

## SECTION 3: QUALITY AND ACCURACY OF FORECAST PERFORMANCE

### 3.1: Approach to Forecasting Assessment

#### Summary

- The Review has assessed Treasury's forecast performance against two desirable properties of forecasts. First, the forecasts should be unbiased, that is the expected forecast error should be zero. And, second, the forecasts should be accurate, that is the actual forecast errors should be minimised to the extent possible.
- To provide a benchmark against which to assess accuracy, Treasury's forecast performance is compared to that of other domestic forecasters and official agencies overseas and also to the performance of naive forecasting rules, based on the past trend behaviour of the forecast series.

#### Description of the data

The Review has focused its assessment of Treasury's macroeconomic forecasts on those series which are most important for revenue forecasting. These series include nominal GDP and the major components of the income measure of nominal GDP, in particular compensation of employees and gross operating surplus. The nominal GDP forecasts are constructed from the real GDP and GDP deflator forecasts, and therefore these series, and the terms of trade, are also assessed. In terms of Treasury's revenue forecasts, the Review has assessed the performance of aggregate taxation revenue and the major heads of revenue. All the analysis presented for revenue is on a cash basis because this is the only method of recognition of revenue that has data back to 1990-91.

The forecasts are assessed over the period 1990-91 to 2011-12, data permitting. The start date for the assessment period was chosen because it coincides with a major structural break in the economy, reflecting the transition of Australia to a low-inflation environment. The forecasts are also assessed over four distinct economic sub-periods that reveal patterns in forecast errors that are obscured over the full sample. These sub periods are: 1990-91 to 1993-94, which includes the early 1990s recession; 1994-95 to 2002-03, which covers a period of relatively stable growth; 2003-04 to 2007-08, which covers the first mining boom; and 2008-09 to 2011-12, which includes the global financial crisis, and the emergence of a second phase of the mining boom.

In principle, the macroeconomic forecast performance could be measured against the ABS's first, or most recent, published outcomes, which are from the June quarter 2012 National Accounts release (at the time of the preparation of this report). The measure of forecast performance depends importantly on the choice of benchmark due to ABS revisions. The Review has compared Treasury forecasts with the most recent estimated outcomes for two reasons. First, the most recent estimated outcomes represent the ABS's current best estimates of the true outcomes. And, second, Treasury's revenue mapping models use the most recent estimates of the nominal economy in order to forecast taxation revenue and hence it is these estimates that are most important for revenue forecasting purposes.<sup>1</sup>

#### Measures of forecasting performance

There are many approaches to measuring forecast performance. The Review bases its assessment of Treasury's forecast performance upon two desirable properties of forecasts. First, the forecasts should be unbiased, that is the expected forecast error should be zero. And, second, the forecasts should be

---

<sup>1</sup> One disadvantage of this approach is that ABS revisions can reflect changes in the definitions of series, including as the result of the adoption of more recent international benchmarks for national accounting statistics.

accurate, that is the actual forecast errors should be minimised to the extent possible.<sup>2</sup> It draws upon metrics that have been commonly used in such analysis, and are easy to interpret. These metrics are the mean error and the mean absolute error (or in percentage points, the mean absolute percentage error).

The *mean* error measures the *bias* of the forecasts. A positive (negative) number indicates that, on average, the forecast has tended to be higher (lower) than the outcome. All other things equal, a figure closer to zero indicates a better forecasting performance. The *mean absolute* error measures the *accuracy* of the forecasts, as it measures the average distance between the forecast and the outcome, which is the size of the typical error. All other things equal, a smaller number indicates a better forecasting performance.

Formally, the metrics are calculated as:

$$\text{Mean error} = \frac{\sum_{i=0}^{n-1} (f_{t-i}^* - f_{t-i})}{n}, \text{ and Mean absolute percentage error} = \frac{\sum_{i=0}^{n-1} |f_{t-i}^* - f_{t-i}|}{n},$$

where:  $f_t^*$  and  $f_t$  are the forecast and actual growth rates for the series being assessed.

The main alternative metric of forecasting performance is the root-mean-squared-error, which places greater weight on large forecast errors. Most studies, such as Zarnowitz (1991)<sup>3</sup> for the United States and Holden and Peel (1988)<sup>4</sup> for the United Kingdom, find that conclusions are insensitive to the choice of measure.

As with any statistical assessment of forecast performance there are limitations in the interpretation of these metrics. In particular, a small sample size reduces the reliability of sample averages as a few large errors can have an unduly large influence. Hence it is necessary to base conclusions on tests of statistical significance. These measures also need to be interpreted in light of the average growth rate of the series being forecast — a 1 percentage point mean error (or bias) in annual growth forecasts for a series that grows on average by 40 per cent per annum is a very different performance to the same mean error in a series that grows on average by 2 per cent.

### Forecast comparisons and their limitations

To provide a benchmark against which to assess accuracy, Treasury's forecasts are compared with those of selected domestic and official agencies overseas. In terms of domestic forecasters, Treasury's macroeconomic forecasts are compared with those produced by the Reserve Bank of Australia (RBA), Deloitte-Access Economics (Access) and Consensus Economics. Its revenue forecasts are compared with those produced by Access. Both sets of forecasts are compared to those produced by official agencies in the United States, Canada, the United Kingdom and New Zealand. Treasury's forecasts are also compared with those generated by a naive forecasting rule, which assumes that the series being forecast simply continues to grow at its recent average observed rate (one, three, five and ten year moving averages of the forecast series were considered).

Forecast comparisons provide insight although they need to be carefully interpreted. In particular, different agencies tend to finalise their forecasts at different times. A forecast prepared at a later time is likely to have an information advantage. This could reflect the receipt of additional official statistics or knowledge of a new macroeconomic development, for example consider the difference between the

2 The serial correlation of the forecast errors and the success rate in identifying the direction of changes in GDP growth are also considered, although the Review has not reported on these metrics in detail.

3 Zarnowitz, V. 1991, 'Has macro-forecasting failed?', *NBER working papers*, no 3867.

4 Holden, K. and Peel, D. 1988, 'A comparison of some inflation, growth and unemployment forecasts', *Journal of Economic Studies*, 15(5), pp. 5-21.

macroeconomic outlook the month before, and the month after, the collapse of Lehman Brothers in September 2008. Forecast comparisons are also sensitive to the chosen sample period.

### Challenges to preparing forecasts

Forecasting errors are inevitable, even with the most rigorous forecasting framework and procedures. Forecasting is an inherently difficult exercise and errors arise from many sources. Models — which describe behavioural economic relationships — are always simplifications of the modern complex economy. Coefficient estimates — which provide an assessment of the strength of economic relationships — may be imprecise, particularly in the face of continual structural change. Exogenous assumptions, such as the exchange rate, or the international economic outlook, might turn out to be wrong. More often than not, there are shocks to the economy which were not anticipated at the time of the forecasts. The official statistics are also subject to revision.

Many of these forecasting errors are unavoidable. That said, a forecasting methodology that draws upon the range of available information, and processes that information efficiently, should help to minimise forecasting errors.

## 3.2: Treasury's Macroeconomic Forecasting Performance

### Summary of macroeconomic forecasting performance

- Treasury's forecasts of nominal GDP growth exhibit little evidence of bias over the past two decades; although, with the benefit of hindsight, forecast errors have been correlated with the economic cycle. Hence, Treasury has tended to underestimate growth during economic upswings and overestimate growth during economic downturns.
- Treasury's macroeconomic forecasts have been reasonably accurate. Treasury's forecast performance has been comparable with that of other domestic forecasters. Its forecasts are comparable with, or better than, those of official agencies overseas. They also compare favourably with statistical benchmarks generated by a naïve trend forecasting rule.
- Within these general findings, however, Treasury's forecasts exhibit periods of quite high accuracy, interspersed with occasional periods of large outliers.
- Treasury's forecasts of GDP deflator growth are less accurate than those of real GDP growth. In particular, there were extended periods in the 1990s where outcomes were overestimated and in the 2000s where outcomes were underestimated. In recent years, this has substantially reflected the difficulty of forecasting commodity prices.

### Nominal GDP

Treasury's forecasts of nominal GDP growth exhibit little evidence of bias over the past two decades, with the mean Budget forecast error being insignificantly different from zero (Table 3.1). Over this period, Treasury's forecasts have been reasonably accurate, exhibiting a mean absolute percentage error (MAPE) of 1.6 percentage points across Budget forecast rounds.

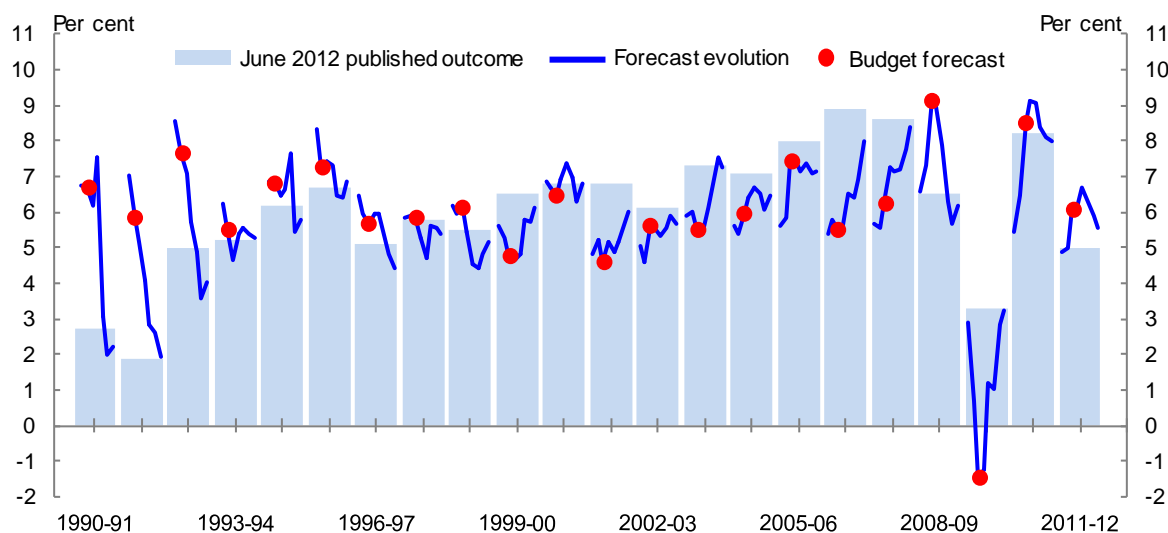
**Table 3.1: Performance of Nominal GDP Growth Forecasts against Most Recent Estimated Outcomes**

	1990-91 to 2011-12		1990-91 to 1993-94		1994-95 to 2002-03		2003-04 to 2007-08		2008-09 to 2011-12	
	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points
All forecast rounds	-0.3	1.2	1.4	1.7	-0.3	0.7	-1.5	1.5	-0.2	1.3
Budget (a)	-0.1	1.6	2.7	2.7	-0.2	0.8	-1.8	1.8	-0.2	2.2
MYEFO (b)	0.0	1.3	2.0	2.0	-0.4	1.0	-1.3	1.3	0.5	1.5

(a) March forecast round for the financial year starting in the July of the same year. Budget forecast from 1996-97.

(b) September forecast round for the financial year which started two months earlier. MYEFO forecast from 1998-99.

That said, an examination of the patterns in forecast errors in Table 3.1, and Figure 3.1, reveals a more variable performance across economic sub periods, with the forecast errors being correlated with the economic cycle, with the benefit of hindsight. In particular, Treasury overestimated nominal GDP growth in the early 1990s (1990-91 to 1993-94), as the recession at that time was not forecast, nor was the speed of the transition to a low inflation environment. It also underestimated nominal GDP growth during Mining Boom Mark I (2003-04 to 2007-08), with broadly offsetting effects over the full sample.

**Figure 3.1: Evolution of Nominal GDP Growth Forecasts**

The patterns in forecast errors in recent years reflect the challenges of forecasting two major economic developments. The first of these relates to the rapid rates of industrialisation in Asia, particularly in China, which increased worldwide demand for natural resources (Mining Boom Mark I). Treasury underestimated the extent of the resultant sharp and sustained rise in commodity prices through the mid-2000s, which led to an underestimation of Australia's terms of trade and, in turn, nominal economic outcomes.

The second relates to the impact of the global financial crisis (GFC), and its aftermath, on the Australian economy. Treasury did not predict the onset of the GFC in 2008-09, and subsequently overestimated its effect on growth in 2009-10. This saw large forecast errors generated in 2008-09 and 2009-10. In particular, in the 2009-10 Budget, at the height of a period of significant global and domestic pessimism, Treasury forecast a recession in 2009-10 that did not eventuate.

These episodes are discussed in more detail in Section 4. These patterns in forecast errors are apparent in subsequent figures and tables, below.

## Real GDP

Treasury's forecasts of real GDP growth also exhibit little evidence of bias, with the mean Budget forecast error being insignificantly different from zero (Table 3.2 and Figure 3.2). Its real GDP growth forecasts have been quite accurate, with the MAPE generally remaining within a range of ½ to 1 percentage point. Treasury's forecasting performance has been less accurate in recent years than over the full sample period, reflecting greater volatility in real GDP growth as a result of the impact of the GFC, and its aftermath, on the Australian economy.

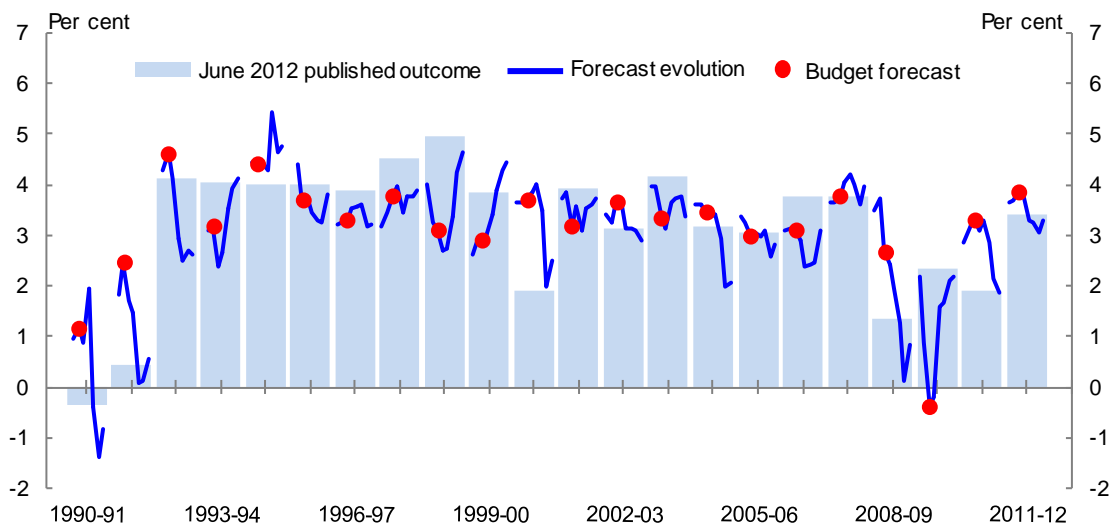
**Table 3.2: Performance of Real GDP Growth Forecasts against Most Recent Estimated Outcomes**

	1990-91 to 2011-12		1990-91 to 1993-94		1994-95 to 2002-03		2003-04 to 2007-08		2008-09 to 2011-12	
	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points
All forecast rounds	-0.1	0.7	0.0	0.9	-0.2	0.7	-0.3	0.5	0.1	0.9
Budget (a)	0.0	0.9	0.8	1.2	-0.3	0.9	-0.2	0.4	0.1	1.5
MYEFO (b)	-0.1	0.9	0.2	1.5	-0.3	0.9	-0.3	0.5	0.3	0.7

(a) March forecast round for the financial year starting in the July of the same year. Budget forecast from 1996-97.

(b) September forecast round for the financial year which started two months earlier. MYEFO forecast from 1998-99.

**Figure 3.2: Evolution of Real GDP Growth Forecasts**



These findings contrast with those of a recent study by Frankel (2011) of official government real growth rate (and budget balance) forecasts between 1985 and 2009 in 33 countries. That study found that official agency forecasts tended to have a positive average bias; are more biased in booms (and are even more biased at the three-year horizon than at shorter horizons). The data for Australia indicate little bias in all these respects compared with other countries.

The different volatility of the various expenditure components of GDP makes some easier to forecast than others (Table 3.3). Not surprisingly, Treasury's forecasts of the most volatile expenditure components tend to be the least accurate, with the largest MAPEs. Treasury has had the greatest difficulty in accurately forecasting business and dwelling investment, with the former, as an import-intensive component of GDP, also having an impact on the accuracy of the imports' forecasts.

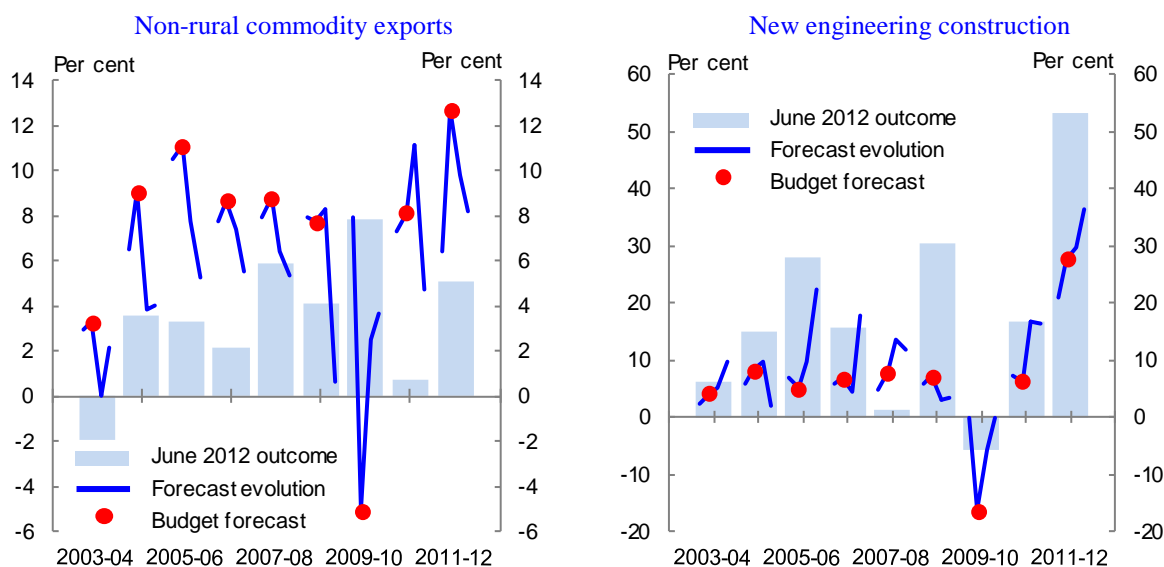
**Table 3.3: Performance of GDP Expenditure Component Growth Forecasts (1998-99 to 2011-12, All Forecast Rounds)**

	Mean error % points	MAPE % points	Standard deviation of series(a)	Share of economy(b) %
Household Consumption	-0.1	0.8	1.3	56
Public Final Demand	-0.2	0.9	1.6	19
Exports	1.2	2.7	5.7	20
Imports	-1.1	3.5	3.0	-21
Business Investment	-2.4	4.7	8.1	15
Dwelling Investment	-0.5	4.9	9.6	6
GDP	-0.1	0.7	1.0	100

(a) Standard deviation of series growth rates, from 1998-99 to 2011-12.

(b) Average share of economy, from 1998-99 to 2011-12.

An examination of the mean forecasting errors of the expenditure components of GDP indicates that Treasury has overestimated exports growth in recent years, and underestimated business investment and, in turn, imports growth. In particular, since the beginning of Mining Boom Mark I, Treasury has consistently overestimated growth in non-rural commodity exports (Figure 3.3). These forecasts are heavily influenced by mining company's stated targets, which have consistently exceeded actual outcomes, in part due to the impact of natural disasters and infrastructure bottlenecks. Treasury has also been overly pessimistic forecasting business investment, particularly the mining-boom related surge in engineering construction (Figure 3.3).

**Figure 3.3: Evolution of Non-rural Commodity Exports and Engineering Construction Growth Forecasts**

### GDP deflator

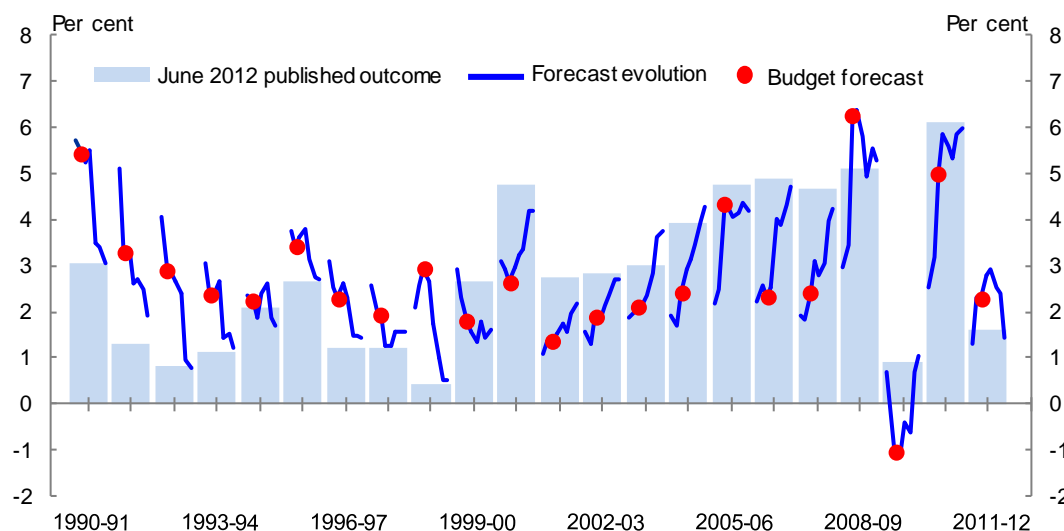
Treasury's forecasts of GDP deflator growth have been less accurate than Treasury's forecasts of real GDP growth. In particular, GDP deflator growth was consistently overestimated in the 1990s, although the size of the forecast error fell on average through the decade (Table 3.4 and Figure 3.4). As discussed, this reflects the recession in the early 1990s, which was not forecast, nor was the durability of the transition to a low-inflation environment. In contrast, over the period from the early 2000s through to the GFC, GDP deflator growth was consistently underestimated, as discussed, due to Treasury underestimating the extent and duration of the sharp rise in Australia's terms of trade as a result of Mining Boom Mark I. These episodes have had broadly offsetting impacts on the mean forecast error over the full sample

**Table 3.4: Performance of GDP Deflator Growth Forecasts against Most Recent Estimated Outcomes**

	1990-91 to 2011-12		1990-91 to 1993-94		1994-95 to 2002-03		2003-04 to 2007-08		2008-09 to 2011-12	
	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points
All forecast rounds	-0.1	1.1	1.4	1.4	-0.1	0.9	-1.1	1.2	-0.4	1.1
Budget (a)	-0.1	1.4	1.9	1.9	0.0	1.2	-1.5	1.5	-0.3	1.2
MYEFO (b)	0.1	1.1	1.8	1.8	-0.1	0.9	-1.0	1.0	0.1	1.0

(a) March forecast round for the financial year starting in the July of the same year. Budget forecast from 1996-97.

(b) September forecast round for the financial year which started two months earlier. MYEFO forecast from 1998-99.

**Figure 3.4: Evolution of the GDP Deflator Growth Forecasts**

These observations lead the Review to recommend that:

#### **Recommendation 7:**

Treasury should invest relatively more resources in understanding and forecasting GDP deflator growth and its components, in particular, commodity prices, and hence in nominal GDP growth.

#### **Comparison with other domestic forecasters**

Treasury's forecasting performance is compared with that of Access and the RBA in Table 3.5 at various forecasting horizons. As discussed in Section 3.1, the Review acknowledges the difficulty of drawing exact like-with-like forecast comparisons. Forecasting institutions run on different timetables, and forecasts made later will naturally have an advantage over those made earlier for a given reference period. For example, the timing of Treasury forecasts has tended to be optimised around the release of National Accounts data, whereas for the RBA they are more likely to be optimised around the release of CPI data. This would contribute to the configuration of relative results for the two sets of forecasts. Results are likely to be sensitive to the choice of sub-periods. To help to reduce informational advantages relating to the timing of the preparation of forecasts, the results for the RBA and Access in Table 3.5 are based on forecasts containing the same National Accounts information as Treasury's forecasts.

Treasury's forecasting performance for the core macroeconomic series have been comparable with that of Access and the RBA over the past two decades (Table 5). The differences in forecasting



accuracy across agencies are small and not statistically significant at the 10 per cent level.<sup>5</sup> That is to say, the differences could not be distinguished from random noise. Consistent with this finding, the ranking of forecasters varies across macroeconomic series and forecasting rounds. The variation in rank suggests that comparisons of relative forecast accuracy will be sensitive to the sample period. Due to data limitations, RBA forecasts for the GDP deflator, nominal GDP and the terms of trade are only available since 2000, and so are not shown in the table. Over this shorter sample, RBA forecast accuracy was not significantly different to that of Treasury.

**Table 3.5: Performance of Access, the RBA and Treasury Forecasts (MAPE)**

To end of financial year:	MYEFO 7 quarters	Budget 5 quarters	MYEFO 3 quarters	Budget 1 quarter	Average forecasts
1993-94 to 2011-12					
<i>Real GDP</i>					
Access	0.8	0.9	0.8	0.5	0.7
RBA	1.0	0.9	0.7	0.4	0.8
Treasury	0.7	0.9	0.8	0.5	0.7
1996-97 to 2011-12					
<i>CPI (tty)</i>					
RBA	1.1	1.0	0.6	0.1	0.7
Treasury	0.9	1.1	0.8	0.3	0.7
<i>GDP deflator</i>					
Access	1.8	1.5	0.8	0.6	1.2
Treasury	1.7	1.4	1.0	0.5	1.1
<i>Nominal GDP</i>					
Access	1.7	1.6	1.0	0.7	1.3
Treasury	1.3	1.5	1.3	0.7	1.2
<i>Terms of trade</i>					
Access	7.3	6.7	4.3	1.8	5.0
Treasury	5.4	4.4	2.7	1.0	3.4

Note: the differences in the results between agencies are statistically insignificant.

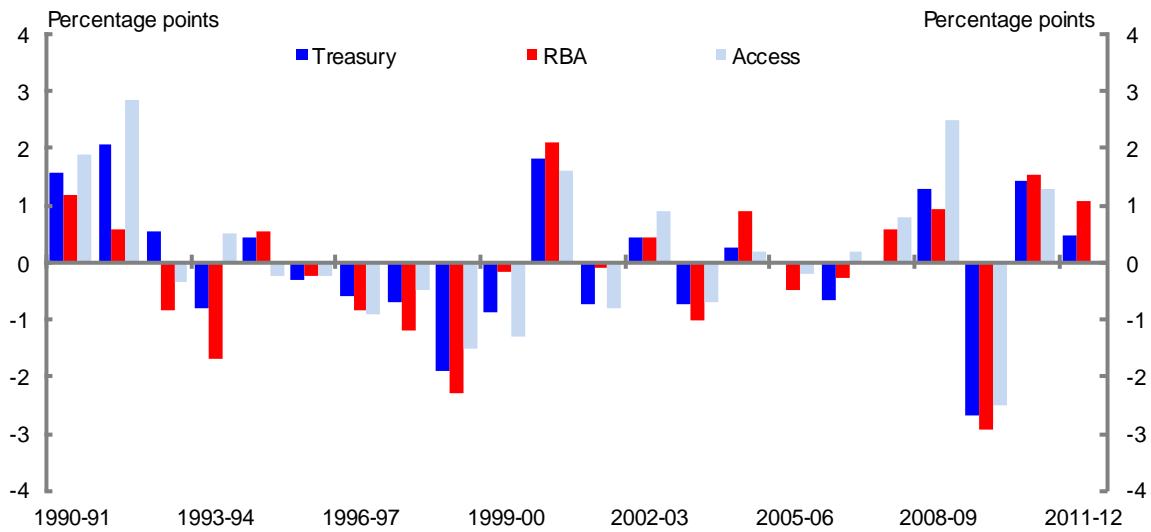
This assessment is supported by examination of the patterns of forecast errors across agencies. Figure 3.5 shows the patterns in forecast errors across agencies for real GDP growth for the Budget forecast round (five quarters before the end of the financial year). The striking feature of this chart is the similarity of the forecast errors, with the small variation across agencies contrasting with the significant variation in errors across time. It may also be interesting to note that the ranking of forecasters has no persistence but changes almost every year, consistent with the large random element in measures of forecast accuracy.

The patterns in forecast errors across agencies for nominal GDP growth and terms of trade growth for the Budget forecast round are shown in Figures 3.14 and 3.15 in the Appendix to this section.

<sup>5</sup> The statistical significance between forecasts was econometrically tested using a Diebold-Mariano test. In general, the test requires the forecast errors of two forecasts  $\{y_{it}\}_{t=1}^T$  and  $\{y_{jt}\}_{t=1}^T$  for a key variable  $\{y_t\}_{t=1}^T$ . These errors are typically denoted as  $\{e_{it}\}_{t=1}^T$  and  $\{e_{jt}\}_{t=1}^T$ . An absolute loss function  $d_t = |e_{it}| - |e_{jt}|$  is then constructed and regressed against a constant, with the null hypothesis of equal forecast accuracy (or zero loss) tested using the Diebold-Mariano test statistic.

$$\text{DM statistic} = \bar{d}_t / \sqrt{\text{var}\left(\frac{\bar{d}_t}{T}\right)}, \text{ where } \bar{d}_t \text{ is the sample mean loss differential.}$$



**Figure 3.5: Comparison of Budget Forecast Errors for Real GDP Growth**

It is noteworthy that the MAPEs for Treasury's MYEFO forecasts (seven quarters before the end of the forecast year) tend to be smaller than those for the Budget forecasts (five quarter before the end of the financial year), despite the latter having more information. This is because the Budget data include a large error relating to the forecast of a domestic recession in 2009-10 due to the GFC, which did not eventuate. The corresponding forecast at MYEFO (two quarters earlier) did not forecast a recession in 2009-10.

The performance of Treasury's forecasts of real GDP growth is also comparable to those of Consensus Economics (see Table 3.6).

**Table 3.6: Performance of Consensus, the RBA and Treasury Forecasts for Real GDP growth: 2000 to 2011: Calendar Year: MAPE**

To end of calendar year:	Budget 7 quarters	MYEFO 5 quarters	Budget 3 quarters	MYEFO 1 quarter	Average forecasts
<i>Real GDP</i>					
Consensus	0.70	0.72	0.81	0.59	0.70
RBA	1.42	0.78	0.84	0.56	0.90
Treasury	0.87	0.66	0.67	0.49	0.67

Note: the differences in the results between agencies are statistically insignificant.

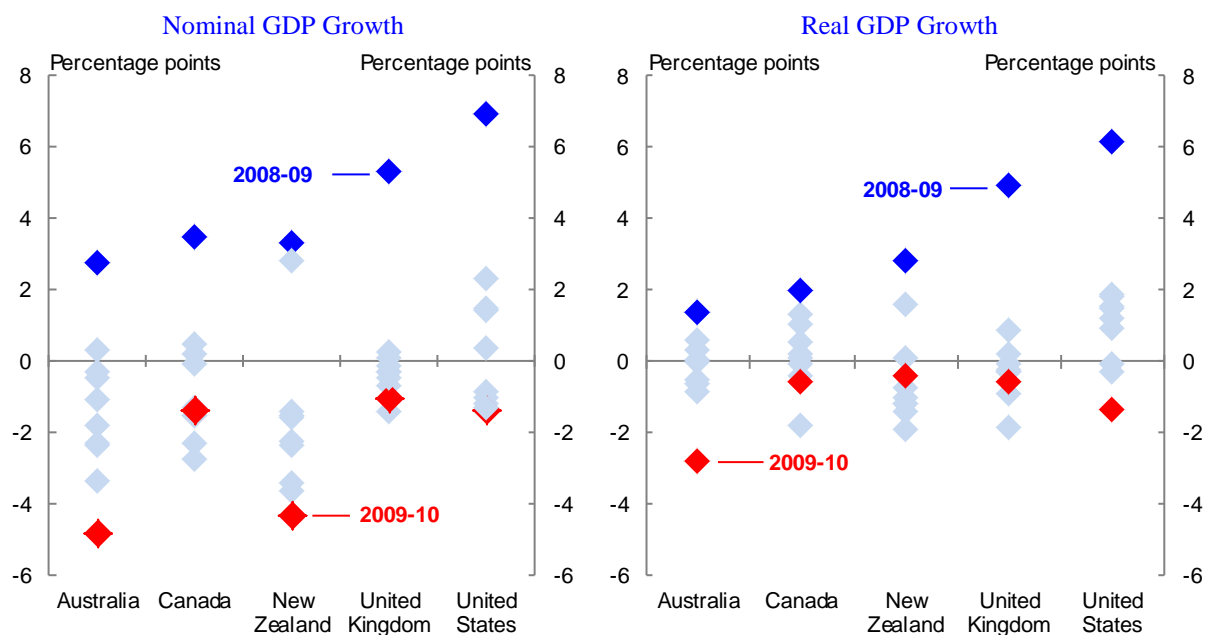
### Comparison with official agencies overseas

Treasury's forecast performance has been comparable with, or better than, the performance of official agencies overseas over the past decade, although some caution is required in making cross country comparisons over a period as short as ten years, and given that official agencies prepare forecasts at different times in the year (Figure 3.6). In particular, Australia's official forecasts of nominal GDP growth outperform those of New Zealand, but are statistically insignificant from those of Canada, the United Kingdom and the United States. Australia's official forecasts of real GDP growth are statistically insignificant different from those of Canada, New Zealand, the United Kingdom and the United States.

It is interesting to note that all official agencies missed the onset of the GFC and tended to overstate its effect on activity in 2009-10, albeit to varying degrees (Figure 3.6). As a result, all official

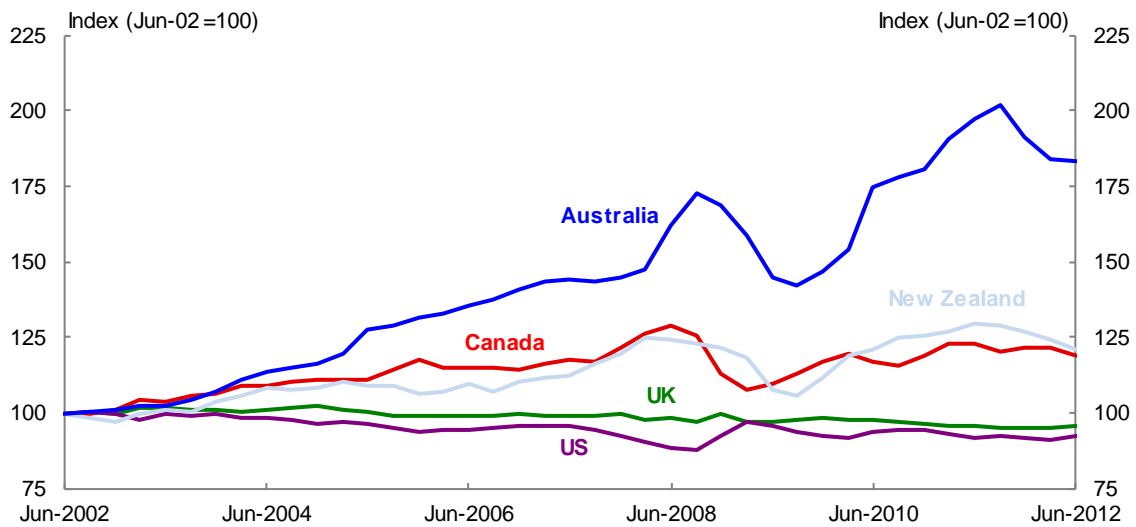
agencies tended to overestimate economic growth outcomes in 2008-09 and underestimate outcomes in 2009-10.

**Figure 3.6: International Comparison of Budget Forecast Errors across Official Agencies: 2001-02 to 2010-11**



Australia’s Budget is published in early May, two months before of the start of the Budget financial year; the United Kingdom’s Budget is published in March, one-month before the start of its Budget March year; Canada’s Budget is published between January and March, within its Budget calendar year, New Zealand’s Budget is published in May, two months before the start of its Budget June year; and the United States’ Budget is published in February eight months before the start of its Budget September year.

It is also worth noting that over this period domestic forecasters faced a dramatic rise in Australia’s terms of trade of almost 200 per cent, which, as discussed, has contributed to the nominal GDP growth forecast errors. This is in stark contrast to the experience of the other countries surveyed (Figure 3.7). This would have contributed to the configuration of the relative results for the international forecast comparison, as would the relative severity of the impact of the GFC on the domestic economies of the countries surveyed (Figure 3.6).

**Figure 3.7: Terms of Trade, by Country**

### Comparison with naïve trend forecasts

Treasury's real and nominal GDP growth forecasts tend to outperform trend estimates for the Budget forecast round (one quarter out from the start of the financial year) (Table 3.7). An examination of the data over sub periods reveals, perhaps unsurprisingly, that it is more difficult to outperform trend estimates during periods of relative economic stability. In contrast, during periods of economic volatility, forecasters can more rapidly incorporate information reflecting a changing environment, for example, the transition to the low inflation environment in the early 1990's, than backward-looking trend forecasts.

**Table 3.7: Performance of Treasury Budget Forecasts and Naïve Trend Forecasts against Most Recent Estimated Outcomes (1990-91 to 2011-12)**

	Nominal GDP		GDP Deflator		Real GDP	
	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points
Budget (a)	0.0	1.6	0.0	1.4	0.0	0.9
Trend forecasts 1 yr	0.5	2.2	0.3	1.7	0.1	1.4*
Trend forecasts 3 yr	0.8	2.2	0.5	1.6	0.2	1.3*
Trend forecasts 5 yr	1.0	2.2	0.6	1.5	0.3	1.3*
Trend forecasts 10 yr	1.4	2.7*	1.1	1.8	0.3	1.1

(a) March forecast round for the financial year starting in the July of the same year. Budget forecast from 1996-97.

\*Trend MAPEs significantly differ from Budget at the 10% level.

## 3.3: Treasury's Revenue Forecasting Performance

### Summary of revenue forecasting performance

- Treasury's taxation revenue forecasts have exhibited little evidence of bias over the full sample, although this conceals sustained periods where Treasury has under, or over, forecast revenue, with broadly offsetting effects overall, as was found with the macroeconomic forecasts.
- Taking into account the high degree of difficulty inherent in preparing revenue forecasts, Treasury's forecasts are on the whole reasonably accurate. They are comparable with those of

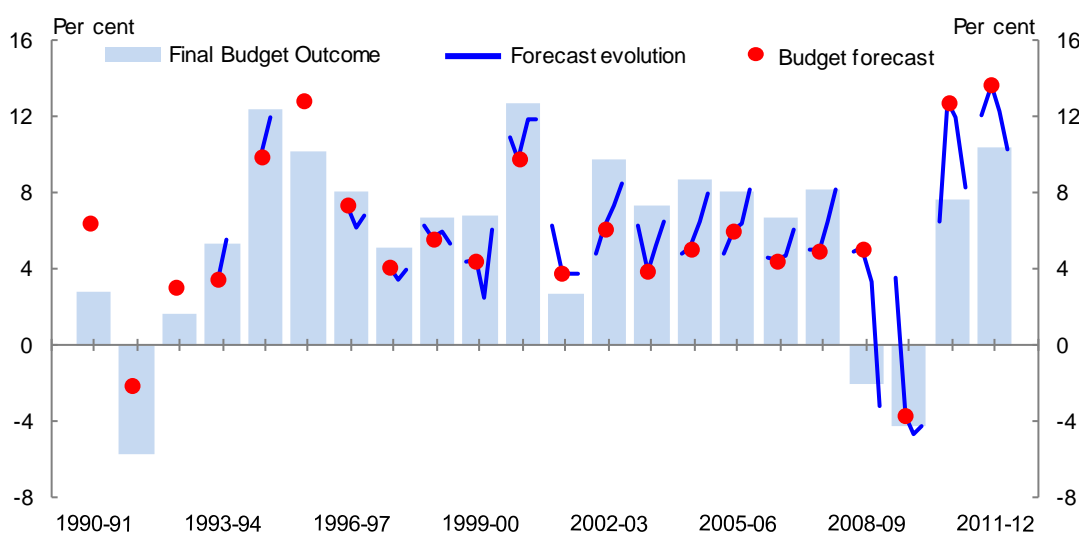
Access and are comparable with, or better than, those of official agencies overseas. They easily outperform statistical benchmarks.

- The heads of revenue with the largest forecast errors in recent years have been company tax and capital gains tax. These are two of the most volatile heads of revenue, and they have been particularly challenging to forecast in the aftermath of the GFC.

### Taxation Revenue

Treasury’s forecasts of taxation revenue growth have exhibited little evidence of bias over the past two decades, with the average Budget forecast error being insignificantly different from zero over this period. Taking into account the high degree of difficulty inherent in preparing revenue forecasts, Treasury’s forecasts are on the whole reasonably accurate (Figure 3.8 and Table 3.8). However, within this timeframe, there have been periods when revenue has been persistently underestimated (in particular, the period from the early 2000s until the GFC) and periods where revenue has been persistently overestimated (the early-1990s recession, and the recent period since the GFC).

**Figure 3.8: Evolution of Taxation Revenue Growth Forecasts**



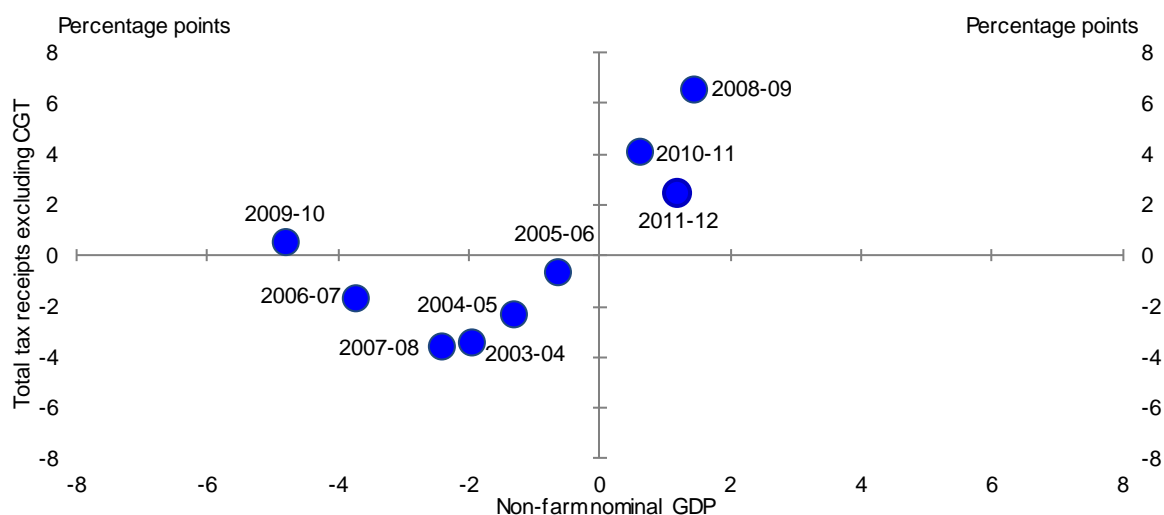
A major contributor to the errors in the taxation revenue forecasts have been the errors in the macroeconomic forecasts. However, there are additional sources of error that reflect taxation-specific factors, such as errors in assumptions made about the timing of the receipt of taxation revenue. As a consequence, the revenue forecast errors are generally larger than those of the macroeconomic forecasts. Over the past 22 years of Budget forecasts, the MAPE for the taxation revenue forecasts has been around 1 percentage point higher than the MAPE for the nominal GDP forecasts (Table 3.8). That said, overall, the revenue forecast errors have been reasonably well correlated with the nominal economy forecast errors, with a correlation coefficient of 0.6 observed over the past decade, although there are some notable outliers (Figure 3.9).

**Table 3.8: Performance of Taxation Revenue and Nominal GDP Forecasts**

	1990-91 to 2011-12		1990-91 to 1993-94		1994-95 to 2002-03		2003-04 to 2007-08		2008-09 to 2011-12	
	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points	Mean error % points	MAPE % points
Budget forecasts (a)										
Nominal GDP	-0.1	1.6	2.7	2.7	-0.2	0.8	-1.8	1.8	-0.2	2.2
Total Tax Revenue	-0.1	2.7	1.7	2.6	-1.2	2.0	-2.9	2.9	4.0	4.0
MYEFO forecasts (b)										
Nominal GDP	0.0	1.3	2.0	2.0	-0.4	1.0	-1.3	1.3	0.5	1.5
Total Tax Revenue	N/A	N/A	N/A	N/A	-1.5	1.8	-1.9	1.9	2.8	3.0

(a) Budget forecast for the financial year which starts in July (two months later). In 1990-91 to 1993-94 and 1996-97 the Budget was published in August and so it is the Budget forecast for the financial year which had started one month earlier.

(b) MYEFO forecast for the financial year which started in July (around four months earlier) and is available from 1996-97. Prior to 1996-97, the September round forecast is used for Nominal GDP (no taxation revenue forecasts are available).

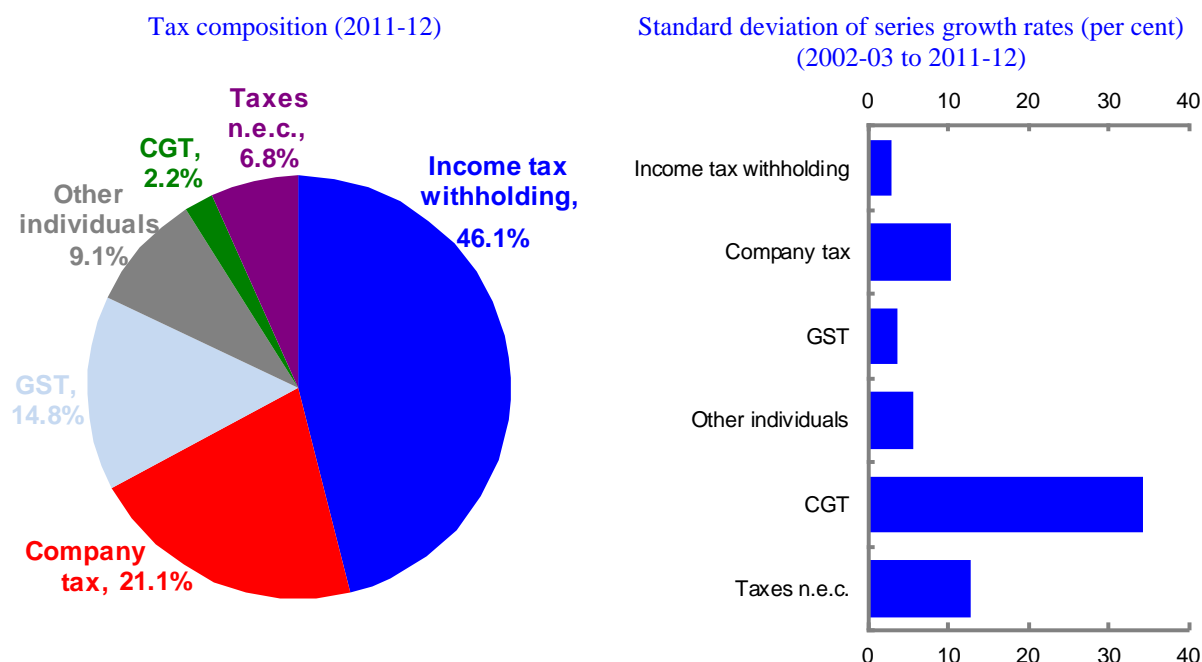
**Figure 3.9: Correlation between Budget Forecast Errors for Growth in Nominal GDP and Taxation Revenue**

### Forecast Errors by Head of Revenue

It can be revealing to break down the total taxation revenue forecast error into contributions from individual heads of revenue in order to better understand the source of the forecast error.

The contribution of an individual head of revenue to the overall taxation revenue forecast error depends upon its share of the tax base (its relative importance), and the error in the forecasts for that head of revenue. The left-hand panel of Figure 3.10 shows the shares of the tax base for each of the main heads of revenue, while the right-hand panel shows the standard deviation of the historical growth rates for each head of revenue. Those with a larger standard deviation are more volatile and will often be more difficult to forecast.

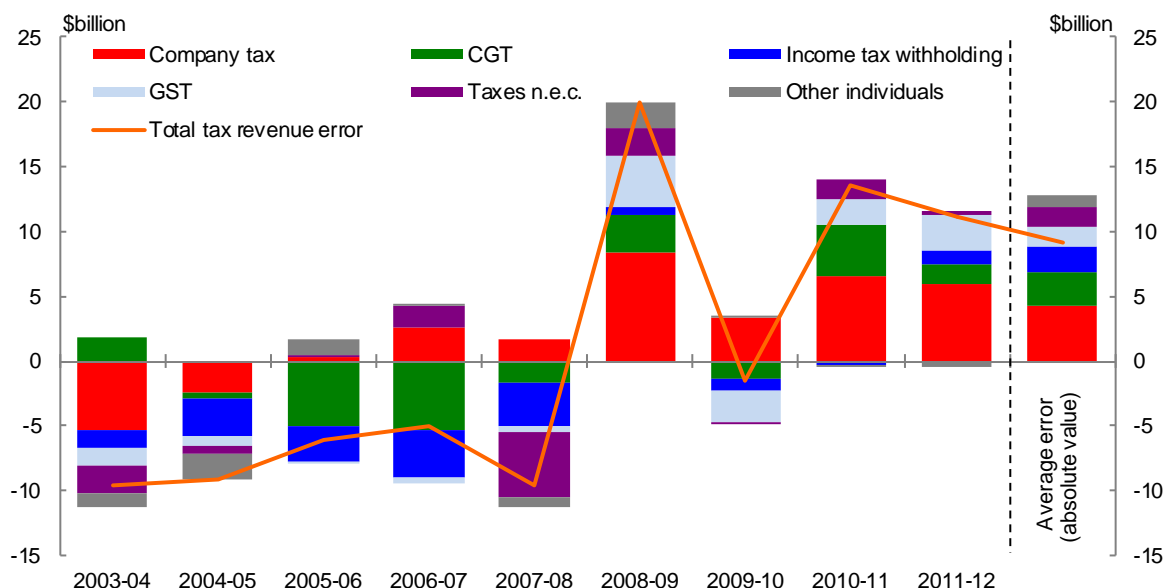
**Figure 3.10: Australian Government Heads of Revenue: Descriptive Statistics**



Note: Taxes n.e.c. includes superannuation taxes; petroleum resource rent tax, fringe benefits tax, excise and customs duty and miscellaneous indirect taxes.

The contributions of the major head of revenues to the total Budget taxation revenue forecast error since the start of Mining Boom Mark I are shown in Figure 3.11.

**Figure 3.11: Contribution to Revenue Error by Head of Revenue (Budget forecasts, 2003-04 to 2011-12)**



As expected, the largest sources of forecast error come from the more volatile heads of revenue, namely company tax and capital gains tax. In particular, the severity of the fall in company tax receipts following the onset of the GFC was not predicted (receipts fell by 12.1 per cent in 2009-10, compared with a forecast fall of only 2.4 per cent). The rebound in company tax receipts since the

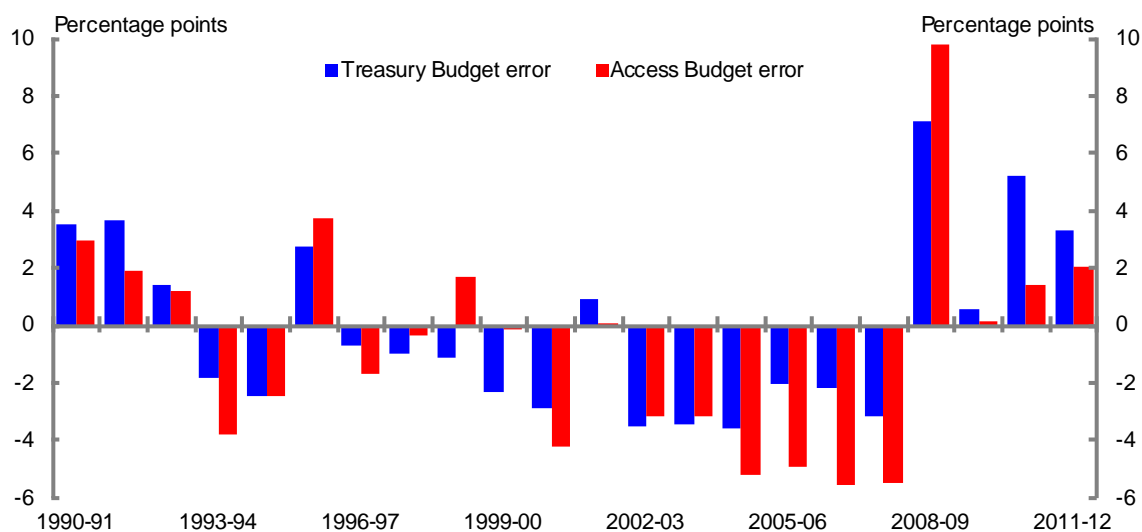
GFC has also failed to meet expectations. The reasons for these errors are explored further in Section 4 of the report, and include higher-than-expected depreciation and royalty deductions associated with the mining boom.

The capital gains tax errors reflect the underestimation of asset price growth during Mining Boom Mark I followed by the general overestimation of asset price growth since the GFC. Capital gains tax revenue has been particularly volatile over this period, with annual outcomes ranging from growth of 60 per cent in 2004-05 to a contraction of 43 per cent in 2009-10. Capital gains tax revenue has been held down by the significant realisation of capital losses incurred during the GFC. These issues are explored further in Section 4.

### Comparison with other domestic forecasters and official agencies overseas

Deloitte Access Economics ('Access') is the only other forecaster of the Australian economy that has published a history of taxation revenue forecasts that is sufficiently long for the purposes of forecast comparison. Treasury's forecasts of taxation revenue have been comparable with those of Access Economics over the past two decades. The differences in forecasting accuracy between Treasury and Access are small and were found not to be statistically significant (at the 10 per cent level). This assessment is supported by an examination of the patterns in forecast errors across agencies in Figure 3.12. A feature of Figure 3.12 is the similarity of the forecast errors, with the small variation across agencies contrasting with the significant variation in errors across time, as was found with the macroeconomic forecast comparison.

**Figure 3.12: Treasury and Access Budget Forecasts of Taxation Revenue Growth**



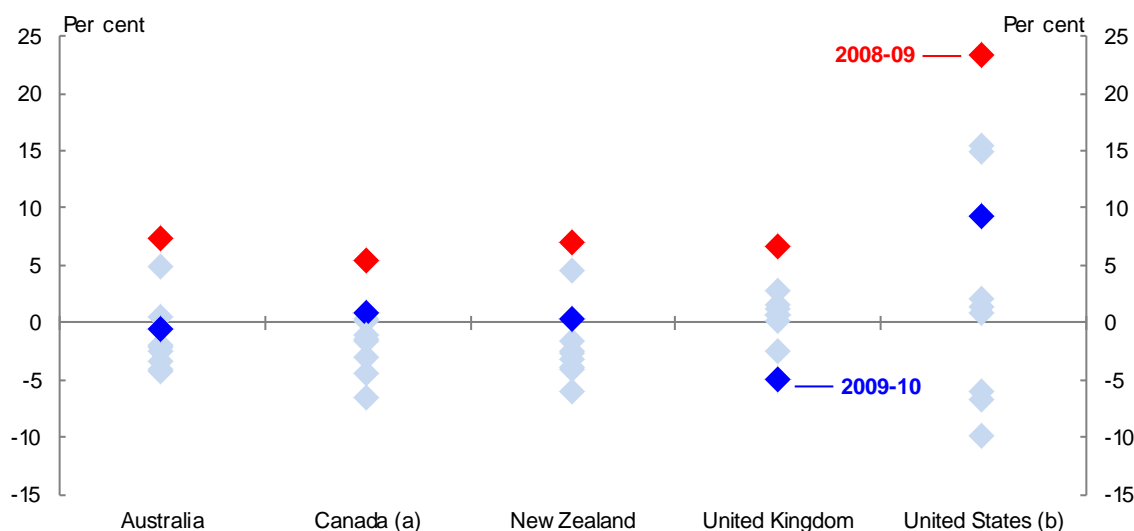
Note: Access forecasts are on an accrual (not cash) basis from 1999-00, and are compared with Final Budget Outcomes on an accrual basis. Access forecasts are generally taken from the May *Budget Monitor* (for Budget), which is typically published around one week ahead of the Budget. Adjustments have been made to Access' forecasts to help to reduce Treasury's information advantage. First, Access does not have information about new policy measures in the relevant update. Access' forecasts have been modified for the policy costings used by Treasury in the relevant update. Second, Access does not have the most up-to-date tax collections information at the time it prepares its revenue forecasts. For example, at Budget, Treasury knows tax collections for the current year up to the end of April. Access will be working off tax collections data from at least one month earlier. There are often significant errors in Access' estimates for the current (base) year. In order to abstract from the base year errors potentially related to different access to information, the comparison has been prepared on a growth rate basis. In fact, the errors for the base year estimates are not correlated with the errors for the Budget year forecasts — the correlation coefficient between the current year errors and Budget year errors (in per cent of level) is 0.04. Nevertheless, the comparison has been prepared in this way to reduce the possibility that the results are driven by Treasury's information advantage.

Compared with the forecasts of official agencies overseas, Treasury's forecasts display less bias than a number of the surveyed agencies over the past decade and, in terms of accuracy, Treasury's



forecasts are comparable with, or better than, those of the surveyed agencies (Figure 3.13). In particular Australian Government Budget forecasts display less bias than those made by official agencies in Canada, New Zealand and the United States. The differences in forecasting accuracy between Australia and official agencies overseas were found not to be statistically significant at the 10 per cent level, except for the United States.

**Figure 3.13: International Comparison of Budget Forecast Errors across Official Agencies: Taxation Revenue Growth: 2001-02 to 2010-11**



(a) Canada excludes 2002-03 as the data is not available. (b) Adjusted for post-Budget changes to policies.

Note: There is a lag between the publication of budget forecasts and the commencement of each country's respective fiscal year. Australia, Canada and New Zealand have a two month lag; the United Kingdom a one month lag; and the United States (as discussed above) an eight month lag.

That said, it should be noted that the United States' Budget forecasts are made further in advance of the beginning of the financial year than in other countries, which may reduce forecast accuracy. In particular, the United States' Budget forecasts are published around eight months prior to the beginning of the financial year, compared with one to two months for the other countries surveyed. Taxation revenue growth has also been more volatile in the United States than in the other countries surveyed, which also makes it harder to forecast.

Unsurprisingly, all international agencies significantly over-predicted taxation revenue growth during 2008-09, the year of the onset of the GFC (Figure 3.13). The pattern in 2009-10 is less clear. Australia, Canada and New Zealand made quite accurate forecasts, while the United Kingdom overestimated the impact of the GFC on taxation revenue and the United States underestimated its impact.

### Comparison with naïve trend forecasts

Treasury's taxation revenue forecasts outperform trend estimates for the Budget forecast round (one quarter out from the start of the financial year) (Table 3.9). The differences in the MAPEs between naïve trend forecasts and Treasury's forecasts are all statistically significant. Underlying (policy adjusted) taxation revenue series are used for this analysis, which remove the advantage that Treasury's forecasts would otherwise have over the trend forecasts relating to the impact of new policy on the outcomes, as well as the construction of the trend estimates. For example, the Treasury forecasts would factor in the introduction of the GST in 2000-01, whereas naïve forecasts based on trends in headline taxation revenue would not capture this new policy. Subsequent to the introduction of the GST, naïve trend forecasts based upon headline taxation revenue would be biased upwards reflecting the introduction of the GST.

**Table 3.9: Performance of Treasury Budget Forecasts and Naïve Trend Forecasts against Estimated Underlying Taxation Revenue Growth Outcomes (1990-91 to 2011-12)**

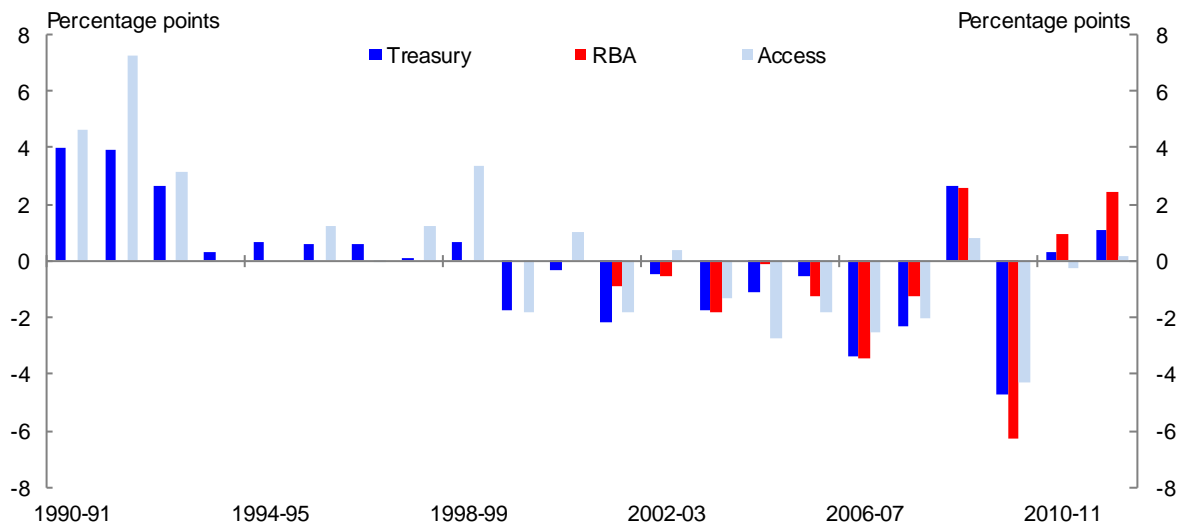
	Underlying Taxation Revenue	
	Mean error	Mean Absolute Percentage Error
	% points	% points
Budget	-0.1	2.4
Trend forecasts 1 yr	-0.2	4.9
Trend forecasts 3 yr	0.8	4.8
Trend forecasts 5 yr	0.8	4.7
Trend forecasts 10 yr	1.9	4.2

Note: Trend is defined as the one, three, five and 10 year moving average annual growth rate in underlying taxation revenue. Underlying taxation revenue growth is calculated by adjusting headline revenue for changes to policy between years.

Treasury's forecasts tend to perform better than trend estimates during periods of economic volatility, particularly during the two downturns in taxation revenue (the early 1990's recession and the GFC). This is because Treasury's forecasters can more rapidly incorporate information relating to these downturns, and the subsequent bounce back in taxation revenue during the recovery phase, than backward-looking trend estimates.

## APPENDIX: MACROECONOMIC FORECAST COMPARISON WITH OTHER DOMESTIC FORECASTERS

**Figure 3.14: Comparison of Budget Forecast Errors for Nominal GDP Growth**



**Figure 3.15: Comparison of Budget Forecast Errors for Terms of Trade Growth**

