

Better than a Coin Toss? The Thankless Task of Economic Forecasting¹

Address by Mr GR Stevens, Deputy Governor, to the Economic Society of Victoria and the Australian Industry Group 'Economic Focus – Australia's Prospects', Melbourne, 17 August 2004.

One of my best forecasts was made by accident. In mid 1999, at one of the regular Parliamentary hearings we have each six months, I was asked about the prospects for the rate of unemployment, which at the time had been fluctuating around 7¹/₂ per cent for several months.² My answer was that unemployment would fall to 'the low sevens, 7 per cent, something like that' by the end of that year. A decline of up to half a percentage point in half a year was, I thought, a reasonably bold forecast; I didn't think it was likely to go below 7 per cent at that stage. But the Hansard reporters recorded my words as 'below 7', not 'the low sevens'. Hence I was recorded as in effect saying that the unemployment rate would decline quite quickly, and before long have a six before the decimal point for the first time in about a decade. I was worried at the time that this seemed much too bold, but the unemployment rate did indeed fall below 7 per cent around the end of 1999. Hence I am happy to have that forecast on the record, even though I didn't actually intend to make

it. Perhaps that says that chance plays as big a role in forecasting as it seems to in many other areas of life.

These days I am more of a consumer of economic forecasts than a producer of them, and while I suppose a former forecaster never entirely loses interest in the forecasting process, it is in the capacity of user that I have come here today. Hence I do not propose to make any forecasts – there are presumably more than enough to choose from as a result of the conference. I will rather offer a few observations about the general processes of forming and using forecasts.

Making economic forecasts remains an occupational necessity, but something of a chore, for many economists including those giving policy advice. For those receiving advice and charged with the responsibility of helping to make decisions, key issues remain deciding how much to stake on a particular view of the outlook, and how to think about the consequences of the forecast and associated policy being wrong. For both producers and users of forecasts, it is also worth looking back at forecast errors – not to berate the forecasters, but rather to see what we can learn from those errors about the way the economy works.

1. Jonathan Kearns provided invaluable assistance in preparation for this speech.

2. Because of data revisions, the unemployment rate for May 1999 is today recorded at 7.0 per cent. But this was originally published as 7.5 per cent.

Evidence on Accuracy

It has long been understood that economic forecasts are not all that good. Most elements of the round-up that I gave five years ago³ still seem apposite.

First, forecasts *are* better than a coin toss – that is, an economic forecast can more often than not be expected to outperform a random process or some very simple extrapolative rule – though often not by all that much. This is not true, however, for some financial variables, where the economics profession's forecasting embarrassment is greatest.

There is some evidence that, in Australia, forecasts improved in the past decade. Table 1 shows that the average absolute error of one-year-ahead forecasts for both GDP

growth and CPI inflation in *The Age* Survey from 1994 to 2003 declined to just over half what it had been in the preceding 10 years. Of course, that period has been one of much reduced volatility in the economy, a fact that has been noted before.⁴ So maybe it was just easier to make forecasts in that period, and the real test will come when the economy enters rougher waters.

A crude way of assessing this would be to see whether a comparison of *The Age* forecasts with those from a naïve forecast rule – that the future value is the same as the current one – revealed an improvement. The second column gives the forecast errors for such a naïve rule. The Theil statistic in the final column shows the ratio of the two errors. 'Good' forecasts have a value less than unity – that is, the forecasters add value in the sense of lowering forecast errors compared with the naïve rule.

Table 1: Forecast Errors^(a)
Percentage points

	Australian CPI inflation			Australian GDP growth			
	Absolute forecast error	Naïve rule error	Theil statistic	Absolute forecast error	Naïve rule error	Theil statistic	
1984–1993	1.77	2.61	0.68	1984–1993	1.81	2.68	0.67
1994–2003	0.92	1.63	0.56	1994–2003	0.88	1.11	0.79

Source: *The Age* Survey

	US CPI inflation			US GDP growth			
	Absolute forecast error	Naïve rule error	Theil statistic	Absolute forecast error	Naïve rule error	Theil statistic	
1984–1993	0.67	0.82	0.82	1984–1993	1.12	2.05	0.55
1994–2003	0.60	0.62	0.96	1994–2003	1.35	1.47	0.92

Source: *Survey of Professional Forecasters*

(a) Mean absolute forecast errors for one-year-ahead forecasts.

3. Stevens (1999).

4. My thumbnail sketch is in a paper with David Gruen in Gruen and Stevens (2000). A more sophisticated analysis is Simon (2001).

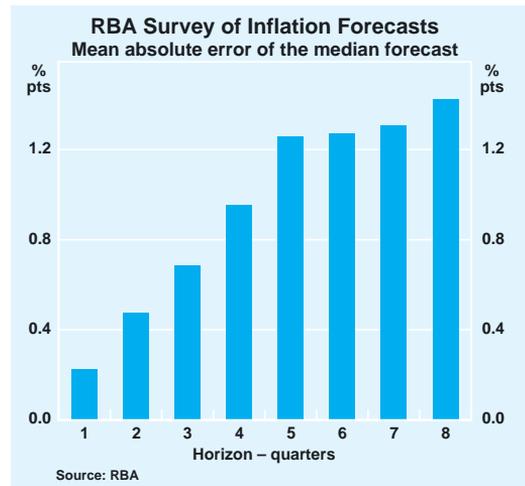
While *The Age* panel's performance improves a lot in the past decade, so does that of the naïve forecast rule. The Theil statistic suggests that the forecasters were adding some value in both periods, but with no major changes between the two.

So we shouldn't get too carried away by lower forecast errors in recent years. To give Australian forecasters their due, however, arguments that a more stable economy is easier to forecast presumably apply just as much to the US economy (at least for most of the period), but the evidence from the US *Survey of Professional Forecasters* is that there was no absolute improvement in forecast accuracy over the past decade. The extent of value added by the forecasters, as measured by this simple test, actually declined.

A second key finding is that the accuracy of forecasts tends to decline somewhat as the forecasting horizon lengthens. For most countries, the accuracy of inflation forecasts is superior to that of growth forecasts at short horizons. I conjecture that this reflects the facts that inflation has a fair bit of inertia. GDP growth, on the other hand, has much less inertial behaviour and its measurement is probably subject to more sampling error. Hence, forecasts one or two quarters ahead for inflation are pretty good compared with growth forecasts. But over longer horizons where inertia weakens, this advantage for the inflation forecasters diminishes.

The RBA has been compiling a survey of private forecasts of inflation for about a decade. Participants – which include some of you here – are asked for a forecast of the CPI over a one- and two-year horizon. We have eight years of data from about 15 forecasters which enables us to make some observations about the way accuracy diminishes with horizon. As expected, the mean absolute error of the year-ended forecasts increases quickly out to about a year (Graph 1). In part this is mechanical as the number of quarters actually being forecast rises from one to four, but the

Graph 1



underlying quarterly errors probably get a bit larger too. But between a five-quarter and an eight-quarter horizon, there isn't that much loss of accuracy. So those comfortable with a horizon of just over a year shouldn't have too much trouble accepting a two-year forecast. That said, the confidence interval for this set of forecasts is still fairly wide.

Third, extreme movements are rarely well-forecast. Late in 2000, Consensus forecasts for US GDP growth in 2001 were about 3½ per cent. But as we now know, the US economy experienced a recession in 2001, and recorded year-average expansion in GDP of about 1 per cent. This was not well-anticipated by the forecasters. Recessions seldom are.⁵ Other major events – like financial crises – have likewise usually not been well-predicted even though in most cases, with hindsight, several warning signs can be seen to have been flashing.

Fourth, structural shifts which are not business cycle events, but which have profound implications for the course of the economy over the medium term, are not well-forecast either and often are not even recognised for some time after their emergence. The rise in US productivity growth in the 1990s is a case in point. (So

5. In fact, given the amount of noise in statistics, combined with their publication lag, it is quite common that professional observers cannot detect (from the figures anyway) that the economy is even in recession until the contraction has been going on for some time. The NBER recession dating committee, for example, did not declare the March 2001 peak in the US economy until November that year.

was the slowing in productivity growth in the mid 1970s.) The permanent downshift in Australia's inflation rate in the 1990s is another.

Finally, it is hard to find evidence that any one forecaster is consistently superior. Indeed, one study of Australian forecasters suggests that outstanding performance in one year has a high likelihood of being followed by very poor performance in the next.⁶ Most studies find that averaging a panel of forecasters will give a better consistent forecast than using any individual forecaster.

These points are all familiar, and leave all of us with much about which to be modest. But I don't wish to denigrate forecasters. The effort to make a forecast, to articulate it, and to describe how and why it might be wrong has some value. We just need to keep in mind that numerical forecasts are not much more than opinion formed (hopefully) within a coherent and disciplined framework. They are not guarantees of performance, and should always be accompanied by a discussion of risks. That discussion is likely to be at least as useful as the point estimates themselves.

Use of Models versus Judgement

Let me turn now to some questions to do with the formation of forecasts. One of the perennial ones is the respective roles of formal models and subjective judgement.

It seems to me obvious that we need both. Any judgemental forecast embodies some notion of how the economy works, unless the numbers really are drawn from a dart board. Most forecasters make some effort to ensure their forecasts for different variables are consistent with each other, and tell some sort of story that can be related to presumed behaviour. That is to say, they have a model of sorts, even if a fairly informal one.

Econometric models are a more formal way of representing the relationships in the historical data. It is usually helpful occasionally to confront the notions in our heads with the data to see if there is any validation for our prejudices.

That said, formal models come, or should come, with various usage warnings. To begin with, there seems to be some evidence that simple models often are more useful than more complex ones, perhaps because they are more robust and so less likely to come unstuck due to structural change, etc. Because their workings are more transparent, users may also be able to spot problems more easily when they start to break down. Simplicity, of course, has to be traded off against the general principle that the economy has many complex interactions, which simple models can miss. But, in general, complex is not always better, especially for short-term forecasting.⁷

Second, some modelling techniques which are thought to be best practice for describing history may not be optimal for forecasting purposes. A case in point is the use of cointegration models, where the deviation from an estimated long-run equilibrium level can be a powerful factor affecting short-run forecasts of changes, as the model wants to move the dependent variable towards the supposed long-run equilibrium. But if there has been a level shift in the equilibrium relationship, such a forecast will be highly misleading, and probably less accurate than a forecast from a model in differences, even though the latter is often considered theoretically less pure.⁸

Another warning is that many models which are in use today are not directly estimated from the data. Many are 'calibrated' – which is to say that certain properties such as means and variances are standardised against the actual data – but that is not the same as testing hypotheses embodied in the model against the data. These models often have very strong theoretical properties, to which most of us

6. Norman (2001).

7. See Hendry and Clements (2003).

8. For a more detailed discussion of this sort of issue in forecasting using models, see Sims (2002) and Hendry and Clements (2003).

would sign up as general propositions, but which can drive the behaviour of the model over the horizon relevant for business cycle forecasting to a substantial degree. Such models have their place, particularly for long-term simulation exercises. But, in my opinion, their use for short-term forecasting in a policy context needs particular care.

In the end, we will probably get the most useful forecasts by combining stable, simple models that capture empirically the most important macroeconomic dynamics in the economy, with judgement informed by the vast array of non-model, and sometimes non-quantitative, information about the current state of the economy which is available in the plethora of partial indicators (not all of which are published by the official statistical agencies).

A finding in the US literature is that the Federal Reserve staff forecasts compiled in the Green Book outperform both private forecasts, particularly for inflation, and pure model forecasts (Romer and Romer 2000). Sims (2002), reviewing this evidence and confirming the finding, attributes a good deal of the performance improvement to the effort to get a more accurate estimate of the current state of the economy, so reducing the errors in the very early period of the forecast. In other words, the judgement of specialist data watchers, combined with insights of well-understood models of both the formal and informal variety, works better than any single technique. I find this a plausible conclusion.

Learning

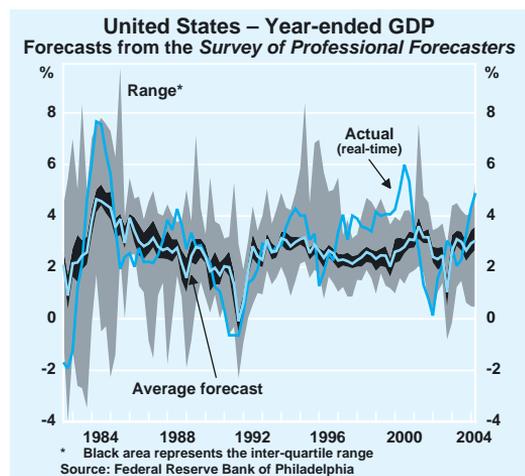
To this we can add that learning is crucial, which is to say that observing the pattern of forecast errors and seeking to draw conclusions for our 'model' of the economy and therefore its future behaviour can, hopefully, improve future forecasts. This is a rather Bayesian idea: we can't know what the economy's parameters are, and we should not

view them as set in concrete anyway – they are subject to variation. One starts with some priors about what these parameters are, i.e. how the economy works. These priors are then confronted by a data sample, and the result is the posterior distribution of parameter values – that is, better-informed guesses about the way things work. As time goes by and new data become available, this working hypothesis of the economy's properties and likely future behaviour is updated. While I could not say that we implement this idea rigorously quarter by quarter in practice, I find it quite appealing as a way of conceptualising both the forecasting of the economy and the conduct of policy.

Let us then look at two examples of forecast errors, and see what they teach us. The first chart is for year-ahead forecasts of US GDP growth, from the *Survey of Professional Forecasters* (Graph 2). The shaded area is the range of forecasts, with the inter-quartile range – the middle 50 per cent of forecasts – in a darker colour. The average forecast is the line in the middle of this range. The forecasts are plotted forward by one year, so that they can be compared with actual year-ended growth, the dark blue line in the graph.

For several years in the second half of the 1990s, virtually all forecasters persistently underestimated the pace of US growth. Time does not permit a detailed decomposition of

Graph 2



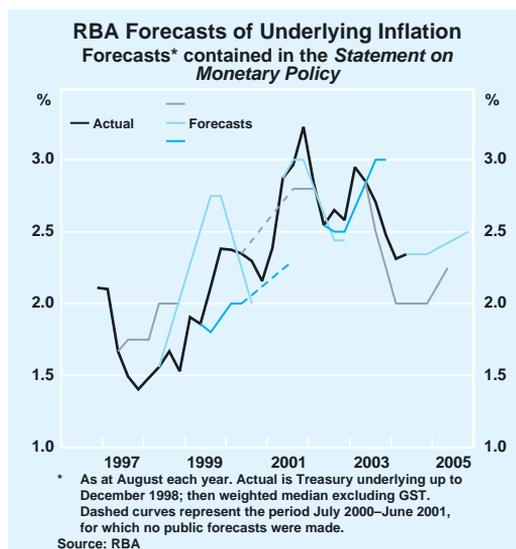
the errors into their various causal components. Suffice to say that, as is now well-known, productivity growth in the US picked up, and so therefore did the US economy's potential growth rate, at least for a period of several years. If we looked as well at forecasts of US inflation, we would find that unexpectedly high growth was not generally accompanied by unexpectedly high inflation. So strong demand growth was met with rapid supply growth. In other words, a high-level treatment of the forecasting errors points us to the productivity story. (Of course, having understood that did not make forecasters much better at predicting the 2001 recession. The 'new economy' was as prone to cyclical setbacks as the old.)

The second chart shows some forecasts closer to home: those enunciated by the RBA in the *Statement on Monetary Policy* in the late 1990s (Graph 3). The chart shows underlying inflation as measured either by the Treasury underlying series or by the median CPI change (since 1998).⁹ Starting from the middle of each year, it shows the RBA's outlook as set out in the *Statement on Monetary Policy* (or the previous quarterly article on 'The economy and financial markets' before the *Statements* became quarterly) which appeared in August.

Several points are of interest. First, through 1997 and 1998, inflation was below the 2–3 per cent target, but was expected to rise over the ensuing period. The rise in inflation did eventually occur, but took longer than originally expected. What was going on here?

One important feature of the behaviour of inflation in the late 1990s and the early part of this decade was that changes in the exchange rate had less effect over a one- to two-year horizon than previous experience

Graph 3



had suggested. We began to detect this as time went by, and accordingly lowered our estimates of the short- and medium-term impact of exchange rate changes on the CPI.

For 2000 and 2001, forecasts tended to be a little on the low side. We believed that inflation was generally tending to increase and this turned out to be right, but the trend was ultimately a bit stronger than forecasters expected.¹⁰

The downward move in inflation since the peak has proceeded in two phases. The forecasts seemed to have had errors on both sides during that period. The most recent forecasts, as you know, suggest that inflation will remain about where it is at present for a few quarters and then move up during 2005, as the effects of the earlier rise in the exchange rate gradually wane. These forecasts are a little higher than ones made earlier in the year partly because the exchange rate is not as high as it was then.¹¹

9. The inflation target is, of course, for the CPI (since 1998). But the main forecasting approach is to forecast a measure of underlying inflation, then add known or assumed 'special factors', mainly oil prices or tax changes, to get the CPI forecast. So for the purposes here, it is most useful to consider the forecasts for underlying inflation.
10. Forecasting was very difficult indeed in the period around the time of the GST, when the price level was due to show a substantial, but once only, rise over several quarters, with the exact quarterly profile highly uncertain. We made no public forecast of a time path through the period from 1 July 2000 to June 2001 (hence the dashed curves), but made forecasts of where inflation would settle thereafter.
11. In fact, differences in the exchange rate outcome from what was assumed are often a significant contributor to forecast errors. A full treatment of forecast performance, for which there is not space here, would need to take that into careful account.

So looking back over this period of seven years we find that, early on, our inflation forecasts tended to persistently be a bit high (a pattern observed for a few years prior to 1997 as well). Lessening the expected impact of exchange rate changes seems to have helped to improve the forecasts. There appears, though, to have been some residual tendency to underestimate how far inflation would go in a new direction once it had turned.

I don't have a graph which characterises as neatly the forecast errors on growth, but I think it is well-known that the Australian economy has over the same period surprised on the upside more often than on the downside. What have we learned from that? One lesson is that the economy's improved inherent flexibility has helped it cope with shocks which in previous times would probably have derailed growth.

Another conclusion, at least on my own part, is that the structural change in household balance sheets which has been under way since about 1995 has consistently been an expansionary factor. Let me mount one of my hobby horses for a moment here. Reverting back to the discussion of models of the economy, informal or formal, this points to an important gap in knowledge. Conventional models of the macro economy are long on detail about demand, output gaps, inflation and so on, but relatively underdeveloped in the financial sphere. But with agents facing fewer and fewer capital market imperfections, the relationships between asset price changes and balance sheet adjustments seem to be of increasing importance to the course of the economy over time. This is where more attention needs to be focused by modellers and forecasters, as well as policy-makers.

Forecasts and Decision-making

Turning to the role of forecasts in making decisions, the task of preparing some sort of forecast is one that, while thankless,

nonetheless must be performed. Decisions based on looking in the rear-view mirror are unlikely to be optimal; trying to look forward, as difficult as that is, should help to achieve better outcomes. This is particularly the case when the decision is one, as in monetary policy, whose effects take a long time to show up, but I think this point generalises to decisions in, say, investment management.

Decision-makers will therefore want to have not just a set of numbers but a sensible story about the future. Regardless of how a forecast was arrived at – from a formal model, judgement, some combination – it is more useful to decision-makers if its main features, and the factors driving them, can be fairly simply explained. The plausibility test of a view is most effectively applied when the story is kept as straightforward as possible (though not more so). A forecast is even more useful if the forecaster can identify key assumptions, be they about exogenous variables or the structure of the economy, and what the consequences would probably be were those assumptions to be astray.

We can, of course, look to the statistical properties of econometric models to give some indication of the size of likely confidence intervals around a central forecast, and that is a reasonable place to start a discussion of uncertainty and risks. Consumers of forecasts should routinely be told something about the size of past errors.

But anyone looking to make a decision on the basis of a set of forecasts – like an investment manager, or an economic policy-maker – looks for some judgement about possibilities to which history may not be a good guide. What if the underlying structure of the economy is changing – i.e. the model parameters are shifting? Suppose, for example, the responsiveness of inflation to exchange rate changes is less than it used to be? What if the economy is more responsive than it used to be to recent changes in interest rates? For that matter, what if it is less responsive? If an outcome is different from the central forecast for some key variable, is it more likely to be higher or lower?

Decision-makers, in other words, are interested in forming a judgement about the *balance of risks*. That judgement can only be subjective, to be sure – but that is where the experience of a good forecaster is most valuable. The decision-maker will then want to combine that subjective assessment of the risks with some sense of the relative costs and benefits of the possible outcomes, and decide which risks they should most avoid, and which they are prepared to run.

This much is, I hope, well understood these days: the question asked of forecasters shouldn't just be 'what's the number?' It should also be 'how could you be wrong?' Hopefully you will probe some of these sorts of questions later today.

Forecasters also need to cast their minds a bit further afield than just the next quarter or two. Perhaps this is where many private forecasters, with a horizon driven by the short-term demands of the financial markets, part company with official forecasters who do not, and should not, share that imperative to focus so heavily on the next figure. The conduct of policy needs to adopt a medium-term focus, and possibly more so than in the past. The economy's heightened short-term stability and flexibility quite possibly delays the consequences of inappropriate policy settings. But it surely does not eliminate them. Some of the recent issues with which we have had to deal, moreover, like the run-up in housing prices and credit, play out over a longer time horizon than the typical short-term forecast covers. So we could easily have been lulled by a reassuring one-year forecast into ignoring problems that would be likely to build up beyond that horizon.

Realistically, forecasters cannot be expected confidently to predict how some of these longer-term dynamics will play out. But there

is little doubt that such things are becoming more important. Forecasters will be more useful the more they can help decision-makers to think through what *might* happen, even if they cannot say with precision what *will* happen.

Conclusion

One of the old forecasters' clichés is to say that, on any particular occasion, there is more than the usual degree of uncertainty. Most forecasting meetings I recall seemed to start that way – even in periods which, we now know with hindsight, turned out to have been remarkably stable by historical standards. Most of the time, such comments are surely an exaggeration. It might be more correct to say that at any one time there are undoubtedly new sources of uncertainty, which people find hard to quantify because they have no historical comparison to go by. In this sense, forecasting has grown no easier despite the advances in statistical and analytical technology over the years.

Given the above, it would seem that it is very important to keep in mind the idea that any forecast is a working hypothesis, based on an understanding of the economy's properties which is likely to continue evolving. Further, a simple focus on a single central forecast alone is unlikely to be as useful as an approach that contemplates one or more alternative possibilities, and helps decision-makers think through their implications.

I wish you all the ideal combination of insight and luck in your forthcoming deliberations.

References

Gruen D and GR Stevens (2000), 'Australian macroeconomic performance and policies in the 1990s', in D Gruen and S Shrestha (eds), *The Australian economy in the 1990s*, Proceedings of a Conference, Reserve Bank of Australia, Sydney, pp 32–72 (also available at <<http://www.rba.gov.au/PublicationsAndResearch/Conferences/2000/GruenStevens.pdf>>).

Hendry DF and MP Clements (2003), 'Economic forecasting: some lessons from recent research', *Economic Modelling*, 20(2), pp 301–329.

Norman NR (2001), 'Measuring the accuracy and value of Australian macroeconomic forecasts since 1990', media release, 16 November.

Richardson D (2001) 'Official economic forecasts – How good are they?', Parliament of Australia, Department of the Parliamentary Library, Current Issues Brief 17.

Romer CD and DH Romer (2000), 'Federal Reserve information and the behavior of interest rates', *American Economic Review*, 90(3), pp 429–457.

Simon J (2001), 'The decline in Australian output volatility', Reserve Bank of Australia Research Discussion Paper No 2001-01 (also available at <<http://www.rba.gov.au/PublicationsAndResearch/RDP/RDP2001-01.html>>).

Sims CA (2002), 'The role of models and probabilities in the monetary policy process', *Brookings Papers on Economic Activity*, 2, pp 1–61.

Stevens GR (1999), 'Economic forecasting and its role in making monetary policy', Reserve Bank of Australia *Bulletin*, September, pp 1–9 (also available at <http://www.rba.gov.au/PublicationsAndResearch/Bulletin/bu_sep99/bu_0999_1.pdf>). ✎