delta\text{rtwi}_{t-5} + \phi_1 \text{soi}_{t-1} + \varepsilon_t$
1981:Q4 to 1985:Q3	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \beta_5 \text{gap}_{t-5} + \gamma_1 \text{rshare}_{t-1} + \delta_3 \Delta\text{tot}_{t-3} + \lambda_2 \Delta\text{rtwi}_{t-2} + \lambda_5 \Delta\text{rtwi}_{t-5} + \phi_1 \text{soi}_{t-1} + \varepsilon_t$
1985:Q4 to 1987:Q2	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \beta_5 \text{gap}_{t-5} + \gamma_1 \text{rshare}_{t-1} + \delta_3 \Delta\text{tot}_{t-3} + \lambda_1 \Delta\text{rtwi}_{t-1} + \phi_1 \text{soi}_{t-1}$ $\psi_2(\text{rcash}_{t-2} + \text{rcash}_{t-3}) + \varepsilon_t$
1987:Q3 to 1989:Q1	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \beta_5 \text{gap}_{t-5} + \gamma_1 \text{rshare}_{t-1} + \lambda_1 \Delta\text{rtwi}_{t-1} + \phi_1 \text{soi}_{t-1}$ $\psi_2(\text{rcash}_{t-2} + \text{rcash}_{t-3}) + \varepsilon_t$
1989:Q2 to 1993:Q2	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \gamma_1 \text{rshare}_{t-1} + \lambda_1 \Delta\text{rtwi}_{t-1} + \phi_1 \text{soi}_{t-1}$ $\psi_2(\text{rcash}_{t-2} + \text{rcash}_{t-3}) + \varepsilon_t$
1993:Q3 to 1995:Q2	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \gamma_1 \text{rshare}_{t-1} + \phi_1 \text{soi}_{t-1} + \psi_2(\text{rcash}_{t-2} + \text{rcash}_{t-3}) + \varepsilon_t$
1995:Q3 to 2002:Q4	$\text{gap}_t = \alpha + \beta_1 \text{gap}_{t-1} + \gamma_1 \text{rshare}_{t-1} + \phi_1 \text{soi}_{t-1} + \psi_2(\text{rcash}_{t-2} + \text{rcash}_{t-3} + \cdots + \text{rcash}_{t-7}) + \varepsilon_t$

Note: Start of sample for all regressions is 1961:Q2.

Appendix C: Implications of our Forecast Combination Results for Model Specification

In Section 4, when real-time combination tests involving two sets of forecasts showed that neither fully encompassed the other, we used the rolling weights estimated from these tests to form a new, combined set of forecasts. An alternative response, however, has been proposed in the literature. This approach, due to Chong and Hendry (1986), argues that the failure of either model to encompass the other should be taken to indicate that one should *adapt* one of the models, building into it elements from the other, until it does fully encompass the latter. It is therefore interesting to ask: do our real-time forecast combination results in Section 4 have implications for the optimal specifications of our Phillips curves?[37]

The main option that forecasters in our setting might investigate would be adding extra (possibly longer) lags of inflation to their Phillips curves in each period, given the strong relative real-time performance of our AR models of inflation. However, the scope for such re-specification is, in practise, limited. To begin with, the general Phillips curve specification used in Gruen *et al* (2002) already allowed for long lags of inflation, as part of an information set strictly richer than that available to our AR models.[38] Moreover, a real-time aspect of the specification search process for our Phillips curve-based models makes the use of encompassing tests of forecast performance, for assessing model mis-specification, problematic.

Specifically, for each data vintage our procedure for estimating potential output creates a variable, the output gap, which (partly by construction) is found to be a highly significant explanator of inflation, but which is itself influenced by the specification of that vintage's Phillips curve. This makes it difficult to amend our real-time Phillips curve specifications to encompass our AR models – say, by adding additional lags of inflation – since they involve a variable, the output gap, whose construction in each period is itself intertwined with their specification.

[37] That said, the residuals shown in Figure 7 of Gruen *et al* (2002) for our final vintage Phillips curve display few of the traditional signs of misbehaviour relating to equation mis-specification. For example, they exhibit no significant autocorrelation (at least after excluding from consideration the period associated with the first OPEC oil price shock in early 1974).

[38] Note, however, that we do impose a constraint of long-run verticality on our Phillips curves, which we do not correspondingly impose in our AR models.

Appendix D: Data Sources and Definitions

Many of the data series used in this paper were also used in Gruen *et al* (2002). Definitions and sources for these series were outlined in Appendix D of that paper. In this appendix we describe only those series used here which were not used in Gruen *et al*, or whose precise definition or source have been altered slightly.

Nominal cash rate and 90-day bank bill rate

The nominal cash rate is the quarter average of monthly data for the interbank overnight rate, for the period from July 1998 onwards, and for the 11am call rate, up to June 1998 (both available from Table F.1 of the RBA *Bulletin*, 'Interest Rates and Yields – Money Market'). For the 90-day bank bill rate we use the average level of this series over the last week of each quarter, also available from Table F.1 of the RBA *Bulletin* [issue 1998:M7].

Weighted average interest rate on overdraft advances

Constructed from RBA *Bulletin* Table 'Advances Classified by Interest Rates' [1963:M9–1988:M11]. Interest rates are assumed to be the midpoint of each category (e.g., 'more than 5% but less than $5\frac{1}{2}$%' is taken to be $5\frac{1}{4}$%), except for end categories, where the relevant boundary is used (e.g., '5% and less' is taken to be 5%). These are then weighted by their share of total advances outstanding.

Bond market inflation expectations

The bond market inflation expectations series is as constructed in Gruen *et al* (2002), except that, rather than using the end-quarter value of the nominal 10-year bond yield, we use the average of the end-month values in each quarter (from Table F.2 of the RBA *Bulletin*, 'Capital Market Yields – Government Bonds').

De-trended real share accumulation index

We first construct a nominal share accumulation index for Australia (including both capital gains and the re-investment of dividends). From 1992:Q3 onwards this is the quarter average of the daily close of the Standard and Poor's (S&P)/Australian Stock Exchange (ASX) 200 accumulation index, available from Datastream (code: ASX200(RI)). For the period 1990:Q1–1992:Q2, we back-cast the S&P/ASX 200 index using changes in the quarter average of the daily closing values of the old ASX All Ordinaries Index (since renamed the ASX Share

Price Index), again available from Datastream (code: AORDASX(RI)). For the periods 1980:Q1–1989:Q4, 1970:Q1–1979:Q4, and 1959:Q1–1969:Q4, we repeat this back-casting procedure, but now using in turn: the quarter average of the end-month values of the ASX Share Price Index (available from the 'S&P/ASX 200 accumulation index' series in Table F.7, 'Share Market', of the RBA *Bulletin* – see table notes); data on the MSCI national net index (official) for Australia, available from <http://www.msci.com>; and finally, linearly interpolated data for the annual All Ordinaries Accumulation Index from Foster (1997), Table 3.18a, available at <http://www.rba.gov.au/Statistics/op8_index.html>.

To construct a real version of this nominal index, we then deflate the series by the 2002:Q4 vintage median consumer price index. Finally, to de-trend the series we use a Hodrick-Prescott (H-P) filter with smoothness parameter $\lambda_{HP} = 1\ 600$. Note that this approach to de-trending ignores real-time considerations (since an H-P filter applied to vintages of our real share accumulation index would yield estimates for the trend level of this series different from those obtained by filtering the full series up to December quarter 2002). However, neglecting this real-time complication has little effect on our results.

Real trade-weighted exchange rate

We use an index of the Australian dollar against a trade-weighted basket of major-trading-partner currencies, adjusted for domestic and foreign consumer prices. Available from <http://www.rba.gov.au/Statistics/>.

Southern Oscillation Index

The Southern Oscillation Index is available from the Commonwealth Bureau of Meteorology at <http://www.bom.gov.au/climate/current/soihtm1.shtml>.

Statutory Reserve Deposits ratio

The Statutory Reserve Deposits (SRD) ratio is the daily average from the RBA *Bulletin* Table 'Ratios of Selected Assets to Total Deposits' [1965:M9–1984:M7]. These original data were seasonally adjusted using the X12 econometric package.

Terms of trade

Data are from National Income, Expenditure and Product (ABS Cat No 5206.0), December quarter 2002.

References

Atkeson A and LE Ohanian (2001), 'Are Phillips curves useful for forecasting inflation?', *Federal Reserve Bank of Minneapolis Quarterly Review*, 25(1), pp 2–11.

Bean C (2003), 'Asset prices, financial imbalances and monetary policy: are inflation targets enough?', in T Richards and T Robinson (eds), *Asset prices and monetary policy*, Proceedings of a Conference, Reserve Bank of Australia, Sydney, pp 48–76.

Beechey M, N Bharucha, A Cagliarini, D Gruen and C Thompson (2000), 'A small model of the Australian macroeconomy', Reserve Bank of Australia Research Discussion Paper No 2000-05.

Bullock M, D Morris and G Stevens (1989), 'The relationship between financial indicators and economic activity: 1968-1987', in I Macfarlane and G Stevens (eds), *Studies in money and credit*, Proceedings of a Conference, Reserve Bank of Australia, Sydney, pp 53–85.

Chong YY and DF Hendry (1986), 'Econometric evaluation of linear macroeconomic models', *Review of Economic Studies*, 53(4), pp 671–690.

de Brouwer G and NR Ericsson (1998), 'Modeling inflation in Australia', *Journal of Business and Economic Statistics*, 16(4), pp 433–449.

Diebold FX and JA Lopez (1996), 'Forecast evaluation and combination', in GS Maddala and CR Rao (eds), *Handbook of statistics*, 14, North-Holland, Amsterdam, pp 241–268.

Fisher JDM, CT Liu and R Zhou (2002), 'When can we forecast inflation?', *Federal Reserve Bank of Chicago Economic Perspectives*, 1Q, pp 30–42.

Foster RA (1997), *Australian economic statistics 1949-50 to 1996-97*, Reserve Bank of Australia Occasional Paper No 8.

Gruen D, A Pagan and C Thompson (1999), 'The Phillips curve in Australia', Reserve Bank of Australia Research Discussion Paper No 1999-01.

Gruen D, T Robinson and A Stone (2002), 'Output gaps in real time: are they reliable enough to use for monetary policy?', Reserve Bank of Australia Research Discussion Paper No 2002-06.

Holden K and DA Peel (1990), 'On testing for unbiasedness and efficiency of forecasts', *Manchester School of Economic and Social Studies*, 58(2), pp 120–127.

Nelson E and K Nikolov (2001), 'UK inflation in the 1970s and 1980s: the role of output gap mismeasurement', Bank of England Working Paper No 148.

Orphanides A, RD Porter, D Reifschneider, RJ Tetlow and F Finan (1999), 'Errors in the measurement of the output gap and the design of monetary policy', Federal Reserve Board Finance and Economics Discussion Series No 99-45.

Orphanides A and S van Norden (2002), 'The unreliability of output-gap estimates in real time', *Review of Economics and Statistics*, 84(4), pp 569–583.

Orphanides A and S van Norden (2003), 'The reliability of inflation forecasts based on output gap estimates in real time', Centre for Interuniversity Research and Analysis on Organizations Technical Report No 2003s-01.

Reifschneider DL, DJ Stockton and DW Wilcox (1997), 'Econometric models and the monetary policy process', *Carnegie-Rochester Conference on Public Policy*, 47, pp 1–37.

Sims CA (2002), 'The role of models and probabilities in the monetary policy process', *Brookings Papers on Economic Activity*, 2, pp 1–40.

Stevens G (2001), 'The monetary policy process at the RBA', Speech to the Economic Society of Australia Victorian Branch, 10 October, Melbourne.

Stock JH and MW Watson (1999), 'Forecasting inflation', *Journal of Monetary Economics*, 44(2), pp 293–335.

Stone A and S Wardrop (2002), 'Real-time national accounts data', Reserve Bank of Australia Research Discussion Paper No 2002-05.

West K (2000), 'Encompassing tests when no model is encompassing', NBER Technical Working Paper No 256.