Improving Forecasting Accuracy at the OBR: Submission to the Treasury Committee Enquiry 'The OBR: 15 years on'

Jennifer L. Castle, David F. Hendry and Andrew B. Martinez
Climate Econometrics, Nuffield College, University of Oxford, UK

January 2026

Highlights:

- The OBR has made systematic large forecast errors since 2010
- Our approach corrects large forecast errors occurring after unexpected shifts
- It can be implemented rapidly even if updating a forecasting system is difficult
- Illustrated here by improved forecasting of UK productivity relative to OBR

As specialists in forecasting theory and practice, we have studied why OBR's forecast errors were systematically large, can explain why, and have developed methods that could substantially improve their accuracy.

The Office for Budget Responsibility (OBR) is tasked with producing twice-yearly independent economic forecasts for the government which are adopted as the official government forecasts by the Chancellor of the Exchequer. As such, their forecasts wield a lot of power, being used to assess whether the fiscal rules on borrowing and debt as specified by the government are likely to be met over the 5-years' ahead forecast period. If the OBR forecasts suggest that the fiscal rules will be breached, the Chancellor adjusts the fiscal plans over the budget period to meet the constraints (or risk breaching the rules and lose credibility). The OBR forecasts, therefore, are critical to fiscal policymaking in the UK, determining the headroom available for the government to implement its preferred policy. Small changes in the forecasts over a 5-year horizon can result in frequent and significant policy adjustments.

There are many arguments as to whether such a system of fiscal rules evaluated on the basis of uncertain forecasts is a sensible system or not: see, e.g., Clift (2023). The introduction of the OBR (formally becoming a statutory body in 2011) was intended to improve transparency, accountability and credibility in policy making. However, our submission focuses on forecast accuracy, which can have unintended consequences for the credibility of fiscal policy making as forecasts that substantially deviate from the outturns question the very premise of the specified policy. The fiscal headroom is dependent on how accurate the forecasts are over the medium

to long-run (3-5 years), and small revisions to the forecasts can lead to large policy adjustments which may be based on erroneous assumptions, poor models, or inaccurate data. Hence, the OBR forecasts have come under scrutiny recently given their central role in fiscal policy making. The central component of the OBR forecasts is a large-scale macroeconomic model, originally developed by the Treasury in the 1970s but then handed over to the OBR since 2010. It is a traditional macroeconomic model in that it describes the economy in a set of accounting identities, behavioural equations and technical relationships: see OBR (2013). It has the advantage that it is not a DSGE model, unlike the Bank of England's COMPASS model: see Hendry and Muellbauer (2018) for a critique of DGSE models for forecasting. Furthermore, judgmental adjustments are made to the model-based forecasts. The OBR publishes an annual forecast evaluation report which assesses the forecast accuracy of its projections, and in 2025 the OBR started to compare their forecasts with external forecasts. The historical record of the OBR is that real GDP growth and productivity tended to be overestimated, and inflation was difficult to forecast with the OBR underestimating its volatility.

In this submission, we compare the forecasts of the OBR for productivity against the smoothed robust forecasts proposed by Martinez et al. (2022).

OBR productivity forecasts

Figure 1(A) shows the OBR one quarter to 20-quarter-ahead forecasts of UK productivity from 2010. These are persistently well above actual productivity, as if OBR expected the trend to return to the pre-financial crisis rate. The systematic forecast failure emphasises that the trend shift was ignored for many years.

Panel (B) contrasts those OBR forecasts with smoothed robust forecasts based on a method developed by Martinez et al. (2022), denoted SRW, which is the acronym for Smooth Random Walk. The SRW smooths over the previous two years to estimate its mean forecast. This forecasting device improves on the original robust approach in Hendry (2006), who demonstrated how to reformulate the forecasts produced by econometric models (of the kind OBR uses) to rapidly adjust for shifts.

Figure 2 plots the resulting root mean-square forecast errors (RMSFEs), a conventional measure of forecast accuracy. RMSFEs calculate the square roots of the average squared differences between forecasts and outcomes, so are in the same units as the standard error of the fitted equation. Figure 2(A) presents the pre-pandemic results which is essentially a replication of Martinez et al. (2022) and shows that the SRW has substantially better performance after roughly 5 quarters ahead. Figure 2(B) extends the results based on the old definition through 2025. While the forecasts now perform similarly till around 10 quarters ahead, there are increasingly large differences between OBR and SRW for longer-ahead forecasts. Figure 2(C) shows a similar pattern using the new forecast definition since 2022. Overall, the results highlight that forecasts naturally become more uncertain further into the future, unavoidable when multi-period ahead forecasts are required, as in a policy context, but the panels in Figure 2 highlight the increasing forecast failure of the longer-ahead OBR forecasts compared to SRW.

Explaining OBR forecast failures and how to avoid them

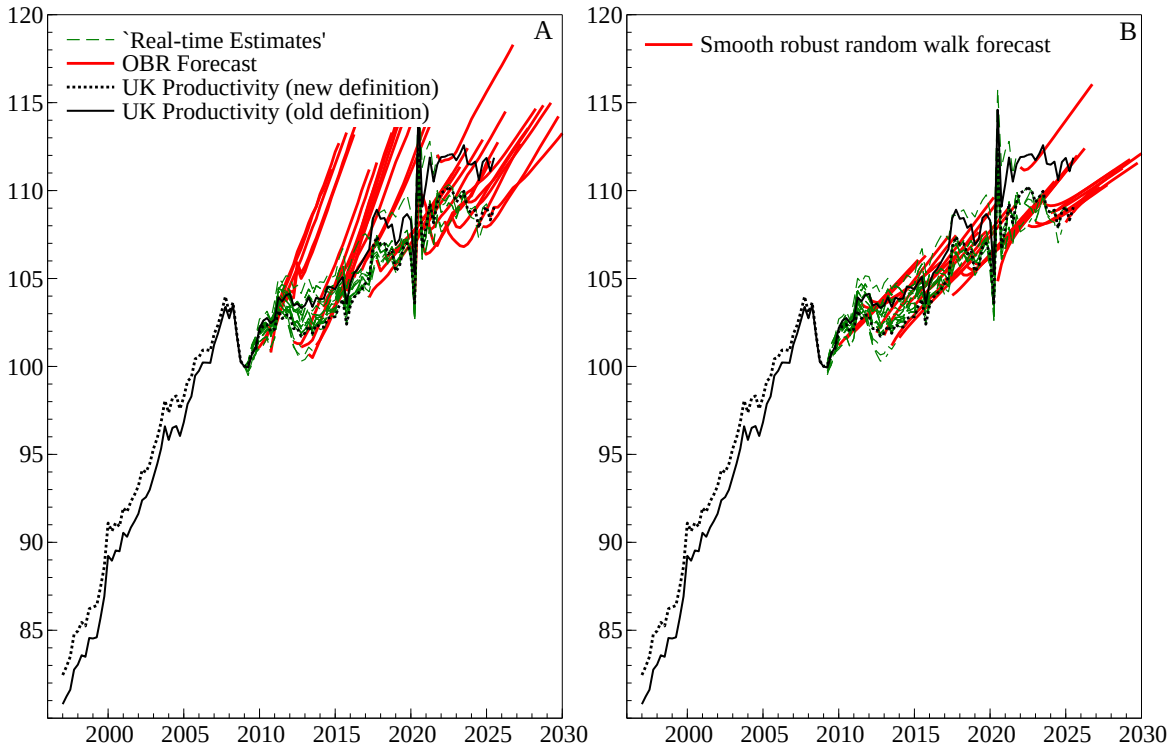

Figure 1: (A) OBR one quarter to five-year-ahead forecasts of UK productivity (output per hours worked); (B) Our corresponding smoothed robust forecasts. The productivity data are the final measured values, whereas 'real time' shows what was calculated at the time of their forecasts. The spikes are around the COVID-19 lockdowns.

A maquette of the large dynamic empirical econometric system used by the OBR, which nevertheless captures the key features of their approach for forecasting, is a single equation autoregressive model of the scalar process that generated the data. Such a simple model has an intercept, a feedback coefficient and a random unpredictable error. The parameters are unknown but can be estimated from past data, reasonably accurately when the parameters stay constant.

At each time T, called the forecast origin, forecasts from the estimated model can be produced from 1-step ahead to a horizon H periods in the future. Such forecasts will be reasonably close to the future outcomes when the process is unchanged. However, if a major shock like a financial or energy crisis occurs and shifts the mean (or trend) but the model is not updated, then forecast errors will converge on the change in the mean. When the mean falls, the model will repeatedly forecast upwards to 'correct' its trajectory back to the old mean (or trend), precisely the pattern seen in Figure 1(A).

When forecasts for based on a large interdependent system, shifts can alter many

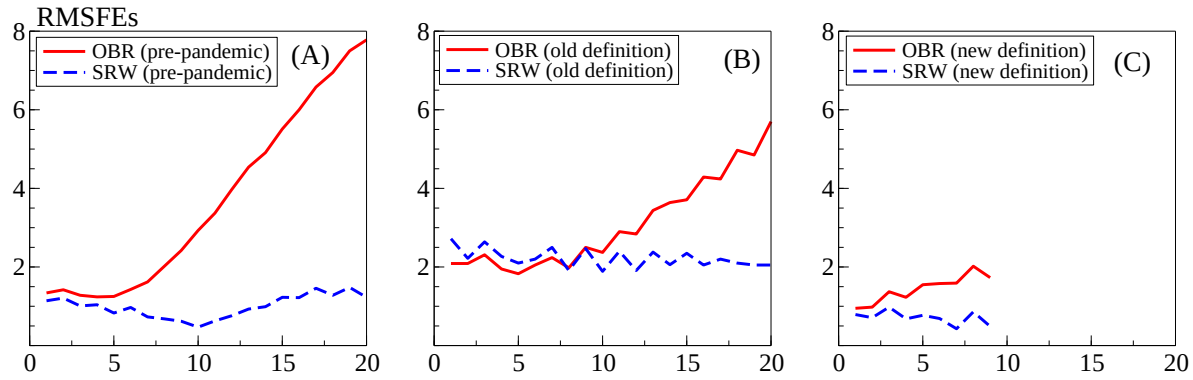

Figure 2: RMSFEs of OBR (solid line) and smoothed robust (denoted SRW, dashed line) one quarter to five-year-ahead forecasts of UK productivity. Panel (A): 2010Q1–2020Q1. Panel (B): 2010Q1–2025Q3 based on the original UK productivity measure. Panel (C): 2022Q3–2025Q3 based on the revised UK productivity measure.

of its equations, making updating such a model difficult. It may take many periods to collect enough information to discover what parameters changed and longer to obtain good estimates of their new values. The introduction of smooth robust forecasts can produce improved accuracy quite rapidly without needing to fix the model as seen in Figure 1(B). Consequently, we urge the OBR to carefully evaluate what features of their model lead to such poor long-run forecasts, and compare their forecasts to a smooth robust procedure for insights into where the inbuilt correction in their model is proving counter productive. Basing economic policy on a bad model and inaccurate forecasts is clearly inadvisable. However, the forecasts at least can be rapidly improved even when there is a sequence of shifts, as occurred for UK inflation over 2020–2024 with Covid-19 lockdowns, the introduction of effective vaccines, supply chain disruption, and Russia's invasion of Ukraine leading to an energy crisis and fossil fuel price explosions. Castle et al. (2025) record the important improvements that are possible in forecast accuracy by a related approach.

I. Appendix: data revision notes

The OBR produced their forecasts of productivity based on Real Gross Value Added excluding oil and gas (ONS code KLS2) and total hours worked (ONS code YBUS) through to March 2022. Starting in November 2022, the OBR changed the measure of productivity that they were forecasting, using Real Gross Domestic Product (ONS code ABMI) and total hours worked (ONS code YBUS).

Martinez et al. (2022) effectively assumed that YBUS was never revised and only used the latest version of total hours worked. However, total hours worked do get revised, albeit the revisions are much smaller than for output. In the above analysis, we use real-time historical vintages of YBUS in addition to both of the output measures and constructe real-time data vintages for both measures of productivity.

Our real-time forecasts of productivity are based on OBR's definition of productivity at each point in time, so we generate real-time forecasts through March 2022 based on the old definition, then based on the new definition starting in November 2022. We evaluate the forecasts based on the latest vintage of the respective definition (i.e., forecasts generated prior to November 2022 use errors for the latest vintage of the old definition; and for all forecasts since November 2022 errors use the latest vintage of the new definition.

Acknowledgements We are pleased to acknowledge financial support from Nuffield College. email: jennifer.castle@magd.ox.ac.uk, david.hendry@nuffield.ox.ac.uk and amartinez@american.edu. All calculations and graphs used *PcGive* (Doornik and Hendry, 2021) and *Ox Professional* (Doornik, 2018).

References

Castle, J. L., J. A. Doornik, and D. F. Hendry (2025). Could the Bank of England have avoided Mis-Forecasting UK Inflation during 2021–23? *International Journal of Forecasting 42*, 13–21. https://doi.org/10.1016/j.ijforecast.2025.07.001.

Clift, B. (Ed.) (2023). *The Office for Budget Responsibility and the Politics of Technocratic Economic Governance*. Oxford: Oxford University Press.

Doornik, J. A. (2018). *Object-Oriented Matrix Programming using Ox* (8th ed.). London: Timberlake Consultants Press.

Doornik, J. A. and D. F. Hendry (2021). *Empirical Econometric Modelling using PcGive: Volume I.* (9th ed.). London: Timberlake Consultants Press.

Hendry, D. F. (2006). Robustifying forecasts from equilibrium-correction models. *Journal of Econometrics 135*, 399–426. https://doi.org/10.1016/j.jeconom.2005.07.029.

Hendry, D. F. and J. N. J. Muellbauer (2018). The future of macroeconomics: Macro theory and models at the Bank of England. *Oxford Review of Economic Policy 34*, 287–328. https://academic.oup.com/oxrep/article/34/1-2/287/4781814.

Martinez, A. B., J. L. Castle, and D. F. Hendry (2022). Smooth robust multi-horizon forecasts. *Advances in Econometrics 43A*, 143–165. https://doi.org/10.1108/S0731-90532021000043A008.

OBR (2013). The macroeconomic model. Briefing paper no. 5. https://obr.uk/docs/dlm_uploads/Final_Model_Documentation.pdf.