REAL-TIME NATIONAL ACCOUNTS DATA

Andrew Stone and Sharon Wardrop

Abstract

Quarterly national accounts data are amongst the most important and eagerly awaited economic information available, with estimates of recent growth regarded as a key summary indicator of the current health of the Australian economy. Official estimates of quarterly output are, however, subject to uncertainty and subsequent revision. Hence, the official estimates of quarterly national accounts aggregates, with which policy-makers must work, may in practice be an inaccurate guide to their 'true' values, not just initially but even for some time after the event.

In this paper we examine over 120 vintages of Australian GDP data to provide an historical assessment of the scale and persistence of real-time errors in the measurement of actual output. The issue of whether it is possible to obtain reliable real-time estimates of the output gap is addressed in detail in a companion paper (Gruen, Robinson and Stone 2002).

Table of Contents

REAL-TIME NATIONAL ACCOUNTS DATA

Andrew Stone and Sharon Wardrop

1.   Introduction

For analysts of macroeconomic trends and prospects, quarterly national accounts data are amongst the most important and eagerly awaited economic information available. Recent quarterly growth figures are regarded as a key summary indicator of the current state of the economy. Moreover, they are an essential ingredient in estimation of the output gap, or current level of actual output relative to a theoretically assessed potential level. This gap represents an important indicator of the extent of excess (or insufficient) demand within the economy – with implications for future inflation and for the optimal trajectory of monetary policy.

Like many economic statistics, however, official estimates of quarterly output are subject to uncertainty and subsequent revision, as further data come to hand and a more detailed reconciliation of different component indicators or survey measures becomes possible. It is therefore important to appreciate that the 'true' values for the national accounts data can never be known precisely. The official estimates of these quantities may, in practice, be an inaccurate guide to the 'true' state of the economy, not just initially but even for some time after the event.

The critical question is: how significant is this problem of initial mismeasurement and subsequent revision of national accounts data, both in scale and persistence? We shall refer to this issue henceforth as the 'real-time problem' for actual output data, reflecting as it does the difficulty which confronts analysts of having to reach judgements based only on the information available in real time, rather than with the benefit of hindsight.

The next two sub-sections offer further background on two principal aspects of the 'real-time problem', while the structure of the remainder of the paper is set out in Section 1.3.

1.1 Output Growth Estimates as a Real-time Economic Indicator

Given the importance attached to quarterly national accounts releases, and the widespread use of such data, it is perhaps surprising how little attention seems to be paid to the fact that these data are subject to (sometimes substantial) revision. While most commentators are undoubtedly aware that alterations are made to such data, few attempts seem to have been made to analyse methodically the scale or pattern (if any) of these adjustments. Likewise, little work seems to have been done on either the implications of possible future changes for how we should interpret the data in the most recent accounts, or the implications of the alterations to historical data incorporated in these accounts for forecasting models estimated on these historical data.

Within the Reserve Bank of Australia (RBA) a notable exception to this was the paper by Gray (1983) on 'Revisions to National Accounts'. In this paper he analysed, *inter alia*, the average scale of revisions to several main national accounts aggregates, the persistence of initial mismeasurement of these aggregates, and the relative degree to which the quarterly, semi-annual and annual growth rates of these aggregates tended to be affected by revisions. In addition, he extended his analysis to a selection of 10 of the major components of the income and expenditure sides of the accounts, and discussed the chief sources of revisions to both component and aggregate series.

A further exception to the general tendency to neglect revisions was the paper by Rossiter and Brown (1989) from the Australian Bureau of Statistics (ABS). In this paper they described preliminary results from modelling of the revisions to various current price, seasonally unadjusted national accounts aggregates (specifically gross domestic product (GDP), gross farm product and private final consumption expenditure). Besides extending previous work on the analysis and modelling of revisions to economic time series, a principal objective of their study was 'to assist the development of improved methodologies for preparation of the Australian National Accounts (ANA)'. One interesting question flowing from this latter objective is whether such analysis, together with improvements in data collection and reconciliation, may have resulted in a gradual decline in the average degree of uncertainty surrounding estimates of GDP growth. We return to this issue later in this paper.

1.2 Output Gaps and Policy Rules in Real Time – The US Debate

A second reason for the importance of real-time issues for monetary policy-makers stems from the periodic revival of interest in the debate between rules-based and discretionary monetary policy, coupled with the fact that various well-known monetary policy rules base their recommended level of the policy instrument on estimates of the output gap.

In recent years this debate has attracted renewed interest, principally sparked by a series of papers by Taylor (1993, 1998) concerning US monetary policy since the 1960s. Taylor's conclusions, however, have been the subject of vigorous debate. This debate has centred around the informational assumptions underpinning his analysis, and the implications of 'real-time' informational limitations for both the conduct of monetary policy and the fair counterfactual analysis of such policy.

In his 1998 paper 'An Historical Analysis of Monetary Policy Rules', Taylor studied actual Federal Reserve behaviour in setting monetary policy over the previous four decades. He then compared the outcomes achieved for output and inflation over this period with counterfactual simulations in which monetary policy was instead set according to a simple rule, specifically his own well-known 'Taylor rule'.

Under this rule the policy interest rate is adjusted linearly in response to current deviations of inflation from some specified target level, π^*, and of actual output from an estimated potential level. Specifically, Taylor's rule (with the standard choice of coefficients) is given by:

$$R_t = R_t^* + \frac{1}{2}\left(\pi_t - \pi^*\right) + \frac{1}{2} gap_t \qquad (1)$$

where R_t denotes the policy interest rate; R_t^* represents the neutral level of R_t (or the sum of the current annual inflation rate, π_t, and the neutral real interest rate, r^*); π^* represents the target annual inflation rate; and gap_t is the output gap in period t.

Taylor's analysis led him to make a striking claim, namely that policy run according to his simple rule would, on average, have produced superior macroeconomic outcomes to those achieved by actual policy over the past 40 years. Having previously observed the similarity between the policy prescription of his simple rule and actual Federal Reserve behaviour in the late 1980s and early 1990s, he observed that:

> A comparison of policy rules and economic outcomes points to the rule the Fed has been using in recent years as a better way to run monetary policy than the way it was run in earlier years. (Taylor 1998, p 35)

Moreover, he claimed that for the US both the 'great inflation' of the late 1960s and 1970s and the severe recession of the early 1980s, as well as a period of monetary tightness in the early 1960s, reflected deviations of monetary policy from that prescribed by the Taylor rule, and could be classified as 'policy mistakes'.[1,2]

Taylor's analysis appeared to provide powerful evidence in support of, if not the mechanical adoption of policy rules such as Taylor's rule by central banks, at least an enhanced role for such rules in informing the setting of monetary policy. His results, however, have been criticised, in particular by Orphanides (2000), on the grounds that the informational assumptions underpinning them were unrealistic.

Clearly, the practical operation of a policy rule involving the output gap, such as Taylor's rule, requires the policy-maker to feed in data on both actual and potential current output. Orphanides observed, however, that in Taylor's work the actual output data used were the latest available historical time series, and therefore included all revisions made to each quarter's data since their first release. None of this information on subsequent revisions, however, would have been available at the time to policy-makers setting short-term interest rates, whether by rule or by

---

[1] Other authors, such as Clarida, Gali and Gertler (1998) and Judd and Rudebusch (1998) have drawn similar conclusions, albeit carefully qualified in light of the problem of interpreting counterfactual regression results based on latest available rather than real-time data.

[2] In fairness, in relation to the early 1980s 'policy mistake', Taylor (1998, p 39) notes that: 'it should be emphasised that this period occurred right after the end of the 1970s inflation and the higher interest rates than recommended by the policy rules may have been necessary to keep expectations of inflation from rising and to help establish the credibility of the Fed…In my view this period has less claim to being a 'policy mistake' than the other two periods'.

discretion. In the event that revisions were substantial, this might mean that policy set according to Taylor's rule using 'real time' actual output data might be significantly different from that based on the same rule using latest available historical data. To this extent, the improved outcomes found by Taylor under his rule-based policy might not represent the superiority of his rule over actual policy. Rather, they might simply represent the benefit of hindsight, contained in the knowledge both of future economic developments and of revisions to past economic data, in assessing the true state of the economy at critical junctures.

Orphanides's charge of failing to distinguish between real time and latest available historical data in relation to *actual output* is clearly a serious one. Potentially more important again, however, is the same charge in relation to *potential output*, which Orphanides suggested may also be levelled against Taylor's analysis.

Orphanides argued that potential output, being unobservable, is particularly prone to prolonged and substantial mis-estimation. Moreover, as a purely theoretical construct, the scope for errors in its real-time estimation is much greater than for the corresponding measurement of actual output. The risk of such mis-estimation would be especially great where a shift occurred in long-term trend productivity growth, since such a trend shift would likely be difficult for policy-makers to detect for many years, obscured as it would be by cyclical variations and statistical noise.

Pursuing this point, Orphanides undertook a careful study of Federal Reserve papers over much of the past 40 years, which contained real-time estimates of the output gap, prepared either by the Fed or the Council of Economic Advisers. Drawing upon this, he argued that the estimates of the output gap used in Taylor's work, based as they were upon a recent vintage of historical data, differed dramatically from those which were available in real time. This was especially true over the 1970s, where the trend decline in US productivity growth from the late 1960s onwards was not fully apparent to US policy-makers for nearly a decade.

When Orphanides re-ran Taylor's counterfactual simulations using his real-time data, he found that under this constraint the Taylor rule actually performed slightly worse, on average, than did actual policy over the period 1966 to 1993.[3]

Taylor (2000) has, in turn, disputed Orphanides' results, arguing in particular that the real-time estimates of potential output used by Orphanides were not accepted by serious economic analysts at the time – especially in those periods where they implied large output gaps which seemed at odds with other indicators of the state of the economy. Nevertheless, in the case of the US data, Orphanides's work served to highlight how substantial this real-time informational issue could be. This may, in turn, have helped to revive interest in the significance of real-time issues for numerous other economic questions, from general macroeconomic forecasting to microeconomic issues such as banking supervision.[4]

1.3 The Structure of this Paper

Clearly, it is important for analysts to be aware of the scale, frequency, persistence and pattern (if any) of uncertainty surrounding the estimates of national accounts aggregates released each quarter by the ABS – and of the implications of this uncertainty for real-time estimates of the output gap. This paper should be viewed as the first of two addressing these general issues for the case of Australia.

A companion paper (Gruen *et al* 2002) addresses the problem of constructing real-time potential output estimates, and in so doing attempts to analyse whether, for Australia, Orphanides's real-time critique need be truly fatal to the use of output-gap estimates as a guide to the setting of monetary policy. In this paper,

---

[3] It should be noted, of course, that the importance of an awareness of the implications of informational limitations for monetary policy was not a new insight on Orphanides's part. As he himself pointed out:

> As early as 1947, Milton Friedman had sharply criticised reliance on unrealistic informational assumptions for Keynesian prescriptions to maintain 'full employment'. (Orphanides 2000, p 4)

[4] An illustration of the growing interest in the issue of real-time informational limitations is the establishment by the Federal Reserve Bank of Philadelphia of a quarterly Real Time Data Set for Macroeconomists (RTDSM) – see Croushore and Stark (2000). For an example of the potential importance of real-time issues in microeconomic analyses, see for instance Gunther and Moore (2000), who examine the performance of early warning models used in the monitoring of the financial condition of US banks.

however, we confine our attention to a purely descriptive examination of the likely scale of the real-time issue for Australia, focusing just on actual output data.

In Section 2 we begin by outlining the availability of real-time actual output data for Australia. Section 2.1 provides a brief run-down of the data we have obtained for various alternative broad measures of output, while Section 2.2 provides a historical rationale for our choice of two real-time actual output measures upon which to focus in the remainder of the paper.[5]

The main content of this paper is then presented in Section 3, where we undertake a basic inspection of the scale of changes over time to our chosen measures of actual output. This is done to provide a feel for the magnitude of the real-time issue facing policy-makers. Purely from such a simple inspection, without any sophisticated model-based analysis, a number of important conclusions can still be drawn regarding the likely seriousness of real-time informational limitations for assessing the current health of the economy.

2. Data Issues

2.1 The Availability of Real-time Output Data

Quarterly current price actual output data for Australia have been published by the ABS in timely fashion since late 1960, while corresponding constant price (or, from September quarter 1998 onwards, chain volume) data have been produced

---

5 The majority of these raw data were assembled and provided to us by the Analytical Services Branch of the Australian Bureau of Statistics, without whose assistance this project would not have been possible. The data, however, represent unofficial ABS statistics, as they have not been subjected to the usual quality assurance checks which apply to all published ABS releases. While some reconciliation checks against ABS publications have been undertaken, a full consistency check of the data has not been carried out. Responsibility for any data errors rests solely with the authors. The real-time data on the two GDP measures ultimately studied in this paper (GDP(E) and a hybrid measure discussed in more detail in Section 2.2) are available from the authors upon request.

since the early 1970s.[6] For the current study we have assembled real-time data on a variety of alternative actual output measures, principally various GDP measures.[7]

To describe in greater detail the extent of real-time data assembled for each of these alternative output measures, it suffices to do so for one of them, say GDP(E).[8] For GDP(E) a total of 109 different vintages of data have been obtained, commencing in 1974:Q4 and running through to 2001:Q4 (see Table 1). The latter 14 of these, which post-date the switch in international standard for national accounting methodology from the United Nations' 1968 *System of National Accounts* (SNA68) to the *System of National Accounts 1993* (SNA93),[9] contain data all the way back to 1959:Q3. Among the remaining 95 vintages of data, however, the situation is more varied, and is summarised in Table 1.

As noted above, the position with regard to real-time data availability is similar for the other two main alternative measures of actual output for which a comparable number of vintages are available, namely GDP(I) and GNFP (see also Table 2).

---

[6] Strictly speaking, these early GDP estimates were produced by the forerunner to the ABS, the Commonwealth Bureau of Census and Statistics (CBCS). The CBCS started publishing annual constant price GDP estimates for Australia in 1963, but the first quarterly (income-based) estimates published were for the September quarter 1971.

[7] Precisely, the series collected are: expenditure and income-based gross domestic product (GDP(E) and GDP(I)) and gross non-farm product (GNFP); production-based and average gross domestic product (GDP(P) and GDP(A)) from the late 1980s onwards; and finally, since the shift to chain-volume gross domestic product measures from September quarter 1998 onwards, the official ABS gross domestic product series (GDP). Data have been obtained on both a current and constant price basis, and in both original and seasonally adjusted terms.

[8] The situation is similar for GDP(I) and GNFP, while for GDP(P) and GDP(A) we have assembled real-time databases of all vintages of these series since their first production in the late 1980s. Likewise, a complete real-time database has been assembled of all vintages of the official ABS chain-volume GDP series, which has been produced since September quarter 1998.

[9] SNA 1993 was prepared and published jointly by the Commission of the European Communities (Eurostat), the IMF, the OECD, the United Nations and the World Bank.

Table 1: Real-time GDP(E) Data

Number of vintages	Corresponding quarters	Description
29	1974:Q4–1980:Q1, 1985:Q1, 1985:Q3, 1986:Q3, 1987:Q1, 1987:Q3, 1988:Q1 and 1988:Q3	Contain only 14 data points
6	1980:Q2–1981:Q3	Contain data from 1966:Q3 onwards – i.e., between 56 and 61 data points
20	1981:Q4–1984:Q4, 1985:Q2, 1985:Q4, 1986:Q1, 1986:Q2, 1986:Q4, 1987:Q2 and 1987:Q4	Contain data from 1969:Q3 onwards – i.e., between 50 and 74 data points
3	1988:Q2, 1989:Q1 and 1989:Q2	Contain data from 1974:Q3 onwards – i.e., between 56 and 60 data points
37	1988:Q4 and 1989:Q3–1998:Q2	Contain data from 1968:Q1 onwards – i.e., between 84 and 122 data points
14	1998:Q3–2001:Q4	Contain data from 1959:Q3 onwards – i.e., between 157 and 170 data points

2.2   Which Output Measure to Focus On?

The availability of real-time data for various alternative measures of actual output raises the question: *what is the most appropriate measure upon which to focus?* We select two series on which to concentrate, expenditure-based GDP and a hybrid GDP measure spliced together from a succession of different GDP series. By way of background to these choices, Table 2 provides a brief history of the production of quarterly GDP statistics by the ABS since these were first published in December 1960.

Table 2: Measures of GDP

Timeline for production of quarterly GDP measures

1960	Publication of current price GDP(E) and GDP(I) estimates commences in ABS Cat No 5206.0 (National Income and Expenditure)
1971	Constant price GDP(I) estimates first included in ABS Cat No 5206.0
1974	Constant price GDP(E) estimates first included in ABS Cat No 5206.0
1988	Publication of GDP(P) estimates commences in ABS Cat No 5222.0 (Gross Product, Employment and Hours Worked)
1988–1991	Gradual move to concurrent publication of ABS Cat Nos 5206.0 and 5222.0, allowing simultaneous production of GDP(E), GDP(I) and GDP(P) measures
1991	Introduction of concurrent publication of the average GDP measure, GDP(A), together with GDP(E), GDP(I) and GDP(P)
1998	Shift by the ABS to chain volume rather than constant price estimates in the national accounts, and introduction of an official chain-volume GDP series

Notes: References to constant price GDP estimates, for vintages prior to the introduction of chain-linking in 1998, denote figures prepared using average prices in a given base year. This base year was typically updated roughly every 5 years. Constant price GDP(I) was estimated by summing the deflated components of GDP(E) and then adding the statistical discrepancy deflated by the GDP(E) deflator.

In deciding which domestic output series to study, we are primarily interested here in whichever series was regarded by analysts at the time as representing the best available output indicator. Prior to 1991 this appears to have been GDP(I), which was regarded as a more reliable measure of the true state of the economy than GDP(E). As Budget Statement 2, Budget Paper No 1 of the 1980-81 Budget put it:

> In Australia, the income-based estimate of GDP has traditionally been used as the basic measure of product growth, with the statistical discrepancy being assigned to the expenditure side of the accounts.

With the introduction of a GDP(P) measure in 1988 it became possible to construct an average GDP measure, GDP(A), from the three alternative income, expenditure and production-based GDP estimates. An ABS analysis, released with the June quarter 1990 national accounts, concluded that:

> All in all, the facts seem to point to GDP(A) as being the best indicator of short-term movements [of gross domestic product]. (Aspden 1990, p 65)

However, at this time GDP(P) was still only being published with a lag relative to GDP(E) and GDP(I), preventing sufficiently timely construction of GDP(A) for it to become the preferred real-time focus of policy-makers. Not until 1991, when concurrent publication of all four GDP measures (E, I, P and A) was achieved, do we therefore regard analysts as switching their primary interest to GDP(A).

Finally, GDP(A) then appears to have remained the principal output measure for analysts until the shift from constant price to chain-volume estimates in the national accounts in September quarter 1998. This shift saw the simultaneous abolition of the GDP(A) terminology and its replacement by an official chain-volume GDP series as the new main focus for analysts.

This history then explains the latter of our two choices of real-time GDP series upon which to concentrate in Section 3, namely the hybrid of GDP(I) until September quarter 1991, GDP(A) from December quarter 1991 to June quarter 1998, and chain-volume GDP thereafter. As noted above, we have chosen to supplement examination of this hybrid series with study also of the real-time behaviour of GDP(E) estimates. The reason for this choice is that, although the hybrid series just described probably represents the main focus of analysts throughout the past three decades among alternative output measures, the bulk of output *forecasting* conducted by the RBA and Treasury over this period has been expenditure-based, reconstructing aggregate output from forecasts of expenditure components in the national accounts such as household consumption, business investment and so forth. Given that policy decisions are likely also to have been influenced, at any particular moment, by forecasts of future output, it is therefore also of interest to have a feel for how significant the real-time problem has been for GDP(E), since errors in the measurement of actual GDP(E) would likely have flowed through to errors in forecasts of this quantity.

3.   The Scale of the Real-time Issue for Australian Output Data

Analysis of the various vintages of output data assembled provides us with a picture of the degree to which assessments of output growth are sometimes altered over time. We find that it has not been uncommon for estimates of the growth of our hybrid measure of GDP described in Section 2 to be significantly amended

over subsequent years – although there is evidence that the scale of these alterations may be smaller now than in the past.[10] A similar story holds for GDP(E) – see Appendix A.

3.1   The Extent of Changes to Output Estimates

To identify the likely scale of the real-time issue for Australia it is useful to start with a straightforward comparison of what different hybrid GDP vintages have implied about output growth in selected, fixed periods. This provides a first impression of the extent to which hybrid GDP estimates are sometimes adjusted over time.

Specifically, we begin by plotting the initial estimate for hybrid GDP growth for a given quarter and comparing this to the estimates for this quarterly growth rate published in subsequent quarters. Figure 1 shows these initial and subsequent estimates of quarterly hybrid GDP growth for both December quarter 1992 and June quarter 1989. As subsequent analysis shows (see Figure 3), the extent of re-assessment over time for these two quarters is not atypical, and other quarters exhibiting more extreme adjustments over time could have been selected.[11]

Figure 1 highlights the extent to which estimates of quarterly GDP growth can change, for quite some time after the event. For example, the initial estimate for December quarter 1992 was a rise in output of 0.7 per cent. Two years later, the same quarter was estimated to have experienced output growth of 1.2 per cent, while the current (2001:Q4) estimate is 1.7 per cent. A similar picture emerges for

---

[10] In what follows, we do not refer to changes in estimates of output growth, based on our hybrid GDP series, as revisions. This is because such changes will not necessarily reflect revisions in the technical sense in which that term is used in the national accounts, namely changes to the underlying data or seasonal factors. This will be particularly so where early GDP(I)-based estimates of output growth are compared with later growth estimates based on either GDP(A) or chain-volume GDP.

[11] A particularly dramatic example is the September quarter 1974 – still in the early days of production of seasonally adjusted constant price GDP(I) estimates for Australia, and just prior to the commencement of concurrent publication of GDP(E) estimates. Output in this quarter was initially reported to have contracted by 2.8 per cent, two quarters later this had been revised to a 3.9 per cent contraction, and two years after the fact (in the September quarter 1976 accounts) the contraction was still estimated to have been 1.2 per cent. In the December quarter 2001 national accounts, however, output in that quarter is estimated to have actually grown by 1.3 per cent!

June quarter 1989. In this case, the initial estimate implied quarterly growth of 0.6 per cent. However, within three quarters this had been revised up to growth of 2.0 per cent, the same figure at which it currently stands, notwithstanding changes in the interim which saw estimates fall as low as 0.9 per cent and rise as high as 2.2 per cent.

Abstracting from the statistical detail of these changes, the cautionary implication for policy-makers is clear: to the extent that inputs like the economy's growth rate are used in the policy formation process, the substantial inherent uncertainty surrounding such data needs to be recognised. This uncertainty complicates the task of responding to the 'true' state of the economy, and suggests the dangers of trying to 'fine-tune' the economy's performance.

**Figure 1: Changing Estimates of Quarterly Output Growth**
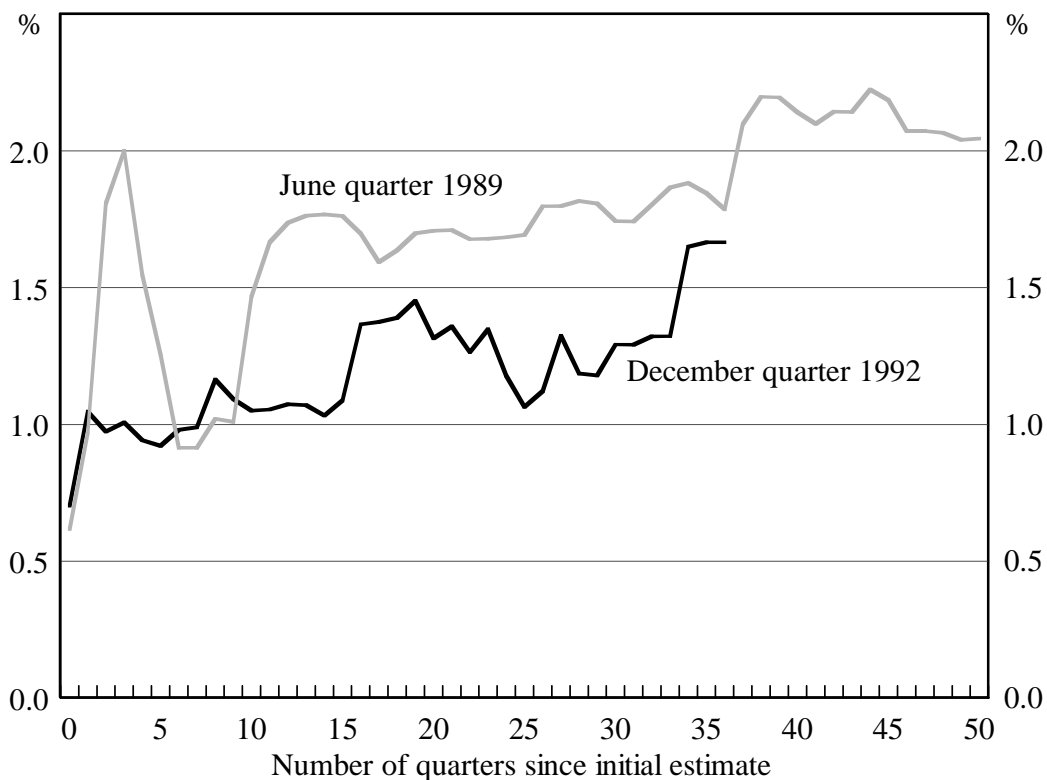
Note: Results shown are for our hybrid GDP measure (see Section 2.2).

Striking as the above examples are, it could be that studying the scale of changes to quarterly GDP growth rates gives an exaggerated impression of the problem facing policy-makers.[12] Amongst other things, this would be so if large adjustments to individual quarterly GDP growth rates mainly just reflected 'shifting' of the timing of activity between quarters. Since both policy setting and forecasting typically take into account more history than just the last quarter, it may be more relevant to study, say, changes to annual growth rates of GDP. Figure 2 shows the alterations made to the four-quarter-ended growth rate of hybrid GDP over time for the same two quarters shown in Figure 1.

Figure 2: Changing Estimates of Four-quarter-ended Output Growth
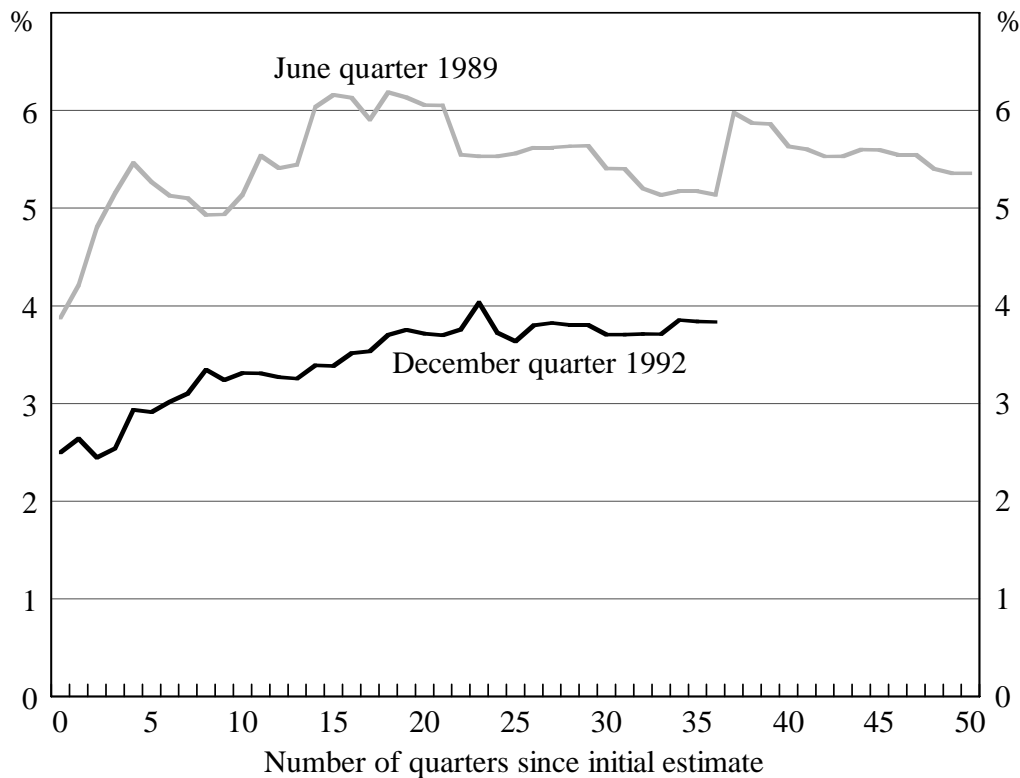
Note: Results shown are for our hybrid GDP measure (see Section 2.2).

---

[12] In each quarterly release of the national accounts the ABS includes a brief discussion of the problems associated with estimating national accounts components for periods of less than a year, and explicitly cautions that these should be taken into account in interpreting estimates over such short periods.

As Figure 2 illustrates, focusing on the changes to four-quarter-ended rather than quarterly GDP growth rates does not necessarily significantly alter the extent, or persistence, of these changes. It is still the case that estimates of four-quarter-ended growth to both December quarter 1992 and June quarter 1989 were substantially and regularly amended for a considerable time after the initial assessment for each period was released.[13] Clearly, while remaining the best guide available, such data can in isolation sometimes lead to serious misperceptions about where the economy is and, in turn, where it is going.[14]

3.2   The Frequency and Persistence of Changes to Output Estimates

Having identified that estimates of GDP growth have, at least for some periods, been subject to substantial alteration for quite some time after they were first released, the natural next question is: *how common are such instances of large mismeasurement of output growth, and for how long have any such errors tended to persist?*

To examine this, we take the December quarter 2001 estimates for hybrid GDP to represent the 'true' measure of hybrid GDP growth over time. Then, to begin with, we calculate the difference between this 'true' measure of quarterly hybrid GDP growth and the *first published estimate* for each quarter, which we refer to as an

---

[13] Again, subsequent analysis (see Figure 4) demonstrates that the two example periods shown in Figure 2 are not atypical. Furthermore, Appendix A illustrates that a similar story holds for revisions to the expenditure-based measure of GDP, GDP(E).

[14] Of course, while shifting focus from quarterly to four-quarter-ended growth rates helps to overcome the timing issue associated with simple 'shifting' of the recording of activity between quarters, it does not completely eliminate this aspect of the revisions process. To further assist in assessing the current state of the economy, it may also be useful to look at still longer-run growth averages (e.g. two-year-average growth rates), as well as at the ABS' trend growth estimates.

'error'. This gives us an idea of the magnitude of the contemporaneous errors made in each quarter in measuring quarterly hybrid GDP growth.[15,16]

Figure 3 shows the results. Quite a number of the initial estimates of quarterly hybrid GDP growth have differed by more than 2 percentage points from the 'true' value for the quarter (although none have done so since the early 1980s), and the mean absolute size of the errors made in the initial estimates of this growth, over the full sample of data vintages from September quarter 1971 to December quarter 2001, is 0.75 percentage points. Interestingly, the errors seem to have become generally less extreme over time, and especially in the 1990s, probably reflecting

---

[15] These 'errors' likely give a somewhat exaggerated measure of the true degree of uncertainty surrounding output growth estimates in real time. This is because analysts would have had other sources of data on the state of the economy available to them in real time, allowing adjustments to be made, at the margin, to the picture of the economy implicit in the hybrid GDP data upon which we are focusing. Most notably, GDP(E) data would have performed such a function between 1974:Q4, when such data were first published, and the introduction of the GDP(A) measure. Nevertheless, these 'errors' should still give a reasonable guide to the rough scale of the real-time problem which has confronted analysts over the past three decades.

[16] A technical issue also arises as to whether the latest available (2001:Q4) chain-volume GDP data provide the most appropriate measure of the 'true' growth of output for quarters far back in history. In part, this relates to the change made in 1998 to the national accounting system under which subsequent accounts have been prepared, and the difficulty of accurately adjusting estimates of real output for periods back in the 1960s, 1970s and 1980s to reflect these changes. It also relates to the fact that the ABS was incorporating a number of amendments to its historical seasonally adjusted GDP estimates around this time (to address, for example, issues regarding the benchmarking of GDP movements in certain years, the removal of residual seasonality in the time series, and so forth). To overcome this problem one could instead use the last set of accounts prepared under the old system of national accounting (SNA68), namely those from June quarter 1998, as the best available measure of the 'true' growth of output over history. However, this would raise a separate issue relating to the appropriateness of the price data used in those accounts to estimate real output growth for quarters far back in time. To avoid altogether the problems associated with having to select a data vintage to represent the 'true' growth of output over time, an alternative approach would be to assess the frequency and persistence of mismeasurement of output growth by comparing initial growth estimates in each quarter with those made (say) one year, three years and five years after the event. A fuller discussion of this issue, and of the results of this alternative approach, is provided in Appendix B.

steady improvement in the collection, processing and reconciliation of the data used to estimate GDP.[17,18]

Figure 3: Errors in the Contemporaneous Measure of GDP
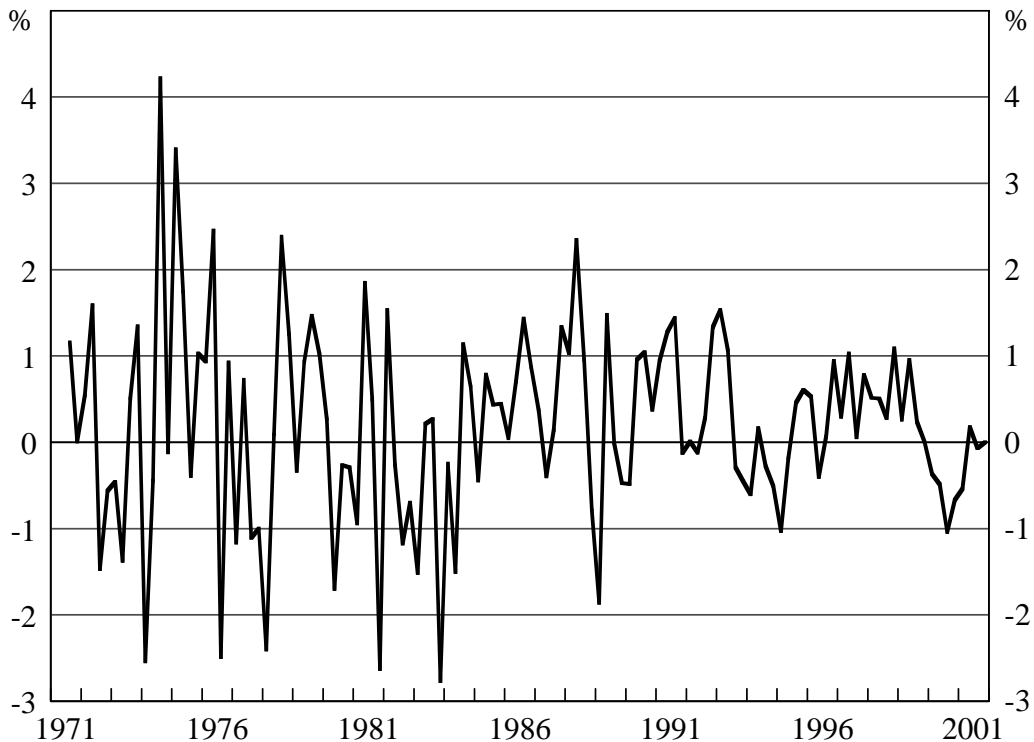
Quarterly growth rates
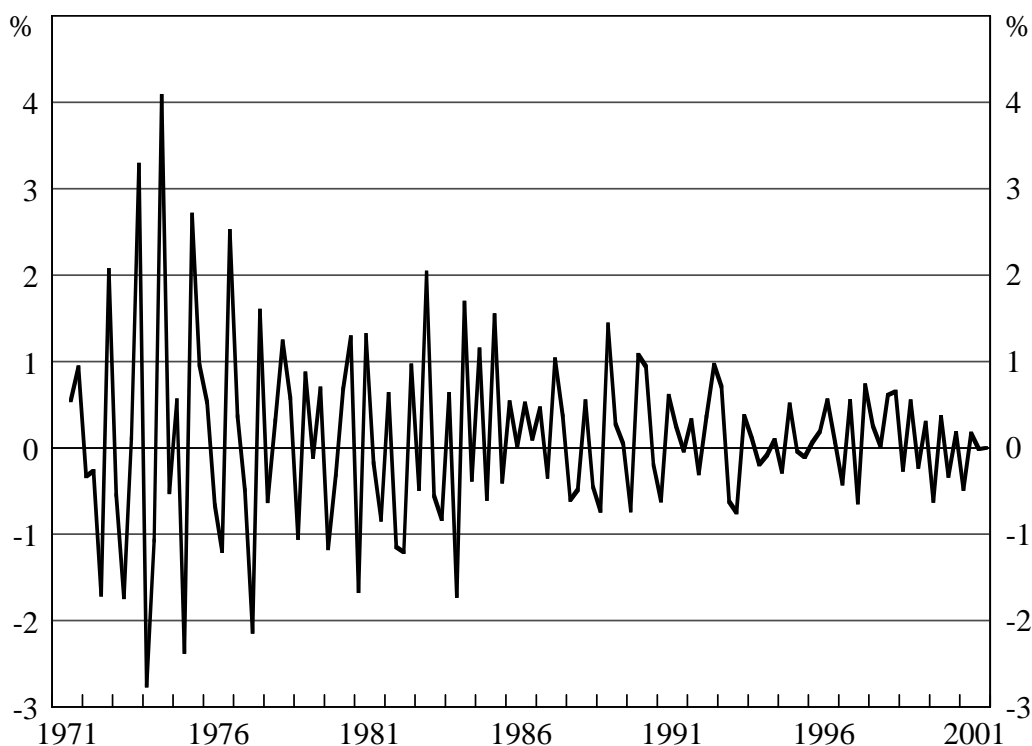
Note: Results shown are for our hybrid GDP measure (see Section 2.2).

While the errors in initial estimates of quarterly hybrid GDP growth have sometimes been very substantial, Figure 3 illustrates that there has been little bias in these initial estimates. Over the past three decades the average amount by which initial ABS estimates of quarterly growth are now reported to have underestimated actual quarterly growth is only 0.1 percentage points.

---

[17] A similar pattern is observed for the errors in the contemporaneous measurement of GDP(E) – see Appendix A. The generally smaller scale of errors in the contemporaneous measurement of quarterly GDP growth for both measures over the 1990s may also, in part, simply reflect the lesser time elapsed for changes to be made to these data than to earlier data.

[18] If we focus only upon the period from March quarter 1990 onwards the mean absolute size of the errors falls significantly, to a little under 0.4 percentage points.

Figure 4: Errors in the Contemporaneous Measure of GDP

Four-quarter-ended growth rates

Note: Results shown are for our hybrid GDP measure (see Section 2.2).

Again, the issue arises of whether it may be misleading to focus on quarterly growth rates, with the high degree of negative autocorrelation evident in the errors in Figure 3 suggesting that simple timing adjustments may account for much of the uncertainty surrounding such growth rates. Figure 4, however, illustrates that initial mismeasurements on a scale relevant to analysts still appear to have occurred regularly, even when four-quarter-ended GDP growth rates are examined. The difference between our best guess as to the 'true' value of four-quarter-ended hybrid GDP growth and its first published estimate has been more than 2 percentage points on 10 occasions (although all occurred prior to the 1990s), and the mean absolute size of the contemporaneous errors made in estimating such

growth is 0.9 percentage points over the full set of data vintages from 1971:Q3 to 2001:Q4.[19]

It is also of interest to examine whether output mismeasurement tends to be quickly identified and remedied, or instead tends to persist for a long time, only slowly dissipating. To investigate this issue, Figure 5 shows a panel graph of the difference, at each quarter, between the 'true' value of four-quarter-ended hybrid GDP growth to that quarter, and the estimates of that quantity provided contemporaneously (top panel – note that this replicates Figure 4), and two, four and eight quarters afterwards (bottom three panels).

Figure 5 shows that discrepancies between the 'true' and estimated values for four-quarter-ended hybrid GDP growth, on a scale significant to analysts, have frequently remained, even eight quarters after release of the initial estimate. While the errors in estimation of four-quarter-ended growth do, on average, tend slowly to decrease over time, the mean absolute size of these errors falls only from 0.9 percentage points for both the contemporaneous estimates and those made with a two-quarter lag, to 0.8 percentage points for those made with a four-quarter lag and just over 0.7 percentage points for those made eight quarters after the event. Thus, the mean absolute size of the errors made in measuring four-quarter-ended growth eight quarters after the event is only around 20 per cent smaller than the mean absolute size of the initial errors made.

---

[19] The bias in these errors is 0.2 percentage points (again towards initial underestimation of four-quarter-ended growth, on average). Over the period from 1990:Q1 onwards the degree of bias is little changed, but the mean absolute error falls to just under 0.6 percentage points.

Figure 5: Errors in Four-quarter-ended GDP Growth

As measured at various time horizons
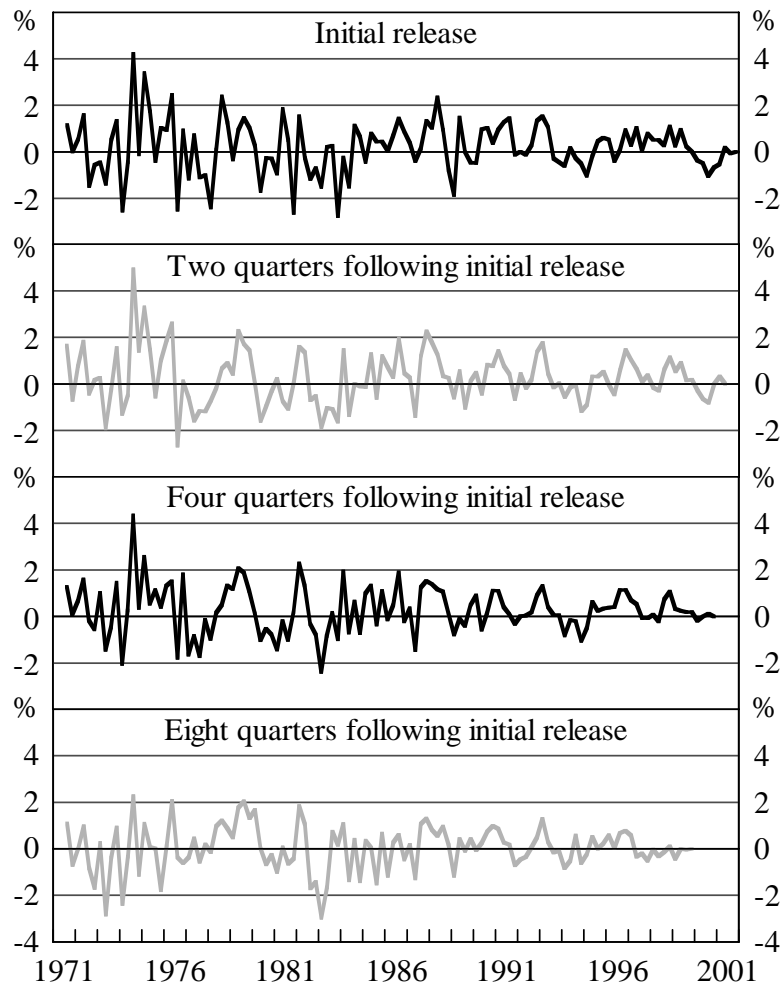
Note: Results shown are for our hybrid GDP measure (see Section 2.2).

4.   Conclusion

The difficulty of obtaining accurate estimates of quarterly national output within a few months is evident from the immense scale of the collection, analysis and reconciliation task involved in assembling these data. It is not surprising, therefore, that over the past three decades changes to output growth estimates on a scale significant to analysts have not been uncommon – although there is evidence that the degree of uncertainty surrounding initial output estimates may be lower now

than in the past.[20] Initial mismeasurement has also frequently been quite persistent, often largely remaining even several years after the period being measured.

These observations have an obvious cautionary implication for the operation of monetary policy. As Orphanides (2000) put it for the US case, they indicate 'the profound importance of appreciating the information problem for successful policy design'. Clearly, it is vital to remember that our uncertainty about the economy derives not merely from the difficulty of forecasting the future paths of key variables such as aggregate output, but also of knowing where those variables are now and where they have been in the recent past.

---

[20] This would be consistent with the significant improvements which have been incorporated in the compilation of the national accounts over the past decade. Most notably, these have included the switch to chain-linking in 1998 and the introduction of benchmarking of the accounts to annual supply-use tables.

Appendix A: Revisions to Expenditure-based GDP Estimates

This appendix replicates, for GDP(E), the analysis of our hybrid measure of GDP carried out in Section 3. In addition to switching our focus to GDP(E), one further change has been made in the treatment below, relative to that in Section 3. This is that, in this appendix, we use the June quarter 1998 national accounts data as giving our best measure of the true growth rates of GDP(E) over history, rather than the December quarter 2001 vintage of data used earlier for our analysis of hybrid GDP. We choose this quarter as it is the last vintage of GDP(E) data produced under the old SNA68 system of national accounting, prior to the

Figure A1: Changing Estimates of Quarterly GDP(E) Growth
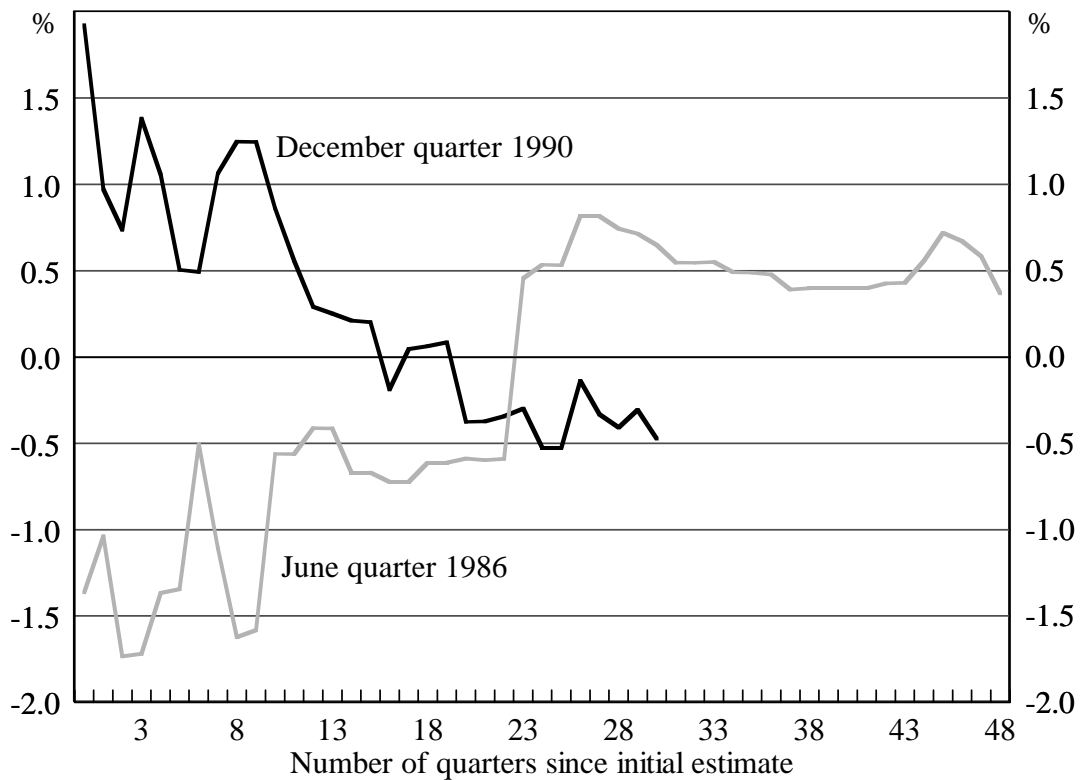
Note: This figure mirrors the analysis in Figure 1 in the main body of the paper.

changeover to SNA93. Using this vintage of GDP(E) data to represent the 'true' measure therefore incorporates the maximum possible set of revisions to each earlier quarter's estimates while still ensuring that comparisons between this and earlier data vintages involve a comparison of 'like with like'.[21]

**Figure A2: Changing Estimates of Four-quarter-ended GDP(E) Growth**
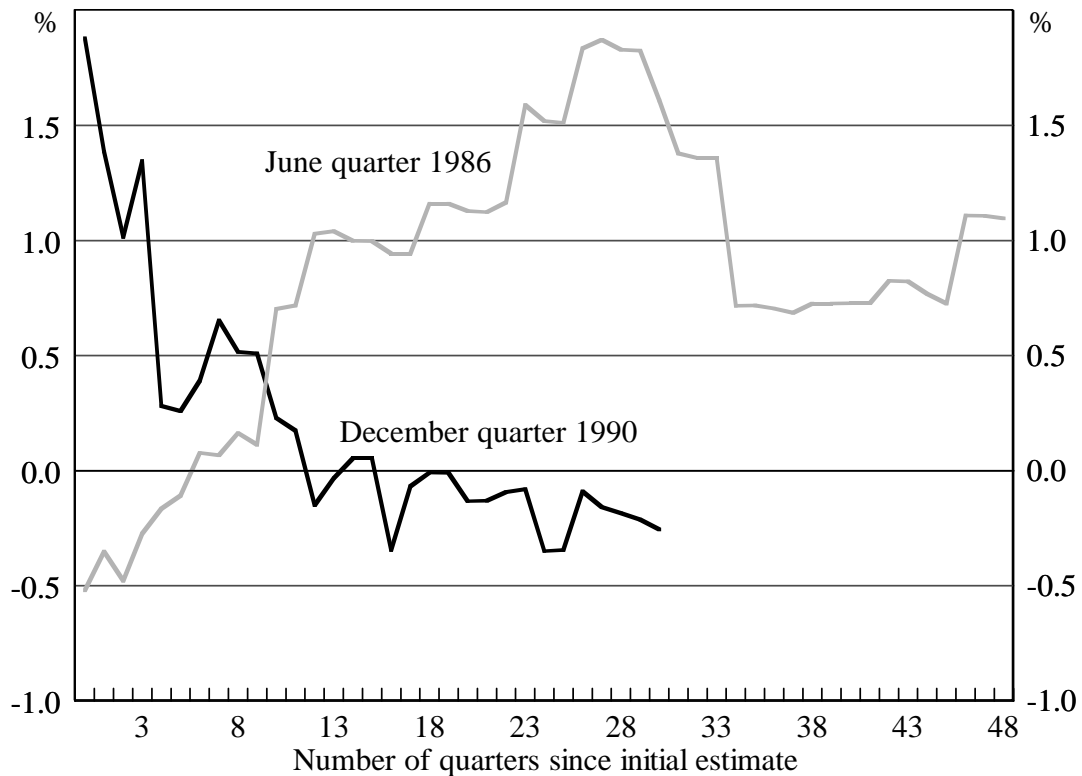
Note: This figure mirrors the analysis in Figure 2 in the main body of the paper.

---

[21] An alternative would be to use the GDP(E) series from the December quarter 2001 national accounts as giving the 'true' measure of GDP(E) growth. This would have the advantage of allowing for an additional 14 vintages worth of data revisions, but at the cost of no longer having all the GDP(E) vintages under comparison having been prepared under a common system of national accounting (see also the discussion in Appendix B). It transpires that adopting this alternative choice makes essentially no difference to the results described hereafter.

As Figures A1 through A5 illustrate, the results for the analysis using GDP(E) are similar to those obtained for our hybrid measure of GDP. This being the case, our general conclusions remain unchanged. Indeed, simple comparison of Figures A1 to A5 with Figures 1 to 5 confirms that the real-time measurement problem is, in fact, somewhat greater for GDP(E) than for hybrid GDP, consistent with the preference of analysts and the ABS for first GDP(I), and then GDP(A), as more reliable measures of output than GDP(E).

Figure A3: Errors in the Contemporaneous Measure of GDP(E)

Quarterly growth rates
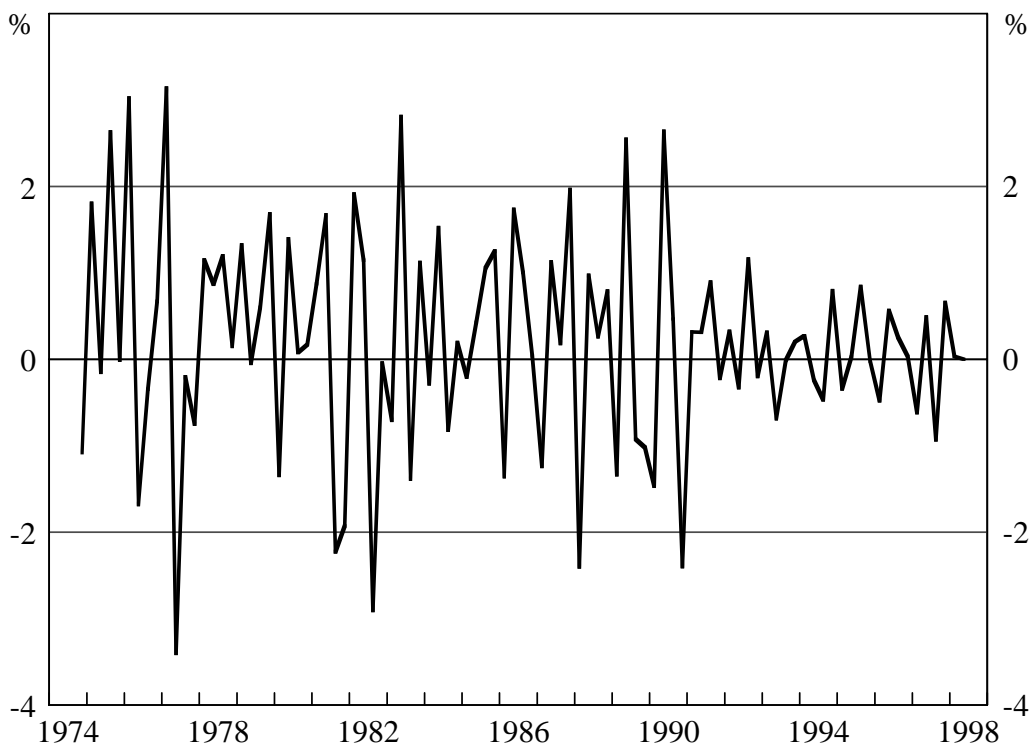
Note: This figure mirrors the analysis in Figure 3 in the main body of the paper.

Note that for Figures A1 and A2 we have chosen to display the successive vintages of data for different quarters from those used in Figures 1 and 2. The quarters selected, June quarter 1986 and December quarter 1990, have been chosen to illustrate the scale of changes to quarterly and four-quarter-ended GDP(E) growth rates not infrequently encountered over time. As Figures A3 and A4 demonstrate, quarters displaying more extreme changes than these two could once again have been selected.

Figure A4: Errors in the Contemporaneous Measure of GDP(E)

Four-quarter-ended growth rates
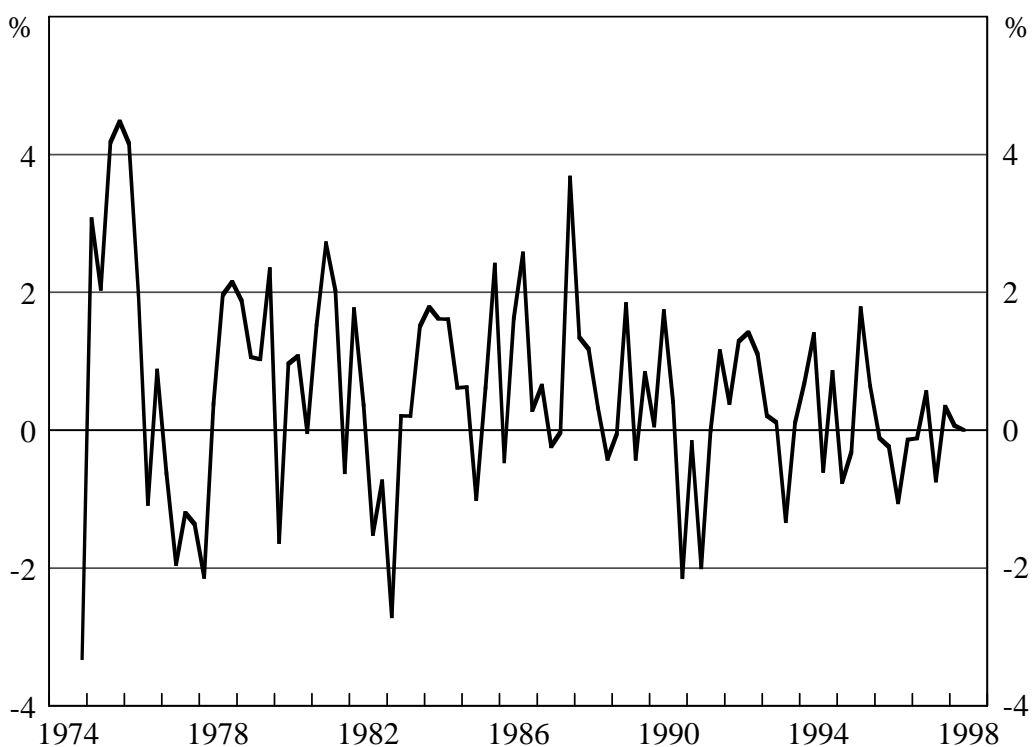
Note: This figure mirrors the analysis in Figure 4 in the main body of the paper.

Figure A5: Errors in Four-quarter-ended GDP(E) Growth

As measured at various time horizons
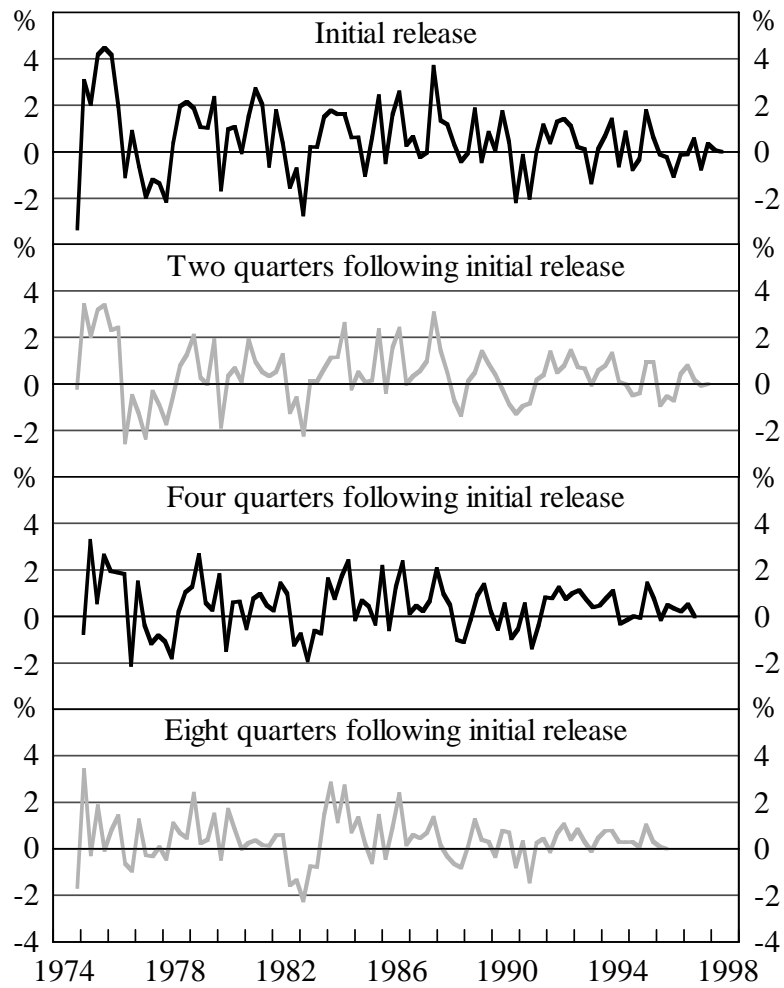
Note: This figure mirrors the analysis in Figure 5 in the main body of the paper.

Appendix B: An Alternative Approach to Assessing the Frequency and Persistence of Alterations to Output Growth Estimates

In Section 3.2 we examined the frequency and persistence of initial mismeasurement of output growth by comparing the initial estimates of quarterly or four-quarter-ended growth for each quarter in history with the estimate for that quarter contained in the latest available (2001:Q4) chain-volume GDP data. This assumes that this latest data vintage provides the best available measure of the 'true' growth of real output in each quarter over history. For periods far back in time, however, a technical issue arises as to whether this is an appropriate assumption.

This relates to the change made in 1998 to the national accounting system under which subsequent accounts have been prepared, and the difficulty of accurately adjusting estimates of real output for periods back in the 1960s, 1970s and 1980s to reflect these changes (so as to give a time series on a consistent accounting basis). To avoid this problem one could instead use the last set of accounts prepared under the old system of national accounting (SNA68), namely those from June quarter 1998, as the best available measure of the 'true' growth of output over history. However, this would raise a separate issue relating to the appropriateness of the price data used in those accounts for obtaining estimates of real output growth for quarters far back in time.

This latter issue relates to the fact that the constant price GDP estimates contained in the June quarter 1998 accounts were only directly computed using average 1989/90 prices (the then price base year) for the most recent 5–10 years. Prior to this, estimates based on earlier base years were then spliced on at 5–10 year intervals to give a continuous historical time series, with this splicing done at a disaggregated level (and therefore possibly affecting growth rates at the aggregate GDP level). Growth rates based on these estimates, for periods prior to the late 1980s, may thus possibly be a less accurate guide to the 'true' real growth occurring in the economy at that time than earlier constant price estimates prepared using the appropriate earlier price base year.

To avoid altogether these various problems associated with having to select a recent data vintage to represent the 'true' growth of output over time, a different approach would be to assess the frequency and persistence of mismeasurement of output growth by comparing initial growth estimates in each quarter with those made (say) one year, three years and five years after the event.[22] The latter horizon broadly reflects the period over which formal revisions are typically made to GDP estimates – by which we mean revisions relating to the incorporation of improved underlying data (for example, resulting from new information about the timing of particular items of activity, from more accurate taxation data which only becomes available with a lag of several years, or from updated census data with which to adjust relevant GDP components, available at five-year intervals).

To determine whether this alternative approach to assessing the scale and persistence of changes to output growth estimates alters the broad picture described in Section 3.2, Figure B1 shows the extent to which the initial estimates of four-quarter-ended growth in our hybrid GDP measure have, over history, been amended one year, three years, and five years after the fact.

With regard to the persistence of initial mismeasurement, changes to the assessment of four-quarter-ended growth made in the first year following the period under consideration have, on average, been somewhat larger than those made over the next two years. These, in turn, have typically been somewhat greater than those made between three and five years after the event. The differences, however, are not dramatic.

---

[22] Note that this alternative approach does not aim to abstract completely from adjustments related to constant price base year changes, when examining the alterations to output growth estimates for quarters far back in history. For example, consider estimates of four-quarter-ended growth to the March quarter 1985. For our hybrid GDP measure the initial estimate for this quarter is of four-quarter-ended growth in GDP(I), measured on a constant price basis using average 1979/80 prices. Three years later, in the March quarter 1988 accounts, the estimates of growth for this period shift to using average 1984/85 prices, which we would expect would give a more accurate indication of actual real growth in the year to March quarter 1985 than estimates based on 1979/80 prices. In assessing the scale of the real-time problem facing analysts in early 1985, we thus do not wish to abstract from the change made in the March quarter 1988 accounts to the constant price base year being used. Rather, we only wish to abstract from subsequent further price base year changes (e.g. to average 1989/90 prices in the December quarter 1992 accounts), which might arguably have made the estimates for output growth in the year to March quarter 1985 less, rather than more, accurate.

Figure B1: Errors in Four-quarter-ended GDP Growth

Alternative methodology
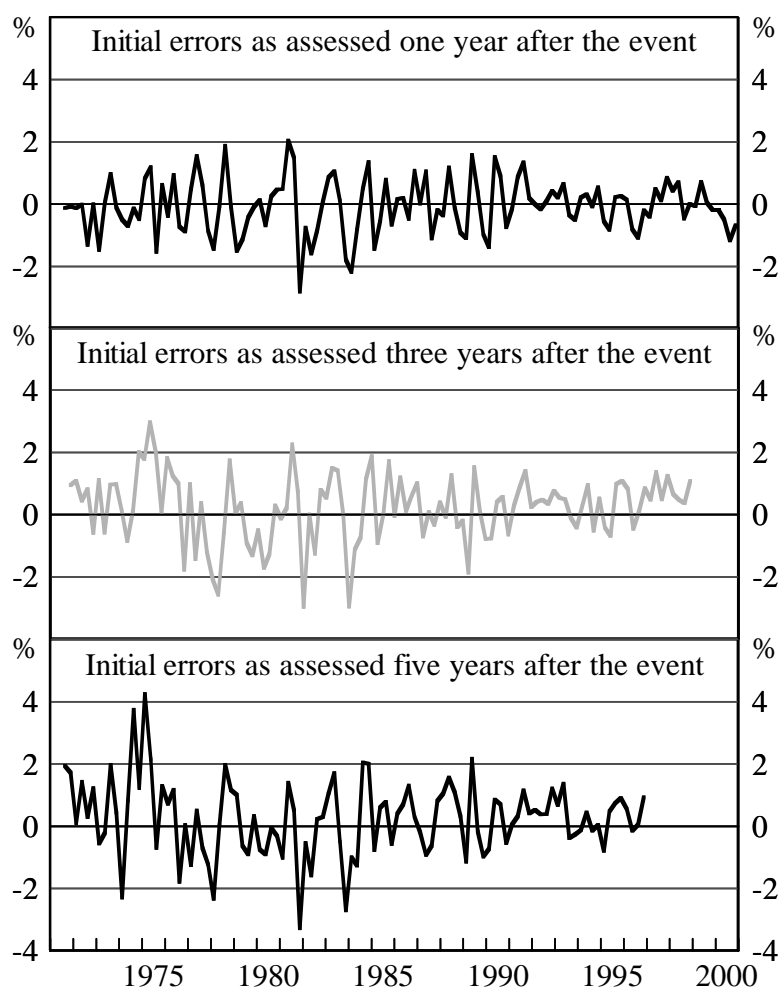
Overall, the results of this alternative analysis confirm the general conclusions found earlier. There appears to be sufficient uncertainty surrounding initial estimates of GDP growth that analysts need to keep this in mind when treating such estimates as a barometer of the current state of the economy. At the same time, there does not appear to be any bias in these estimates, and there is evidence that, with the improvements in national accounting methodology and data collection/reconciliation over the past three decades, the scale of the real-time problem may be smaller now than it has been in the past.

References

Aspden C (1990), 'Which is the Best Short-Term Measure of Gross Domestic Product? A Statistical Analysis of the Short-Term Movements of the Three Measures of Gross Domestic Product and Their Average', Australian Bureau of Statistics Cat No 5206.0 (Australian National Accounts: National Income and Expenditure), June quarter, pp 57–65.

Clarida R, J Gali and M Gertler (1998), 'Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory', NBER Working Paper No 6442.

Croushore D and T Stark (2000), 'A Real-Time Data Set for Macroeconomists: Does Data Vintage Matter for Forecasting?', Federal Reserve Bank of Philadelphia Working Paper No 00-6.

Gray B (1983), 'Revisions to National Accounts', Reserve Bank of Australia, mimeo.

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.

Gunther J and R Moore (2000), 'Early Warning Models in Real Time', Federal Reserve Bank of Dallas, Financial Industry Studies Working Paper No 1-00.

Judd J and G Rudebusch (1998), 'Taylor's Rule and the Fed: 1970–1997', Federal Reserve Bank of San Francisco *Economic Review*, 3, pp 3–16.

Orphanides A (2000), 'The Quest for Prosperity Without Inflation', European Central Bank Working Paper No 15.

Rossiter P and N Brown (1989), 'The Analysis and Modelling of Revisions to the Australian National Accounts: Some Preliminary Results', paper presented to the Economic Society of Australia Annual Conference, Adelaide, 10–13 July.

Taylor JB (1993), 'Discretion versus Policy Rules in Practice', *Carnegie-Rochester Conference Series on Public Policy*, 39, pp 195–214.

Taylor JB (1998), 'An Historical Analysis of Monetary Policy Rules', NBER Working Paper No 6768.

Taylor JB (2000), 'Comments on Athanasios Orphanides' The Quest for Prosperity Without Inflation', PowerPoint presentation slides, available at <http://www.stanford.edu/~johntayl/PptLectures>.