

# Treasury Forecasts: the Tendencies and Consequences of Inaccuracy

by Liam Ward-Proud

## INTRODUCTION

In each Budget, the treasury makes numerous forecasts. Percentage GDP growth is usually one of the most widely reported of these figures, and is used as a basis for future debt and borrowing forecasts. The accuracy of these forecasts is therefore a matter of high importance. This briefing aims to use simple mathematical and statistical analysis to assess the accuracy of such growth predictions, based on a sample of thirty forecasts.

Each spring Budget contains one forecast for each of the three following calendar years; Table 1 compares the different forecasts made by the treasury with the actual outplay of events each year.

TABLE 1	Percentage GDP Growth (1), (2)									
	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000
Type I Forecast	-3.5	2	3	2.25	3.25	3.25	2.25	2.25	2.5	3
Type II Forecast	2.5	2.75	3	2.75	3.25	3.25	3.25	2.5	2.5	2.5
Type III Forecast	2.75	3	2.5	2.75	3.25	2.75	2.5	2.5	3	2.5
Outturn	-5	0.75	3	2.75	1.75	3	2.25	1.75	2.25	3

## TERMINOLOGY

The following definitions are my own:

- “Type I Forecasts” are forecasts made in the spring budget for the ensuing year only. For example, the Type I Forecast for 2003 was made in the 2003 Budget.
- “Type II Forecasts” are forecasts made in the spring budget for the next calendar year only. For example, the Type II Forecast for 2003 was made in the 2002 Budget.
- “Type III Forecasts” are forecasts made in the spring budget for the year after next. For example, the Type III Forecast for 2003 was made in the 2001 Budget.
- “Outturn” is simply the actual growth for the year, the figure for percentage GDP growth for the previous year contained in each Budget.

GDP growth is measured as the percentage increase on last year's GDP. Type I Forecasts are therefore calculated using an outturn figure, last year's growth, whereas Type II forecasts are calculated as percentage growth on the growth forecast for the ensuing year. Type III Forecasts are therefore calculated as percentage growth on the growth forecast in the Type II Forecast. The forecast for percentage GDP growth three years away relies on all intermediate forecasts, so the uncertainties rack up.

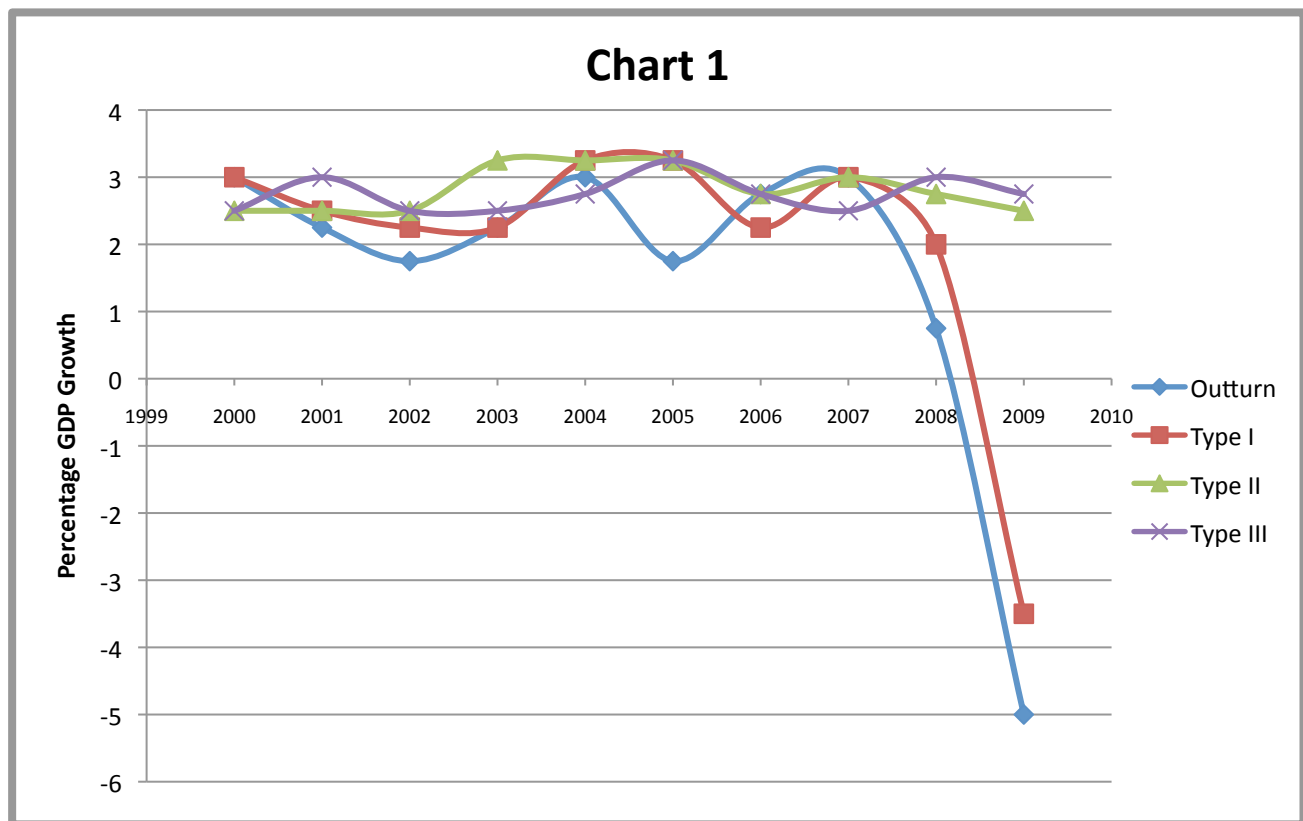


Chart 1 plots the three types of forecast against the outturn growth rate. From the chart, it is possible to point to a few overall trends that can be studied in greater detail. The trends to be studied include the correlation between forecast and outturn figures, the extent to which forecasts are more accurate in times of lower economic volatility and the extent to which the forecasts may be biased towards overestimation.

With some basic statistical tools, it is possible to analyse the data with the aim of coming to some conclusions about the tendencies of Budget forecasts for GDP growth.

## MEASURING MOVEMENTS

One way of assessing the overall accuracy of the forecasts is to look for a correlation between the forecast and the outturn. Pearson's correlation coefficient measures how well two sets of data correspond to each other in relative terms, producing a number between -1 and 1 to represent the correlation between the two sets of data (3).

A correlation between two variables does not imply causation. The degree of correlation will illustrate the extent to which each type of forecast tends to 'co-vary' with the growth outturn.

<b>TABLE 2</b>	<b>Correlation Coefficient (4)</b>
<b>Type I vs. Outturn</b>	0.97
<b>Type II vs. Outturn</b>	0.36
<b>Type III vs. Outturn</b>	-0.15

The correlation coefficient for Type I Forecasts is relatively high, suggesting that the numbers tend to move in sync with actual GDP growth. In many cases of forecasting it is easier to predict over a short timescale than a longer one, explaining the differences in accuracy between the different types of forecasts. Also, by the time of the Budget (late March/early April), the treasury usually has estimates for the growth experienced in the first quarter of the year, reducing the timescale even further.

The correspondence between forecast and reality seems to diminish when the forecaster is attempting to predict events that are temporally further afield. The Type II Forecast correlation coefficient seems to suggest little or no correlation at all, while for the Type III Forecasts, the correlation is closer to negative than positive. Once again, it must be stated that the correlations (whether negative or positive) do not suggest that forecasts are causally linked to the outturn.

The correlation coefficient for Type III Forecasts is particularly alarming. That it is negative is not so significant, since the figure is close to zero. But the correlation is not even close to being a strong, positive one, signalling that the movements of the two sets of numbers have almost nothing to do with each other.

Comparing the correlation of the forecasts and outturns is interesting in assessing how well the forecasts and outturn track each other as a group, but it is possible to look in closer detail at the distribution of the differences between each forecast and the outturn. It is possible to go into more detail as to the distributions of the individual forecasting errors.

## **ACCURACY AS A FUNCTION OF STABILITY (MORE CORRELATION)**

A feature of much of the debate on the issue of forecasting is that predictions seem to underestimate the possibility of sudden, high-impact leaps in both positive and negative directions. The sample data can be used to investigate this idea by comparing the percentage change in GDP growth (5) of the outturn with the absolute forecast error (6) of the corresponding forecast.

If, as one would suspect, economic volatility has a relationship with forecast errors, the two figures will have a positive correlation.

TABLE 3	Absolute Forecast Error									
	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000
Type I	1.5	1.25	0	0.5	1.5	0.25	0	0.5	0.25	0
Type II	7.5	2	0	0	1.5	0.25	1	0.75	0.25	0.5
Type III	7.75	2.25	0.5	0	1.5	0.25	0.25	0.75	0.75	0.5
Growth Rate	7.67	0.75	-0.09	-0.57	0.42	-0.33	-0.17	0.22	0.25	-0.5

Table 3 compares the growth rate, a measure of economic volatility, with the forecast error. If, as one would suspect, higher economic volatility tends to be inaccurately forecasted, there will be a positive correlation between the growth rate and the absolute forecast error. Pearson's correlation coefficient can be used again here:

TABLE 4	Correlation Coefficient (4)
Growth Rate vs. Absolute Forecast Error - Type I	0.62
Growth Rate vs. Absolute Forecast Error - Type II	0.98
Growth Rate vs. Absolute Forecast Error - Type III	0.99

The correlation is moderately positive for Type I forecasts and very strong for Types II and III. This supports the hypothesis that treasury forecasts are more likely to be accurate when the GDP Growth rate is low and steady, while the forecasts are much more likely to be inaccurate when the growth rate is higher, especially for Type II and III forecasts. This suggests that treasury forecasts are not likely to accurately predict any large changes in GDP in either direction.

Bearing in mind that the treasury bases forecasts for borrowing and debt levels on forecasts for GDP growth, this is particularly worrying. Forecasts for 'public sector net debt' and 'public sector net borrowing' extend up to four years after the spring budget; the figures are presented as a percentage of future GDP. The analysis suggests that the spring budget is highly unlikely to predict any dramatic changes in GDP, meaning that future forecasts for debt and borrowing can be along way off target. If, as happened in 2008 and 2009, the long-term forecasts for GDP growth are well above the outturns, government borrowing and spending plans will have to be completely revised in light of GDP outturns. Government fiscal plans, as well as plans for reducing the nation's structural deficit and national debt will be completely scuppered because of the failure to adequately plan for a slowdown in growth.

Forecasting is hard, and almost always inaccurate. Accordingly, any forecasts that are of crucial importance to future borrowing and spending figures, such as GDP growth, should be revised downwards for purposes of future spending and borrowing plans. This means the government will have less money to promise for future projects, but it would help guard against the large, unforeseen increases in borrowing and debt levels.

The treasury recognises this, stating:

*“The fiscal projections continue to be based on the assumption that trend output growth will be 1/4 percentage point lower than the centre of the forecast range, which is designed to add caution;” Budget 2010, page 182.*

The data, however, suggests that this safeguard is not nearly cautious enough. The difference of a quarter of a percent of GDP is negligible compared to the perceived errors, which are in some cases over seven and a half percent of GDP above the actual outturn. There is no systematic way of setting a better adjustment, but a quarter of a percent does not seem to be in any way an adequate safeguard against inaccurate forecasts.

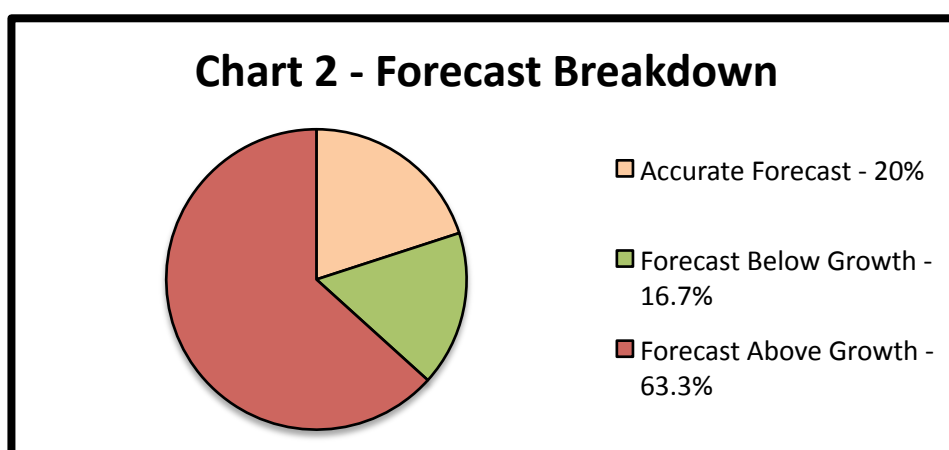
## OVERESTIMATION, TESTING THE HYPOTHESIS

A danger identified so far, it seems, is that ‘overestimates’ of growth can lead to unrealistic assumptions about the health of the economy, which can counteract attempts at fiscal discipline. The contention that overestimation is more likely than underestimation can be tested against the sample data.

All forecasts will be either accurate, an underestimate or an overestimate. By subtracting the outturn figure from the forecast figure, it is possible to see into which category each individual forecast falls. ‘+1.25’, for example, means that the treasury forecast for percentage GDP growth was 1.25% of GDP higher than the outturn figures. This data is represented in the table below.

TABLE 5	Forecast Minus Outturn									
	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000
Type I	+1.5	+1.25	0	-0.5	+1.5	+0.25	0	+0.5	+0.25	0
Type II	+7.5	+2	0	0	+1.5	+0.25	+1	+0.75	+0.25	-0.5
Type III	+7.75	+2.25	-0.5	0	+1.5	-0.25	+0.25	+0.75	+0.75	-0.5

Table 5 presents the actual difference between the forecast and the outturn, without taking absolute figures. This allows us to measure the distribution of forecast errors between overestimates, accurate forecasts and underestimates. The distribution can be seen more clearly in Chart 2:

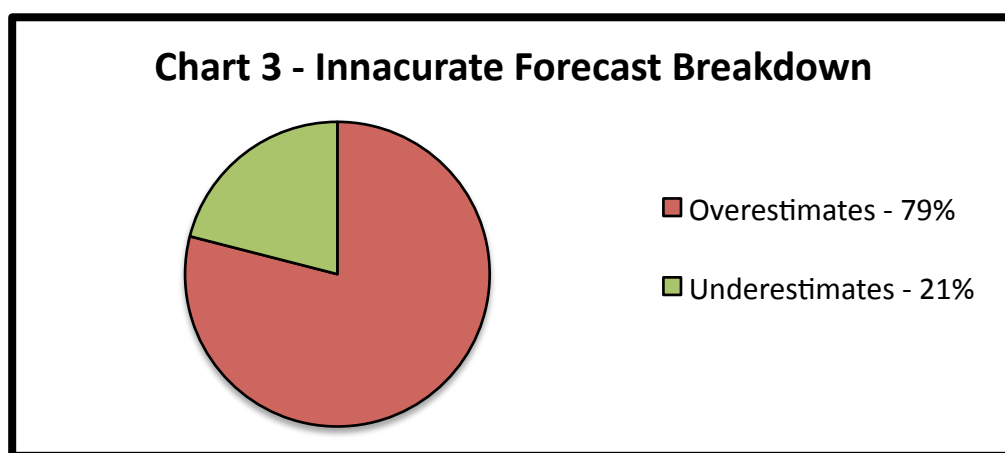


From the data it seems that treasury forecasts tend to predict higher growth than the outturn figures more often than not. For these ten years, six out of ten one-year treasury forecasts exceeded actual growth outturns; seven out of ten two-year forecasts and six out of ten three year forecasts exceeded the growth outturn. In total, of the thirty forecasts sampled, 63.3% exceeded actual growth outturns, 20% were accurate, and 16.7% underestimated actual growth.

Inaccuracy is to be expected, but the data suggests that treasury forecasts are far more likely to be overestimates than accurate or underestimates.

The treasury claims that the process is not biased in either direction; the sampled data allows us to test this claim. Let us call this claim the 'unbiased forecast hypothesis', implying that the probability of an overestimate is equal to that of an underestimate. This assertion can be tested and compared with the sample figures.

It may help to think about the issue in the following way. The accurate forecasts are not so relevant to this hypothesis; there are enough inaccurate forecasts for the makeup to be studied, so the accurate ones can be removed for now. If the forecasting process were not biased in either direction, one would expect the remaining inaccurate forecasts to be distributed evenly between overestimates and underestimates. But this is clearly not the case; in fact, 79% of the non-accurate forecasts are overestimates while 21% are underestimates. Chart 3 presents this imbalance:



## HYPOTHESIS TESTING

There is a clear asymmetry between the underestimates and overestimates, suggesting a bias towards overestimation. Working out the probability that an unbiased process would produce such results can test this potential bias. (7)

If the treasury forecasts are truly unbiased, the probability that the forecasting errors turned out as they did should be high enough that we are not inclined to reject this hypothesis. As it turns out, testing this

hypothesis against the data produces a probability so small that there is little choice but to reject the hypothesis.

In fact, there is only a 0.33% (8) chance of an unbiased process producing such an asymmetric outcome. The extremely low probability leads us to firmly reject the hypothesis that Budget percentage GDP growth forecasts are unbiased and allows us to suggest, given that 79% of inaccurate forecasts in the sample are overestimates, that the forecasts tend to be overestimates of actual growth.

There seems to be a systemic bias toward overestimating percentage GDP growth. Given that the treasury seems to be less likely to predict large swings in GDP well, and that the process seems biased towards overestimation, the ramifications for fiscal discipline are serious. The figure of one quarter of a percentage point used by the treasury to add caution should be greatly increased.

## **A WORD ON THE SAMPLE**

It is vital to remember the conclusions reached are based on a sample, and, as such, are subject to suspicion regarding any idiosyncratic contingencies of the sampled time-series. The temptation to automatically generalise from this sample to all types of forecasts must be resisted.

However, the sample of 30 forecasts, including three types of forecast, is just about large enough to draw inferences from; any smaller, and the inferences may be slightly dubious. It is also notable that, 2009 discounted, the sampled period is one of the most stable periods of growth in the country's history. If forecasting is so inaccurate in this period, it seems safe to assume that the perceived biases will hold beyond the sampled time period. Indeed, the current treasury forecast for 2010 growth is arguably looking increasingly unrealistic in light of the poor 0.2% growth for the first quarter.

The strength of this sample is that it includes three qualitatively different types of forecasts, meaning that the conclusions can be postulated more generally than otherwise.

The fact that inaccuracies exist does not constitute a criticism of current treasury staff. I am not arguing that, based on these figures, the current treasury is uniquely bad at forecasting. I have little confidence that any group of people in the same situation would be able to do better.

## **CONCLUSION**

Treasury growth forecasts have some serious flaws. The key findings of this research suggest that far more caution should be used when using such forecasts, and are summarised below.

- Type I Forecasts correlate fairly well with growth outturns. Forecasting errors are evident, but relatively lower than other types of forecast. Type I Forecasts tend to be the most accurate of the three budget forecasts, probably because of the shorter time-scale for uncertainty.
- Type II forecasts do not correlate very well with growth outturns and can have relatively high absolute forecasting errors. Type II Forecasts are thus considerably less accurate than Type I, probably because of the longer time frame for uncertainties and the fact that the percentage growth figure is based itself on a forecast. The inaccuracies of the Type I forecast therefore carry over into Type II.
- Type III forecasts do not correlate with growth outturns and also have high absolute forecasting errors. Type III forecasts thus tend to be very inaccurate. The inaccuracies of the two intermediate forecasts will carry over into the Type III, causing both absolute accuracy and correlation to be very low.
- Each type of forecast tends to be more accurate when the GDP growth rate is lower and far less accurate when the growth rate (a rough measure of volatility) is higher. The strong positive correlations between growth rate and absolute forecast error for Type II and III forecasts suggests that the treasury is very unlikely to foresee large swings in GDP growth, undermining not only the relevance of the forecasts to these high-consequence events but also the ability of the treasury to forecast and control borrowing and debt levels. This should be adjusted for by increasing the margin between GDP forecast and the estimate used for fiscal planning to a figure higher than a quarter-percent.
- A consistent bias towards overestimation is evident in the sampled forecasts. 79% of non-accurate forecasts are overestimates. The probability of this outcome being produced by a 'neutral', non-biased process is low enough to reject the idea that the process is unbiased and, by implication, assert that there is strong evidence for a systemic bias towards overestimation. Government projections should adjust in order to take into account this evident bias, especially in times of greater economic volatility.

That final point is of particular importance. To take a recent example, at the time of the 2010 Budget, the government pledged to 'halve the deficit in four years'. This means that by the 2014 budget, the difference between spending and receipts would have to be at least as low as 5.9% of GDP. Notice that this figure could mean drastically different amounts according to the actual GDP for the year in question. When GDP is lower than expected, demand for public services is still the same, as are the budgetary requirements of defence, policing and other government departments. So the amount of money the government is required to spend will stay broadly the same, but it will have less money than predicted as a percentage of GDP. Depending on what happens to GDP, that 5.9% could be far more than the government needs in order to meet its current spending commitments, or far less. We have seen that it is likely to be less, given that the government will probably overestimate GDP growth; this



will result in unforeseen and unplanned cuts in public spending, or simply the abandonment of the target for reducing borrowing.

Now, imagine that the treasury had used very pessimistic forecasts for GDP growth. If, as is likely, GDP growth turns out to be lower than the headline forecast, borrowing and spending plans can be broadly stuck to because they will have been projected using a low figure. If growth is higher than expected, the government will be able to pay off some of the national debt. I think it is clear that spending and borrowing projections should be based on a GDP growth forecast that is adjusted down by far more than a quarter percent.

## **NOTES**

1. Figures adapted from Budget Reports going back to 1998, available on HM Treasury Website.
2. The treasury sometimes gives a range for % growth in their forecasts. Where this is the case, I have used the midpoint of the upper and lower bounds of the forecast range to get one number representative of the range. The range is never more than 0.5%.
3. The correlation coefficient will be close to 1 when the two sets of numbers 'co-vary' well; in other words when one variable goes up, the corresponding figure from the second set will move up to a similar degree. The correlation coefficient will be close to -1 when the two sets of data vary in opposite directions; in other words when a figure from the first set goes up, the corresponding figure from the second set will go down. A correlation coefficient close to 0 would suggest that the two sets of data are uncorrelated, that the fluctuations are inconsistent and tend to cancel each other out.
4. All correlation coefficients are rounded to two decimal places.
5. The percentage change in GDP growth for a given year is calculated as last year's percentage growth minus the given year's growth, all divided by last year's percentage growth. It is here used as a measure of economic volatility.
6. The absolute forecast error is calculated as the forecast growth minus the outturn growth. The absolute value of the figure is taken, ensuring all the figures are positive. The absolute forecast error is a measure of the amount by which the forecast was wrong, disregarding whether it was an overestimate or an underestimate.
7. The 'unbiased forecast hypothesis' can be tested against the sample data to assess its plausibility. The fact that inaccurate forecasts can only be overestimates or underestimates permits modelling by the binomial distribution. If treasury forecasts truly were unbiased, the

expected probability of getting an overestimated forecast would equal that of getting an underestimate - both would be 0.5. The number of trials here is twenty-four, being the amount of non-accurate forecasts. The hypothesis is that the treasury forecasts are unbiased and from this we derive the parameters, presented as:

$$F \sim \text{Bin}(24, 0.5)$$

That is to say, the forecasts are distributed binomially with 24 sampled forecasts and the probability of an overestimate at 0.5. The probability derived thus is 0.003305376, using the given parameters, modelled Binomially and calculated using Excel's BINOMDIST function.