Nuffield College
UNIVERSITY OF OXFORD

ECONOMICS DISCUSSION PAPERS

Paper No: 2025-W01

Forecasting UK Inflation over 2021-2024

By Jennifer L. Castle, Jurgen A. Doornik and David F. Hendry

January 2025

Forecasting UK Inflation over 2021–2024

Jennifer L. Castle[a], Jurgen A. Doornik[b] and David F. Hendry[b*]
[a] Climate Econometrics and Calleva Project, Magdalen College, University of Oxford, UK;
[b] Climate Econometrics, Nuffield College, University of Oxford, New Road, OX11NF, UK.

January, 2025

Abstract

UK annual consumer price inflation rose rapidly from mid 2021, peaking over 9% in late 2022 after a series of essentially unpredictable shocks led to large forecast errors by the Bank of England. A sequence of increasingly large same-sign 1-step-ahead forecast errors as the forecast origin advances are most likely due to a sudden trend shift. We show that a small number of impulse indicators acting as intercept corrections to set forecasts back on track can be quickly tested for replacing by a broken linear or log-linear trend, illustrated by forecasting the UK's inflation.

JEL classifications: C2, C5, J3.
Keywords: Rapid Shift Detection, Trend Breaks, Impulse Indicators, Intercept Corrections, UK Inflation.

1 Introduction

After a series of essentially unpredictable shocks from the ending of COVID-19 pandemic lockdowns, supply chain disruption then the energy crisis caused by Russia's invasion of Ukraine, UK annual inflation measured by the Consumer Prices 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 (2024) shows that standard forecasting benchmarks like a random walk and a scalar autoregression had 1-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%: see (Coroneo, 2024, Table 1).

A sequence of increasingly large one-sided 1-step ahead forecast errors as the forecast origin advances suggests a trend change. Forecasts can be 'put back on track' by impulse indicators acting as intercept corrections (ICs) at the forecast origin as the value of the impulse indicator is the forecast error at that time point (see Clements and Hendry, 1996, although those authors use step indicators as ICs). Such forecast errors could be due to large outliers or mis-measurements, a step shift in the mean of the process, or a trend break. When accurate forecasting is the objective, in addition to rapidly detecting a shift, a forecasting model for the post-shift period is required. To isolate the source of the succession of such forecast errors, and so capture any sudden rapid shifts, we use a deterministic-trend model and test if the first few significant ICs can be eliminated when replaced by a broken linear or log-linear trend, as well as encompassing a step shift.

*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 Neil Ericsson. All calculations and graphs used *PcGive* (Doornik and Hendry, 2021) and *Ox Professional* (Doornik, 2018). email: jennifer.castle@magd.ox.ac.uk, jurgen.doornik@nuffield.ox.ac.uk and david.hendry@nuffield.ox.ac.uk

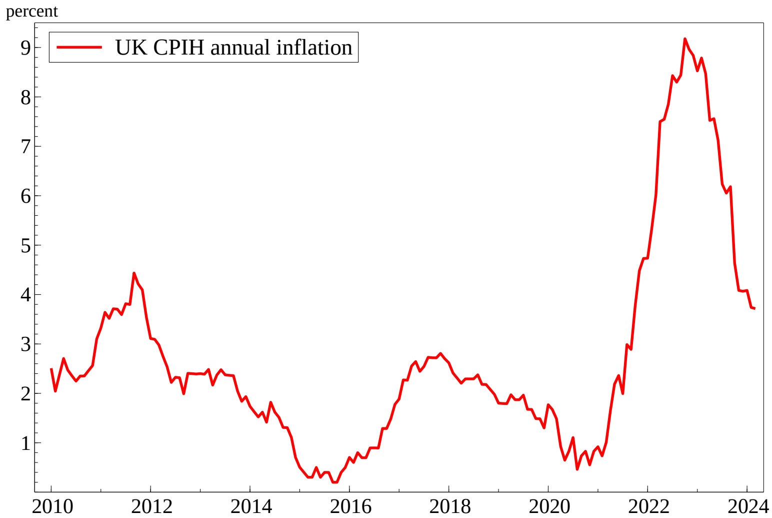

Figure 1: UK CPIH annual inflation, 2010–2024.

Using impulse indicators to correct forecast origin mis-forecasts has three advantages. First, since they act as one-off ICs, the next forecast commences from the forecast origin data value with unchanged parameter estimates, which will lead to further noticeable forecast errors if there is a location shift or trend shift, but not from one-off outliers or measurement errors which such ICs fix. Second, successive large forecast errors reveal that the current model is inadequate and needs updating, but the few new observations available from which to do so are usually insufficient for adapting large systems. However, extending a deterministic-trend model by a broken linear or log-linear trend is easily implemented. Finally, the adequacy of either of those trend extensions can be tested by the resulting insignificance of the ICs and an encompassing test against the other trend assumption: see Castle et al. (2024) for the underlying theory, and a discussion of other approaches to break detection. Because the new and old trends will diverge increasingly, ever larger forecast errors will result if not corrected. Consequently, despite having few (only 2 or sometimes 3) post-break observations, the new trend can be estimated reasonably accurately and so continue to forecast adequately until another shift occurs.

Since our (ex post) forecasts can be tested against the actual outcomes, we can evaluate the approach for the recent surge in UK inflation, acting as an investigator who sequentially forecasted many steps ahead. On finding a sequence of large, same-signed 1-step ahead forecast errors despite correcting such errors using impulse indicators as intercept corrections, the forecaster seeks the earliest date each sequence of ICs can be replaced by 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, 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 annual inflation forecasts can be derived. The forecasts are respectably accurate till the following shift, and one set is able to predict 17 months ahead across peak inflation and its ensuing slowdown.

The structure of the paper is as follows. Section 2 briefly describes indicator saturation estimators, then Section 3 detects and forecasts after sudden shifts in annual UK inflation over 2021(3)–2024(3). Section 4 simulates rapid detection of a deterministic trend break and Section 5 concludes. The Appendix analyses the properties of forecasting after one large forecast error.

2 Indicator saturation estimation

Indicator saturation estimators (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:

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

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 must use very tight significance levels α to control the probability of retaining irrelevant indicators, particularly if impulses, steps and trends are searched jointly, denoted super-saturation (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).

3 Modelling 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 modelling the log of monthly Consumer Prices 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 equation (1) where all trends have been scaled by 100:

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

$$\begin{aligned}
\widehat{\sigma} &= 0.20\% \;\; \mathsf{R}^2 = 0.999 \;\; \mathsf{F}_{\mathsf{ar}}(7,116) = 2.986^{*} \;\; \mathsf{F}_{\mathsf{arch}}(7,121) = 0.78 \;\; T = 2010(1)\text{-}2021(3) \\
\chi^2_{nd}(2) &= 2.22 \;\; \mathsf{F}_{\mathsf{Het}}(20,114) = 1.46 \;\; \mathsf{F}_{\mathsf{reset}}(2,121) = 0.07 \;\; \mathsf{F}_{\mathsf{Chow}}(1,123) = 10.5^{**}
\end{aligned}$$

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.[1] 10 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 $\widehat{\beta}_1 = 0.071$.

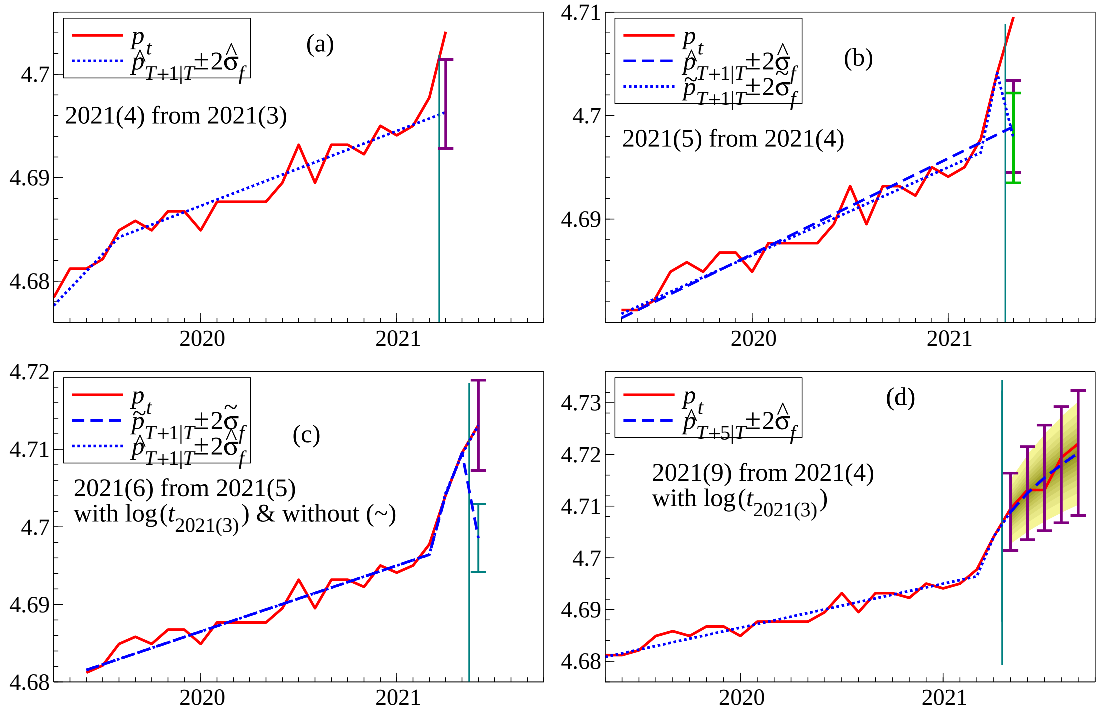

Figure 2: Fitted & actual values for $\widehat{p}_t$ with 1-step ahead forecast for: (a) 2021(4) from 2021(3) from (1); (b) 2021(5) from 2021(4) without $\mathsf{I}_{2021(4)}$ ($\widehat{p}_{T+1|T}$, dashed) and with ($\widetilde{p}_{T+1|T}$, dotted); (c) 2021(6) from 2021(5) (i) with $I_{2021(5)}$ added to ((1): $\widetilde{p}_{T+1|T}$, dashed) and (ii) after adding $\log(t_{2021(3)})$ ((2): $\widehat{p}_{T+1|T}$, dotted); (d) forecasting by (2) to 2021(9) from 2021(4) with both error bars and fans.

Figure 2(a) shows the fitted and actual values and the 1-step ahead forecast by (1) for 2021(4) from 2021(3) and (b) 2021(5) from 2021(4) with $1_{2021(4)} = 0.0078$ ($\widetilde{p}_{T+1|T}$: $\mathsf{F}_{\mathsf{Chow}}(1, 123) = 29.0^{**}$) and without ($\widehat{p}_{T+1|T}$: $\mathsf{F}_{\mathsf{Chow}}(1, 124) = 22.7^{**}$). 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 intercept correction (denoted by $\mathsf{I}_{2021(4)}$), 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, emphasising the trend shift.

Next, panel (c) shows forecasting 2021(6) from 2021(5) after also adding to (1) (i) $1_{2021(5)}$ with $\mathsf{t} = 5.4$ and (ii) the log of the broken linear trend $\log(t_{2021(3)})$, (which is zero initially as $t_{2021(3)}$ starts at unity). The latter eliminates the two impulse indicators for 2021(4) and 2021(5)

[1] 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), $\mathsf{F}_{\mathsf{arch}}$ tests autoregressive conditional heteroscedasticity (see Engle, 1982), $\mathsf{F}_{\mathsf{het}}$ tests residual heteroskedasticity (see White, 1980), $\chi^2_{\mathsf{nd}}(2)$ tests non-Normality (see Doornik and Hansen, 2008), $\mathsf{F}_{\mathsf{reset}}$ tests non-linearity (see Ramsey, 1969), and $\mathsf{F}_{\mathsf{chow}}$ tests parameter constancy (see Chow, 1960) over the forecast period. One star indicates test significance at 5%, two at 1%.

and is recorded in (2) with an insignificant Chow test. Finally (d) extends the forecast horizon to 2021(9) with a multi-step RMSFE = 0.15% and $\mathsf{F}_{\mathsf{Chow}}(5, 123) = 0.52$, so (2) forecasts better out of sample than the in-sample fit, despite the coefficient of $\log(t_{2021(3)})$ being estimated from just 2 observations.

$$\begin{aligned}
\widehat{p}_t \;=\;& \underset{(0.019)}{4.45} - \underset{(0.039)}{0.27}\,\tau_{2010(11)} + \underset{(0.028)}{0.28}\,\tau_{2011(4)} + \underset{(0.01)}{0.083}\,\tau_{2013(4)} - \underset{(0.03)}{0.22}\,\tau_{2014(10)} \\
& - \underset{(0.19)}{0.80}\,\tau_{2015(4)} + \underset{(0.18)}{0.70}\,\tau_{2015(5)} - \underset{(0.02)}{0.21}\,\tau_{2016(2)} + \underset{(0.17)}{1.04}\,\tau_{2018(12)} \\
& - \underset{(0.23)}{1.3}\,\tau_{2019(1)} + \underset{(0.06)}{0.39}\,\tau_{2019(4)} + \underset{(0.006)}{0.071}\,t + \underset{(0.003)}{0.010}\,\log(t_{2021(3)}) \qquad (2)
\end{aligned}$$

$$\begin{aligned}
\widehat{\sigma} \;=\;& 0.20\% \;\; \mathsf{R}^2 = 0.999 \;\; \mathsf{F}_{\mathsf{ar}}(7, 116) = 2.98^{*} \;\; \mathsf{F}_{\mathsf{arch}}(7, 122) = 0.78 \;\; T = 2010(1)\text{-}2021(4) \\
\chi^2_{nd}(2) \;=\;& 2.28 \;\; \mathsf{F}_{\mathsf{Het}}(20, 114) = 1.46 \;\; \mathsf{F}_{\mathsf{reset}}(2, 121) = 0.07 \;\; \mathsf{F}_{\mathsf{Chow}}(1, 123) = 0.52
\end{aligned}$$

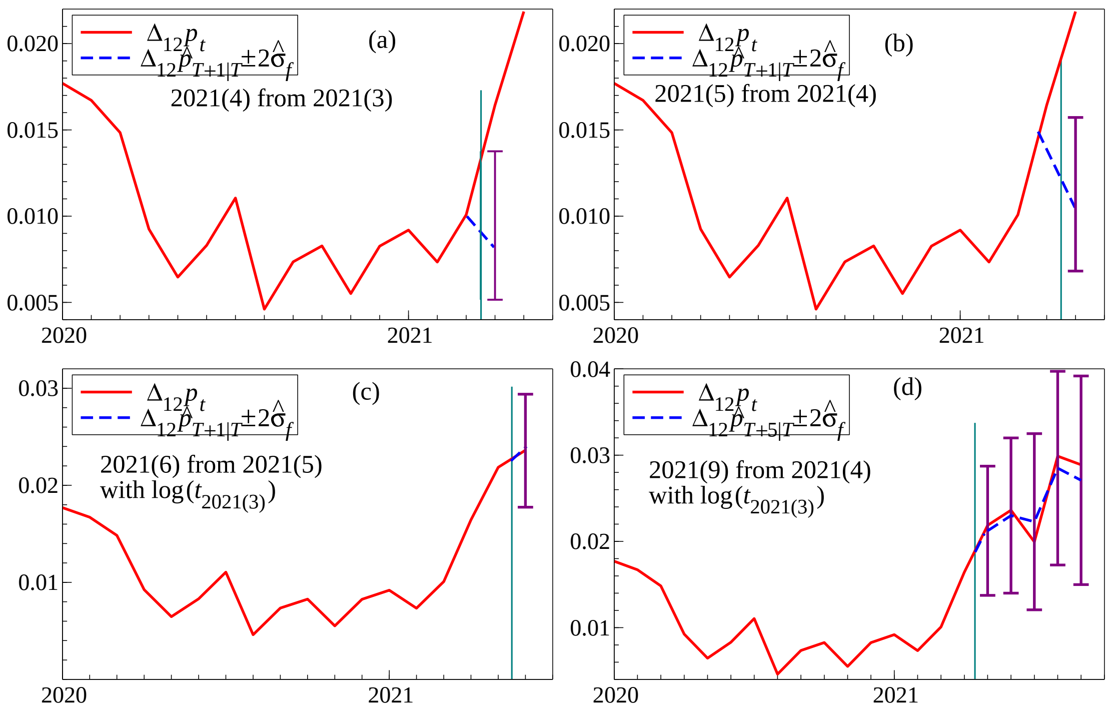

Figure 3: Actual and derived forecast values for $\widehat{\Delta_{12}p_t}$ and the 1-step ahead forecasts for: (a) 2021(4) from 2021(3); (b) 2021(5) from 2021(4); (c) 2021(6) from 2021(5) after adding $\log(t_{2021(3)})$ in (3); (d) forecasting 2021(9) from 2021(4) with $\log(t_{2021(3)})$.

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 re-centered on annual changes. The outcomes are shown in Figure 3. The forecasts (a)–(d) are derived from those in Figure 2. That the first two are below the previous outcome is all too common with equilibrium correction models.

3.1 The next break

We continue to follow the forecaster making 1-step ahead forecasts as the forecast origin now advances from 2021(9). The model is (2) augmented by $\log(t_{2021(3)})$, now estimated up to 2021(9) with $\widehat{\sigma} = 0.20\%$. As $\mathsf{F}_{\mathsf{Chow}}(1, 128) = 12.38^{**}$ when forecasting 2021(10) from 2021(9), a second break has happened as seen in Figure 4(a). This is confirmed when next forecasting 2021(11) from 2021(10) in Figure 4(b). 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 deliver $\mathsf{F}_{\mathsf{Chow}}(1, 128) = 44.8^{**}$, confirming it is not simply a step shift.

From Figure 4(b), 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 4(c) with $\mathsf{F}_{\mathsf{Chow}}(1, 129) = 0.34$ and $\widehat{\sigma} = 0.20\%$. Finally, Figure 4(d) shows multi-step forecasts from 2021(11) to 2022(3), with $\mathsf{F}_{\mathsf{Chow}}(5, 128) = 1.81$ and RMSFE = 0.40%.

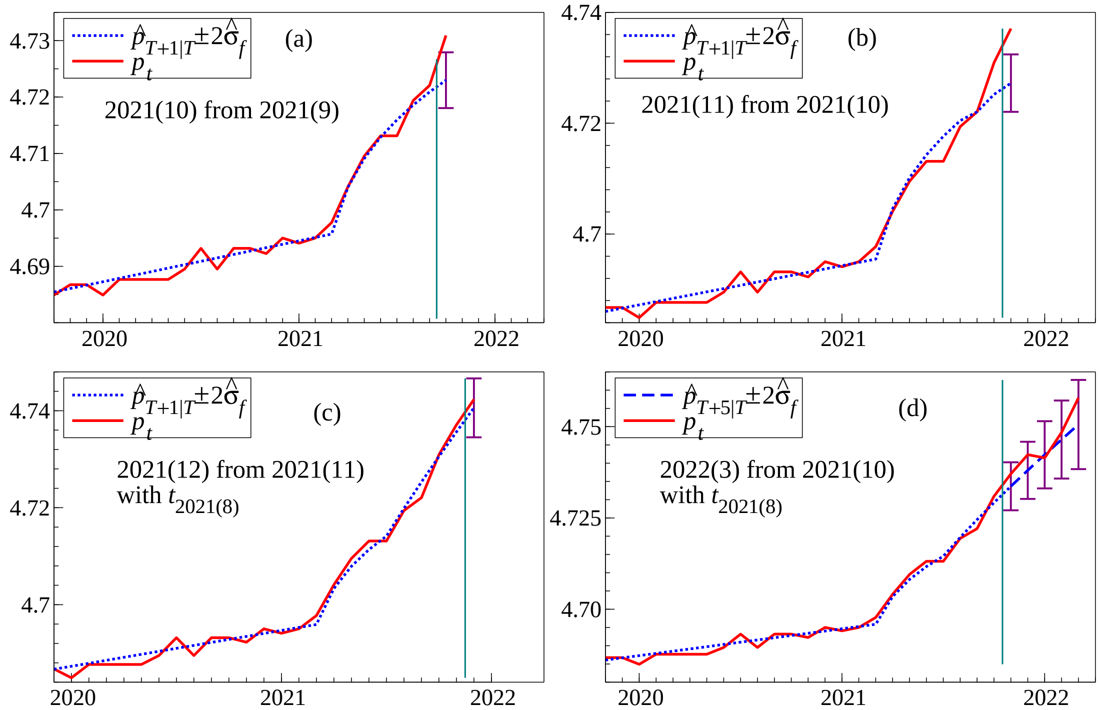

Figure 4: Fitted and actual values for $\widehat{p}_t$ and the 1-step ahead forecast for: (a) 2021(10) from 2021(9); (b) 2021(11) from 2021(10); (c) 2021(11) from 2021(10) after adding the linear trend $t_{2021(8)}$ to (2); (d) forecasting from 2021(10) to 2022(3) with $t_{2021(8)}$.

3.2 Not another break!

Forecasting 2022(4) from 2022(3) leads to another significant failure with $\mathsf{F}_{\mathsf{Chow}}(1, 133) = 63^{**}$ as seen in Figure 5(a). By the time this shift could have been observed, Russia's invasion of 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) (Panel (b)) leading to $\mathsf{I}_{2022(4)}$ & $\mathsf{I}_{2022(5)}$ with ts of 7.9 and 8.0.

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

$$
\begin{aligned}
\widehat{\sigma} &= 0.21\%\ \ \mathsf{R}^2 = 0.999\ \ \mathsf{F}_{\mathsf{ar}}(7,126) = 3.53^{**}\ \ \mathsf{F}_{\mathsf{arch}}(7,134) = 1.21\ \ T = 2010(1)\text{–}2022(4) \\
\chi^2_{nd}(2) &= 3.83\ \ \mathsf{F}_{\mathsf{Het}}(26,121) = 3.25^{**}\ \ \mathsf{F}_{\mathsf{reset}}(2,131) = 0.13\ \ \mathsf{F}_{\mathsf{Chow}}(1,132) = 0.04
\end{aligned}
$$

Adding the next broken log-linear trend $\log(t_{2022(3)})$ eliminates the ICs and delivers the outcome in (3) yielding $\mathsf{F}_{\mathsf{Chow}}(1,132) = 0.042$ for 2022(5), shown in Figure 5(c), greatly reducing the next forecast error, as $\mathsf{F}_{\mathsf{Chow}}(1,133) = 0.16$ for 2022(6). The model continues to forecast reasonably accurately through to 2023(9), which is 17 periods ahead, and although p_t is slightly overpredicted from 2023(7) leading to $\mathsf{F}_{\mathsf{Chow}}(17,133) = 3.96^{**}$, the RMSFE is 0.47%, and the tracking is close as seen in Figure 5(d).

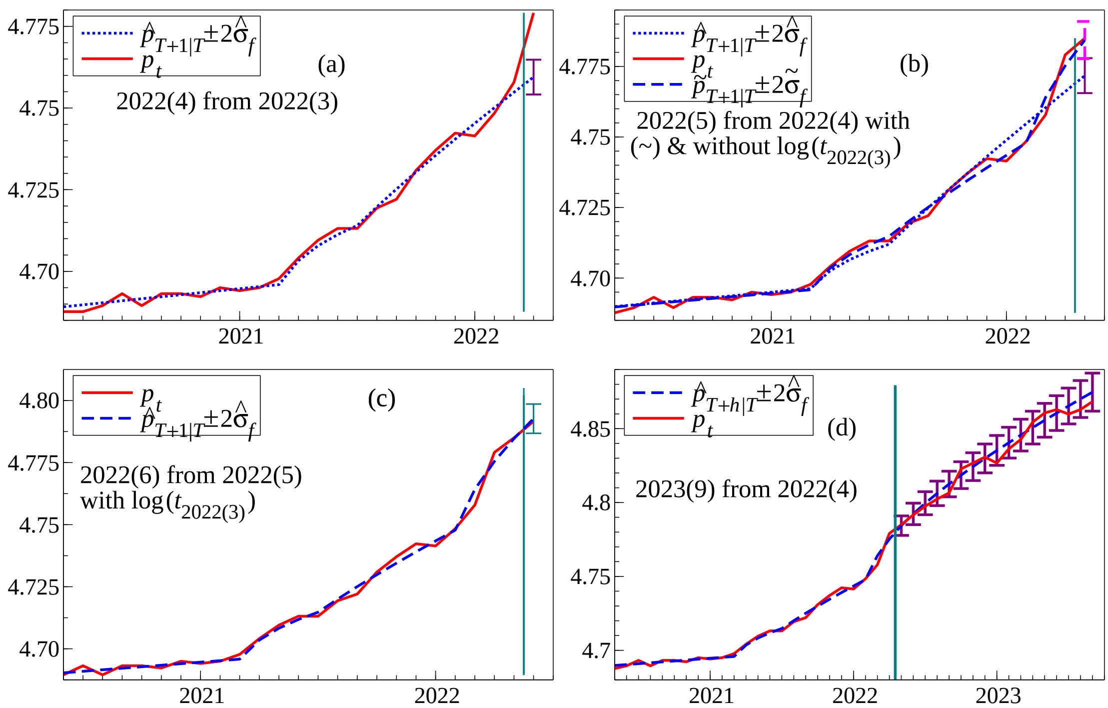

Figure 5: Fitted and actual values for $\widehat{p}_t$ and the 1-step ahead forecast for: (a) 2022(4) from 2022(3); (b) 2022(5) from 2022(4); (c) 2022(6) from 2022(5) after also adding $\log(t_{2022(3)})$ to (2); (d) forecasting from 2022(4) to 2023(9) by (3).

Figure 5(d) with the 17-steps ahead forecasts for 2022(5)–2023(9) confirms that medium-term forecasts can be usefully accurate despite the most recent broken trend being estimated from just

2 observations, obviously conditional on no new breaks occurring.[2]

3.3 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 17-steps ahead forecasts for p_t show that is indeed possible. Figure 6 plots all the sets of multi-step ahead forecasts for UK prices and Figure 7 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.

3.4 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 $\mathsf{F}_{\mathsf{Chow}}(1, 150) = 7.3^*$ revealing another trend break as the increases in p_t slowed. Once again, a broken log-linear trend from 2023(10) produces $\mathsf{F}_{\mathsf{Chow}}(5, 150) = 1.1$ over 2023(11)–2024(3) although fitted to just one non-zero observation as shown in Figure 6. That coefficient estimate is -0.0105 (0.004), and estimated till 2024(3) is -0.011 (0.0013), hence the accurate forecasts.

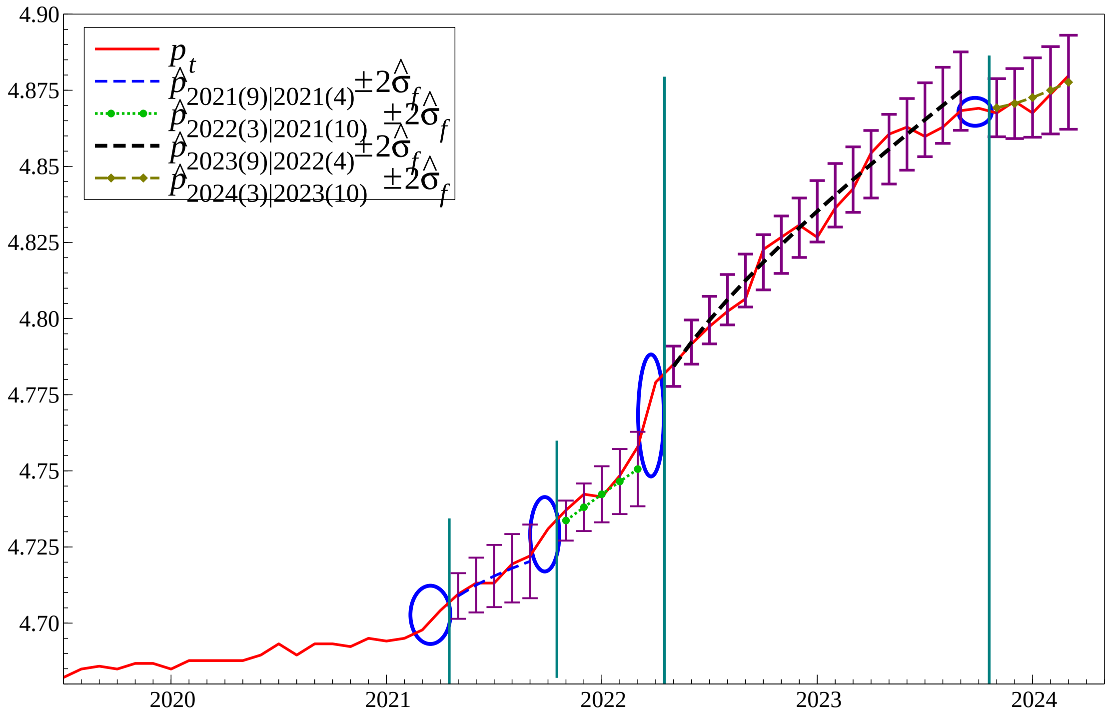

Figure 6: Multi-steps ahead forecasts of p_t over the inflation upsurge and fall from 2021(4)–2024(3), spanning the four episodes shown by vertical lines with ellipses highlighting the breaks.

[2] 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 (3).

The Bank started raising interest rates from 0.1% in February 2022, continuing to raise in small steps till 5.25% in August 2023, yet our accurate 17 month ahead (albeit ex post) forecasts to 2023(9) were made in 2022(4) prior to most changes. Had the method in Castle et al. (2024) been available before 2022(4), the same forecasts would have been made: one wonders what the MPC would have made of them. The possible implications range from 'the impacts of interest rates on inflation are slow and small'; 'the inflationary surge was mainly global energy price driven, so domestic policy had little effect'; to 'interest rate increases slowed a linear trend price increase to a log-linear, happening to match our forecasts'. Hendry and Muellbauer (2024) estimate the impacts on UK inflation in 2022 of import price inflation, energy shortages and price rises of 170%, suggesting they accounted for about $3/4$ of the peak 9% rise in CPIH, and all these influences dropped considerably in 2023 so UK inflation rose more slowly.

Figure 7 collects our whole period multi-step ahead monthly forecasts and outcomes over 2021(4)–2024(3) for annual inflation ($\Delta_{12}p_t$), including the large forecast errors that prompted the additions of broken trends (two vertically-aligned error bars are without and with ICs).

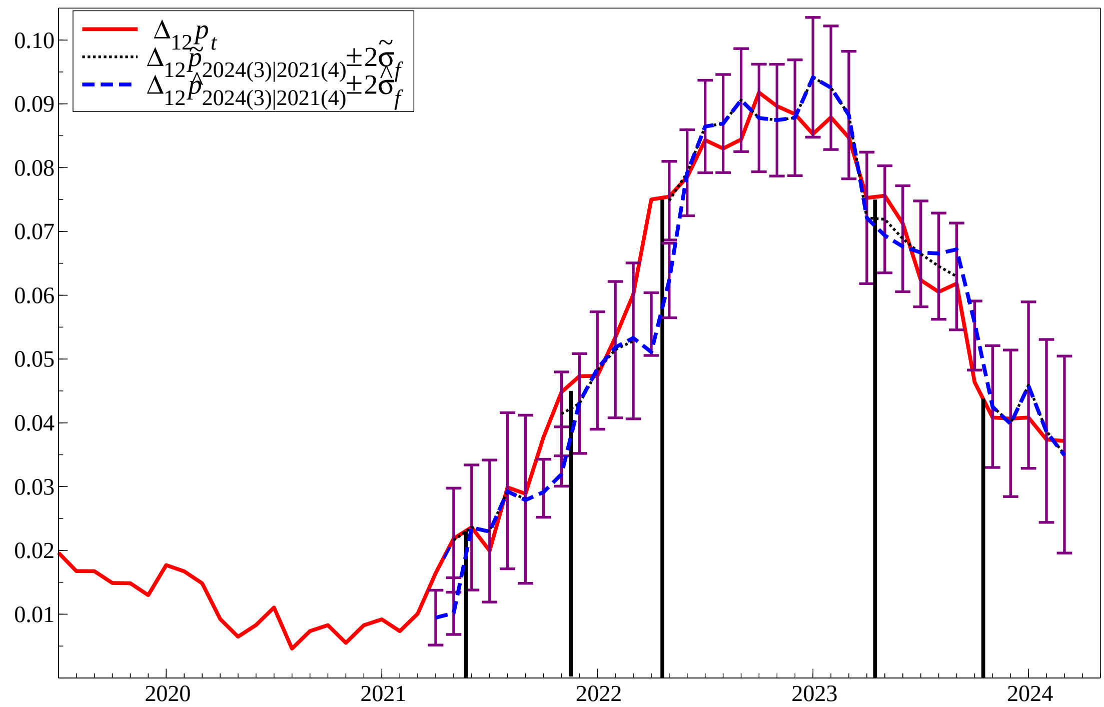

Figure 7: Multi-steps ahead forecasts and outcomes over 2021(4)–2024(3) for $\widetilde{\Delta_{12}p}_{T+h|T}$ (dashed) & sub-period forecasts $\widehat{\Delta_{12}p}_{T+h|T}$ (dotted) shown with error bars.

An alternative comparison is with how the basic model would have forecast in the absence of our approach to rapidly correcting after shifts. Figure 8 records whole period outcomes using 12 month ahead forecasts for first three episodes and 6 and 5 months ahead for last two. Forecasting beyond the next break naturally leads to significant forecast errors but that could not be known till after the shifts and Figure 7 highlights the advantages of then rapidly correcting from the initial 1-step ahead large forecast errors. Moreover, not correcting would have been a very bad strategy, and although we have used the deterministic trend equation as the illustration in Figure 8, similar patterns of failure would have occurred for the typical benchmarks that Coroneo (2024) considered.

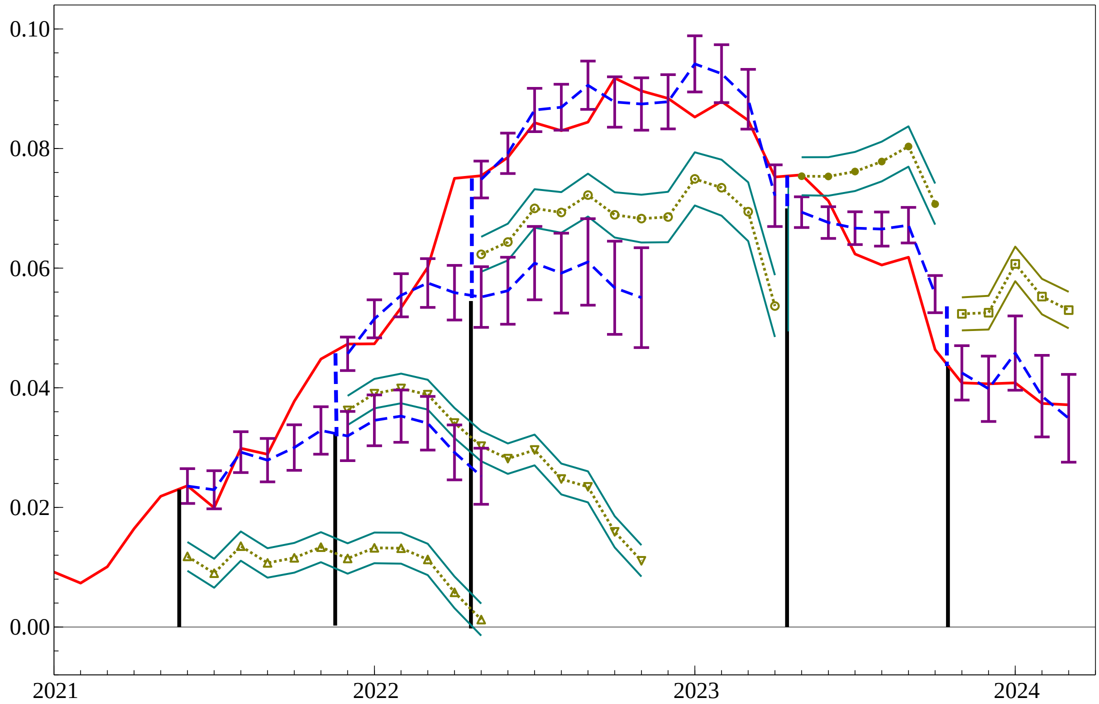

Figure 8: 12 month ahead forecasts for first three episodes and 6 & 5 months ahead for last two, immediately after shifts (vertical lines) (i) without broken trends (green, dotted lines with symbols, always very poor) and (ii) with broken trends (poor after first 2 breaks).

4 Simulating rapid detection of a deterministic trend break

We undertake a Monte Carlo simulation to evaluate the detection and forecast performance of the real-time forecasting procedure, both under the null of no break and under the alternative of a 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 $t = 1, \ldots, T + H$. All the simulations are based on $M = 10000$ replications.

The DGP is given by:

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

where $\boldsymbol{\tau}_T = (-79, -78, \ldots, -1, 0, 0, \ldots, 0)'$ but the full sample trend is $\boldsymbol{t} = (1, 2, \ldots, T + H)'$. 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 $\lambda + \rho$.

We estimate the model recursively for $t = 1, \ldots, T + h$ over $h = 1, \ldots 5$:

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

and record 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 $t = 1, \ldots, T+5$ is equivalent to recursively testing the individual indicators as the window increases from $T+1$ to $T+5$.

Figure 9 records the proportion of replications in which the impulse indicators are significant at a 1% significance level as the sample is increased from $t = 1, \ldots, T+1$ to $T+5$. When $\lambda = 0$, the trend in the forecast period is the same as the in-sample period, so is under 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 λ^*, determines the detectability of shifts.

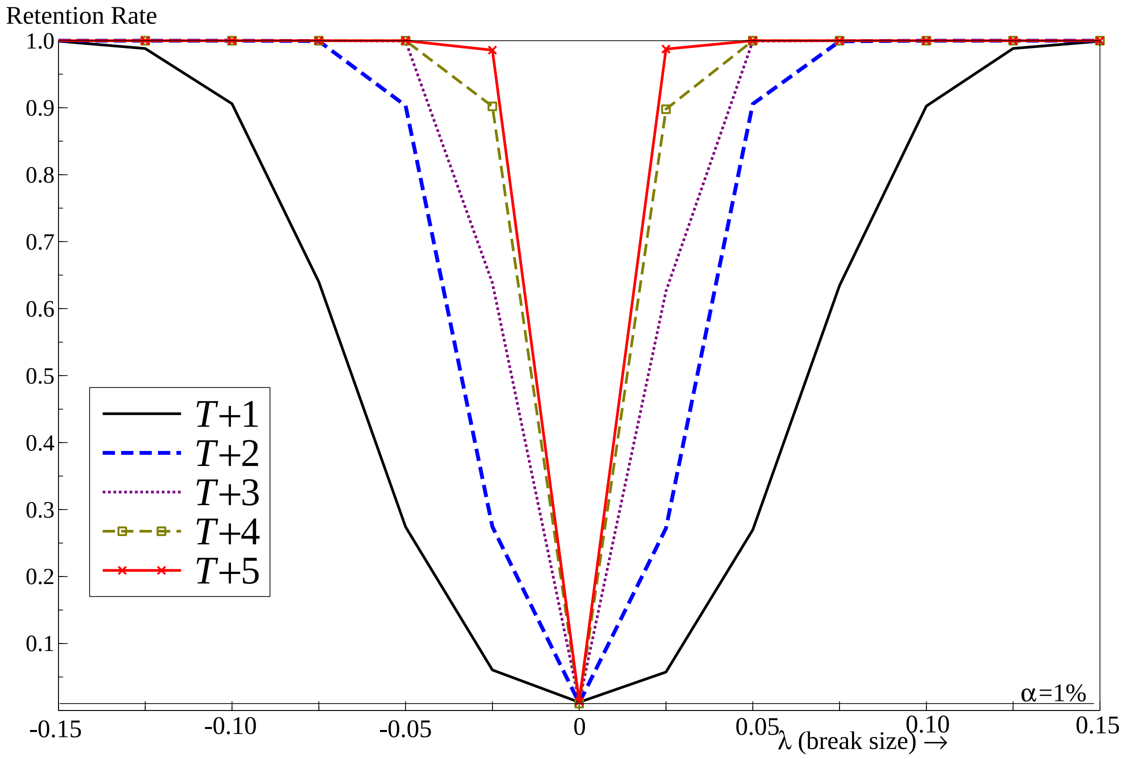

Figure 9: Proportion of impulse indicators that are significant at 1% for a break in trend (4).

After just two observations the retention rate of the impulse indicators increases rapidly as the trend break gets larger, and 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.[3]

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:

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

$$y_t = \beta_0 + \beta_1 t + \beta_2 \log(t^*_{T+h}) + \sum_{j=2}^{5} \beta_{\iota_j} 1_{T+j} + \nu_t \tag{7}$$

[3] 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.

where $t_{T+h} = h$ for $h = 1, \ldots, H$ and 0 for $1, \ldots, T$, and $\log(t^*_{T+h}) = \log(t_{T+h} + 1)$ 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 10 records the results, where panel (a) records the retention rate of the included trend, panel (b) records the average T-value of the trend (using T as a 't' value given the other uses of 't') and panel (c) records the retention rate of the impulse indicators when the trend is included. 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. Under the null of no break the retention probabilities are close to the chosen significance level. The trends are highly significant, even after just two observations, if the break is moderately large, shown by the average $\mathsf{T}-$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 misspecification is revealed by the 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 1% of the draws, the significance level of the exclusion test.

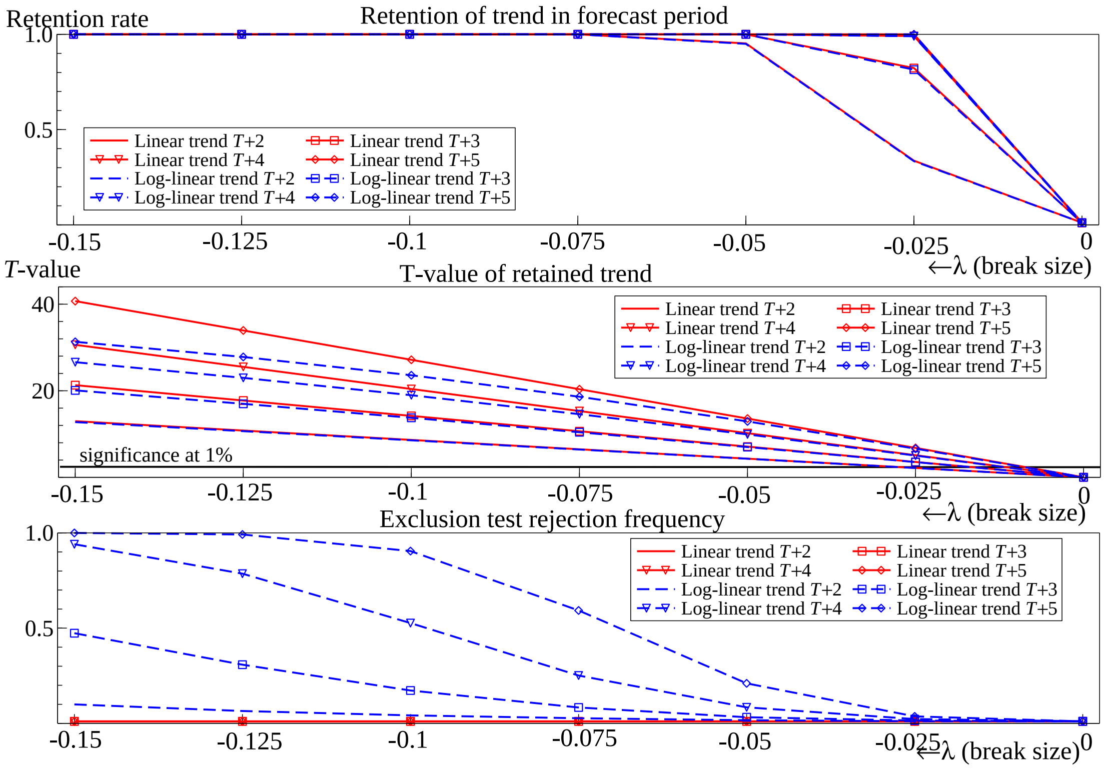

Figure 10: (a) the proportion of trends that are significant at a 1% significance level; (b) the average T-statistic of the trend; (c) the proportion of replications in which the p-value of the exclusion test over the ICs is less than 0.01, i.e. the number of draws in which ICs are significant in addition 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 $\mathsf{I}_{T+2}, \ldots, \mathsf{I}_{T+5}$.

The forecast performance of the approach is assessed by comparing five alternative forecasting approaches. These include (i) ignoring the break, i.e. (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, i.e. (6) with $\beta_{\iota_j} = 0,\ \forall j$;
(iv) using a log-linear trend from the forecast origin, i.e. (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 $t = 1, \dots, T_*$ and forecasts are produced for $T_* + 1, \dots, T_* + 5$. All models are identical for the initial recursion as the break in trend occurs at $T + 1$. Thus the models are estimated over $t = 1, \dots, T+1$ and dynamic forecasts are produced for $T+2, \dots, T+6 = T_*+1, \dots, T_*+5$. 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 $t = 1, \dots, T+2$ and dynamic forecasts are obtained for $T+3, \dots, T+7$, etc. up to an in-sample period $t = 1, \dots, T+4$ with forecasts over $T+5, \dots, T+9$. 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 with $|\lambda^*| \geq 2$ when $\sigma_\epsilon = 0.025$.

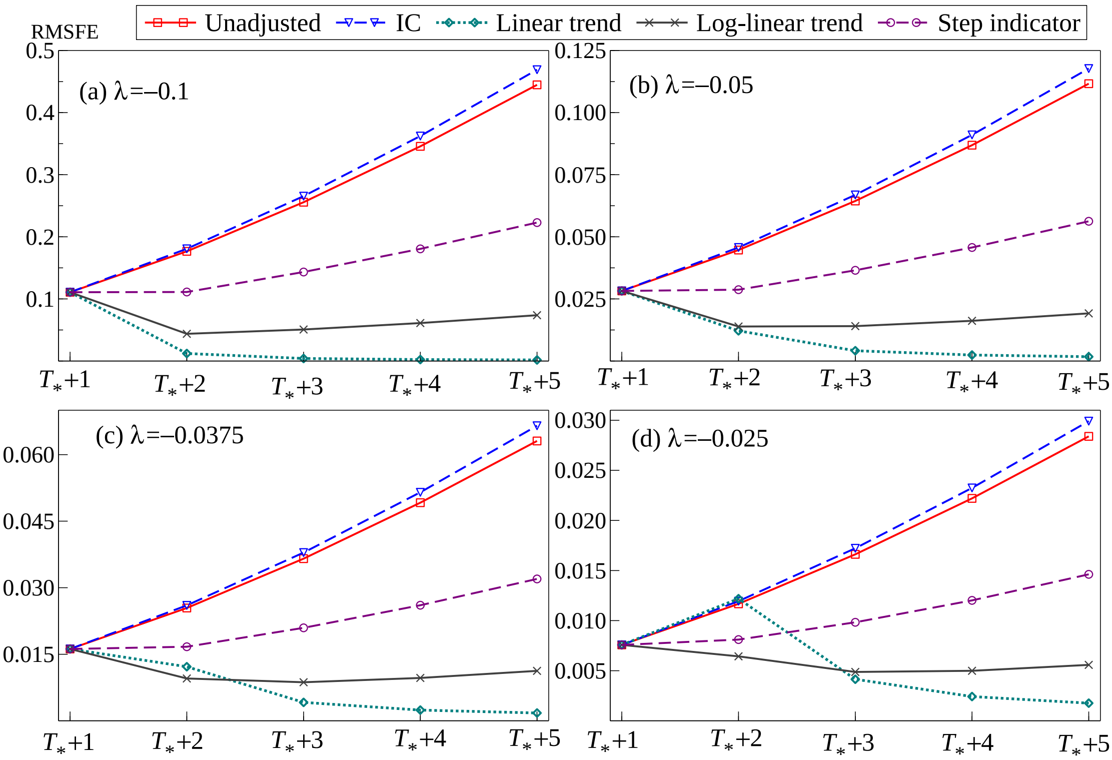

Figure 11: RMSFEs for a sequence of 5 dynamic forecasts commencing at $T_* = T, \dots, T+4$, where the break in linear trend of magnitude λ occurs at $T + 1$.

The RMSFE results are reported in figure 11 for the four different break magnitudes.[4] When the break is large, just one observation is sufficient to estimate the linear trend and correct the forecasts, with the linear trend dominating the forecast performance of breaks of $\lambda = -0.05$ or larger in absolute value. Even if the break is of moderate size, the linear trend dominates after just 2 observations. The mis-specified log-linear trend does not remove the bias but it is still effective at reducing RMSFE relative to an unadjusted model. The intercept correction to set the forecasts back on track slightly worsens the forecast error relative to doing nothing. 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 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 linear and log-linear trend would mitigate this when the break form was uncertain.

[4]The RMSFE scale in figure 11 differs across break sizes to highlight differences between models. MEs and results for a range of other break sizes are available on request.

5 Conclusion

A sudden unanticipated upsurge 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 is as follows.

1] Once significant 1-step ahead forecast errors are detected, add impulse indicators acting as intercept corrections (ICs) to the model. The ICs ensure that the forecasting model is unchanged, hence maintain the previous trend, so rapidly reveal departures from any new upswing, and help discriminate trend breaks by a test from location shifts, outliers or measurement errors (which ICs 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 impulse indicators, 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 modelled 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). 1-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 by a log-linear trend commencing in 2022(3) enabled 17-steps ahead forecasting from 2022(5) to 2023(9) as annual inflation first peaked then fell. Modelling the final shift from 2023(10) again proved usefully accurate till our sample end, but would certainly fail as annual inflation stabilized. The 4 essentially unpredictable trend shifts are clearly visible in Figure 7, and were followed by significant forecast errors, but we experienced only seven large errors overall when rapidly detecting breaks, rather than long periods of systematic forecast failure.

Such an approach could 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 when the next failure will occur. When the model and data do not have a trend but a new one commences with 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.

Appendix: Forecasting after one large forecast error

Castle et al. (2024) analyse forecasting after trend breaks based on two or three observations and show its feasibility: here we consider doing so after a single post-break observation. Let $\widehat{\varepsilon}_{T+1|T} = y_{T+1} - \widehat{y}_{T+1|T}$ be a large 1-step ahead forecast error which happens following the deviation of a new log-linear trend from a previous linear trend at time T where:

$$y_{T+1} = \delta (T+1) + \psi \log \left(\mathtt{t}_{\{t>T\}} \right) + \varepsilon_{T+1} \tag{8}$$

when $\mathtt{t}_{\{t>T\}}$ is a linear trend equal to $(t - T + 1)$ for $t \geq T$ and zero otherwise, and $\varepsilon_t \sim \mathsf{IN}\left[0, \sigma_\varepsilon^2\right]$.

Since the break is not known at T, the next unadjusted forecast is (denoted by $\widehat{\ }$):

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

with the forecast error:

$$\widehat{\varepsilon}_{T+1|T} = \left(\delta - \widehat{\delta}\right)(T+1) + \psi \log(2) + \varepsilon_{T+1} \quad (9)$$

as $\mathsf{t}_{\{t>T\}} = 2$ at $T+1$. When $\mathsf{E}[\widehat{\delta}] = \delta$, neglecting the variance component from $(\delta - \widehat{\delta})$ as trend coefficient estimates have variances $\mathsf{O}\left(T^{-3}\right)$, then $\mathsf{E}[\widehat{\varepsilon}_{T+1|T}] = \log(2)\psi$ and $\mathsf{V}[\widehat{\varepsilon}_{T+1|T}] = \mathsf{E}[(\widehat{\varepsilon}_{T+1|T} - \mathsf{E}[\widehat{\varepsilon}_{T+1|T}])^2] = \sigma_\varepsilon^2$ with a mean-square forecast error (MSFE) of $\sigma_\varepsilon^2 + (\log(2)\psi)^2$.

The next forecast is made after adding the IC $\mathsf{I}_{\{T+1\}} = \widehat{\varepsilon}_{T+1|T} 1_{\{T+1\}}$ to set the forecast back on track at the origin, where from (9), the fitted value at $T+1$ becomes (denoted by $\widetilde{}$):

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

so:

$$\mathsf{I}_{\{T+1\}} = \left[(\delta - \widehat{\delta})(T+1) + \boldsymbol{\psi}\mathbf{log}(\mathbf{2}) + \varepsilon_{T+1}\right] 1_{\{T+1\}}$$

and hence $\mathsf{I}_{\{T+1\}}$ 'captures' the trend shift $\psi\log(2)$, which is the source of the large forecast error. However, since $1_{\{T+1\}} = 0$ at $T+2$, adding it still leads to the next 1-step forecast:

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

and hence another large forecast error, as seen above.

If instead, the forecaster had added the broken log-linear trend $\log\left(\mathsf{t}_{\{t>T\}}\right)$ 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 its coefficient would be $\mathsf{I}_{\{T+1\}}$ scaled by $\log(2)$ with $\mathsf{I}_{\{T+1\}}/\log(2) = \overline{\psi}$ so:

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

leading to:

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

where:

$$\mathsf{E}\left[\psi - \overline{\psi}\right] = \psi - \mathsf{E}\left[\mathsf{I}_{\{T+1\}}\right]/\log(2) = 0$$

so now $\mathsf{E}\left[\overline{\varepsilon}_{T+2|T+1}\right] = 0$, with a MSFE of σ_ε^2. For the final break in our sample, $\mathsf{I}_{2023(10)} = \overline{\psi} = -0.0073\ (0.003)$ whereas the log-linear trend coefficient was estimated as $-0.0105\ (0.004)$ which is $\mathsf{I}_{\{2023(10)\}}/\log(2)$. The resulting MSFE was approximately $\widehat{\sigma}_\varepsilon^2$ and stayed at that level for the remainder of the forecast horizon. Two caveats are that the data 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 the IC, hence the need for two forecast errors to test the break type.

References

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 (2024). Forecasting after the start of a trend break. Working paper, Nuffield College, Oxford University.

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.

Coroneo, L. (2024). Forecasting for monetary policy. *International Journal of Forecasting*, Submitted.

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, Washington, D.C. https://eesp.fgv.br/sites/eesp.fgv.br/files/file/Neil_Ericsson.pdf.

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.

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. and J. A. Doornik (2014). *Empirical Model Discovery and Theory Evaluation*. Cambridge, Mass.: 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 forecasting record? *Economic Observatory*. https://www.economicsobservatory.com/why-did-the-bank-of-england-need-a-review-of-its-forecasting-record.

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.

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.

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.