

THE PREDICTIVE POWER OF POLITICAL PUNDITS: PRESCIENT OR PITIFUL?

Abstract

*Although Australian political pundits frequently make predictions about the future, little systematic evidence exists about the accuracy of these predictions. To assess the predictive power of experts, we survey the transcripts of two well-known political programs – *Insiders* and *Meet the Press* – and record all falsifiable forecasts. Looking at the three months prior to both the 2007 and 2010 federal elections, we are struck by the paucity of falsifiable predictions, with most pundits heavily qualifying their predictions (so that they can never be said to be wrong). In 32 hours of television, we identify 20 falsifiable forecasts in our sample, of which we judge thirteen to be correct. We conclude with some suggestions for political talk shows, and for political scientists seeking to better analyse expert predictions.*

How accurate are pundits at predicting the future? A series of studies analysing US political analysts has found that their forecasts are frequently no more accurate than the toss of a coin (e.g. see Gardner, 2010; Silver, 2012; Tetlock, 2005). As one reviewer summed up the evidence: ‘Human beings who spend their lives studying the state of the world, in other words, are poorer forecasters than dart-throwing monkeys.’ (Menand, 2005, summarising Tetlock, 2005) To date, most of the research on the accuracy of political forecasts has focused on the United States. However, there are reasons to believe that Australian political analysts might be more accurate. Australians’ reputation for directness might lead one to expect more frequent predictions, and a reputation for cynicism might lead one to expect more accurate predictions.¹ So it seems reasonable to explore whether Australia’s political pundits perform any better than their US counterparts.

In this article, we present new empirical results on predictions made by political pundits appearing on two of Australia’s most prominent Sunday morning political shows – *Insiders* on ABC and *Meet the Press* on Channel 10.² We searched for falsifiable predictions (that is, those that have the potential to be proven wrong) in the three-month period before the 2007 and 2010 federal elections. To preview our results, we were struck by the paucity of such predictions, which suggests that Australian pundits are hesitant to make predictions that could definitively be proven wrong. In the two leading political talk shows that we analysed, less than 0.2 per cent of airtime was used to make falsifiable predictions. We identified 20 falsifiable predictions in our sample, and judged thirteen to be correct – an accuracy rate of 65 per cent.

In the next section of the article, we review the relevant literature, followed by a discussion of our methodology. We then present results on the scope and accuracy of

predictions from Sunday morning political talk shows before discussing the implications of our findings, and speculating as to why on-air falsifiable predictions are so rare. The article concludes with some suggestions for researchers and political pundits.

Past evidence on future predictions

In the field of prediction accuracy, the seminal study is Tetlock (2005), whose two-decade-long study of expert political judgement assessed 27,450 judgements about the future from 284 experts (who were guaranteed anonymity) in fields such as political science, economics and journalism. Participants were surveyed privately, and asked to assign probabilities to future outcomes (usually three possible futures), which were then compared with actual results to determine scores for what Tetlock terms ‘calibration’ and ‘discrimination’. Calibration refers to an expert’s ability to assign probabilities: ‘Do events they assign 0.10 or 0.50 or 0.90 probabilities materialise roughly 10 per cent or 50 per cent or 90 per cent of the time?’ A discrimination score requires an expert to ‘assign larger probabilities to things that subsequently happen than to things that do not’.

These expert predictions were then compared with various benchmarks, including predictions of dilettantes, extrapolation algorithms and the assumption that all possible outcomes should be given equal weight (which Tetlock calls the ‘chimp’ approach). Experts out-performed the hypothetical chimpanzee on *discrimination* (unsurprising, given that there is no more undiscriminating strategy than to assign an equal probability to each outcome). However, they under-performed the chimp on *calibration*, exhibiting larger forecast errors. When scores for calibration and discrimination were combined, humans emerged only slightly ahead of the chimp approach.

Although Tetlock’s results paint an unflattering picture for the general predictive power of experts, his key finding was that some experts did better than others. He concluded that ‘what people think’ (such as their political ideology) is far less important for their judgement than ‘how people think’. Using the terminology of philosopher Isaiah Berlin, Tetlock drew a distinction between ‘hedgehogs’ and ‘foxes’: ‘intellectually aggressive hedgehogs knew one big thing and sought, under the banner of parsimony, to expand the explanatory power of that big thing to “cover” new cases; the more eclectic foxes knew many little things and were content to improvise ad hoc solutions to keep pace with a rapidly changing world’. In distinguishing between different groups of experts, Tetlock’s findings were clear: foxes consistently out-performed hedgehogs on measures of political judgement.³

For the most part, the remainder of the literature on prediction looks at forecasts made in the mass media. Silver (2012) looks at panellists on the US television show *The McLaughlin Group*, and finds that exactly equal numbers of political predictions were correct and incorrect. Gardner (2010) provides numerous examples of failed expert predictions in areas such as economics, politics and information technology. Yet, although many expert predictions are barely better than a coin toss, consumers of predictions keep going back for more – people still tune in to hear which party will win the next election, and corporations continue to engage forecasting firms even though their previous predictions were inaccurate. Gardner attributes this behaviour to our psychological craving for certainty in an uncertain world. Humans do not like the idea of events being random, and we want people who can peer into the future to tell us what will happen. In this market for predictions, it pays for purveyors of predictions to be confident. In Gardner’s words, ‘confidence convinces’. Our attraction to certainty in an uncertain world and our preference for confident experts expose us to a paradox that both Tetlock and Gardner lament: the experts most likely to be sought after by the media for predictions are those most likely to get it wrong – hedgehogs.

Other researchers have reached similar conclusions. Surowiecki (2004) notes that physicians, nurses, lawyers, engineers, entrepreneurs and investment bankers are over-confident about their predictions (see also Odean, 1998; Speirs-Bridge et al., 2010). Eighty-two per cent of young US drivers put themselves in the top 30 per cent of their group (Svenson, 1981). Similarly, new business owners think their business is twice as likely to succeed than the typical business of its kind (Cooper et al., 1988).

Economic forecasters have been found to do a poor job of identifying turning points (Blake et al., 1986; Pasaogullari and Meyer, 2011; Sherden, 1998). Indeed, Berg and Pattillo (2000) tested 'early warning system' models created before the Asian financial crisis of 1997–99, and found that more than 50 per cent of predicted crises did not eventuate. As Paul Samuelson famously put it: 'the stock market has called nine out of the last five recessions'. Conversely, forecasters in a December 2007 *Wall Street Journal* poll put the chance of recession at only 38 per cent; as the data would subsequently show, the US economy was already in recession at that time (Silver, 2012). Haldane (2012) points out that the traditional model of forecasting bank failure (risk-based capital ratios) was a poorer predictor than the simpler measure of leverage ratios. Among experts, weather forecasters are one of the few groups whose predictions are accurate: it rains on 30 per cent of the days when forecasters predict a 30 per cent chance of showers (Silver, 2012; Surowiecki, 2004).

In the psychology literature, Kahneman and Klein (2009) conclude that the experience of an expert is a poor guide to the accuracy of their predictions. As Kahneman (2011: 221) puts it: 'The question is not whether these experts are well trained. It is whether their world is predictable.' Soll and Larrick (2009) point out that the average of two experts' predictions is typically more accurate than either taken individually, yet we tend to prefer to choose between predictions rather than average them. In economic forecasting, a substantial literature exists on the optimal approaches to pooling forecasts (e.g. see Clemen, 1989; Hendry and Clements, 2004).

Yet, despite the large international literature on forecasting, relatively few Australian studies have investigated the forecast accuracy of our experts. An exception to this is Leigh and Wolfers (2006), who in their survey of election forecasting methods refer to a survey of ten experts published six days before the 2004 federal election. In that survey, three thought Mark Latham would win, while seven thought John Howard would win, but with a smaller majority than in 2001. None forecast the true result – a Howard victory with an increased majority.

Methodology

In choosing which predictions to include, we take as our starting point the notion of falsifiability. This has its origins in Popper (1959), who argued that the notion of 'falsification' was essential to scientific progress. For a prediction to be falsifiable, there must be some scientific evidence that could prove it wrong. For example, Popper pointed out that the statement 'All swans are white' is falsifiable by the discovery of a black swan, while the statement 'White swans exist' is not falsifiable.

In the cast of political pundits, we searched for falsifiable predictions made by pundits on the state of the world at some time in the future. In an election context, such predictions can be simple – for example, Labor will win the 2010 election but with a reduced majority. The predictions made also needed to be testable in a contemporary setting. For example, a hypothetical prediction made in 2007 that the Australian Labor Party would win three consecutive federal elections is falsifiable, but the answer is unlikely to be available to us until the end of 2013. Predictions that equivocated on all the various scenarios that could result from an election campaign were not falsifiable, and we do not use them in our analysis.

We define a pundit as any person involved in an interview or discussion on *Insiders* or *Meet the Press* who is not a politician (excluding the ‘Talking Pictures’ segment on *Insiders*). In the case of *Insiders*, pundits consisted of press gallery journalists such as Malcolm Farr and Lenore Taylor, and commentators such as Gerard Henderson and Andrew Bolt. In the case of *Meet the Press*, while several journalists do appear on the show, they are there primarily as fellow interviewers. As such, the pundits in our sample included strategists Bruce Hawker (Labor aligned) and Ian Kortlang (Coalition aligned), as well as pollsters Andrew Catsaras and David Briggs.

In order to test the accuracy of the collective *Insiders* and *Meet the Press* punditry, we scoured transcripts in search of predictions in the three-month period prior to the 2007 and 2010 federal elections. This included sixteen *Insiders* episodes in 2010 and thirteen *Insiders* episodes in 2007. The extra episodes in 2010 are attributable to the period of negotiation between the independents and the leaders of both major parties that resulted from the hung parliament. Transcripts for *Meet the Press* during the 2007 election are not available (Paul Bongiorno informs us that they were lost in a website transition), so our data only include episodes during the 2010 election campaign. There were fifteen *Meet the Press* episodes in total, but only six contained appearances by political pundits (in their role as experts rather than interviewers).

Table 1: Identities of pundits in pre-election programs

Panellist	Fox or hedgehog?*	Appearances as pundit	
		Insiders	Meet the Press
Paul Kelly**	Hedgehog	13	0
Andrew Bolt	Hedgehog	10	0
Lenore Taylor	Fox	9	0
Misha Schubert	Fox	7	0
Malcolm Farr	Fox	6	0
Karen Middleton	Fox	6	0
David Marr	Hedgehog	6	0
Piers Akerman	Hedgehog	6	0
Brian Toohey	Hedgehog	5	0
George Megalogenis	Hedgehog	5	0
Fran Kelly	Fox	5	0
Glenn Milne	Hedgehog	5	0
Gerard Henderson	Hedgehog	4	0
Dennis Atkins	Fox	3	0
Annabel Crabb	Fox	3	0
Laura Tingle	Fox	2	0
Kerry-Anne Walsh	Fox	2	0
Phillip Coorey	Fox	1	0
Michael Stutchbury	Hedgehog	1	0
Jim Middleton	Fox	1	0
Bruce Hawker	Hedgehog	0	3
Ian Kortlang	Hedgehog	0	2
Andrew Catsaras	Fox	0	1
David Briggs	Fox	0	1

* Categorisation of foxes and hedgehogs is subjective (see text for definitions).

** Paul Kelly appeared in all thirteen *Insiders* programs prior to the 2007 election as a correspondent. He was not part of the traditional ‘panel’.

Source: Authors’ analysis of 2007 and 2010 transcripts for *Insiders* and *Meet the Press*.

Table 1 shows a combined list of *Insiders* and *Meet the Press* punditry during the period. Note that Fran Kelly, Malcolm Farr and Glenn Milne also appeared on *Meet the Press* during this period, but not as pundits. Following Tetlock (2005), we classified Australian pundits as foxes (who draw on many eclectic traditions of thought, seek opinions that contradict their own and are self-critical when proven wrong) or hedgehogs (who know one big thing, prefer to express problems in simple terms, express strong confidence in their forecasting abilities and rarely apologise for erroneous forecasts). Naturally, such a classification is subjective, and reflects our personal opinions only. For example, as *Insiders* has three panellists on each program, tallying up the appearances by pundits with Paul Kelly's appearances as a correspondent gives a cumulative total of 100 effective (panellists appeared on multiple occasions) appearances over 29 episodes. For *Meet the Press*, there were seven effective appearances by pundits over six episodes (there were fifteen *Meet the Press* episodes during the three-month period before the 2010 election, but only six contained pundits). With *Insiders* running for approximately 60 minutes and *Meet the Press* for approximately 30 minutes, this equates to approximately 1920 minutes (32 hours) of footage across the two shows. While it is true that, in both shows, not all time is spent talking to pundits (there are interviews with politicians and other segments), it is nonetheless a substantial amount of data.

Assessing prediction accuracy

The number of falsifiable predictions was, in a word, underwhelming. After scouring hundreds of pages of transcripts (and contending with the imperfections of closed caption-based text) we were able to find 20 falsifiable predictions in approximately 32 hours of footage on *Insiders* and *Meet the Press*. The breakdown was fifteen predictions in 29 hours of *Insiders* and five predictions in six hours of *Meet the Press*. This equates to one prediction every 96 minutes across the two programs. The fifteen predictions had an average length of approximately three seconds, and a cumulative length of 45 seconds. Put another way, only 0.04 per cent of the available time on *Insiders* was assigned to falsifiable predictions. For *Meet the Press*, 0.14 per cent of available time was allocated to falsifiable predictions.

Table 2 shows, in reverse chronological order, every falsifiable prediction that we could find in the sample data. Each prediction has been labelled either right or wrong, and has been assigned a subjective difficulty grading (easy, medium or hard), which again reflects the personal opinions of the authors.

Of the fifteen *Insiders* predictions, nine of them occurred prior to the 2010 election, and six were made prior to the 2007 election.⁴ Four of the nine 2010 predictions were related to the timing of the federal election; two of the six *Insiders* predictions before the 2007 election were based on the election date. Thus, of the fifteen predictions made in 29 episodes, 40 per cent of them related to the timing of the 2007 and 2010 elections. Other prediction topics in 2010 included the mining tax, Kevin Rudd's leadership of the ALP and Tony Crook's party preference (who he would support following the hung parliament). In 2007, predictions were made around the future use of polling, John Howard's position as prime minister (the likelihood of a pre-election challenge from Peter Costello) and the likely winner of the federal election. *Meet the Press* offered no falsifiable predictions from journalists, so we had to rely on predictions made by strategists Bruce Hawker and Ian Kortlang, as well as pollster Andrew Catsaras. In six hours of footage, this yielded five predictions – hardly an abundant supply, but a higher strike rate than *Insiders*.

Of the 20 predictions used in our sample, we judged thirteen to be correct. These correct predictions differed in complexity and time: some would take months to be proved correct, whereas others were provable within a matter of days. Given the

Table 2: Falsifiable predictions on *Insiders* and *Meet the Press*

Date	Program	Pundit/s	Prediction	Correct?	Difficulty?
5-Sep-2010	<i>Insiders</i>	Fran Kelly and Glenn Milne	Tony Crook will support the Coalition	Right	Easy
22-Aug-2010	<i>Insiders</i>	Andrew Bolt	Kevin Rudd is finished in the ALP	Wrong	Medium
22-Aug-2010	<i>Meet the Press</i>	Ian Kortlang	A minority Liberal government is the more likely scenario	Wrong	Medium
15-Aug-2010	<i>Meet the Press</i>	Andrew Catsaras	ALP to win the election with 76-78 seats	Wrong	Hard
18-Jul-2010	<i>Insiders</i>	Dennis Atkins	Kevin Rudd will recontest Griffith	Right	Easy
4-Jul-2010	<i>Insiders</i>	Phillip Coorey	2010 election date – 21 August 2010	Right	Hard
4-Jul-2010	<i>Insiders</i>	Malcolm Farr	2010 election date – late August	Right	Medium
4-Jul-2010	<i>Insiders</i>	Andrew Bolt	2010 election date – 28 August 2010	Wrong	Hard
20-Jun-2010	<i>Insiders</i>	George Megalogenis	2010 election date – August or September	Right	Medium
13-Jun-2010	<i>Insiders</i>	David Marr	Government can turn public opinion on mining tax around	Right	Medium
13-Jun-2010	<i>Insiders</i>	Lenore Taylor	Government can turn public opinion on mining tax around if concessions are made	Right	Medium
6-Jun-2010	<i>Meet the Press</i>	Bruce Hawker	Greens vote to reduce from 16 per cent with movement back to ALP on election day	Right	Medium
6-June-2010	<i>Meet the Press</i>	Bruce Hawker	Lindsay Tanner and Anthony Albanese to both hold their seats	Wrong	Medium
6-Jun-2010	<i>Meet the Press</i>	Bruce Hawker	Kevin Rudd to lead ALP to 2010 election	Wrong	Medium
30-Sep-2007	<i>Insiders</i>	Paul Kelly	Howard to move soon for election in November	Right	Easy
23-Sep-2007	<i>Insiders</i>	Gerard Henderson	We will focus less on opinion polls in three years' time	Wrong	Medium
16-Sep-2007	<i>Insiders</i>	Fran Kelly	Election to be called before next sitting period	Right	Easy
9-Sep-2007	<i>Insiders</i>	Paul Kelly	John Howard to lead Coalition to 2007 election	Right	Medium
9-Sep-2007	<i>Insiders</i>	Andrew Bolt	Costello will either be PM now or never. He won't hang around	Right	Hard
2-Sep-2007	<i>Insiders</i>	Piers Akerman	Labor is likely to win the election	Right	Easy

Source: Authors' analysis of 2007 and 2010 transcripts for *Insiders* and *Meet the Press*.

distinct difference in punditry across the two shows (*Insiders* primarily made up of journalists and *Meet the Press* made up of strategists and pollsters), the accuracy of the predictions will be detailed separately.

As mentioned previously, what was striking about *Insiders* projections was the prevalence of predictions based on the timing of the election. Six of the fifteen predictions were directly related to when Prime Minister Howard in 2007, and Prime Minister Gillard in 2010, would call the next federal election. The pundits were quite successful in predicting the election time in 2010: Phillip Coorey correctly predicted the exact date of the election (21 August 2010) on the *Insiders* program on 4 July 2010; Malcolm Farr, a fellow panellist, was less specific but anticipated an election date in ‘late August’; and Andrew Bolt on the same program said ‘August 28 is probably my call’. On a previous *Insiders* episode on 20 June 2010, George Megalogenis – in response to a question on election timing – said ‘I would be thinking August or September’.

In the lead-up to the 2007 election, Fran Kelly said – in reference to when John Howard would call the election – during the episode that aired on 16 September 2007: ‘I think he will call it perhaps before the next sitting period.’ At that stage the next sitting was scheduled to begin on 15 October 2007. Mr Howard called the election a day earlier. On 30 September 2007, journalist Paul Kelly said ‘I would anticipate Howard would move fairly soon for an election in November.’ The election date was 24 November 2007.

Of the six election-timing predictions above, we judged five to be correct. Clearly, all five predictions are not the same: predicting an election within a given month is far less specific than a particular date. Also, much like weather forecasting, one would expect a better prediction closer to the eventual date (Gardner, 2010: 245). So while Phillip Coorey made a perfect prediction about the timing of the election, Andrew Bolt’s incorrect answer of 28 August 2010 is a victim of its own specificity. These factors have been taken into consideration when difficulty ratings have been assigned to each prediction.

Other correct predictions were Piers Akerman’s suggestion on 2 September 2007 that Labor would win the next election (qualified with Jim Middleton’s assumption that everything would not go perfectly for the Coalition in the next five weeks); Paul Kelly’s prediction on 9 September 2007 that John Howard would lead the Coalition to the 2007 election (there were rumours that treasurer and deputy leader Peter Costello would challenge); Dennis Atkins’ prediction on 18 July 2010 that Kevin Rudd would recontest the seat of Griffith in the 2010 poll; and the joint prediction by Glenn Milne and Fran Kelly on 5 September 2010 that Tony Crook would support the Coalition rather than the incumbent Labor government. Mr Crook declared his support for the Coalition two days later.

The above correct predictions were all made over a short timeframe. *Insiders* provided us with three longer-term projections that have proved to be correct. In urging Peter Costello to challenge John Howard for the leadership of the Coalition, Andrew Bolt said on *Insiders* on 9 September 2007 that ‘Peter Costello is either going to be PM before the election or never’. He went on to suggest that Peter Costello would not ‘hang around’ in the event of a Coalition defeat. Mr Costello declined the opposition leadership after Labor won the 2007 election and retired from politics in 2009. The other two correct long-term predictions were made by David Marr and Lenore Taylor on 13 June 2010. This program went to air in the midst of negotiations between the Labor government and the mining industry over the proposed Resource Super Profits Tax (RSPT). David Marr made the prediction that, in time, public support for the tax

would improve: 'I think time could be on the government's side. I think, with time, more people will see that this tax is not the demon tax of the mining companies' representation.' Lenore Taylor also offered a similar sentiment: 'My view is that the government can turn the debate around, but only if they make concessions on the tax.' At the time of writing, public support for the Minerals Resource Rent Tax was strong.

While incorrect predictions do provide a temptation to embarrass the offending pundits, it is important to state that our intention is not to 'show up' or 'catch out' political pundits. Aside from the already mentioned incorrect prediction on election timing from Andrew Bolt, there were two other incorrect predictions on *Insiders*. On 23 September 2007, Gerard Henderson said: 'Whatever the outcome of the election, I suspect next time around in three years' time there will be less focus on opinion polls.' While a statement like this is difficult to prove, the subsequent election in 2010 was widely derided for being overly focused on opinion polls at the expense of public policy. The other incorrect *Insiders* prediction was made by Andrew Bolt on 22 August 2010 (the day after the 2010 federal election) on the topic of Kevin Rudd's future: 'I know it's crazy, but you mentioned Kevin Rudd, I think he's finished in the ALP.' Kevin Rudd was later appointed Foreign Minister after Julia Gillard formed a minority government on 7 September 2010.

Meet the Press provided five falsifiable predictions, of which one proved to be correct and four incorrect. The sole correct prediction was made by Bruce Hawker on 6 June 2010 after The Greens were polling with a primary vote of 16 per cent in the most recent Newspoll. He said: 'I would be surprised if it stayed all the way to polling day ... I think there will be a movement back to Labor.' The Greens went on to record a primary vote of 12 per cent at the 2010 election, and Labor's vote improved to 38 per cent on election day, from 35 per cent in the Newspoll survey. Hawker made two incorrect predictions during the same episode: that both Lindsay Tanner and Anthony Albanese would hold their seats, which were under pressure from The Greens (Tanner retired from politics, but Labor did lose his former seat of Melbourne at the general election; Albanese won); and that Kevin Rudd would lead the Labor Party to the 2010 election (he was replaced as leader by Julia Gillard on 24 June 2010). After prompting from host Paul Bongiorno, Coalition strategist Ian Kortlang predicted on 22 August 2010 (the day after the federal election) that a minority Liberal government was more likely. Andrew Catsaras was the only pundit to offer a falsifiable prediction on the number of seats to be won: he suggested on 15 August 2010 that 'Labor is most likely to win the election' with 76–78 seats. His next most likely scenario was a hung parliament, the eventual result.

Discussion

What is striking about the predictions in the data is the lack of predictions based around policy. Among the key issues in 2007 were climate change, industrial relations and tax policy. In 2010, climate change policy was again prominent, as well as broadband, immigration and the mining tax. Yet the only falsifiable predictions made about policy were those by David Marr and Lenore Taylor, both on the topic of the proposed mining tax in June 2010. Instead, predictions tended to be based more around politics rather than policy. Indeed, the timing of the election, which accounts for 40 per cent of all *Insiders* predictions, seems at face value to be a relatively trivial issue. Strategic arguments can be made about why the timing of the election might matter, but whether or not the 2010 election was held on 21 August or 28 August seems insignificant.

A recent example of this type of analysis was apparent after Prime Minister Gillard announced in a policy speech at the National Press Club on 30 January 2013 that the federal election would be held on 14 September 2013. This announcement dominated news coverage in subsequent days, even though there had been a high likelihood that the election would ultimately take place in August or September anyway. This seemingly disproportionate interest in election dates may be evidence of a general tendency of pundits and the media to focus on the political ‘horse race’ instead of policy detail. During the 2010 election campaign, for example, more substantive predictions could have centred on which policies taken to the election by both parties would survive afterwards, or whether or not key policies would eventually pass both houses of parliament after the election.

In the case of *Meet the Press*, the strategists (Hawker and Kortlang) only made falsifiable predictions that were sympathetic to their party’s cause.⁵ One would expect Mr Hawker to predict that Labor’s vote would recover, that Lindsay Tanner and Anthony Albanese would hold their seats, and that the Labor leader at the time, Kevin Rudd, would go on to lead the party at the next election. Statements to the contrary would have been far more newsworthy. Similarly, Ian Kortlang may have greatly annoyed Coalition parliamentarians if he had predicted that the Labor Party was likely to form a minority government. Four predictions from strategists cannot be used to garner any meaningful conclusions regarding predictive accuracy, but to the extent that strategists do offer predictions based purely on partisanship, it is questionable why their predictions are sought at all.⁶ These partisan commentators are the ultimate hedgehogs: no matter the context, they almost invariably predict the outcome beneficial to their political party. If we ourselves can predict what they will say, how useful are their predictions? In a small sample such as ours, the paucity of data means that definitive conclusions should not be drawn. That is not to say that impressive predictions were not made: Phillip Coorey’s precise prediction of the 2010 federal election date was notable, given that there were then at least a dozen possible election dates. Andrew Bolt’s prediction about Peter Costello’s political future was also surprisingly prescient, given the conventional wisdom at the time that Costello would take on the Liberal Party leadership in the event of a Coalition defeat. His announcement the day after the 2007 election that he was not seeking the party leadership was viewed as a shock. Yet we have too few predictions per pundit to crown anyone the nation’s best predictor. Similarly, it is unsurprising that we see no clear evidence in our data for the predictive superiority of foxes or hedgehogs.

One of the most noteworthy things we observed in the data was the general unwillingness of leading pundits to make falsifiable predictions. Rather than making predictions that gave conclusive right/wrong responses, pundits tended to hedge their statements, using phrases such as ‘maybe’, ‘chance’, ‘probably’, ‘potentially’, ‘I don’t know’, ‘could be’, ‘my guess’, ‘who knows?’, ‘not sure’ and ‘wait and see’.

Pundits’ timidity could be a symptom of their experience – perhaps being all too aware of old political adages such as ‘expect the unexpected’.⁷ A case in point was the ongoing analysis provided by Paul Kelly on *Insiders* during the 2007 election campaign. On 9 September 2007, Kelly correctly predicted that John Howard would lead the Coalition to the 2007 election. A week later, he reaffirmed his belief that Howard would stay, but with a qualification:

John Howard did have the opportunity to go, he’s declined that, he stared down the party. Howard imposed his will on the party. He’s utterly convinced he’s the best person to lead the government into the election and I think that it’s very

hard, it's very difficult to blast out an incumbent prime minister. History tells us this. Having said all that, you can't rule out ... you can't rule out the possibility that something might happen, this is a highly volatile situation.

A statement like this offers the perfect hedge: Howard is likely to lead the Coalition, but you can't rule out the possibility that something might happen.

Kelly hedged at other times during the campaign. On 14 October 2007 (acknowledging that the election would be called that day), Kelly said: 'The polls are very much against John Howard and the government ... Labor and Kevin Rudd are the favourites. They still need to win 16 seats. That will not be easy given John Howard's record, his experience and his tenacity.' And with six days to go before the election, on 18 November 2007, Kelly remained non-committal: 'At this stage, one has to say that Kevin Rudd must be favourite to win this election. I don't think the contest is over. Labor needs to get a very substantial national swing to win, it needs that swing in the right place. Labor can still get 51 per cent of the vote and lose.' The above statements by Paul Kelly indicate his sentiment, but there is no falsifiable prediction. Every statement contains a hedge: you can't rule out something happening with the Liberal Party leadership; Labor is the favourite but you can't write off Howard; the contest isn't over and Labor needs a big swing.

As one of Australia's most experienced political commentators, Kelly is probably more aware than most of the role that luck plays in politics.⁸ Other pundits were similarly unwilling to make falsifiable predictions. In response to a question about the Labor leadership on 13 June 2010, Lenore Taylor – appearing on *Insiders* – agreed with a news article being discussed suggesting nobody was organising a leadership challenge and that Julia Gillard didn't want it, before saying: 'I think that's right, but whether it remains right is another question.' Another example of vagueness was Dennis Atkins' comments on *Insiders* on 18 July 2010: 'It's a bit early to talk about wild seats but I'll throw one in. The seat of Canning in Western Australia, a 4.4 per cent Liberal seat but the Labor Party very cleverly pre-selected a former state minister (Alannah MacTiernan) who at the last state election got a swing to her. Her state seat is within the seat of Canning. Watch Canning.' What Atkins meant by 'watch Canning' is open to interpretation. Did he mean MacTiernan would win Canning? Or did he merely mean that MacTiernan would outperform her party in Western Australia? The result was that MacTiernan lost the seat but there was a swing towards her of 2.62 per cent, compared with a general swing against Labor in Western Australia of 3.15 per cent. Was Atkins right or wrong? The vagueness of the prediction means we cannot say.

Why, then, were falsifiable predictions so infrequent in the data? Perhaps our pundits are aware that even though they possess more knowledge of political issues than laypeople, luck plays a major role. Another possibility is that blogs and social media may dissuade pundits from making decisions for fear of being mocked in the Twittersphere and blogosphere. Or do our pundits – many of whom value their objectivity – feel that falsifiable predictions may be misconstrued as partisan support? That is, will a prediction suggesting a Labor election win, or a high primary vote for The Greens, be interpreted as a promotion? All of these issues may be at work, but what was noticeable throughout the data was that Australia's political pundits tended to behave more like foxes than hedgehogs. It would be easy to look at Paul Kelly's hedges as a simple case of fence sitting, but Kelly is an experienced journalist who knows that politics is complex and uncertain, and that there are no golden rules which can be followed to predict election outcomes. Nevertheless, while you can make a

defensive case for the timidity of Australia's political pundits, the question of how accurate Australia's political pundits are at predicting the future remains unresolved.

To answer that definitively, we need more data. A nice start would be for *Insiders* host Barrie Cassidy to insist on 'falsifiable predictions' in his final segment, rather than the current 'observations or predictions' from the pundits. While such an innovation would be helpful, assessing whether a prediction has proven true or false is likely to provoke objections from the pundits involved. As Gardner (2010) notes, when predictions go awry, pundits often invoke an array of excuses to exonerate themselves. Such excuses include the 'I was almost right' defence, and the 'wait and see' defence. In our analysis, had David Marr and Lenore Taylor been incorrect about public opinion changing on the mining tax, they could easily have invoked the 'wait and see' option to defend themselves.

For this reason, we advocate a Tetlock-style experiment involving Australia's leading political pundits. Such an experiment substantially removes the ability of pundits to invoke the common defence mechanisms when their predictions go astray. An appropriate study would require our pundits to make falsifiable predictions about the political world in the future. Predictions could cover a range of topics such as the party likely to win the next federal election, the next leaders of each party, and the passage of key legislation. There would be a sufficient number of questions to ensure that lucky hits do not skew the results. We need more evidence than our analysis could provide to properly separate flukes from prescience. Furthermore, an anonymous survey like Tetlock's would help determine whether our pundits as a group are worth the airtime they are afforded. An open survey could also give us the ability to rank our political pundits based on their predictive powers.

Conclusion

Kahneman (2011) tells the story of his time as a psychologist in the Israeli Army, where he was assigned to evaluate candidates for officer training based on their performance in a team exercise. As he tells the story, he and his fellow evaluators usually had little difficulty agreeing on whether a candidate would 'make it' as an officer. Then, every few months, they had a feedback session in which they could compare their assessments with those of the commanders who had been monitoring the officers. The feedback session showed that their ability to forecast ability was negligible – not much better than blind guesses. Yet the team continued to assess new candidates as before, unshaken in the confidence of its predictive accuracy. Kahneman then writes (2011: 211):

What happened was remarkable. The global evidence of our previous failure should have shaken our confidence in our judgments of the candidates, but it did not. It should also have caused us to moderate our predictions, but it did not. We knew as a general fact that our predictions were little better than random guesses, but we continued to think and act as if each of our specific predictions was valid. I was reminded of the Müller-Lyer illusion, in which we know the lines are of equal length yet still see them as being different.

Given the inaccuracy of predictions, the optimal situation would probably be to have more discussion about the past and present, and less prognostication about the future. But if we must have expert predictions, we make the following appeal to Australia's political pundits. If you want to play the prediction game, why not subject yourselves to an experiment to determine who among you is the most prescient pundit in the country?

To those pundits who are coy about the idea of being ranked, consider this: should you confound the conventional wisdom and out-perform expectations, you will have every opportunity to gloat. And if you don't exceed expectations, you can take consolation in knowing that your predictive powers are no worse than the world's leading political pundits and public intellectuals. Our prediction is that fewer than half of Australia's pundits would be willing to subject themselves to an open survey in which they were asked to forecast the outcome of a dozen major political events. However, having made such a falsifiable prediction, we should add that we hope to be proven wrong.

Acknowledgement

We are grateful to Paul Bongiorno, Barrie Cassidy, Lenore Taylor, editor Sue Turnbull, production editor Susan Jarvis and two anonymous referees for a variety of helpful comments on earlier drafts. Naturally, they should not be assumed to agree with our conclusions.

Notes

- ¹ This tradition of cynical wit was perhaps best summed up by a 1983 exchange between then Prime Minister Malcolm Fraser and journalist Peter Bowers. In a long-winded answer, Fraser spoke of how Australians liked their horse races but would never gamble on Australia's future by voting Labor. Bowers immediately followed up with: 'You wouldn't want to bet on that, would you?'
- ² The ABC is Australia's publicly owned broadcaster, while Channel 10 is a commercial network. We chose Sunday talk shows because the format typically leads to more prediction-making than other political news formats. We chose *Insiders* and *Meet the Press* because we wanted Sunday talk shows that had covered the most recent two elections, and whose transcripts were readily available to us and other researchers.
- ³ Because Tetlock's empirical research guaranteed anonymity to participants, there is no way of checking his methodology for separating experts into various categories. By contrast, our study relies on publicly available video recordings and transcripts of pundits' predictions. Accordingly, other researchers who wish to check our coding are free to watch the original programs themselves. It is also worth noting that while Tetlock garnered responses via privately conducted surveys of respondents, our analysis is based on public predictions. While our sample size is considerably smaller, it is nonetheless likely that our findings are more directly applicable to political prognostication in the real world.
- ⁴ Only one of the fifteen *Insiders* predictions (Gerard Henderson's prediction on 23 September 2007) was made in the closing segment of the program (that is, the portion that follows *Insiders*' 'talking pictures' segment). This is somewhat remarkable, given that our sample covers 29 programs, each with three pundits. The pundits presumably prepare for this segment prior to the show, but they clearly had a preference for 'observations' over falsifiable predictions.
- ⁵ Tetlock (2005) formally tests the relationship between ideology and forecast accuracy, and finds some evidence that moderates are more accurate than those on the left or right (in Figure 3.1).
- ⁶ This may be yet another case of what Macnamara (2012) documents as the effectiveness of the public relations industry in serving its clients, despite many journalists' misgivings about the PR industry's impact on public communications.
- ⁷ Asked what was most likely to blow a government off course, British Prime Minister Harold Macmillan was said to have replied 'events, dear boy, events' (the veracity of the quote has been disputed).
- ⁸ Tetlock (2005) formally tests the relationship between reputation (as measured by total Google hits) and forecast accuracy, and finds no evidence that more famous pundits differ in their accuracy.

References

Berg, A. and Pattillo, C. 2000, 'The Challenges of Predicting Economic Crises', *International Monetary Fund Economic Issues*, no. 22, pp. 561–86.

Blake, D., Beenstock, M. and Brasse, V. 1986, 'The Performance of UK Exchange Rate Forecasters', *Economic Journal*, vol. 96, no. 384, pp. 986–99.

Clemen, R. 1989, 'Combining Forecasts: A Review and Annotated Bibliography', *International Journal of Forecasting*, no. 5, pp. 559–83.

Cooper, A., Woo, C.Y. and Dunkelberg, W.C. 1988, 'Entrepreneurs' Perceived Chances for Success', *Journal of Business Venturing*, no. 3, pp. 97–108.

Gardner, D. 2010, *Future Babble: Why Expert Predictions Fail – and Why We Believe Them Anyway*, Dutton, New York.

Haldane, A.G. 2012, 'The Dog and the Frisbee', paper presented at the Federal Reserve Bank of Kansas City's 36th Economic Policy Symposium, Jackson Hole, WY, 31 August.

Hendry, D. and Clements, M. 2004, 'Pooling of Forecasts', *Econometrics Journal*, no. 7, pp. 1–31.

Kahneman, D. 2011, *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York.

Kahneman, D. and Klein, G. 2009, 'Conditions for Intuitive Expertise: A Failure to Disagree', *American Psychologist*, no. 64, pp. 512–26.

Leigh, A. and Wolfers, J. 2006, 'Competing Approaches to Forecasting Elections: Economic Models, Opinion Polling and Prediction Markets', *Economic Record*, vol. 82, no. 258, pp. 325–37.

Macnamara, J.R. 2012, 'Journalism and Public Relations: Unpacking Myths and Stereotypes', *Australian Journalism Review*, vol. 34, no. 1, pp. 35–52.

Menand, L. 2005, 'Everybody's an Expert', *New Yorker*, 5 December, pp. 98–101.

Odean, T. 1998, 'Volume, Volatility, Price, and Profit When All Traders are Above Average', *Journal of Finance*, vol. 53, no. 6, pp. 1887–1934.

Pasaogullari, M. and Meyer, B. 2011, 'The Accuracy of CPI Inflation Forecasts', *Federal Reserve Bank of Cleveland Economic Trends*, 7 January.

Popper, K. 1959, *The Logic of Scientific Discovery*, Hutchinson, New York.

Sherden, W. 1998, *The Fortune Sellers: The Big Business of Buying and Selling Predictions*, John Wiley & Sons, New York.

Silver, N. 2012, *The Signal and the Noise: Why So Many Predictions Fail – but Some Don't*, Penguin, New York.

Soll, J. and Larrick, R. 2009, 'Strategies for Revising Judgment: How (and How Well) People Use Others' Opinions', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, no. 35, pp. 780–805.

Speirs-Bridge, A., Fidler, F., McBride, M., Flander, L., Cumming, G. and Burgman, M. 2010, 'Reducing Overconfidence in the Interval Judgments of Experts', *Risk Analysis*, no. 30, pp. 512–23.

Surowiecki, J. 2004, *The Wisdom of Crowds*, Doubleday, New York.

Svenson, O. 1981, 'Are We All Less Risky and More Skilful than Our Fellow Drivers?', *Acta Psychological Review*, vol. 89, pp. 155–91.

Tetlock, P. 2005, *Expert Political Judgment: How Good Is It? How Can We Know?*, Princeton University Press, Princeton, NJ.

Phillip Metaxas is an independent researcher, email: phillipmetaxas@iinet.net.au.

Andrew Leigh is the Federal Member for Fraser. He was formerly a Professor of Economics at the Australian National University.