Journal of Forecasting

WILEY

RESEARCH ARTICLE OPEN ACCESS

A Novel Approach to Forecasting After Large Forecast Errors

Jennifer L. Castle[1] | Jurgen A. Doornik[2] | David F. Hendry[2]

[1]Climate Econometrics and Calleva Project Magdalen College, University of Oxford, Oxford, UK | [2]Climate Econometrics Nuffield College, University of Oxford, Oxford, UK

Correspondence: David F. Hendry (david.hendry@nuffield.ox.ac.uk)

Received: 21 August 2025 | **Revised:** 12 October 2025 | **Accepted:** 16 October 2025

Funding: We are pleased to acknowledge financial support from the Research Council of Norway, project 324472, on "Model invariance and constancy in the face of large shocks to the Norwegian macroeconomic system" and Nuffield College, University of Oxford.

Keywords: impulse indicators | intercept corrections | rapid shift detection | trend breaks | UK inflation

ABSTRACT

A sequence of increasingly large same-sign 1-step-ahead forecast errors are most likely due to a sudden unexpected shift. Large forecast errors can be expensive, but also contain valuable information. Impulse indicators acting as intercept corrections to set forecasts back on track can be quickly tested for replacing outliers, a location shift or broken trend, greatly improving forecast accuracy. The analysis is applied to forecasting the UK's annual consumer price inflation which rose rapidly from mid-2021 to over 9% in 2022 after a series of essentially unpredictable shocks led to large forecast errors by the Bank of England.

JEL Classification: C2, C5, J3

1 | Introduction

There is a large literature on detecting and forecasting after breaks, striving to rapidly detect breaks and modify forecasting devices accordingly, well referenced in the papers in the special issue edited by Giannellis et al. (2025). It has proved difficult to do both so far as large forecast errors could be due to big outliers or mis-measurements, a sudden step shift in the mean of the process, or a trend break. Although such errors can be expensive, they also contain valuable information. In particular, a short sequence of increasingly large one-sided 1-step-ahead forecast errors as the forecast origin advances suggests an unexpected shift, most likely a trend break. Consequently, large forecast errors contain valuable and usable information about their source. Isolating information about the source of forecast errors has broader use in applications such as forecast-extending time series where filters are applied at the end of the sample, precisely when accurate information on structural breaks and outliers is needed (see, e.g., Findley et al. 1998, and its successor X13-ARIMA-SEATS). Exponential smoothing methods and the seasonal adjustment literature rely on the signal regarding the type of break to accurately extrapolate data at the forecast origin: Gardner (2006).

Forecasts can be "put back on track" at the forecast origin by impulse indicators acting as intercept corrections (ICs), as the value of the impulse indicator is the forecast error at that time point (see, e.g., Hendry and Clements 1994). Using impulse indicators to offset forecast origin mis-forecasts has three advantages. First, because they act as one-off ICs, the next forecast commences from the forecast-origin data observation with unchanged parameter estimates, which will lead to further noticeable forecast errors if there is a location shift or trend break, but not from one-off outliers or measurement errors that such ICs fix. Second, successive large forecast errors reveal that the current model is inadequate and needs updating, which is possible

All calculations and graphs used *PcGive* (Doornik and Hendry 2021) and *Ox Professional* (Doornik 2018).

by ICs, but the few new mis-forecast observations available from which to do so are usually insufficient for adapting large systems. Third, it is feasible to test if the first few significant ICs are eliminated when replaced by a step shift or a broken linear or log-linear trend, with an encompassing test against the alternatives, guiding the appropriate correction.

Here, to isolate the source of a succession of such large forecast errors, and so capture sudden rapid shifts, we use a deterministic-trend model for the log-level of a nonstationary time series. Because any broken new trend will diverge increasingly from the previous trend, ever larger forecast errors will result if not corrected. Consequently, despite having few (only 2 or even 1) post-break observations, we show that the new trend can be estimated reasonably accurately and so continue to forecast adequately until another break occurs.

After a series of essentially unpredictable shocks from the ending of the COVID-19 pandemic lockdowns, supply chain disruption, then the energy crisis caused by Russia's invasion of Ukraine, UK annual inflation measured by the monthly Consumer Price Index (here, including owner occupiers' housing costs, CPIH) rose rapidly from mid-2021, peaking over 9% in late 2022. Figure 1 records the annual inflation time series. Coroneo (2025) shows that standard forecasting benchmarks like a random walk and a scalar autoregression had one-quarter ahead root mean square forecast errors (RMSFEs) over 2019.Q1–2023.Q4 of 2.3% and 1.7% when the Bank of England inflation target was 2%.

Because our (ex post) forecasts can be tested against the actual outcomes, we can evaluate the approach for the UK's recent surge in inflation, acting as an investigator who sequentially forecasts many steps ahead at each forecast origin using only available information. On finding a sequence of large, same-signed 1-step-ahead forecast errors despite correcting such errors using impulse indicators as ICs, the forecaster seeks the earliest date each sequence of ICs can be replaced by a step shift, a broken linear or log-linear trend to improve forecasts when a sharp upswing is in progress. Using monthly time series over 2010(1)–2024(3) on the log of the CPIH (see https://www.ons.gov.uk/datasets/cpih01/editions/time-series/versions/48), we find a series of sudden trend shifts, each of which can be detected in turn after a couple of large forecast errors, from which

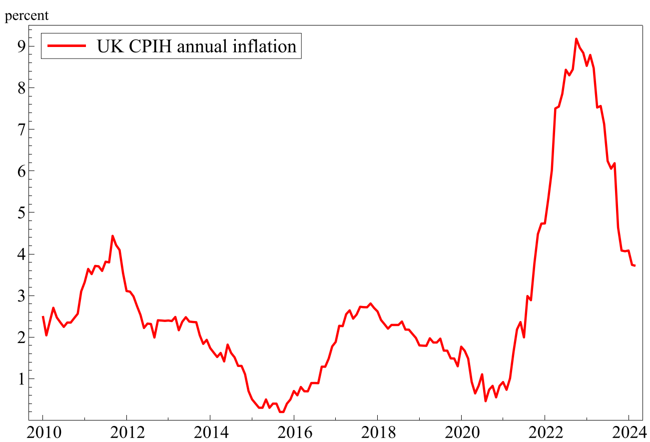

FIGURE 1 | UK CPIH annual inflation, 2010–2024.

corrected annual inflation forecasts can be derived. The resulting forecasts are respectably accurate till the following shift, and one set is able to predict 16 months ahead across peak inflation and its ensuing slowdown.

Immediately after each of the first three unexpected breaks, the forecaster finds that the first of their h-step forecasts is badly wrong, adds an IC, and forecasts h-steps ahead again and after a month discovers that the next 1-step forecast is also badly wrong. Adding another IC for that error confirms two large and increasing same-sign errors, so the forecaster tests to see which of a broken linear or log-linear trend eliminates them most effectively (essentially an encompassing test) and then again forecasts h-steps ahead. So long as no further large forecast errors occur, the forecaster continues with that selected broken trend model. If another break happens, the procedure just described is repeated. To illustrate the power of the approach, the figures in Section 6 also often show what would have happened if the break had been ignored. To see what the forecast would have been for the second month after the break once the broken trend was selected, most of the multi-step forecasts used a coefficient estimated from just the first large forecast error. Given its success, we both selected and estimated a log-linear broken trend from just one large forecast error for the final break, the theory of which is discussed in Section 3. Castle et al. (2025) provide a summary of our resuts in discussing the Bernanke (2024) review of mis-forecasting UK inflation by the Bank of England during 2021–2023.

The structure of the paper is as follows. Section 2 discusses the history of ICs to provide the context as to why such an approach may not have been considered previously. Section 3 explains the valuable information in large forecast errors and analyzes the properties of forecasting by a broken trend after just one large error. Section 4 simulates rapid detection of a deterministic trend break; then Section 5 briefly describes indicator saturation estimators (ISEs) used for model selection. Section 6 applies the analysis to rapidly detecting and forecasting after the sudden unexpected shifts in annual UK inflation over 2021(3)–2024(3). Section 7 concludes.

2 | A Brief History of Intercept Corrections

On a historical note, Suits (1957) discusses the many potential roles of dummy variables in regression analysis, although he does not consider their use as "intercept corrections" for forecasting. Klein (1971) described having used "add factors" based on recent forecast errors when forecasting but does not seem to have investigated why they helped. Young (1979) saw their role as projecting past errors into the future based on apparent model mis-specification or non-constancy of unknown sources which were expected to persist, and Wallis et al. (1986) suggest that such adjustments to model-based forecasts do improve forecast accuracy. At the start of Hendry's career in the 1960s, such devices were also known derogatorily as "ad hoc factors" and even "con adjustments" so created a negative affect towards them that probably discouraged serious study: A Google Scholar search for "intercept correction" yields very few cited works even now. An acceptable euphemism was to refer to the broader concept of judgemental adjustments (see, e.g., Turner 1990, who found mixed success). Ericsson (2017, §3) remembered Hendry being critical at a US Federal Reserve meeting of Peter Hooper adjusting his forecasts with add factors, questioning why do so if the model was good. However, as Ericsson (2017) also notes, advising the 1991 UK Parliament's Treasury and Civil Service Committee on economic forecasting (Hendry 2025, previously unpublished by accident) changed his views, leading to Hendry and Clements (1994) on a theory to explain ICs, and then Clements and Hendry (1996) with an explanation why a step indicator form of IC worked well after a location shift, as well as a chapter in Clements and Hendry (1998). Although postmortems examining forecast errors were common, the information actually contained in large forecast errors to radically improve forecasts after shifts remained hidden in plain sight as we now discuss.

3 | The Valuable Information in Large Forecast Errors

We consider forecasting after a single observation on a trend break to emphasize their high information content. Suppose we have data for $t = 1, \ldots, T, \ldots, T + h$, given initial conditions, where we initially observe the in-sample data $t = 1, \ldots, T$ and forecast sequentially through $T + 1, \ldots, T + h$. The data generation process (DGP) is given by

$$y_t = \delta t + \psi \log\left(\mathrm{t}_{\{t>T\}}\right) + \epsilon_t \quad (1)$$

where $\{\epsilon_t\}$ is an independent normally distributed error with a constant mean of zero and variance σ_ϵ^2, denoted $\epsilon_t \sim \mathrm{IN}\left[0, \sigma_\epsilon^2\right]$, and define $\widehat{\epsilon}_{t+1|t} = y_{t+1} - \widehat{y}_{t+1|t}$ as the 1-step-ahead forecast error. The broken trend $t_{\{t>T\}}$ is unity up to T, then 2, 3, 4, ... from $T + 1$ onwards. Hence, in-sample, there is a constant linear trend, but after T, there is an unanticipated deviation of a new log-linear trend. Although (1) is a simplistic DGP, it highlights the information content of large forecast errors and captures the essential characteristics of many trend breaks observed in time-series data in fields as diverse as climate, weather, economics, and finance. Additional variables $\{x_{i,t}\}$ can be included in the DGP and/or model without problem.

First, consider forecasting at the point of the break, $T + 1$. Figure 2 provides a schematic diagram. Because the break is not known at T, the unadjusted forecast for $T + 1$ (denoted by ^) is with the forecast error:

$$\widehat{\epsilon}_{T+1|T} = \left(\delta - \widehat{\delta}\right)(T + 1) + \psi \log(2) + \epsilon_{T+1} \quad (2)$$

as $\mathrm{t}_{\{t>T\}} = 2$ at $T + 1$. When $\mathrm{E}\left[\widehat{\delta}\right] = \delta$, neglecting the variance component from $\left(\delta - \widehat{\delta}\right)$ as trend coefficient estimates have variances $\mathrm{O}\left(T^{-3}\right)$, then $\mathrm{E}\left[\widehat{\epsilon}_{T+1|T}\right] \approx \log(2)\psi$ and $\mathrm{V}\left[\widehat{\epsilon}_{T+1|T}\right] = \mathrm{E}\left[\left(\widehat{\epsilon}_{T+1|T} - \mathrm{E}\left[\widehat{\epsilon}_{T+1|T}\right]\right)^2\right] = \sigma_\epsilon^2$ with a mean square forecast error (MSFE) of $\sigma_\epsilon^2 + (\log(2)\psi)^2$. An error is large relative to $\widehat{\sigma}_\epsilon^2$, which would have been estimated in-sample, so must be due to $(\log(2)\psi)^2$.

Moving forward one period, as a large forecast error is observed at $T + 1$ the forecaster may think there is an adative outlier and correct the final observation using an IC, $\mathrm{I}_{\{T+1\}} = \widehat{\epsilon}_{T+1|T} 1_{\{T+1\}}$ where $1_{\{T+1\}}$ is an impulse indicator equal to zero except unity at $T + 1$. This will set the final in-sample fitted value (denoted by) equal to the outturn as

$$\widetilde{y}_{T+1|T+1} = \widehat{\delta}(T + 1) + \mathrm{I}_{\{T+1\}} = \delta(T + 1) + \psi \log(2) + \epsilon_{T+1} = y_{T+1}$$

from (1) so:

$$\mathrm{I}_{\{T+1\}} = \left[\left(\delta - \widehat{\delta}\right)(T + 1) + \psi \mathbf{log}(\mathbf{2}) + \epsilon_{T+1}\right] 1_{\{T+1\}} \quad (3)$$

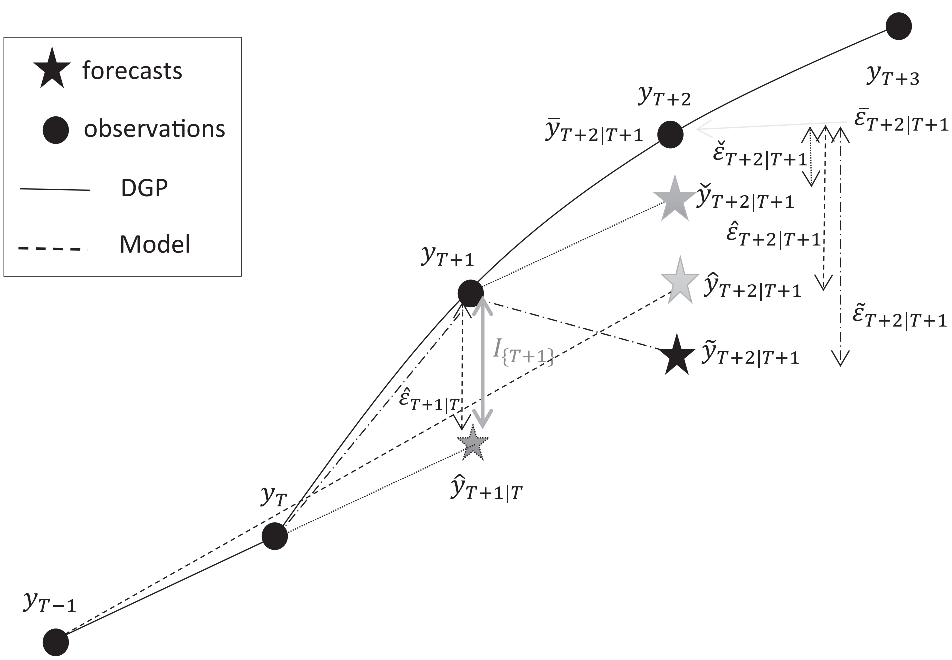

FIGURE 2 | Illustrative diagram of forecasts and forecast errors made when there is a break in trend. The diagram sets the stochastic component to zero.

and hence, $I_{\{T+1\}}$ in (3) "captures" the trend shift $\psi\mathbf{log(2)}$ (in bold just to highlight the crucial term), which is the source of the large forecast error. By "dummying out" the observation at $T+1$, the in-sample trend continues to be extrapolated, shown by the forecast marked by the dark star in Figure 2. Because $1_{\{T+1\}}=0$ at $T+2$, adding it would still lead to the next 1-step forecast:

$$\widetilde{y}_{T+2|T+1}=\widehat{\delta}(T+2)$$

and hence another large forecast error.

If the forecaster does not think the observation at $T+1$ is an adative outlier and chooses to ignore it, they would continue with the deterministic trend model. As this is recursively estimated they would have forecast a slightly smaller forecast error than the intercept corrected model as the realization at y_{T+1} would have biased the full-sample trend estimate, leading to the forecast $\widehat{y}_{T+2|T+1}$. The extent of the bias is a function of T.

Once the forecaster observes y_{T+1} but assumes it is a level shift, they would have adjusted their forecast by extending the IC to a step $S_{\{T+2\}}$, resulting in the forecast:

$$\check{y}_{T+2|T+1}=\widehat{\delta}(T+2)+I_{\{T+2\}}$$

where $I_{\{T+2\}}=\left[\left(\delta-\widehat{\delta}\right)(T+1)+\psi\,\mathbf{log(2)}+\epsilon_{T+1}\right]S_{\{T+2\}}$, recorded by the medium-grey star in Figure 2. The forecast error is reduced but still biased.

If, instead, the forecaster had added the broken log-linear trend $\log\left(t_{\{t>T\}}\right)$ on the assumption that was the problem, and anyway acts as a damped trend for forecasting the log price level (see, e.g., Gardner and Mckenzie 1985), then the forecast would be

$$\overline{y}_{T+2|T+1}=\widehat{\delta}(T+2)+\overline{\psi}\log(3)$$

where $\overline{\psi}=\widehat{\epsilon}_{T+1|T}/\mathbf{log(2)}$, leading to

$$\begin{aligned}\overline{\epsilon}_{T+2|T+1}&=\delta(T+2)+\psi\log(3)+\epsilon_{T+2}-\widehat{\delta}(T+2)-\overline{\psi}\log(3)\\&=\left(\delta-\widehat{\delta}\right)(T+2)+(\psi-\overline{\psi})\log(3)+\epsilon_{T+2}\end{aligned}$$

where

$$\mathrm{E}[\psi-\overline{\psi}]=\psi-\mathrm{E}\left[\widehat{\epsilon}_{\{T+1|T\}}\right]/\log(2)\approx 0$$

so now $\mathrm{E}\left[\overline{\epsilon}_{T+2|T+1}\right]\approx 0$, with an approximate MSFE of σ^2_ϵ.

Applying the method to forecast $\overline{y}_{T+2|T+1}$ to the UK inflation example in Section 6, for the last break, the estimate of $I_{2023(10)}=\overline{\psi}=-0.0073$ (0.003), whereas the final log-linear trend coefficient estimate was -0.0105 (0.004), which is $I_{\{2023(10)\}}/\log(2)$, matching our analytical results.[1] The MSFE was approximately $\widehat{\sigma}^2_\epsilon$ and stayed at that level for the remainder of the forecast horizon. Two caveats to estimating a broken log-linear trend from one large forecast error are that the data generation process needs to have an approximate log-linear trend after the break, and no further breaks occur. Otherwise, if the large forecast error is due to an outlier, measurement error or step shift, forecasts could be worse after such an IC, hence the need for two forecast errors to check the break type, as occurs in most cases below. Even so, determining a viable correction two periods after a trend break would be rapid.

4 | Simulating Rapid Detection of a Deterministic Trend Break

We undertook a Monte Carlo simulation to evaluate the detection and forecast performance of the above real-time forecasting procedure, both under the null of no break and under the alternative of a trend break. In the simulations, there are $T=80$ in-sample observations and H out-of-sample observations to evaluate forecasts. Impulse indicators are denoted I_j, and trend indicators are denoted τ_j ending at time j, whereas t is a linear trend for $t=1,\ \dots,T+H$. All the simulations are based on $M=10{,}000$ replications.

The DGP is given by (here bold denotes vectors)

$$\boldsymbol{y}=\mu\mathbf{1}+\lambda\boldsymbol{\tau}_T+\rho\boldsymbol{t}+\boldsymbol{\epsilon},\qquad \boldsymbol{\epsilon}\sim\mathrm{IN}_T\left[\mathbf{0},\sigma^2_{\boldsymbol{\epsilon}}\boldsymbol{I}\right]\tag{4}$$

where $\boldsymbol{\tau}_T=(-79,-78,\ \dots,-1,0,0,\ \dots,0)'$, but the full sample trend is $\boldsymbol{t}=(1,2,\ \dots,T+H)'$. There is a break in trend that occurs at observation $T=80$, forecasting recursively over the next $H=5$ periods. We set the intercept $\mu=5.5$, the end-of-sample trend $\rho=0.05$, and $\sigma_\epsilon=0.025$, varying λ to give different magnitude trend breaks. DGP coefficients have been chosen to correspond to values that we may observe for annual growth rates when the regressand is in logs so the growth rate after T is ρ; up to T, the growth rate is $\lambda+\rho$.

The model is estimated recursively for $t=1,\ \dots,T+h$ over $h=1,\ \dots\ 5$:

$$y_t=\beta_0+\beta_1 t+\sum_{j=1}^{5}\beta_{I_j}1_{T+j}+\nu_t\tag{5}$$

recording the significance of the impulse indicators' coefficients. As the impulse indicators are orthogonal and dummy out the final observations, estimating the model with 5 impulse indicators over $t=1,\ \dots,T+5$ is equivalent to recursively testing the individual indicators as the window increases from $T+1$ to $T+5$.

Figure 3 records the proportion of replications in which the impulse indicators are significant at a 1% significance level as the sample is increased from $t=1,\ \dots,T+1$ to $T+5$. When $\lambda=0$, the trend in the forecast period is the same as the in-sample period, and thus this case pertains to the null of no break. When $\lambda=-0.025$, the trend initially has a growth rate of 2.5%, which then increases to 5% at $T+1$ onwards. At the extreme of $\lambda=-0.1$, the trend growth rate is initially falling at 5% before reversing and increasing at 5% at $T+1$. The ratio of λ to σ_ϵ, denoted $\lambda^\dagger$, determines the detectability of shifts.

For $|\lambda^\dagger|\geq 4$, the impulse indicators would be retained more than 90% of the time at $T+1$ and similarly for $|\lambda^\dagger|\geq 2$ at $T+2$.

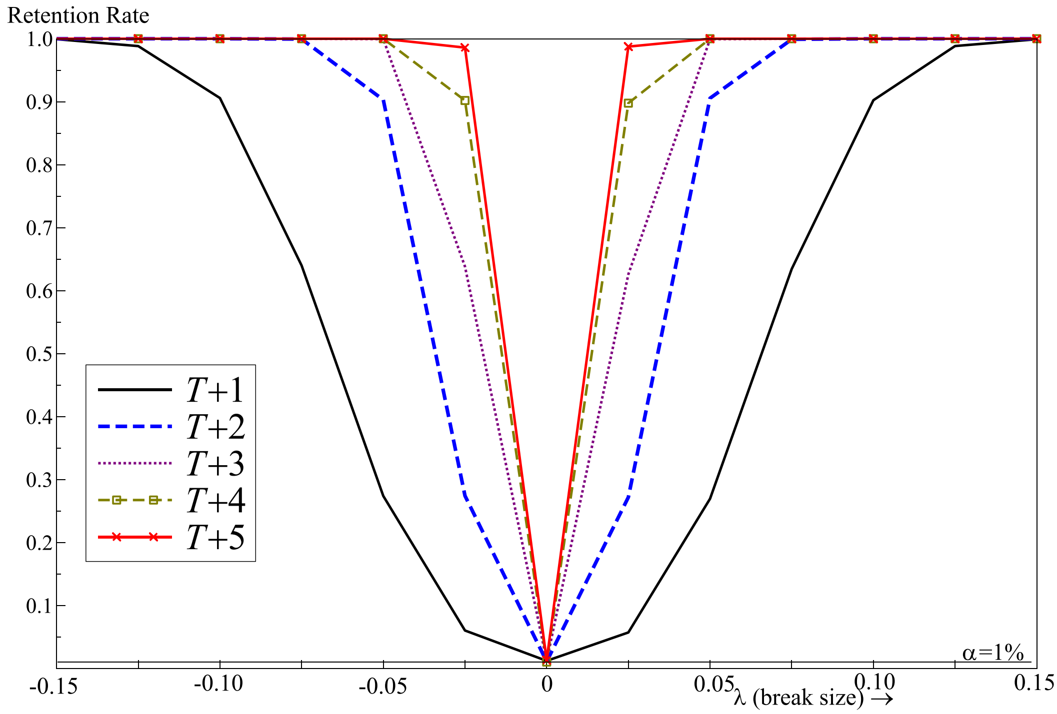

FIGURE 3 | Proportion of impulse indicators that are significant at 1% for a break in trend (4).

By 4 or 5 observations, a break of almost any size is easily detected by the impulse indicators. When there is no break, the impulse indicators are retained at close to the nominal significance level. Detection rates are essentially symmetric in the sign of λ. A less noisy DGP (lower σ_ϵ) will increase the retention probability and vice versa for a given λ magnitude.[2]

We next test at 1% for the significance of including a linear or log-linear trend in the model. Recursive estimation over $t = T+2, \ldots, T+5$ (commencing in $T+2$ to give one post-break in-sample observation at $T+1$) is applied to

$$y_t = \beta_0 + \beta_1 t + \beta_2 t_{T+h} + \sum_{j=2}^{5} \beta_{\iota_j} 1_{T+j} + v_t \quad (6)$$

$$y_t = \beta_0 + \beta_1 t + \beta_2 \log\left(t_{T+h+1}\right) + \sum_{j=2}^{5} \beta_{\iota_j} 1_{T+j} + v_t \quad (7)$$

where $t_{T+h} = h$ for $h = 1, \ldots, H$ and 0 for $1, \ldots, T$ and $\log\left(t_{T+h}+1\right) = \log(h+1)$ for $h = 1, \ldots, H$ and 0 for $1, \ldots, T$. We test for the significance of the last ICs (dropping 1_{T+1} to avoid perfect collinearity with the trend) using an exclusion test.

Figure 4 reports the results, where panel (a) records the retention rate of the included trend, panel (b) records the average t-value of the trend, and panel (c) records the retention rate of the impulse indicators when the trend is included. Under the null of no break, the retention probabilities are close to the chosen significance level. The probability of retaining either the linear or log-linear trend increases rapidly as the break size increases, and there is little difference between the two trend specifications. The approximating trends are highly significant, even after just two observations, if the break is moderately large, shown by the average t-values, and the linear trend dominates the log-linear trend as the DGP is a linear trend. The log-linear trend does a good job of approximating the trend, but its mis-specification is revealed by the additional retention of impulse indicators (panel c), which increases over the longer forecast horizon as the log-linear trend diverges from the linear trend. Under the correctly specified linear trend, the impulse indicators are retained in addition to the trend for just 1% of the draws, the significance level of the exclusion test.

The forecast performance of the approach is assessed by comparing five alternative forecasting approaches. These include:

i. ignoring the break, that is, (5) with $\beta_{\iota_j} = 0, \ \forall j$, denoted "unadjusted";
ii. using an IC for the last in-sample observation to intercept correct at the forecast origin (i.e., not extrapolating the break forward), denoted "IC";
iii. using a linear trend from the forecast origin, that is, (6) with $\beta_{\iota_j} = 0, \ \forall j$;
iv. using a log-linear trend from the forecast origin, that is, (7) with $\beta_{\iota_j} = 0, \ \forall j$; and finally
v. using a step indicator from the forecast origin T_* extrapolated forward.

The forecasting exercise undertakes a series of 5 dynamic forecasts. The models are estimated over $t = 1, \ldots, T_*$, and forecasts are produced for $T_* + 1, \ldots, T_* + 5$. All models are identical for the initial recursion, as the break in trend occurs at $T+1$. Thus, the models are estimated over $t = 1, \ldots, T+1$, and dynamic forecasts are produced for

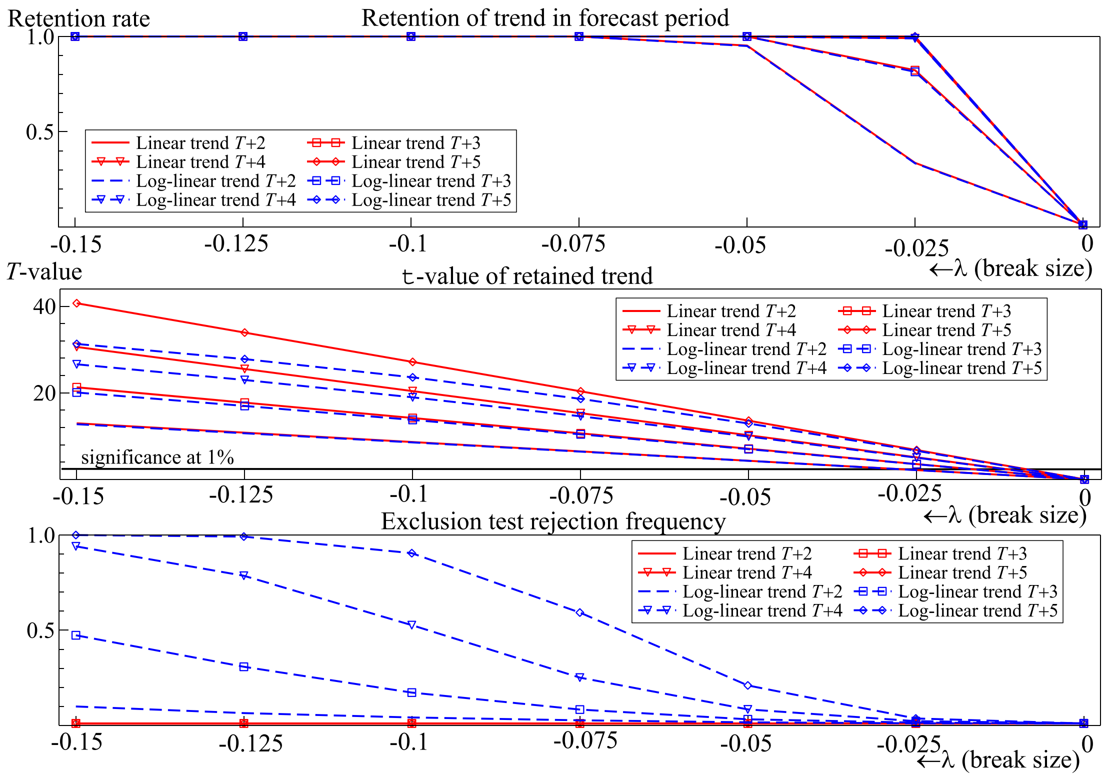

FIGURE 4 | (a) The proportion of trends that are significant at a 1% significance level. (b) The average t-statistic of the trend. (c) The proportion of replications in which the p value of the exclusion test over the ICs is less than 0.01, that is, the number of draws in which ICs are significant in addition to the trend. At $T+2$, the exclusion test applies to just I_{T+2}; at $T+3$, the joint test is for I_{T+2} and I_{T+3}, and for $T+5$, tests for the joint exclusion of $\mathrm{I}_{T+2}, \ldots, \mathrm{I}_{T+5}$.

$T+2, \ldots, T+6 = T_*+1, \ldots, T_*+5$. In this recursion, there is just one observation to estimate the linear trend, the log-linear trend, and the step shift. The models are then estimated over $t = 1, \ldots, T+2$, and dynamic forecasts are obtained for $T+3, \ldots, T+7$, etc. up to an in-sample period $t = 1, \ldots, T+4$ with forecasts over $T+5, \ldots, T+9$. Mean forecast errors (ME) and root mean square forecast errors (RMSFE) are recorded across the 5 dynamic forecasts for each forecast horizon for λ equal to $(-0.1, -0.05, -0.0375, -0.025)$, so only two break examples have $|\lambda^{\dagger}| \geq 2$ when $\sigma_{\epsilon} = 0.025$.

The RMSFE results are reported in Figure 5 for the four different break magnitudes.[3] When the break is large, just one observation is sufficient to estimate the broken linear trend and correct the forecasts, with the broken linear trend dominating the forecast performance for breaks of $\lambda = -0.05$ or larger in absolute value. Even if the break is of moderate size, the broken linear trend dominates after just 2 observations. The mis-specified broken log-linear trend does not remove the bias, but it is still effective at reducing RMSFE relative to an unadjusted model. The IC to set the forecasts back on track slightly worsens the forecast error relative to doing nothing, as expected. The step shift is the wrong model but could be detected recursively. Under the null of no trend break ($\lambda = 0$), there is a cost to using the broken linear trend in RMSFE, particularly after just one observation, which is still present after 4 observations, although the costs are small and averaging across the broken linear and broken log-linear trends would mitigate this when the break form was uncertain.

5 | Indicator Saturation Estimation

ISEs are designed to detect outliers, shifts in means, or breaks in trends (inter alia) at any points in a time series without knowing their numbers, signs, magnitudes, or timings while retaining relevant explanatory variables. The approach adds an indicator variable with the appropriate formulation for every observation in a sample of size T to the set of potential regressors then searches for significant indicators (see Hendry and Doornik 2014). Indicator variables could be impulse indicators (IIS), $1_j = 1$ for $t = j$ and zero otherwise for $j = 1, \ldots, T$; step indicators, $S_j = 1_{t \leq j}$ and zero otherwise (SIS); or trend indicators (TIS; see Walker et al. 2019), which are the cumulation of step indicators:

$$\boldsymbol{\tau}'_2 = (-1, 0, \ldots, 0) \ldots; \boldsymbol{\tau}'_t = (-t+1, -t+2, \ldots, -1, 0 \ldots 0) \ldots; \boldsymbol{\tau}'_T = (-T+1, \ldots, -1, 0).$$

Thus, τ_{date} denotes a trend indicator ending at *date*, t is the full sample linear trend $t = 1, \ldots, T$, whereas t_{date} is a broken deterministic linear trend commencing at *date*. Also, S_{date} denotes a step indicator ending in *date*, and s_{date} is a step shift commencing at *date*. All saturation indicators (1_j, S_j, and τ_j) are designed to be zero in the forecast period.

A tree search algorithm with expanding and contracting block searches allows all indicators to be investigated for possible significance: Castle et al. (2021) provide details of the search algorithm. Given the resulting high dimensionality, selection

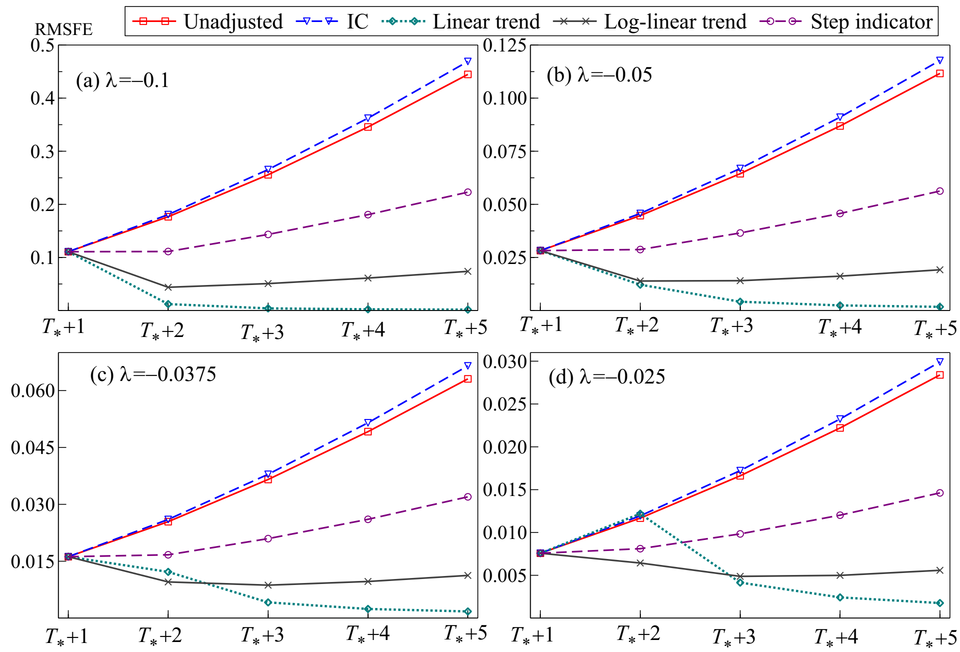

FIGURE 5 | RMSFEs averaged across a sequence of 5 dynamic forecasts commencing at $T, \dots, T+4$, where the break in linear trend of magnitude λ occurs at $T+1$, so the outcomes at T_*+3 average over forecast errors at $T+3, \dots, T+7$.

must use tight significance levels α to control the probability of retaining irrelevant indicators, particularly if impulses, steps and trends are searched jointly, denoted supersaturation (Ericsson 2012). We set $\alpha = 0.0001$ for TIS, $\alpha = 0.005$ for SIS, and $\alpha = 0.01$ for impulse indicators to select these sequentially when $T = 130$ using the results in Hendry and Johansen (2015).

6 | Modeling and Forecasting UK Annual Inflation

The combination of the COVID-19 pandemic, supply chain disruption and the energy crisis caused by Russia's invasion of Ukraine led to several rapid upswings in UK inflation. Here, we investigate how quickly they could have been detected by modeling the log of monthly Consumer Price Index including owner occupiers' housing costs (CPIH), denoted p_t (source: Office of National Statistics) with a data set over 2010(1)–2024(3). We follow a forecaster making 1-step-ahead forecasts as the forecast origin advances each month from 2021(3): This could simply be the first of a multi-step sequence. The initial model explains the log-level p_t by an intercept and linear trend then using TIS selected at $\alpha = 0.0001$ up to 2021(3) as recorded in (8) where all trends have been scaled by 100:

$$\begin{aligned}\widehat{p}_t = \; & \underset{\underset{[0.022]}{(0.019)}}{4.45} - \underset{\underset{[0.036]}{(0.039)}}{0.27}\,\tau_{2010(11)} + \underset{\underset{[0.024]}{(0.028)}}{0.28}\,\tau_{2011(4)} + \underset{\underset{[0.01]}{(0.01)}}{0.083}\,\tau_{2013(4)} + \underset{\underset{[0.03]}{(0.03)}}{0.22}\,\tau_{2014(10)} \\ & - \underset{\underset{[0.25]}{(0.19)}}{0.80}\,\tau_{2015(4)} + \underset{\underset{[0.23]}{(0.18)}}{0.70}\,\tau_{2015(5)} - \underset{\underset{[0.01]}{(0.02)}}{0.21}\,\tau_{2016(2)} + \underset{\underset{[0.08]}{(0.17)}}{1.04}\,\tau_{2018(12)} \\ & - \underset{\underset{[0.10]}{(0.23)}}{1.3}\,\tau_{2019(1)} + \underset{\underset{[0.04]}{(0.07)}}{0.39}\,\tau_{2019(4)} + \underset{\underset{[0.007]}{(0.006)}}{0.071}\,t \end{aligned} \tag{8}$$

$$\widehat{\sigma} = 0.20\% \quad R^2 = 0.999 \quad F_{ar}(7,116) = 2.98^{*}$$

$$F_{arch}(7,121) = 0.78 \quad T = 2010(1)\text{-}2021(3)$$

$$\chi^2_{nd}(2) = 2.22 \quad F_{Het}(20,114) = 1.46$$

$$F_{reset}(2,121) = 0.07 \quad F_{Chow}(1,123) = 10.5^{**}$$

There are no systematic differences between the conventional standard errors and HACSEs, so only the former are reported below and used in calculating forecast standard errors.[4] Ten earlier shifts in UK log price level since 2010 were detected at 0.01%, correcting the overall trend to 0.071 (i.e., 0.85% pa). Because all indicators are zero beyond their dates, the forecasts are $\widehat{p}_{T+h|T} = \widehat{\beta}_0 + \widehat{\beta}_1(T+h)$, where for the forecast for 2021(4) from 2021(3), $\widehat{\beta}_0 = 4.45$ and $\widehat{\beta}_1 = 0.071$. However, $p_t = \beta_0 + \beta_1 t + v_t$ is an equilibrium-correction equation, which can be written as $\Delta p_t = \beta_0 - (p_{t-1} - \beta_1 t) + v_t$ so suffers the pernicious problems of that class, as seen below.

Figure 6a shows the fitted and actual values and the 1-step-ahead forecast by (8) for 2021(4) from 2021(3) and (b) 2021(5) from 2021(4) with $1_{2021(4)} = 0.0070$ ($\widetilde{p}_{T+1|T}$: $F_{chow}(1,123) = 29.0^{**}$) and

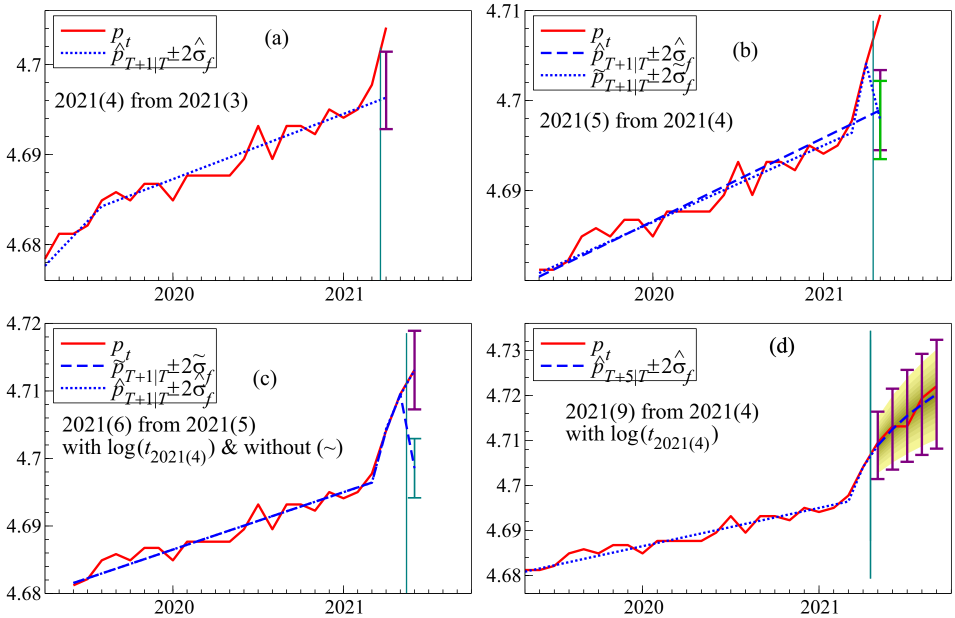

FIGURE 6 | Fitted and actual values for $\widehat{p}_t$ with 1-step-ahead forecast: (a) 2021(4) from 2021(3) from (8); (b) 2021(5) from 2021(4) without $\mathrm{I}_{2021(4)}$ ($\widehat{p}_{T+1|T}$, dashed) and with ($\widetilde{p}_{T+1|T}$, dotted); (c) 2021(6) from 2021(5) (i) with $\mathrm{I}_{2021(5)}$ added to ((8): $\widetilde{p}_{T+1|T}$, dashed) and (ii) after adding $\log(t_{2021(4)})$ ((9): $\widehat{p}_{T+1|T}$, dotted); (d) forecasting by (9) to 2021(9) from 2021(4) with both error bars and fans.

without ($\widehat{p}_{T+1|T}$: $\mathrm{F_{Chow}}(1,124) = 22.7^{**}$). Although the interval forecasts (denoted by bars reporting $\pm 2\widehat{\sigma}_f$) are not based on a congruent model, they offer a guide to the forecast uncertainty assuming no further trend breaks. When $1_{2021(4)}$ is included, it acts as an IC (denoted by $\mathrm{I}_{2021(4)}$), so the next forecast commences from the 2021(4) outcome and hence leads to a similar forecast error but via a large downward forecast as the upswing is in progress, emphasizing the trend shift.

Next, panel (c) shows forecasting 2021(6) from 2021(5) after also adding to (8) (i) $1_{2021(5)}$ with t = 5.4 and (ii) the log of the broken linear trend $\log(t_{2021(4)})$, where $t_{2021(4)}$ is unity up to 2021(3), (so $\log(t_{2021(4)})$ is zero), then 2,3,4, ... from 2021(4) onwards. Adding $\log(t_{2021(4)})$ eliminates the two impulse indicators for 2021(4) and 2021(5) and is recorded in (9) with an insignificant Chow test. Finally, panel (d) extends the forecast horizon to 2021(9) with a multi-step RMSFE = 0.15% and $\mathrm{F_{Chow}}(5,123) = 0.52$, so so (9) forecasts better out of sample than the in-sample fit, despite the coefficient of $\log(t_{2021(4)})$ being selected from just 2 observations and estimated from 1.

$$
\begin{aligned}
\widehat{p}_t = & \underset{(0.019)}{4.45} \; - \underset{(0.039)}{0.27}\,\tau_{2010(11)} + \underset{(0.028)}{0.28}\,\tau_{2011(4)} + \underset{(0.01)}{0.083}\,\tau_{2013(4)} - \underset{(0.03)}{0.22}\,\tau_{2014(10)} \\
& - \underset{(0.19)}{0.80}\,\tau_{2015(4)} + \underset{(0.18)}{0.70}\,\tau_{2015(5)} - \underset{(0.02)}{0.21}\,\tau_{2016(2)} + \underset{(0.17)}{1.04}\,\tau_{2018(12)} \\
& - \underset{(0.23)}{1.3}\,\tau_{2019(1)} + \underset{(0.06)}{0.39}\,\tau_{2019(4)} + \underset{(0.006)}{0.071}\,t + \underset{(0.003)}{0.010}\log\left(t_{2021(4)}\right)
\end{aligned} \tag{9}
$$

$\widehat{\sigma} = 0.20\%$ $\quad \mathrm{R}^2 = 0.999$ $\quad \mathrm{F_{ar}}(7,116) = 2.98^{*}$

$\mathrm{F_{arch}}(7,122) = 0.78$ $\quad T = 2010(1) - - 2021(4)$

$\chi^2_{nd}(2) = 2.28$ $\quad \mathrm{F_{Het}}(20,114) = 1.46$

$\mathrm{F_{reset}}(2,121) = 0.07$ $\quad \mathrm{F_{Chow}}(1,123) = 0.03$

Although we have used p_t to test for and model changes in trend, the forecasts for annual inflation are easily derived from the levels' forecasts as $\widehat{\Delta_{12} p}_{T+h|T} = \widehat{p}_{T+h|T} - p_{T+h-12}$, and will have the same error bars as the log level, now recentered on annual changes. The outcomes are shown in Figure 7. The forecasts (a)–(d) are derived from those in Figure 6. That the first two are below the previous outcome is all too common with equilibrium-correction models.

6.1 | The Next Break

We continue to follow the forecaster noting their 1-step-ahead forecasts as the forecast origin now advances from 2021(9). The model is (9) augmented by $\log(t_{2021(4)})$, now estimated up to 2021(9) with $\widehat{\sigma} = 0.20\%$. As $\mathrm{F_{Chow}}(1,128) = 12.38^{**}$ when forecasting 2021(10) from 2021(9), a second break has happened as seen in Figure 8a. This is confirmed when next forecasting 2021(11) from 2021(10) in Figure 8b. Adding ICs for 2021(10) and 2021(11) yields t values of 3.5^{**} and 5.3^{**}, but forecasting 2021(12) from 2021(11) would have delivered

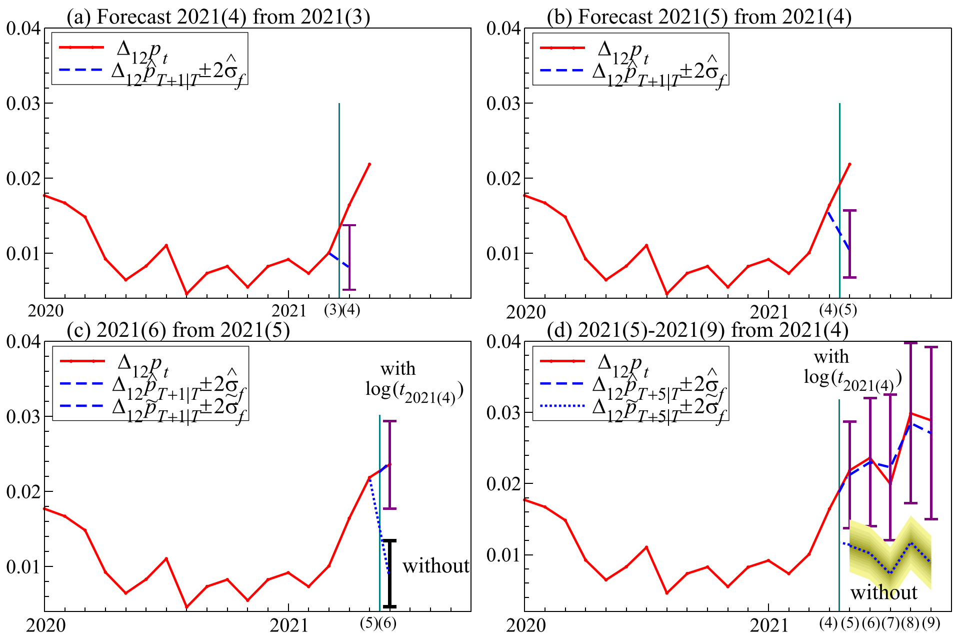

FIGURE 7 | Actual and derived forecast values for $\widehat{\Delta_{12} p_t}$ and the 1-step-ahead forecasts: (a) 2021(4) from 2021(3); (b) 2021(5) from 2021(4); (c) 2021(6) from 2021(5) without and after adding $\log(t_{2021(4)})$ in (7); (d) forecasting 2021(9) from 2021(4) with and without $\log(t_{2021(4)})$.

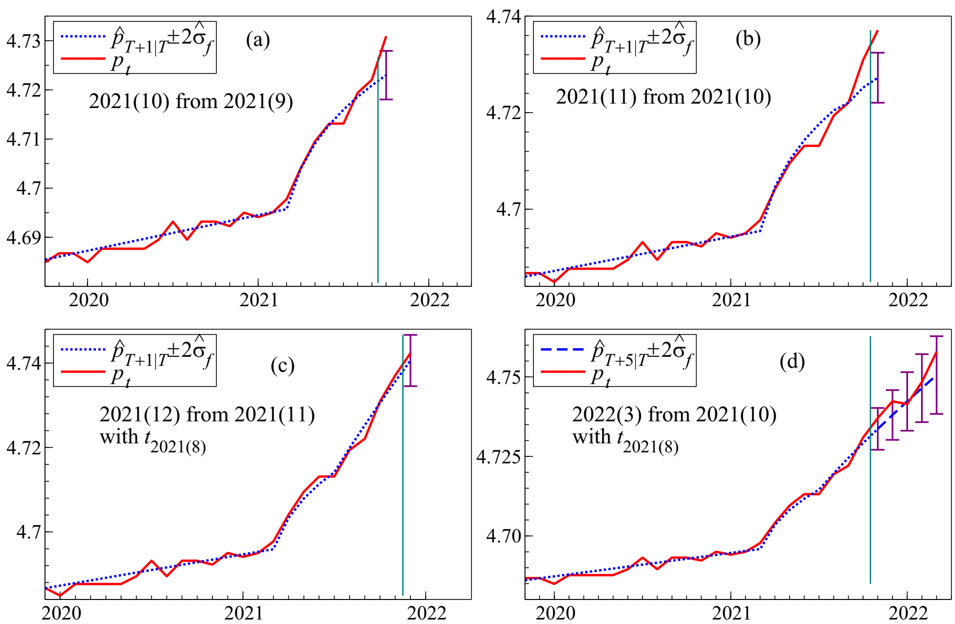

FIGURE 8 | Fitted and actual values for $\widehat{p}_t$ and the 1-step-ahead forecast: (a) 2021(10) from 2021(9); (b) 2021(11) from 2021(10); (c) 2021(11) from 2021(10) after adding the linear trend $t_{2021(8)}$ to (9); (d) forecasting from 2021(10) to 2022(3) with $t_{2021(8)}$.

$F_{Chow}(1,128) = 44.8^{**}$, confirming it is not simply outliers or a step shift.

From Figure 8b, the break probably started in 2021(8), so we created a broken linear trend starting then, denoted $t_{2021(8)}$. Adding it to the model with the 2 impulse indicators made them insignificant with an insignificant Chow test, and eliminating the indicators produced the outcome in Figure 8c with $F_{Chow}(1,129) = 0.34$ and $\widehat{\sigma} = 0.20\%$. Finally, Figure 8d shows multi-step forecasts from 2021(11) to 2022(3), with $F_{Chow}(5,128) = 1.81$ and RMSFE = 0.40%.

6.2 | Not Another Break!

Forecasting 2022(4) from 2022(3) leads to another significant failure with $F_{Chow}(1,133) = 63^{**}$ as seen in Figure 9a. By the time this shift could have been observed, Russia's invasion of

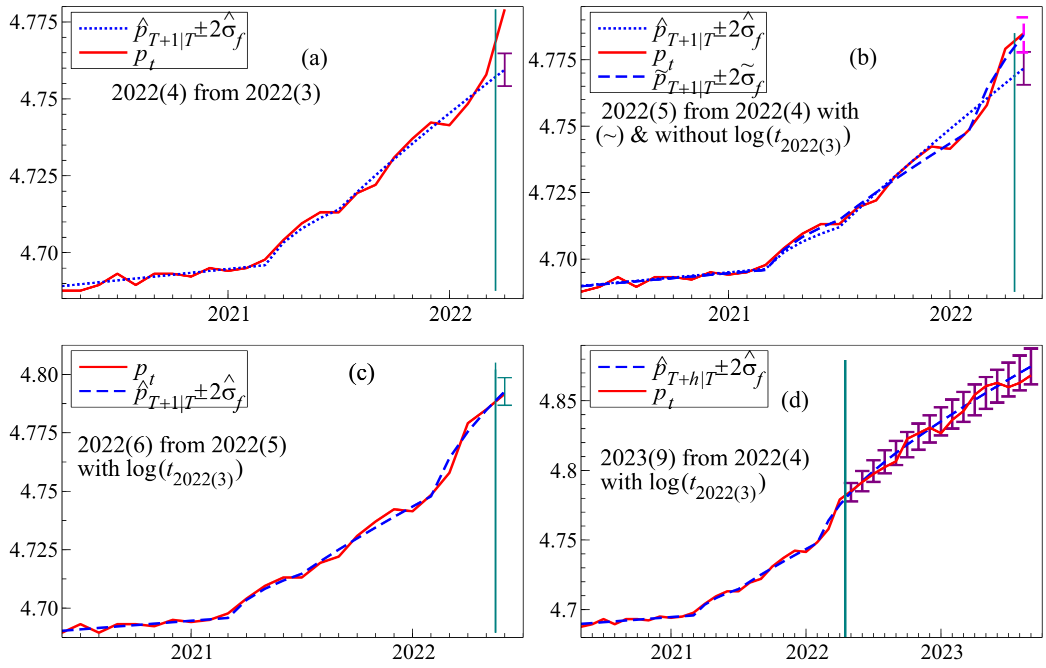

FIGURE 9 | Fitted and actual values for $\hat{p}_t$ and the 1-step-ahead forecast for (a) 2022(4) from 2022(3); (b) 2022(5) from 2022(4); (c) 2022(6) from 2022(5) after also adding $\log(t_{2022(3)})$ to (9); (d) forecasting from 2022(4) to 2023(9) by (10).

Ukraine and the consequent energy crisis and fuel and food price rises had occurred, so such a shift would not be a surprise, confirmed by another large error forecasting 2022(5) from 2022(4) (Figure 9b), leading to $I_{2022(4)}$ and $I_{2022(5)}$ with t values of 7.9^{**} and 8.0^{**}.

Figure 9d with 16-step-ahead forecasts for 2022(6)–2023(9) confirms that medium-term forecasts can be usefully accurate despite the most recent broken trend again being selected from just 2 observations, and estimated from 1, obviously conditional on no new breaks occurring.

$$\begin{aligned}\hat{p}_t = &\underset{(0.020)}{4.45} \underset{(0.041)}{-0.27}\,\tau_{2010(11)} \underset{(0.029)}{+0.28}\,\tau_{2011(4)} \underset{(0.010)}{+0.083}\,\tau_{2013(4)} \underset{(0.032)}{-0.22}\,\tau_{2014(10)} \\ &\underset{(0.20)}{-0.80}\,\tau_{2015(4)} \underset{(0.19)}{+0.70}\,\tau_{2015(5)} \underset{(0.02)}{-0.21}\,\tau_{2016(2)} \underset{(0.18)}{+1.04}\,\tau_{2018(12)} \\ &\underset{(0.24)}{-1.04}\,\tau_{2019(1)} \underset{(0.068)}{+0.39}\,\tau_{2019(4)} \underset{(0.006)}{+0.071}\,t \underset{(0.001)}{+0.010}\log(t_{2021(4)}) \\ &\underset{(0.035)}{+0.28}\,t_{2021(8)} \underset{(0.002)}{+0.017}\log(t_{2022(3)})\end{aligned} \tag{10}$$

$$\hat{\sigma} = 0.21\% \quad R^2 = 0.999 \quad F_{ar}(7,126) = 3.53^{**}$$

$$F_{arch}(7,134) = 1.21 \quad T = 2010(1) - - 2022(4)$$

$$\chi^2_{nd}(2) = 3.83 \quad F_{Het}(26,121) = 3.25^{**}$$

$$F_{reset}(2,131) = 0.13 \quad F_{Chow}(1,133) = 0.04$$

Adding the next broken log-linear trend $\log(t_{2022(3)})$ eliminates the ICs and delivers the outcome in (10), yielding $F_{Chow}(1,133) = 0.042$ for 2022(5), shown in Figure 9c, greatly reducing the next forecast error, as $F_{Chow}(1,133) = 0.16$ for 2022(6). The model continues to forecast reasonably accurately through to 2023(9), which is 17 periods (with 16 ahead), and although p_t is slightly overpredicted from 2023(7) leading to $F_{Chow}(17,133) = 3.96^{**}$, the RMSFE is 0.47%, and the tracking is close as seen in Figure 9d.[5]

6.3 | Now We Go Back Down

We continued the multi-step forecast to 2023(9) in order to check if the approach could capture inflation first peaking then falling. The 16-step-ahead forecasts for p_t show that is indeed possible. Figure 10 plots all the sets of multi-step-ahead forecasts for UK prices and Figure 11 for annual inflation, $\widehat{\Delta_{12}p_{T+h|T}}$. Although the model estimated up to 2022(4) has three broken trends all with positive coefficients, nevertheless the annual inflation forecasts over 2022(5)–2023(9) capture the first eight falls after the downturn in inflation. In a sense, this is partly an artefact of the previous year's inflation being higher than the current one but also requires accurate forecasts of $p_{T+h|T}$. As no intermediate 1-step-ahead forecasts yielded significant errors, we assume the forecaster continues with the same model.

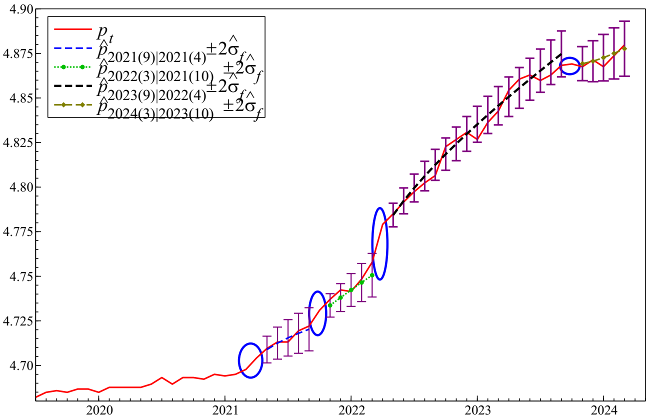

FIGURE 10 | Multi-step-ahead forecasts of p_t over the inflation upsurge and slow down from 2021(4) to 2024(3), spanning the four episodes with ellipses highlighting the break periods with the large forecast errors that were used to estimate the broken trends.

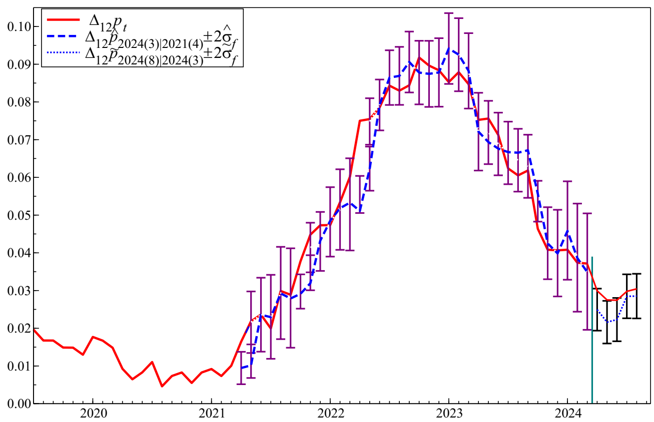

FIGURE 11 | Multi-step-ahead forecasts $\Delta_{12}\widehat{p}_{T+h|T}$ (dashed) and outcomes over 2021(4)–2024(3) (solid) with post-sample forecasts $\Delta_{12}\widetilde{p}_{2024(8)|2024(3)}$ (dotted) all shown with error bars (two vertically overlapping error bars are for forecasts without and with ICs).

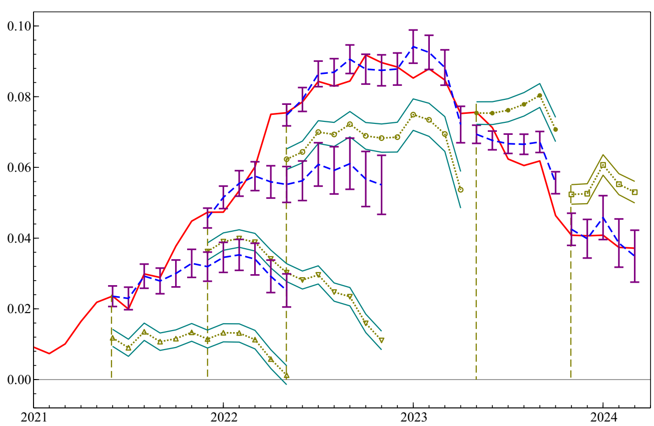

FIGURE 12 | Twelve-month-ahead forecasts for the first three episodes and 6 and 5 months ahead for last two. Vertical lines mark 2 months after shifts: (i) without broken trends (green, dotted lines with symbols, always very poor) and (ii) with broken trends (only poor after next unanticipated break).

6.4 | A Final Shift

Although the error forecasting 2023(9) for 2023(8) is insignificant, when forecasting 2023(10) from 2023(9), a significant forecast error occurs with $F_{Chow}(1,150) = 7.3^*$, revealing another trend break as the increases in p_t slowed. Once again, a broken log-linear trend from 2023(10) produces $F_{Chow}(5,150) = 1.1$ over 2023(11)–2024(3) although fitted to just one nonzero observation as shown in Figure 10. That coefficient estimate is −0.0105 (0.004), and estimated till 2024(3) is −0.011 (0.0013), hence the accurate forecasts.

Figure 11 collects our multi-step-ahead monthly forecasts of annual inflation over 2021(4)–2024(3) ($\widehat{\Delta_{12}p}_{T+h|T}$ with $\pm 2\widehat{\sigma}_f$) shown as dashed and outcomes ($\Delta_{12}p_t$, solid), including the large forecast errors that prompted the additions of broken trends (vertically overlapping error bars are for forecasts without and with ICs). After Castle et al. (2025) was completed, new data became available to test the continuation of the final shift model's forecasts $\widetilde{\Delta_{12}p}_{2024(8)|2024(3)}$ shown as dotted in Figure 11 and, although generally lower than the outcomes, have a RMSFE of 0.43% with $F_{Chow}(5,155) = 1.53$.

The Bank of England started raising interest rates from 0.1% in February 2022, continuing to raise in small steps till stopping at 5.25% in August 2023, yet our accurate 16-month-ahead (albeit ex post) forecasts to 2023(9) were made in 2022(4) prior to most increases. Castle et al. (2025) discuss what the MPC might have made of them (also see Hendry and Muellbauer 2024).

An alternative comparison is with how the basic model would have forecasted in the absence of our approach to rapidly correcting after shifts. Figure 12 records whole period outcomes using 12-month-ahead forecasts for the first three episodes and 6 and 5 months ahead for the last two. Forecasting beyond the next break naturally leads to significant forecast errors, but that could not be known until after the shifts are observed, and Figure 11 highlights the advantages of then rapidly correcting using the initial 1-step-ahead large forecast errors. Conversely, not correcting would have been a very bad strategy, and although we have used a deterministic trend equation as the illustration in Figure 12, similar patterns of forecast failure would have occurred for the typical benchmarks that Coroneo (2025) considered.

7 | Conclusion

A sudden unanticipated shift in a variable will usually create a sequence of large same-sign 1-step-ahead forecast errors as the forecast origin advances. The procedure described in this paper to avoid systematic forecast failure is as follows.

1. Once a significant 1-step-ahead forecast error is detected, add an impulse indicator acting as an IC to the model. The IC ensures that the forecasting model is unchanged; hence, it maintains the previous trend, so it will rapidly reveal departures from any new upswing, and by an encompassing test, it helps discriminate trend breaks from location shifts, outliers, or measurement errors (the last 2 of which impulse indicators can correct).

2. After two (or perhaps three) large increasing same sign 1-step-ahead forecast errors have led to a significant sequence of ICs, a broken linear or log-linear trend can be estimated and tested for its adequacy by replacing the ICs, as well as tested against step shifts and the alternative trend formulation.
3. Despite being selected from just one to three observations, the new broken trends can continue to forecast acceptably accurately further ahead until another trend break occurs.

An application to the upsurge since 2021 in UK annual inflation illustrated this last possibility. The log level of the monthly CPIH was modeled by first applying trend-indicator saturation (TIS) at a 0.01% significance to an equation with an intercept and linear trend fitted to the historical data from 2010(1) to 2021(3). One-step forecast errors in 2021(4) and 2021(5) produced significant ICs, which a log-linear trend starting in 2021(3) replaced and could forecast accurately five months ahead. Then 1-step forecast errors from another break in 2021(8) were handled by a linear trend, forecasting till 2022(3), when the energy crisis occurred, with large forecast errors in 2022(4) and 2022(5). Replacing those with a log-linear trend commencing in 2022(3) enabled 16-step-ahead forecasting from 2022(5) to 2023(9) as annual inflation first peaked and then fell. Modeling the final shift from 2023(10) again proved usefully accurate till our sample end, including "out-of-sample" to 2024(8) after annual inflation stabilized. The 4 essentially unpredictable trend shifts are clearly visible in Figure 10 and were followed by significant forecast errors, but our approach experienced only seven large errors overall by rapidly detecting breaks, rather than long periods of systematic forecast failure.

Such an approach can potentially quickly detect sudden increases and "tipping points" at the start of their evolution, acting both as an early-warning system and providing a glimpse of the road ahead, albeit without knowing why nor when the next failure will occur. If the model and data do not have a trend but a new one commences after a break, a similar approach is effective, but now not including impulse indicators reveals the shift more quickly because otherwise ICs can improve forecasts sufficiently to hide the change. A forecasting agency could publish the IC forecast but record the one without ICs and switch when a new broken trend is detected.

Acknowledgments

We are pleased to acknowledge financial support from the Research Council of Norway, project 324472, on "Model invariance and constancy in the face of large shocks to the Norwegian macroeconomic system" and Nuffield College, and helpful comments from an anonymous referee, Neil Ericsson, and participants at the Second International Conference on the Climate–Macro–Finance Interface at Bayes Business School, University of London.

Conflicts of Interest

Jurgen A. Doornik and David F. Hendry have developed Autometrics, which is included in the OxMetrics software package, and have a share in the returns.

Data Availability Statement

The data and PcGive batch file with output can be downloaded in InfForcErrs25F.zip from https://www.climateeconometrics.org/data/.

Endnotes

[1] Standard errors reported in parentheses.

[2] An equation standard error of 2.5% is similar to that for UK GDP. Simulation results for alternative values of σ_ϵ are available on request.

[3] The RMSFE scale in Figure 5 differs across break sizes to highlight differences between models. MEs and results for a range of other break sizes are available on request.

[4] Coefficient standard errors shown in parentheses, with heteroskedasticity and autocorrelation consistent standard errors (HACSEs) in brackets, $\widehat{\sigma}$ is the residual standard deviation, F_{ar} tests residual autocorrelation (see Godfrey 1978), F_{arch} tests autoregressive conditional heteroscedasticity (see Engle 1982), F_{het} tests residual heteroskedasticity (see White 1980), $\chi^2_{nd}(2)$ tests non-normality (see Doornik and Hansen 2008), F_{reset} tests nonlinearity (see Ramsey 1969), and F_{chow} tests parameter constancy (see Chow 1960) over the forecast period. One star indicates test significance at 5%, two at 1%.

[5] After 3 broken trends to 2022(3), a forecaster might have reselected by TIS, and the resulting model would have forecast slightly more accurately than (10).

References

Bernanke, B. 2024. "Forecasting for Monetary Policy Making and Communication at the Bank of England: A Review." Review, Bank of England, London, UK. https://www.bankofengland.co.uk/independent-evaluation-office/forecasting-for-monetary-policy-making-and-communication-at-the-bank-of-england-a-review.

Castle, J. L., J. A. Doornik, and D. F. Hendry. 2021. "Robust Discovery of Regression Models." *Econometrics and Statistics* 26: 31–51. https://doi.org/10.1016/j.ecosta.2021.05.004.

Castle, J. L., J. A. Doornik, and D. F. Hendry. 2025. "Could the Bank of England Have Avoided Mis-Forecasting UK Inflation During 2021–24?" *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2025.07.001.

Chow, G. C. 1960. "Tests of Equality Between Sets of Coefficients in Two Linear Regressions." *Econometrica* 28: 591–605. https://doi.org/10.2307/1910133.

Clements, M. P., and D. F. Hendry. 1996. "Intercept Corrections and Structural Change." *Journal of Applied Econometrics* 11: 475–494. https://doi.org/10.1002/(SICI)1099-1255(199609)11:5<475::AID-JAE409>3.0.CO;2-9.

Clements, M. P., and D. F. Hendry. 1998. *Forecasting Economic Time Series*. Cambridge University Press. https://doi.org/10.1017/CBO9780511599286.

Coroneo, L. 2025. "Forecasting for Monetary Policy." *International Journal of Forecasting*. https://doi.org/10.48550/arXiv.2501.07386.

Doornik, J. A., and H. Hansen. 2008. "An Omnibus Test for Univariate and Multivariate Normality." *Oxford Bulletin of Economics and Statistics* 70: 927–939. https://doi.org/10.1111/j.1468-0084.2008.00537.x.

Engle, R. F. 1982. "Autoregressive Conditional Heteroscedasticity, With Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50: 987–1007. https://doi.org/10.2307/1912773.

Ericsson, N. R. 2012. *Detecting Crises, Jumps, and Changes in Regime*. Working paper, Federal Reserve Board of Governors. https://eesp.fgv.br/sites/eesp.fgv.br/files/file/Neil_Ericsson.pdf.

Ericsson, N. R. 2017. "Economic Forecasting in Theory and Practice: An Interview With David F. Hendry." *International Journal of Forecasting* 33: 523–542. https://doi.org/10.1016/j.ijforecast.2016.10.001.

Findley, D. F., B. C. Monsell, W. R. Bell, W. R. Otto, and B.-C. Chen. 1998. "New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program (With Discussion)." *Journal of Business and Economic Statistics* 16: 127–177. https://doi.org/10.1080/07350015.1998.10524743.

Gardner, E. S. 2006. "Exponential Smoothing: The State of the Art: Part II." *International Journal of Forecasting* 22: 637–666. https://doi.org/10.1002/for.3980040103.

Gardner, E. S. J., and E. Mckenzie. 1985. "Forecasting Trends in Time Series." *Management Science* 31: 1237–1246. https://doi.org/10.1287/mnsc.31.10.1237.

Giannellis, N., S. G. Hall, G. P. Kouretas, G. S. Tavlas, and Y. Wang. 2025. "Policymaking in Periods of Structural Changes and Structural Breaks." *Journal of Forecasting* 44, no. 3: https://onlinelibrary.wiley.com/toc/1099131x/2025/44/3. [Correction added on 10 November 2025, after first online publication: The publication details of this reference have been corrected in this version.]

Godfrey, L. G. 1978. "Testing for Higher Order Serial Correlation in Regression Equations When the Regressors Include Lagged Dependent Variables." *Econometrica* 46: 1303–1313. https://doi.org/10.2307/1913830.

Hendry, D. F. 2025. "Looking Back to 1991 Economic Forecasting: Introducing Cointegration." *Oxford Bulletin of Economics and Statistics*. https://doi.org/10.1111/obes.70003.

Hendry, D. F., and M. P. Clements. 1994. "On a Theory of Intercept Corrections in Macro-Economic Forecasting." In *Money, Inflation and Employment: Essays in Honour of James Ball*, edited by S. Holly, 160–182. Edward Elgar. https://doi.org/10.4337/9781781958919.00015.

Hendry, D. F., and J. A. Doornik. 2014. *Empirical Model Discovery and Theory Evaluation*. MIT Press. Some chapters on. https://doi.org/10.7551/mitpress/9780262028356.001.0001.

Hendry, D. F., and S. Johansen. 2015. "Model Discovery and Trygve Haavelmo's Legacy." *Econometric Theory* 31: 93–114. https://doi.org/10.1017/S0266466614000218.

Hendry, D. F., and J. N. J. Muellbauer. 2024. "Why Did the Bank of England Need a Review of Its Dorecasting Record?" Economic Observatory. https://www.economicsobservatory.com/why-did-the-bank-of-england-need-a-review-of-its-forecasting-record.

Klein, L. R. 1971. *An Essay on the Theory of Economic Prediction*. Markham Publishing Company.

Ramsey, J. B. 1969. "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis." *Journal of the Royal Statistical Society B* 31: 350–371. https://www.jstor.org/stable/2984219.

Suits, D. B. 1957. "Use of Dummy Variables in Regression Equations." *Journal of the American Statistical Association* 52: 548–551. https://doi.org/10.2307/2281705.

Turner, D. S. 1990. "The Role of Judgement in Macroeconomic Forecasting." *Journal of Forecasting* 9: 315–345. https://doi.org/10.1002/for.3980090404.

Walker, A., F. Pretis, A. Powell-Smith, and B. Goldacre. 2019. "Variation in Responsiveness to Warranted Behaviour Change Among NHS Clinicians: A Novel Implementation of Change-Detection Methods in Longitudinal Prescribing Data." *British Medical Journal* 367: l5205. https://www.bmj.com/content/367/bmj.l5205.

Wallis, K. F., M. J. Andrews, P. G. Fisher, J. Longbottom, and J. D. Whitley. 1986. *Models of the UK Economy: A Third Review by the ESRC Macroeconomic Modelling Bureau*. Oxford University Press.

White, H. 1980. "A Heteroskedastic-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48: 817–838. https://doi.org/10.2307/1912934.

Young, R. M. 1979. "Forecasting the US Economy With an Econometric Model." In *Economic Modelling*, edited by P. Ormerod. Heinemann.