Debt Office Commentary

Forecast evaluation 2013–2018

Mattias Bruér, Senior Economist, Economic Analysis[1]

How accurate are the Debt Office's forecasts in comparison with those of other institutions? In this edition of Debt Office Commentaries, the Debt Office is compared with 17 other institutions during the period 2013 to 2018. The Debt Office's macro forecasts turned out well. According to the mean rank, which summarises the forecasts' accuracy, the Debt Office ends up in fifth place. Viewed on an annual basis, our macro forecasts were most accurate in the years 2013, 2014 and 2017. For the years 2015 and 2018, the macro forecasts were of average accuracy and for 2016 accuracy was poorer than average. The Debt Office's forecasts of the growth of household consumption have consistently found a place in the very top tier. However, there is no equally clear pattern for other macro variables. The Debt Office's GDP forecasts were in the top tier for the years 2013 and 2017 but were in the bottom tier for 2015 and 2016.

Regarding the forecasts of central government finances (fiscal forecasts), there are often only small differences between the forecasts made by the Debt Office, the Ministry of Finance, the National Institute of Economic Research and the Swedish National Financial Management Authority. The Debt Office's fiscal forecasts were successful for 2013 and 2014 and were averagely accurate for 2018. During the period 2015 to 2017, the fiscal forecasts were poorer than average. The results thus indicate that there is a correlation between successful macro forecasts and accurate fiscal forecasts, perhaps with a certain time lag.

What is the significance of future outlook when times are uncertain?

During the post-war period, the economic outlook has never changed at a quicker pace than that occurring at the moment. Both Sweden and the rest of the world are in a synchronised economic crisis in which measures and restrictions to limit the coronavirus pandemic are sharply slowing down activity. In the shadow of tremendous human suffering, existing economic forecasts have needed to be revised and subsequently updated. The forecasts are thus transitory and must be quickly replaced as new information becomes available.

For those of us working with forecasts, the conditions are truly challenging and uncertainty is enormous at the moment. At the same time, economic forecasts are basically always uncertain; as a forecaster you are often surprised and are perpetually learning. It can therefore sometimes be useful to take a step back and examine what went well and what worked less well in previous forecasts and hopefully learn from the findings. Having forecasts at the forefront is particularly important when they form the basis of economic and political decisions. For the Debt Office, accurate and well-founded forecasts are a prerequisite for achieving the overall objective of central government debt management: to minimise the cost of government debt in the long term without the risk becoming too high. The long-term perspective is also one of the reasons why this study focuses on forecasts covering a slightly longer horizon.

In this edition of Debt Office Commentaries we are – for the first time – undertaking a formal and rigorous evaluation of the Debt Office's macroeconomic forecasts and forecasts of central government finances (fiscal forecasts) to see how they relate to those made by other institutions. Economic forecasts are in fact often evaluated; the Riksbank, the Ministry of Finance and the National Institute of Economic Research regularly undertake evaluations in which the forecasters are compared with one another. However, the Debt Office's forecasts have never been included in these comparisons.

How the evaluation is conducted

The analysis is mainly based on data collected by the National Institute of Economic Research. In the study, a comparison is made of the forecasters' full-year forecasts of GDP growth, household consumption, unemployment, employment growth, inflation, payroll, the central government net borrowing requirement, and central government net lending. The net borrowing requirement and net lending are measured in SEK billions. Unemployment is measured as the annual average of the proportion of unemployed. Other variables are measured as the annual average of the annual percentage change.

In total, 17 forecasting institutions are included in the study. Not all forecasters report forecasts for all variables. Which forecasters are evaluated therefore depends on the variable being analysed. For the macroeconomic variables, forecasts are compared that have been made by the Debt Office, (RGK), the Government (Gov), the National Institute of Economic Research (NIER), the Swedish National Financial Management Authority (ESV), Riksbanken (RB), Nordea (Nor), Handesbanken (SHB), Skandinaviska Enskilda Banken (SEB), the Swedish Association of Local Authorities and Regions (SKR), the Confederation of Swedish Enterprise (SN), Unionen (Un), Swedbank (Swed), the Swedish Trade Union Confederation (LO), the European Commission (EU), the Swedish Retail Institute (HUI), the Swedish Public Employment Service (AF), and Danske Bank (DB).

For forecasts of payroll, the central government net borrowing requirement, and central government net lending, the comparison is made between the Debt Office, the Government, the National Institute of Economic Research and the Swedish National Financial Management Authority.

Each analysis is based on the first preliminary outcome of the respective variable. This mainly matters for GDP and household consumption because the national accounts in particular can be subject to extensive revisions.

How the forecasts are evaluated
The difference between outcome and forecast is called forecast error. Using the forecast errors, different evaluation measures can be calculated that partly shed light on different aspects of how well the forecasters have succeeded with the forecasts. In this report we have chosen to present the following evaluation measures:

• **Mean error** – the average of the forecast errors
• **Mean absolute error** – the average of the absolute values of the forecast errors
• **Mean rank** – calculated from aggregated mean absolute errors[2]

Mean errors show whether the outcomes have been systematically overestimated or underestimated in the forecasts, which is also called bias. When the mean error is calculated, forecast errors with different signs cancel each other out. When the accuracy is to be evaluated, we are therefore more inclined to use another evaluation measure, such as mean absolute error.

With regard to the accuracy of the forecasts, we have chosen not to report mean absolute errors explicitly, because the combination of many variables and many forecasting institutions soon makes it difficult to get an overview of the results. Instead, the forecasts are ranked based on the size of the mean absolute errors. Different variables have different degrees of complexity for making forecasts. Therefore, the variables' forecast errors should not be added up or directly compared. When we want to form an opinion about the overall forecasting ability, the mean rank can be calculated instead. The mean rank is not affected by the above problems, but neither does it show if the difference in forecasting ability is large or small.

Forecasts for the coming year are analysed
The forecasters publish their forecasts at different times. In practice, this results in access to different amounts of information, which can affect the results. The more information that is available, the greater should be the accuracy of the forecasts. However, forecasts that stretch longer into the future are less affected by new outcomes and indicators, compared with forecasts for the current year. Therefore, in an attempt to make a fair study, only forecasts made for the next year are evaluated. The forecast horizon will thus be up to 24 months.[3]

---

2 *The mean rank is a better way to aggregate the variables than to calculate the mean of the mean absolute errors of the variables. The reason for this is that mean absolute error can/should not be compared between variables because different variables have different levels of complexity for making forecasts.*

3 *Forecast errors for forecasts done during the current year decrease as the forecast horizon becomes shorter. This result is logical because quarterly outcomes from the national accounts and other statistics are important input for short-term forecasts. For the forecasts for year t+1 however, there is not as clear a trend of forecast errors decreasing as the forecast horizon becomes shorter (see Bergvall, 2005).*

Focusing on forecasts with slightly longer horizons aso has other advantages. When the forecasts form the basis of fiscal and monetary policy decisions, it is the growth prospects of the next few years that are the most relevant, rather than the growth during the current year. For the Debt Office too, forecasts with a longer horizon are more relevant because it takes time to implement changes in central government debt management.[4]

The Debt Office's macroeconomic forecasts

Evaluation of the years 2013−2018

The players' forecasts are evaluated both for the entire period and for each year. By analysing a longer period of time, the impact of chance on the results is reduced and more general conclusions can be drawn about the ability to forecast[5]. Annual evaluations also have advantages. Annual evaluations can show whether the aggregate results are driven by individual forecast years.

If individual years drive the aggregate results, it is often possible to track the results and explain them afterwards, which in the long run can provide deeper insight into the forecasting process. One example is the Debt Office's GDP forecasts in which the aggregate results are driven by large forecast errors during an individual year (see below).

If a player has often made accurate forecasts, a 'forecast of forecast ability' can be that the player's forecasts will aso be successful in the future. One example is the Debt Office's forecasts of household consumption, which have consistently been in the top tier (see below).

GDP has been systematically underestimated

Tables 1 and 2 show the mean errors for the entire period in total and for each year respectively. For GDP, there has been a tendency to underestimate next year's outcome. For the entire period, all forecasters except the Riksbank have underestimated the outcome on average. The Debt Office's GDP forecasts are among those that underestimated this development the most (0.4 percentage points).

The annual compilation shows that the result is largely explained by the fact that growth was greatly underestimated in 2015. Also, for 2016, GDP was considerably underestimated by several forecasters, including the Debt Office.

For household consumption, the forecast errors are in the opposite direction of those for GDP. With the exception of Nordea, all forecasters overestimated consumption on average. However, the Debt Office's forecasts of household consumption are among those with the least mean error. In addition, in terms of the annual development, the Debt Office has generally made small mean errors in its consumption forecast.

---

4 *Using the same logic, t+2 forecasts can be reported, but because the forecast base is significantly greater with t+1 forecasts, we have chosen this delimitation.*

5 *Forecasters tend to underestimate the pace of recovery periods and overestimate it in periods when growth slows, as the force of the turnaround is seldom seen in the first outcome, only first when the national accounts are revised. It is therefore an advantage that the years around the financial crisis (2008−2009) are not included in the study.*

The exception was 2018 when the Debt Office overestimated household consumption by 0.93 percentage points on average. Most other players also grossly overestimated household consumption in 2018. The most important factor in why household consumption was overestimated so much that year is the so-called Bonus Malus system for the taxation of newly registered cars and light vehicles that came into force on 1 July 2018. This affected households' purchases of cars and thus household consumption to a far greater extent than expected.

Employment has been surprisingly strong
The fact that the labour market developed unexpectedly strongly during the analysis period is reflected in the results. Looking at the entire period, the players have collectively underestimated employment growth by 0.6 percentage points on average (see Table 1). It is not individual forecast years that drive the underestimation. On the contrary, employment has exceeded expectations year after year, as Table 2 shows.

At the same time, the forecasts of unemployment have come close to the outcomes. The combination of underestimating employment at the same time as the unemployment forecasts had small mean errors can be explained by the fact that the strong performance of the labour market was also unexpected for the forecasters.

The Debt Office's macro forecasts are well-placed in the ranking
In terms of accuracy, the Debt Office's macro forecasts are well-ranked, except for GDP and employment for which the Debt Office's forecasts, viewed over the entire period, fall below the average (Table 3). We have had the best success with the forecasts of household consumption and CPI, while the unemployment forecasts end up close to the average.

The fact that the Debt Office's GDP forecasts end up among the least accurate is driven by the forecast errors for 2015 when growth was significantly higher than the levels expected by the forecasting institutions. The first preliminary outcome was 4.1 per cent, and the Debt Office's forecasts were between 2.2 and 3.0 per cent. The Debt Office made one of the largest underestimations of growth, which certainly hurt its position in the ranking. The Debt Office also significantly underestimated GDP regarding 2016. For both 2015 and 2016, the Debt Office ended up in 15th place out of 17 when it came to forecasting GDP.

Overall, however, the Debt Office's forecasting ability regarding macro variables has been above average. Based on the mean absolute error, its macro forecasts end up in sixth place when the mean rank is compared with other forecasting institutions (see Figure 1). According to the annual ranking of the macro forecasts, 2013, 2014 and 2017 were particularly successful (see Table 4). On the other hand, accuracy was generally worse in 2016 and in other years our macro forecasts ended up in the middle segment.

Figure 1. Mean rank for macro variables
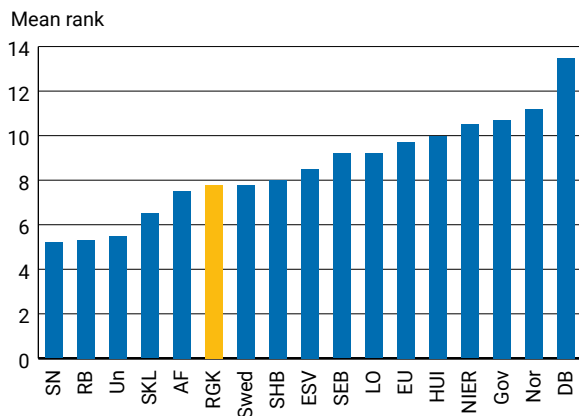
Mean rank

Mean rank is based on the mean of each institution's ranking for the individual macro variables. The results are based on absolute average errors in respect of annual t + 1 forecasts for the period 2013–2018.

The net borrowing requirement and central government net lending

This section analyses the net borrowing requirement and central government net lending for the period 2013–2018. Comparisons are made between the Debt Office's forecasts and those by ESV, NIER and the Government.

Figure 2. Mean error for the net borrowing requirement, t-1 forecasts

SEK Billion

Debt Office — ESV
NIER — Swedish Government
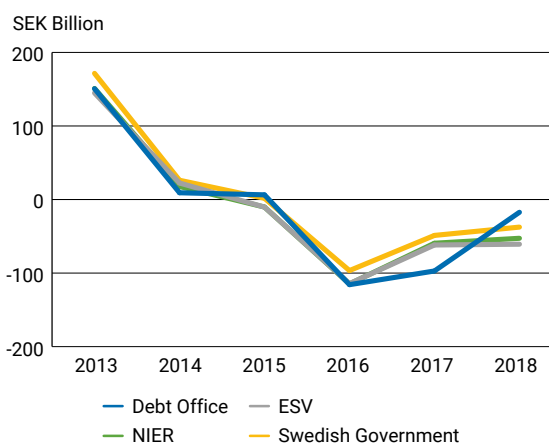
Figure 3. Mean error for central government net lending, t-1 forecasts

SEK Billion

Debt Office — ESV
NIER — Swedish Government
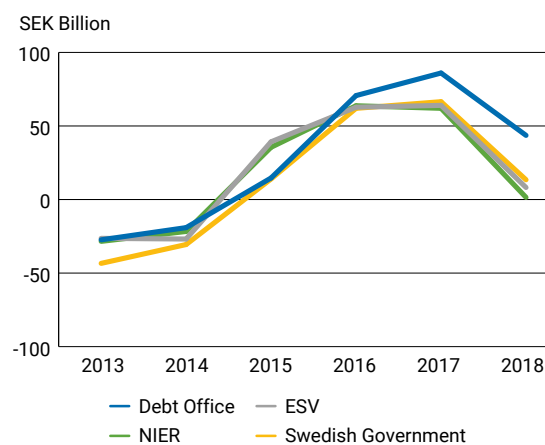
Positive mean errors indicate that the net borrowing requirement/central government net lending has been underestimated.

Net borrowing requirement: large mean errors in individual years cancel each other out over the entire period

Overall, for the period 2013 to 2018, the mean errors in the net borrowing requirement forecasts have been small. On average, the Debt Office, NIER and ESV overestimated the net borrowing requirement by just over SEK 10 billion, while the Government's forecasts showed an underestimation of just under SEK 3 billion.

In individual years, however, the mean errors have been large. The largest error during the period was in 2013. This is mainly explained by loans to the Riksbank totalling SEK 106 billion, of which there was no knowledge in 2012. The Riksbank raised loans through the Debt Office to increase the foreign exchange reserve, which the Riksbank can do at short notice in accordance with the Budget Act.

The mean errors were also large in 2016 and 2017, but then the outcomes were much lower than all players' forecasts. The most important explanation for the Debt Office overestimating the net borrowing requirement in 2016 was capital investments in tax accounts. The Debt Office realised what was happening in tax accounts at the end of 2015, but the capital investments were not included in the forecasts made at the time. In addition, the economy grew significantly faster than expected, which meant that tax income, primarily from consumption and capital, increased. Capital investments in tax accounts contributed to the net borrowing requirement being overestimated once more in 2017. The Debt Office assumed that there would be an outflow from tax accounts, whereas the inflow continued. But the main reason for the overestimation was that the economy continued to strengthen. This meant that tax income was higher, and expenditure slightly lower, than expected.

The forecasts of the net borrowing requirement show great similarities between the institutions, as Figure 1 illustrates. The similarities between the forecasts by ESV and NIER are explained by the fact that these organisations collaborate with regard to tax forecasts. It is really only the Debt Office's forecasts for 2017, and to a lesser extent 2018, that clearly deviate from the assessments made by others. In 2017, the Debt Office's mean error was larger than the other forecasters' and in 2018, the Debt Office's mean error was smaller. The larger deviation for 2017 compared with other players is likely due to the Debt Office making a different analysis of capital investments in tax accounts. The Debt Office analysed tax accounts before other players did and assumed that it was a matter of temporary tax payments that did not depend on the development of the underlying tax bases. This may have led us to first overestimate the effect of tax accounts but then underestimate the impact of the economic climate on tax income. Since we assumed that it was temporary, we expected that payments would be made during 2017 as a result of changed interest terms. Instead, deposits increased, which only became apparent in the final months of 2017.

Central government net lending has been underestimated in recent years

Central government net lending is not affected by lending to the Riksbank or capital investments in tax accounts. Net lending also allocates payments to the time when the economic activity takes place. In general, net lending is therefore considered to be a better indicator of underlying central government finances than the net borrowing requirement/budget balance.

Considering the entire period, the forecasters have underestimated central government net lending on average. The Debt Office's forecasts have had the largest mean error and those of the Government have had the smallest mean error. The annual evaluation shows that net lending was slightly overestimated in 2013 and 2014, while it was underestimated quite considerably for the years 2016 to 2018. The most likely explanation for this is that the forecasters still underestimate fluctuations in the economy and the impact these have on central government finances. Similarly for the net borrowing requirement, forecast errors were greatest during 2016 and 2017.

Figure 2 shows than even for central government net lending, the players' forecasts have had great similarities. What deviates most is that the Debt Office underestimated net lending more than other institutions for 2017 and 2018. The Debt Office also underestimated net lending for 2016, but the differences compared with other forecasters were small. However, the results for 2016 to 2018 are reflected in higher mean absolute errors and mean square errors for the Debt Office for those years.

A correlation between the macro forecasts and the fiscal forecasts

Overall, the Debt Office's macro forecasts are well-ranked. According to the mean rank that summarises the results for the entire period, our macro forecasts end up in sixth place out of 17 institutions. At the same time, in some years the forecasts have been more successful than others. For 2013, 2014 and 2017, the Debt Office's macro forecasts were particularly successful. For the years 2015 and 2018, the macro forecasts were of average accuracy and for 2016 accuracy was poorer than average.

The annual results for the fiscal forecasts show that there is a correlation between the accuracy of the macro forecasts and that of the fiscal forecasts, perhaps with a certain time lag. The Debt Office succeeded relatively well with the fiscal forecasts for 2013, 2014 and 2018 but was less successful for 2016 and 2017. Regarding 2015, the performance was mixed; the Debt Office's forecasts of central government net borrowing requirements were the least accurate, while our forecasts of central government net lending were the most accurate.

The advantage of presenting a ranking as in this study is that the results are easy to interpret and understand. In such a comprehensive study regarding the number of institutions, years and variables, it quickly becomes difficult to get an overview of the results. However, the disadvantage is that the ranking can visually exaggerate the differences between the forecasters. This applies not least with regard to the fiscal forecasts, for which, in several cases, it is not possible to determine with the naked eye which forecasters are most or least accurate.

Appendix 1: Tables

The forecasts of both the Debt Office and other institutions are evaluated in the following tables – in terms of forecast accuracy and the systematic errors (bias) that have occurred. Both macro and fiscal forecasts for the period 2013–2018 are reported here.

Table 1. Mean error for forecasts published t-1 for the entire period 2013–2018

	GDP	unemp	emp	cons	inf
RGK	0.41	-0.15	0.71	-0.16	0.18
NIER	0.08	-0.05	0.61	-0.36	0.30
ESV	0.32	-0.02	0.81	-0.32	0.26
Gov	0.02	0.09	0.57	-0.75	0.30
RB	-0.02	-0.02	0.63	-0.53	0.03
Nor	0.37	-0.18	0.87	0.04	0.32
SHB	0.02	-0.25	0.46	-0.53	0.04
SEB	0.05	0.04	0.58	-0.44	0.38
SKR	-0.19	0.07	0.69	-0.96	0.00
SN	0.12	-0.14	0.64	-0.36	0.20
Un	0.07	-0.17	0.40	-0.13	0.03
Swe	0.02	-0.08	0.57	-0.32	-0.16
LO	0.03	-0.22	0.57	-	0.28
EU	0.12	-0.08	0.48	-0.47	-
HUI	0.00	-0.09	-	-0.56	0.25
AF	0.06	-0.11	0.39	-0.34	0.33
DB	0.60	-0.21	-	0.21	0.87
Mean	**0.12**	**-0.09**	**0.60**	**-0.37**	**0.22**

	pay	BR	NL
RGK	0.41	-11	28
NIER	0.01	-13	20
ESV	0.35	-11	19
Gov	0.09	3	14
Mean	**0.21**	**-8**	**20**

Table 1 shows the mean errors at all the institutions compared. The mean errors are based on forecasts made the year before and are calculated on the first preliminary outcome for each variable. Positive mean errors mean that the forecasts underestimated the outcome on average.

The variables included are GDP, unemployment (unemp), employment (emp), household consumption (cons), inflation (inf), payroll (pay), central government net borrowing requirement (BR), central government net lending (NL).

For the net borrowing requirement and central government net lending, the mean error is stated in SEK billions. For other variables the mean error is stated in percentage points.

Table 2. Annual mean error for the Debt Office's forecasts for next year

	GDP	unemp	emp	cons	inf
2013	-0.10	0.10	0.86	-0.10	
2014	-0.30	-0.10	0.83	-0.03	
2015	1.47	-0.10	0.43	0.07	
2016	0.73	-0.48	0.42	-0.13	
2017	0.20	0.18	0.64	0.20	0.40
2018	0.13	-0.32	0.87	-0.93	0.10

	pay	BR	NL
2013	-0.60	151	-27
2014	0.30	9	-19
2015	0.23	6	15
2016	0.83	-116	71
2017	0.40	-97	86
2018	0.60	-17	44

Table 2 shows how the Debt Office's forecasts turned out (mean error) during the measurement period. Here, each year's forecast is compared with the actual result for the following year.

It is important to distinguish between what mean errors tell us (systematic errors in the forecasts or bias), which is what Table 1 and Table 2 show in different periods, and the accuracy shown in different periods in Tables 3 and 4.

The comparison is based on forecasts for the coming year. The tables show the ranking based on the mean absolute error. MR stands for mean rank and is a way to summarise the results.

Table 3. Accuracy of the forecasts 2013 to 2018

	GDP	unemp	emp	cons	inf	MR
RGK	12	8	12	1	4	7.4
NIER	5	10	9	13	15	10.4
ESV	4	3	15	4	13	7.8
Gov	16	11	5	15	9	11.2
RB	3	2	10	12	2	5.8
Nor	17	9	14	3	11	10.8
SHB	8	12	4	14	8	9.2
SEB	6	5	8	8	14	8.2
SKR	2	1	13	16	3	7.0
SN	1	4	11	5	5	5.2
Un	14	14	1	2	1	6.4
Swe	7	13	7	9	6	8.4
LO	10	16	6		7	9.8
EU	11	15	3	10		9.8
HUI	13	6		11	10	10.0
AF	9	7	2	7	12	7.4
DB	15	17		6	16	13.5

	pay	BR	NL	MR
RGK	4	4	4	4.0
NIER	1	3	3	2.3
ESV	3	2	1	2.0
Gov	2	1	2	1.7

Table 3 shows the ranking of all institutions for each variable for the entire period 2013–2018.

Table 4. Annual ranking of the Debt Office's forecasts for next year

	GDP	unemp	emp	cons	inf	MR
2013	1	3	6	1		2.8
2014	6	4	10	1		5.3
2015	15	3	13	1		8.0
2016	15	13	13	2		10.8
2017	2	8	5	2	11	5.6
2018	9	16	9	6	3	8.6

	pay	BR	NL	MR
2013	3	2	1	2.0
2014	3	1	1	1.7
2015	4	4	1	3.0
2016	4	4	4	4.0
2017	4	4	4	4.0
2018	2	2	4	2.7

Table 4 shows the Debt Office's annual ranking for each variable. Low numbers are better than high numbers. Which forecasting institutions are included in the comparison depends on which variable is evaluated. In some cases, there is no ranking in the tables because not all forecasters have reported forecasts for all variables.

For payroll, the net borrowing requirement and central government net lending, the forecasts from the Debt Office, the Government, the National Institute of Economic Research and the Swedish National Financial Management Authority are evaluated. For other variables, the forecasts from up to 17 institutions are evaluated.