Accepted: 6 May 2024
DOI: 10.1111/ecca.12533

ORIGINAL ARTICLE

Economica LSE

Forecasting the UK top 1% income share in a shifting world

Jennifer L. Castle | **Jurgen A. Doornik** | **David F. Hendry**

University of Oxford

Correspondence
Davide Hendry, Nuffield College, University of Oxford, New Road, Oxford OX1 1NF Email: david.hendry@nuffield.ox.ac.uk

Abstract
UK top income shares have varied hugely over the past two centuries, ranging from more than 30% to less than 7% of pre-tax national income allocated to the top 1 percentile. We build a congruent dynamic linear regression model of the top 1% income share allowing for economic, political and social factors. Saturation estimation is used to model outliers and trend breaks, proxying underlying structural changes driving income inequality in the UK. We use the model to forecast the top 1% income share over the last 15 years, and compare to a range of forecast devices. Despite a well-specified constant parameter model conditioning on significant explanatory variables, the best performing forecasts are obtained from a random walk and a smoothed random walk. These results are explained by the presence of shifts in the income share over the forecast period, resulting in forecasts from equilibrium correction models converging to the wrong equilibrium. Our best prediction for 2026 based on the most recent data from 2021 (a 5-year ahead projection) is that the pre-tax top 1% income share will remain at the most recent realized value of 12.7%, but there is a large degree of uncertainty, with a 95% confidence band ranging from 10% to 15.7%.

1 | INTRODUCTION

This paper has two antecedents.

First, Sir Tony Atkinson, an LSE alumnus, was concerned about income and wealth distributions throughout his career, and would surely have been pleased to include an insightful paper in this centenary volume had he been alive. Atkinson and Piketty (2007, 2010) and Atkinson *et al.* (2011) were all concerned with analysing top income shares and the inequalities that they

This paper is part of the Economica 100 Series. Economica, the LSE "house journal" is now 100 years old. To commemorate this achievement, we are publishing 100 papers by former students, as well as current and former faculty. David Forbes Hendry received his MSc and PhD from the LSE.

revealed, so remembering Tony, we apply recently developed tools to modelling and forecasting the pre-tax UK top 1% income share. Tony was well aware of the complex determinants of top income shares, which are affected differentially over time by wars, energy crises, pandemics, fiscal and monetary policy, legislation, inheritance, demography and longevity, *inter alia*. Not only are the impacts of these variables difficult to model empirically, many are harder to forecast than the income shares themselves, so here we use data-based forecasting methods.[1]

Second, when he was at LSE, Hendry (1974) sought to understand the role of stochastic specification relative to other aspects of model formulation. His empirical system comprised six equations for consumers' expenditure on durable goods (C_d), all other goods and services (C_n), gross domestic fixed capital formation (I), inventory investment (I_v), and imports of goods and services (M), closed by an empirical relation for disposable income (Y_d). The system was acknowledged to be naive, both to minimize the already large computational burden for the time, and to focus attention on the stochastic properties of the error processes. The quarterly data sample was from 1957Q(1) to 1967Q(4), of which the last two observations were used for a test of its forecasts, which chanced to avoid the catastrophic forecast errors he later made for 1968Q(1) and 1968Q(2) (see Ericsson 2017). However, those later errors made him carefully analyse economic forecasting, a topic that he has continued to research with several co-authors, including those for this paper. The implications from five decades of research on modelling and forecasting facing economic shifts are applied below.

In this analysis, we focus on one measure of inequality, namely the pre-tax national income share allocated to the top 1 percentile of the UK population, recorded in Figure 1, although it should be noted that there are many measures of inequality that could be considered alternatively; see Hills *et al.* (2010) and Atkinson and Voitchovsky (2011). There are missing observations, with symbols marking observed data points, and we interpolate using an unobserved components time series model to infill missing data (Koopman *et al.* 2007). The data show periods of increasing and decreasing inequality with some rapid trend changes, posing difficulties for forecasting the future evolution of income inequality.

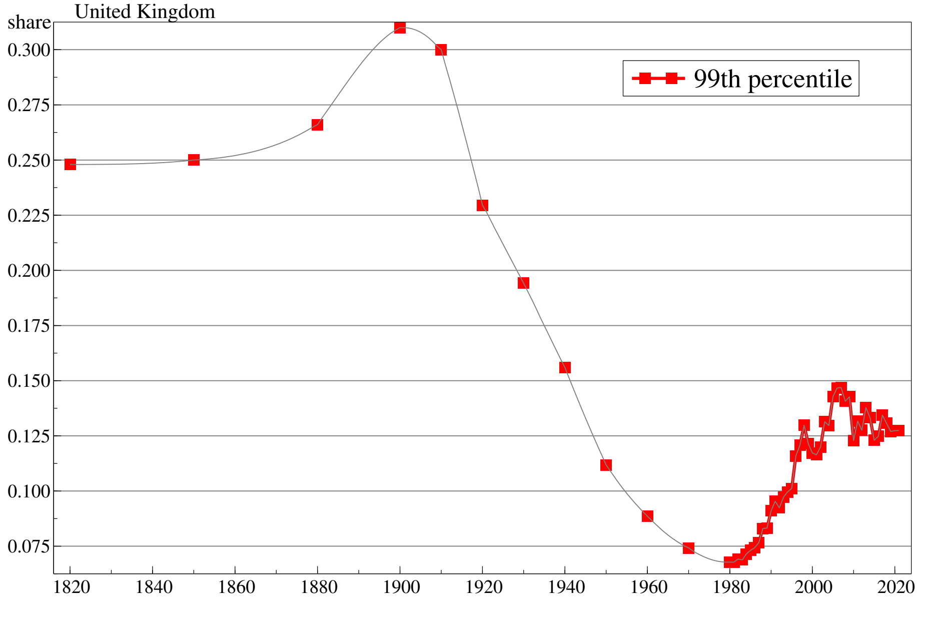

FIGURE 1 Top 1% share of pre-tax national income since 1820 for the UK. *Notes*: Square symbols represent observed data, with interpolated data for missing observations in grey. Source: World Inequality Database, https://wid.world (accessed 2 May 2024).

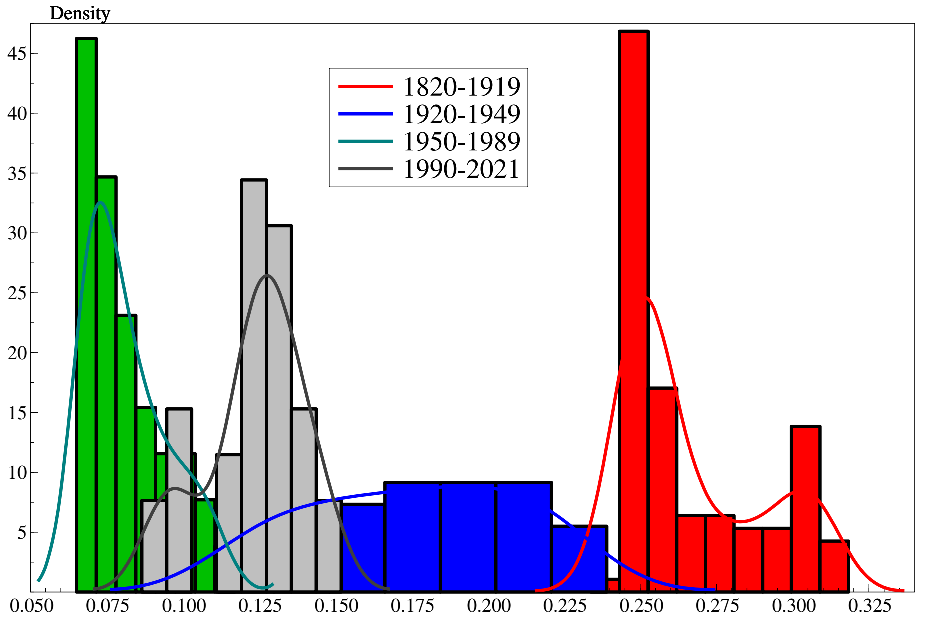

FIGURE 2 Distribution of top 1% share of pre-tax national income for the UK by sub-period.

Modelling and forecasting inequality data is difficult as it is wide-sense non-stationary, namely from evolving stochastic trends and sudden distributional shifts such as those shown in Figure 2. Problems are exacerbated by missing data, the latency in data provision, substantive revisions and occasional changes to data measurement systems. Figure 2 records the distribution of the top 1% income share by sub-periods, albeit using some interpolated data, which reveals significant shifts over time, with no overlap between pre-1920 and the 1950–89 period. Much forecasting theory implicitly assumes a stationary world, perhaps after differencing to remove any unit roots, which is an invalid assumption for income inequality data and its determinants. Hence we investigate forecasting methods that also allow for distributional shifts. Predicting future trends in income inequality matters both for guiding redistributive policy and for potential impacts on economic growth.[2]

There has been interest in forecasting income inequality trends stemming from Piketty (2014). Hubmer *et al.* (2021) provide a theoretical framework to modelling wealth inequality in the USA, motivating a range of potential drivers of inequality. Pierdzioch *et al.* (2022) look at higher-frequency forecasting of UK income inequality, and do find some short-term predictability in various inequality measures. Hood and Waters (2017) provide an example of scenario-based forecasting of UK income inequality. Gindelsky (2018) uses an autoregressive-distributed lag framework to select forecasting models of the US top income share, and finds that naive forecasting devices are hard to beat in terms of forecast performance. Blanchet *et al.* (2022) provide high-frequency income distributions for the USA, but do not produce forecasts. The literature producing medium-term *ex ante* econometric forecasts of top income shares is limited, and one of the contributions of this paper is to rectify this for the UK.

Our objective is to investigate what forecasting methods are useful for such non-stationary time series data with measurement errors and distributional shifts due to political, social, economic and policy changes in an unknown data generating process (DGP). Despite a large body of theory, economists have incomplete knowledge of the DGP driving income inequality, so must search for reasonable empirical modelling approximations. We use automatic model selection methods to build a conditional model of the top 1% UK income share, allowing for a range of economic and political drivers, many of which are significant in explaining movements in income inequality. The underlying DGP of the top income share is complex, so indicator saturation is used to model distributional shifts as well as proxy omitted variables. Three significant trend shifts are integral to the model (1998, 2000 and 2007). The model, estimated over 1975–2016, is well specified and provides a coherent explanation of the data in-sample, but forecast failure over 2017–21 highlights the problems of forecasting income inequality. The best performing forecasting model is a random walk, but there is a large degree of uncertainty in the forecasts.

A pseudo-out-of-sample forecasting exercise is used to evaluate forecast performance, computing sequences of dynamic forecasts over 2002–21. The benchmark models considered include univariate models that estimate either a stochastic or a time-varying deterministic trend. Gindelsky (2018) finds that naive forecasting models often outperform more complex models in the US context, and our findings are similar for the UK. Although the in-sample performances of naive models such as the random walk and smoothed random walk are worse than that of the conditional model, their forecast performance dominates. The explanation for such a result is provided in Section 2, which discusses the provable forecast theorems under wide-sense non-stationary data. Conditional models require forecasts of all their constituent parts, adding considerable additional forecast uncertainty when many of the explanatory variables are harder to forecast than income inequality itself. This result explains why univariate models often dominate in forecasting: system forecasting rarely pays unless variables have closely interacting feedbacks. Furthermore, the extrapolated trend is crucial to forecast performance but is particularly difficult to predict at turning points. Robust forecast devices that are essentially agnostic as to the direction of the trend then perform better. This is why a random walk can be hard to beat. Comprehensive conditional models are needed to understand the evolution of income inequality and inform policy, but for forecasting, an alternative set of models is warranted.

The structure of the paper is as follows. Section 2 briefly reviews the provable theorems about forecasting in stationary processes, and contrasts those with what can be established for wide-sense non-stationary data. Section 3 outlines approaches to addressing non-stationarity in the form of distributional shifts both in-sample using trend-indicator saturation and out-of-sample through robustifying model-based forecasts. Section 4 applies trend-indicator saturation to modelling the UK's top 1% income share, before Section 5 conducts a pseudo-out-of-sample forecasting exercise. The relative merits of modelling trend breaks in-sample and robustifying forecasts out-of-sample are assessed, and we find that the random walk and smoothed random walk forecasts dominate. To explain the forecasting evidence, Section 6 relates these empirical results to the provable forecasting theorems for the wide-sense non-stationary world described in Section 2. Section 7 concludes.[3]

2 | FORECASTING LINEAR STATIONARY AND NON-STATIONARY PROCESSES

A number of well-known theorems can be proved when forecasting future values of a strictly stationary process; see, for example, Clements and Hendry (1998, 1999). Defining a causal model as one that includes variables from the DGP, so its regressors are relevant, we summarize these as follows.

(1) A model matching the DGP with known parameter values will dominate in forecasting.
(2) Adding further variables to the DGP model will not improve forecasts.
(3) Causal models will outperform non-causal models (i.e. models without any relevant variables).
(4) The conditional expectation of the future value delivers the minimum mean square forecast error (MSFE) given the information used.
(5) Misspecified models (that omit relevant causal variables) will have higher forecast error variances than correctly specified ones, but should forecast within their anticipated accuracy.

(6) Long-run interval forecasts are bounded above by the unconditional variance of the process.
(7) Neither parameter estimation uncertainty nor high multi-collinearity between included variables greatly increases forecast error variances except in very small samples.
(8) Misspecification testing to develop congruent models can improve forecast performance.

When the model coincides with a constant mechanism, its causal information is always useful, and produces better forecasts than non-causal information, so adding non-DGP variables produces no improvement. However, including causally relevant information generally improves forecasts when the model is misspecified, although misspecified, misestimated linear models with residual autocorrelation can provide unbiased forecasts when the error process is symmetrically distributed, as Hendry and Trivedi (1972) show in a Monte Carlo method using antithetic variates (see, for example, Hammersley and Handscomb 1964).

An implication of these results is that forecast failure (where forecasts lie systematically outside their anticipated intervals) should not occur in stationary distributions (see, for example, Miller 1978). Sadly, the historical track record of economic forecasting intermittently demonstrates marked failures, often out-performed by 'naive devices' such as shown by Nelson (1972). The explanation is that when variables to be forecast have location shifts and/or broken trends both in-sample and out, then none of the results in (1)–(8) need hold, as can be established by counterexamples leading to their converses, as follows for wide sense nonstationary processes.

(1a) A model of the DGP with known parameters in-sample need not dominate in forecasting.
(2a) Adding non-causal variables to the in-sample model can improve forecasts.
(3a) Non-causal models can outperform correct in-sample causal relationships.
(4a) Current conditional expectations of future values can be badly biased and not minimum MSFE when outcomes are drawn from different distributions.
(5a) The correct estimated in-sample model can deliver worse forecasts than the average of several misspecified predictors.
(6a) Long-run interval forecasts are potentially unbounded.
(7a) Parameter estimation uncertainty can substantively increase interval forecasts, as can changes in correlations between conditioning variables at or near the forecast origin.
(8a) Misspecification testing to develop congruent models need not improve forecast performance.

Stationarity is not a viable assumption in economics, so the implications in (1a)–(8a) are far more realistic for forecasting. Facing a non-constant mechanism almost always creates a misspecified model, in which case is a viable theory of economic forecasting possible? Despite making the relatively weak assumptions that the economy is non-stationary from stochastic trends and unanticipated structural breaks, forecasting from a model that differs from the DGP in unknown ways (which would be corrected if known) and is selected and estimated from unreliable data, a theory can be formulated from which many useful insights can be derived; Clements and Hendry (1998) do so based on a taxonomy of all forecast errors, and show that the theory matches many features of real-world forecasting. These include explaining: why location shifts (broken trends cumulate these) are the most pernicious source of forecast failure; why artificial non-causal variables such as intercept corrections can improve forecasts; why differenced-data predictors can outperform congruent in-sample theory-specified models; and hence why the best forecasting model is not necessarily the best model on which to base policy. As most forecasting models of levels are equilibrium correction, equilibrium-mean shifts are a common source of systematic forecast failure, so cointegration is helpful in forecasting only if equilibrium means remain constant.

3 | HANDLING TREND BREAKS AND SHIFTS IN INCOME INEQUALITY DATA

One of the main objectives of inequality forecasts is to predict the direction of the trend—'is inequality likely to increase, decrease or remain stable?'—as this has significant implications from both an economic and societal viewpoint, and will feed into policy decisions. As Figure 2 demonstrates, distributional shifts are prevalent that will impact forecast performance. Furthermore, shifts that initially occur out-of-sample later become in-sample, so need to be carefully modelled for later forecasts. In this section, we discuss approaches to detect and model in-sample trend breaks (Subsection 3.1) and then adapt the forecasts to possible breaks over an out-of-sample period (Subsection 3.2).

3.1 | Breaks in-sample

In-sample breaks can be either ignored or modelled. Forecast accuracy can improve by ignoring sufficiently small structural breaks as the benefits of modelling breaks may be offset by the lack of precision in post-break parameters along with uncertainty as to the exact break date; see Elliott and Müller (2014). Boot and Pick (2020) propose a test to determine whether modelling a structural break improves forecast accuracy, given the bias variance trade-off. Furthermore, Pesaran and Timmermann (2005) show that forecast errors can be unconditionally unbiased even in the presence of breaks in the autoregressive coefficients as long as the unconditional mean of the process remains unchanged, but as Hendry and Clements (2003) show, this requires offsetting parameter shifts to hold the unconditional mean constant. Breaks in the in-sample trend growth rate are particularly pernicious as misestimated forecast origin values will be extrapolated over the forecast horizon. Therefore detecting and modelling broken trends in-sample is essential to mitigate forecast failure by obtaining an unbiased estimate of the trend to be extrapolated at the forecast origin, especially if no later shifts occur. As an example of how bad not doing so can be, the Office of Budget Responsibility five-year-ahead forecasts of UK productivity had a root mean square forecast error (RMSFE) of 7.8, overpredicting by roughly 8% every forecast for a decade (see Martinez *et al.* 2022).

3.1.1 | Trend-indicator saturation

Trend-indicator saturation (TIS) (see Castle *et al.* 2019) provides econometric technology to detect and extract changing trends and use the most recent of these to forecast. It is an extension of impulse indicator saturation (IIS) (see Hendry *et al.* 2008; Johansen and Nielsen 2009), which is used to detect outliers. An impulse indicator $\iota_{\{t\}}$ takes value 1 for observation t, and 0 otherwise, $t = 1, \dots, T$. All T indicators are selected over using a multi-path search algorithm, with impulse indicators retained at a chosen significance level.[4]

TIS 'saturates' a model of a variable y_t (say) with $T-4$ trend indicators, denoted $\tau_{\{t\}}$, for $t = 3, \dots, T-2$, so we have time series $\tau_{\{3\}} = (-3, -2, -1, 0, \dots, 0)$, $\tau_{\{4\}} = (-4, -3, -2, -1, 0, \dots, 0)$, and so on, where $\tau_{\{T\}} = (-T, -(T-1), \dots, -3, -2, -1)$ replaces the usual full-sample trend t by $t-(T+1)$. Advantages of this specific formulation are that any retained $\tau_{\{s\}}$ correspond to the trends that would have been found by an earlier investigator using a sample up to time s; and in the forecasting context, having obtained a congruent in-sample empirical representation, these earlier $\tau_{\{s\}}$ play no further role. Differences between successive $\tau_{\{j\}}$ act as step changes, so the saturating set can characterize a range of aspects of distributional shifts.

For $t \leq T$, a general trend specification with additional regressors $\mathbf{z}_t$ (bold denotes a vector) is

$$y_t = [\beta_0] + [\beta_1 t] + [\boldsymbol{\beta}_2 \mathbf{z}_t] + \sum_{j=3}^{T-2} \beta_j \tau_{\{j\}} + \epsilon_t, \tag{1}$$

where $\epsilon_t \sim \mathsf{IN}[0, \sigma_\epsilon^2]$ denotes an independently distributed normal random variable with constant mean zero and constant variance σ_ϵ^2 in-sample. As equation (1) has more than T regressors, it cannot be estimated directly, hence a 'machine-learning' multi-path block search algorithm such as *Autometrics* (Doornik 2009) with expanding and contracting phases is used to select indicators that are significant at an appropriate level (say $\alpha \leq \min(0.001, 1/T)$). Tight significance levels are used, given the large number of indicators, but the intercept, trend and any other regressors in equation (1) are retained without selection, shown by $[\cdot]$, called 'fixed' below (see Hendry and Johansen 2015). Castle *et al.* (2019) and Castle and Hendry (2022) provide simulation evidence on the selection performance of TIS under both the null hypothesis of no trend breaks and the alternative of a trend shift, as well as when used for forecasting.

3.2 | Breaks out-of-sample

Even if breaks in-sample have been handled to obtain a congruent model with unbiased parameter estimates, breaks that occur over the forecast horizon will lead to poor forecasts. Unless such breaks are predictable (Castle *et al.* (2011) discuss the huge information requirements), forecasters will inevitably make large forecast errors. The approaches to forecasting breaks out-of-sample fall into two categories. The first are non-linear models in which previous regimes could occur again in the future, driven by exogenous variables, past dynamics or past probabilities of switching regime. Such models include Markov switching models (see, for example, Boot and Pick 2018) and threshold forecasting models (see, for example, Kapetanios 1999; Clements *et al.* 2004). Clements and Krolzig (1998) consider both forms of non-linearity in a forecasting exercise, and find that while forecasting within regimes is improved using non-linear models, predicting regime shifts fares less well in terms of forecast accuracy. The second approach is to assume that structural breaks shift not to a previously experienced regime, but to a new, unanticipated regime. In this setting, non-linear models cannot inform future shifts. Instead, if breaks are unanticipated and not predicted as they happen, then systematic forecast failure can still be mitigated. We adopt the second approach here, using various robust devices that adapt rapidly to shifts, although forecasting top income shares from non-linear models may warrant further research.

3.2.1 | Smoothed robust devices

Once the data sample has extended to include breaks in trend that occurred over the forecast period, there is a class of forecasting devices designed to avoid systematic forecast failure. In contrast to TIS, these robust methods do not aim to detect breaks, but following a break, rapidly update estimates of the long-run mean and growth rate to avoid forecasts continuing to converge to their pre-shift equilibrium means; see Martinez *et al.* (2022). The principle behind such a device is to replace full-sample estimates of equilibrium means with subsample estimates after a break. The robust device can be applied in levels if the process is assumed to be a first-order autoregression, denoted AR(1), or in differences if the process includes a full-sample deterministic trend, where the differenced forecasts are then integrated to compare forecast accuracy in levels.

The random walk is a simple case of a robust device that uses the last in-sample observation as a single point estimate of the equilibrium mean. Robust devices augment this principle by including point estimates of the growth rate via the difference of the dependent variable. The problem with such a predictor is that parameter estimates can be erratic based on a single data point. Smoothed robust devices use local estimators averaging over recent data near the forecast origin to pin down estimates of the equilibrium mean and growth rate, thus avoiding any breaks that contaminate the estimates earlier in the sample. Online Appendix I provides the formal definition of the smoothed random walk (SRW) and smoothed robust (SRB) device used in the empirical example.

3.2.2 | Cardt: calibrated average of rho, delta and thima

One automated forecast device that we use is Cardt (see Castle *et al.* 2021), which stands for calibrated average of rho (ρ, an AR(1)), delta (δ, an AR(1) in differences) and thima (trend halved integrated moving average). The device was developed for short-term forecasting, and performed well in the M4 forecasting competition in which 100,000 different time series over a range of frequencies had to be forecast; see Makridakis *et al.* (2020) and Doornik *et al.* (2020). It was also one of the best predictors of deaths from Covid-19, outperforming relative to the Los Alamos National Laboratory[5] and the Institute for Health Metrics and Evaluation[6] early in the pandemic, when little was known about its evolution (see Doornik *et al.* 2021). The method dampens trends and growth rates, averages across forecasts, and robustifies the forecasts to breaks in the data by 'overdifferencing'. It is particularly useful in avoiding explosive roots, and its robustness properties to changing trends make it a useful benchmark to include in the empirical analysis below.

4 | MODELLING UK TOP INCOME SHARES

There is a large literature on changing trends in both income and wealth distributions (see, for example, Roine and Waldenström 2015), which documents the myriad causal factors driving the trend changes, which are also country-specific. The many potential drivers include technological change, trade and financial globalization, financial deepening, and changes in labour market institutions and education, along with explicit redistributive policies; see Dabla-Norris *et al.* (2015). More specific to the UK are household structures and the tax and benefits system, although we analyse the pre-tax and benefit income data. As data on UK-specific structural changes are limited over the long time series, we use saturation estimation to proxy underlying structural changes. TIS will model economic and policy changes that shift income inequality in the UK that are not captured by the conditioning variables. The approach ensures a statistically adequate dynamic linear regression when there are latent or missing variables, and the break dates can be interpreted *ex post* when they align with economic and policy shifts. This approach has the advantage that it is generic and can therefore be applied to any country without specific contextual knowledge.

Although our objective is not to find a causal model of income inequality but to provide useful forecasts of the time series, we investigate relevant correlated factors, including economic and political measures. For the economic measures, we include per capita GDP growth (Figure 3(a)), real wage growth (Figure 3(b)), the average tax burden (Figure 3(c)), and inflation measured by the implicit GDP deflator (Figure 3(d)). The average tax burden is included despite the inequality measure based on the pre-tax income share as the tax burden can be thought of as a proxy for the preferences of the state with regard to inequality. Five-year historical moving averages of the economic measures are also included (recorded in Figure 3) as inequality may respond slowly to changes in macroeconomic conditions, and the smoothed time series

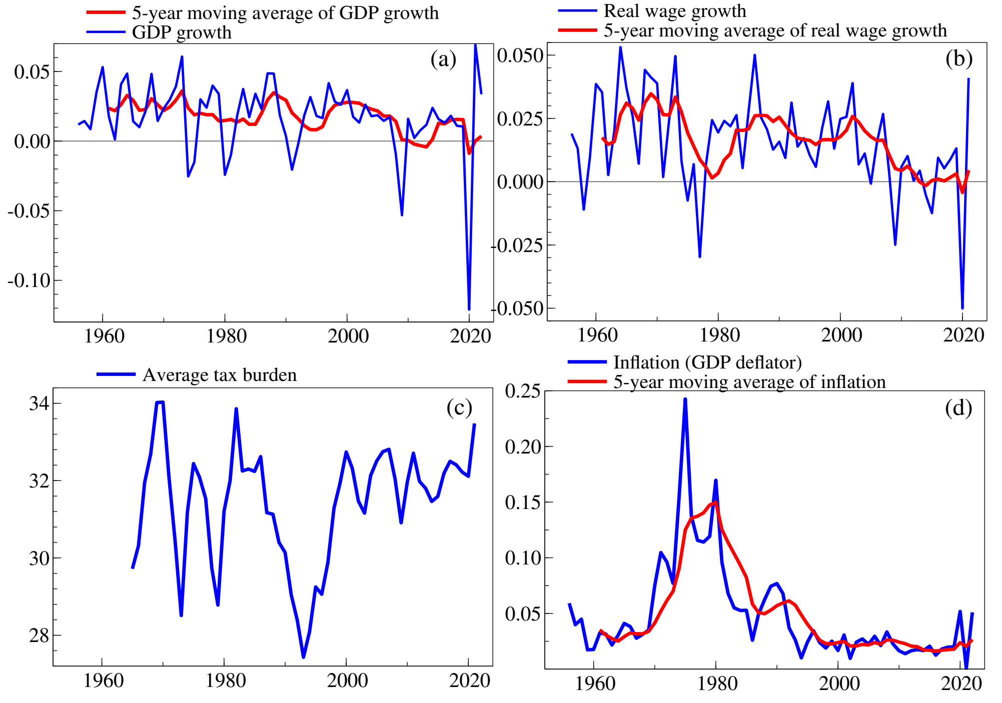

FIGURE 3 Per capita GDP growth, real wage growth, average tax burden and implicit GDP deflator.

are more efficient in capturing slowly evolving dynamics. That said, we allow for both effects in selection.

We also consider globalization and government ideology factors. Figure 4(a) records the KOF Globalisation Index[7] measure for the UK; see Dreher *et al.* (2008) (Potrafke (2015) surveys the literature that uses an early version of the KOF index (https://kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-globalisation-index.html). Figure 4(c) records 26 indices produced by the Global State of Democracy (GSoD) Initiative[8] for the UK, which include measures of the degree of representation, rights, rule of law and participation. Given the range of GSoD indices, the first four principal components are extracted (Figure 4(b)) and are included as a measure of change in government and state ideology. The first four principal components explain 80% of the variation across the indices, although the linear combination means that interpretation of any retained principal components is difficult.

The general model specification allows for all the economic and political explanatory variables along with any relevant dynamics. There are also likely to be outliers and trend breaks, possibly due to measurement error but also because of omitted variables, either latent or for which data are not available that are relevant in the DGP. Saturation estimators are included in the general model to allow for such effects, ensuring a congruent model specification. We jointly apply IIS to detect outliers and TIS to detect shifts in the trend rate of inequality that are not explained by the economic and political variables. Our interest is in forecasting the top 1% income share, so the analysis is undertaken at that level.

Building on equation (1), defining y_t as the pre-tax income share allocated to the top 1 percentile in the UK at time t, the general model formulation is

$$y_t = [\beta_0] + [\beta_1 t] + \beta_2 y_{t-1} + \lambda' \mathbf{z}_t + \sum_{j=3}^{T-2} \gamma_j \tau_{\{j\}} + \sum_{j=1}^{T} \delta_j \iota_{\{j\}} + \varepsilon_t, \tag{2}$$

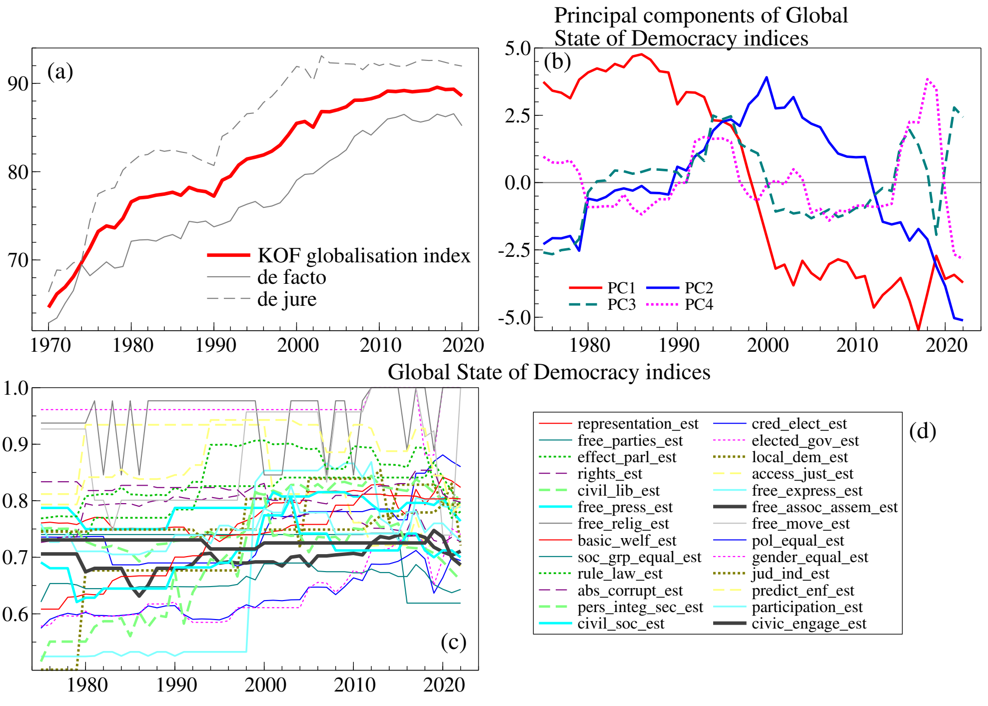

FIGURE 4 Globalization and government ideology factors. *Notes*: (a) The KOF Globalisation Index for the UK. (b) The first four principal components of the 26 GSoD Initiative indices for the UK. (c) The 26 GSoD Initiative indices, with labels in (d). See https://www.idea.int/gsod/gsod (accessed 2 May 2024) for label definitions.

where $\varepsilon_t \sim \mathrm{IN}[0, \sigma_\varepsilon^2]$. Here, $\mathbf{z}_t$ includes GDP growth (Δg_{t-j}), real wage growth ($\Delta(w-p)_{t-j}$), inflation (Δp_{t-j}) and contemporaneous values of their 5-year moving averages $\left(\widetilde{\Delta g}_t^{5yr}, \widetilde{\Delta(w-p)}_t^{5yr}, \widetilde{\Delta p}_t^{5yr}\right)$, the average tax burden (Tax_{t-j}), the change in the KOF Globalisation Index ($\Delta KOFGI_{t-j}$) and the first four principal components of the GSoD indices ($PC_{i,t-j}$), for $i = 1, \ldots, 4$ and $j = 0, 1$. Also, $[\cdot]$ denotes a fixed variable, so selection is not applied to the intercept and full-sample deterministic trend.[9] The retained $\tau_{\{j\}}$ measure subsample deviations from the 'long-run' trend t, which in turn captures unexplained growth.

The general model has more variables (102) than observations (41 for 1976–2016). The procedure to select a specific model from equation (2) proceeds as follows.

1. Fix all conditioning variables dated t and $t-1$ in the model (only t dated moving averages of the economic variables) along with the constant and trend, excluding y_{t-1} (imposing $\beta_2 = 0$), and apply IIS and TIS selection at $\alpha = 0.1\%$.
2. Hold the constant and trend fixed, and apply selection at $\alpha = 1\%$ over the selected regressors from step 1.
3. Add y_{t-1} to the selected model, and test for significance; if found, eliminate any resulting insignificant variables.

The benefit of this procedure is that it avoids estimating a coefficient close to 1 on y_{t-1} at the outset forcing a unit root, which would be a misspecified model but is robust to breaks in trend. It is harder to detect changes in trend when the lagged dependent variable is included because estimates of β_2 are biased towards unity with unmodelled structural breaks. By allowing for structural change initially, both deterministic and stochastic trends are feasible in the general model.

For the in-sample analysis, results are reported for the period 1976–2016, withholding 5 observations over 2017–21 for *ex post* out-of-sample forecasting. In the forecasting exercise in Section 5, the estimation sample is expanded one observation at a time to undertake recursive estimation and forecasting. Estimating the model recursively allows for breaks to be detected after they have moved from the forecast period to the estimation sample, and it ensures that the forecasts are not based on parameter estimates that would not have been feasible at the forecast origin. Income inequality data are often subject to methodological changes and data revisions that could result in model—as well as parameter—non-constancy. An advantage of our automatic selection approach is that models can be updated rapidly as soon as new data arrive or current data are revised, by not only re-estimating the model but re-selecting from equation (2). Significant changes in model specification will point towards methodological changes.

The selected model (denoted *general* below) is[10]

$$\begin{aligned}\widehat{y}_t =& \underset{(0.044)}{0.42} - \underset{(0.071)}{0.63}\, t - \underset{(0.044)}{0.13}\, \Delta(w-p)_{t-1} - \underset{(0.167)}{0.67}\, \Delta\widetilde{(w-p)}_t^{5yr} - \underset{(0.087)}{0.22}\, \widetilde{\Delta p}_t^{5yr} \\ &+ \underset{(0.084)}{0.31}\, \Delta KOFGI_t + \underset{(0.068)}{0.19}\, Tax_{t-1} - \underset{(0.081)}{0.53}\, PC_{1,t} - \underset{(0.120)}{0.39}\, PC_{2,t-1} \\ &- \underset{(0.34)}{2.0}\, \iota_{\{2010\}} + \underset{(0.26)}{1.46}\tau_{\{1998\}} - \underset{(0.27)}{1.49}\tau_{\{2000\}} + \underset{(0.07)}{0.88}\tau_{\{2007\}},\end{aligned} \tag{3}$$

with

$$\begin{aligned}&\widehat{\sigma} = 0.0029, \quad R^2_{adj} = 0.989, \quad F_{ar}(2,26) = 0.12, \quad F_{arch}(1,39) = 0.39, \\ &\chi^2_{nd}(2) = 5.14, \quad F_{het}(22,17) = 1.81, \quad F_{reset}(2,36) = 9.11^{**}, \\ &F_{chow}(5,28) = 13.2^{**}, \quad T = 1975\text{–}2016.\end{aligned}$$

The equation standard error 0.0029 implies significant explanatory power given an unconditional standard deviation 0.028 over the in-sample period (where y_t is the share, so lies in the set $[0,1]$). Excluding the impulse and trend indicators gives equation standard error 0.0081, so the broken trend plays an important role that is not captured by the other variables included in the model. The model passes all diagnostics apart from the RESET test of functional form, and is well specified in-sample, although there is forecast failure in the out-of-sample period. Figure 5 records the model fit along with diagnostic results. Given the importance of the moving averages of inflation and real wages, we test whether non-linear functions—specifically quadratic functions—of the 5-year moving averages of inflation, real wages and per capita GDP growth are significant, but conclude that they do not play a role ($F(3,25) = 0.78$, p-value 0.52), and this agrees with the index test for non-linearity (Castle and Hendry 2010), which suggests that there is no evidence of omitted non-linearity ($F(22,6) = 1.39$, p-value 0.36).

Adding the lagged dependent variable to equation (3) gives p-value 0.62, so it is not significant. Both economic and political variables play a role in explaining the 1% income share. In what follows, we discuss the descriptive model but note that alternative distributional empirical models would be needed to provide causal interpretations to the retained regressors.

Real-wage growth is an important factor, both in the previous year and over a smoother 5-year average, where higher real-wage growth dampens down the share of income accruing to the top 1%, and the effects have been very stable over the last 40 years (Online Appendix III records the recursive parameter estimates). Endogeneity of real-wage growth is avoided as it enters lagged and backward smoothed, despite contemporaneous real-wage growth being included in the initial set of possible regressors. Smoothed inflation is significant but with a counterintuitive sign. As the inflation data are aggregated over all income distributions, it is interpreted as an aggregate effect, and one explanation for the sign is that poorer households tend to be more indebted and inflation reduces the real value of debt. The average tax burden is significant and enters positively.

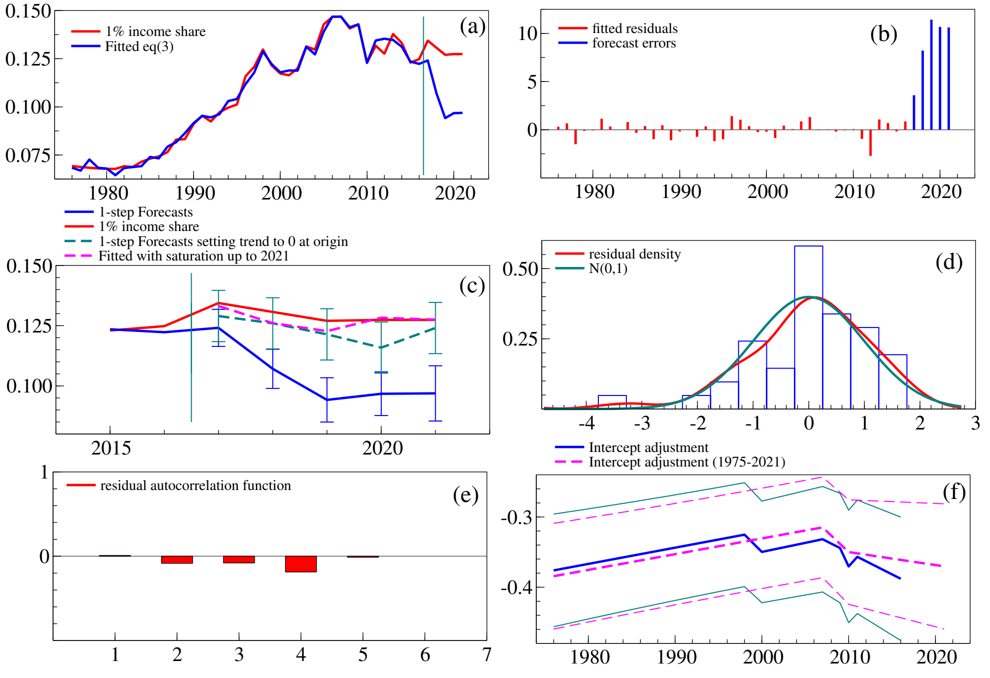

FIGURE 5 Model fit and diagnostic results. *Notes*: (a) Model fit. (b) Residuals. (c) 1-step-ahead forecasts using known values of the exogenous regressors, with forecasts setting the out-of-sample trend to 0 (dashed green line), and in-sample fit with saturation up to 2021 (dashed pink line). (d) Residual density. (e) Residual autocorrelation function. (f) Intercept adjustment from the retained impulse and trend saturation indicators, compared to intercept adjustment applying saturation over full sample (1975–2021).

Political variables are also significant in the model. The first two principal components of the GSoD Initiative account for over 65% of the variation in the indicators.

One impulse indicator (2010) and three trend breaks (1998, 2000 and 2007) are retained, with their linear combination recorded in Figure 5(f), although we do not interpret the intercept adjustment term causally as the saturation estimators can be proxying omitted variables or unmodelled shifts. These play an important role, as without the trend breaks, the equation standard error is almost three times its magnitude with the indicators. However, the extrapolated trend impacts on the forecasts. The model finds a downward shift in the trend of income inequality from 2007, with a blip in 2010. This is extrapolated into the forecast period, and results in large positive forecast errors as income inequality flattens. If the trend was set to 0 at the forecast origin, then the 1-step-ahead RMSFE, assuming known values of future variables, reduces from 2.6 to 0.7 over 2017–21 (dashed green forecasts in Figure 5(c)). Moreover, estimating the selected model over the full sample 1975–2021 but re-selecting the saturation estimators results in a trend break in 2010 rather than the impulse retained in equation (3) (Figure 5(f)). This trend flattens the decline in inequality, and results in an in-sample root mean squared error over 2017–21 of 0.29 (Figure 5(c)). This trend break is crucial to the forecast evolution: assumptions about future trends dominate forecast performance, as discussed in Section 5.

4.1 | Benchmark models

Three benchmark models are considered for forecast comparison with equation (3), including: *Trend*, which is a model with a constant and trend (no saturation or conditioning variables);

sTIS, which applies saturation, fixing the constant and trend (no lagged regressand or conditioning variables); and *ArTIS*, which fixes the constant, trend and lagged regressand (with no conditioning variables).

Assuming a constant trend throughout the sample, denoted the Trend model, we have

$$\widehat{y}_t = \underset{(0.006)}{0.017} + \underset{(0.014)}{0.21}\, t, \tag{4}$$

with

$$\widehat{\sigma} = 0.0106, \quad R^2_{adj} = 0.86, \quad F_{ar}(2,38) = 35^{**}, \quad F_{arch}(1,40) = 24^{**},$$
$$\chi^2_{nd}(2) = 0.13, \quad F_{het}(2,39) = 8.2^{**}, \quad F_{reset}(2,38) = 52^{**},$$
$$F_{chow}(5,40) = 3.7^{**}, \quad T = 1975\text{–}2016.$$

The model exhibits a significant upward trend and fails most misspecification tests, making any inference hazardous.

Next, we apply IIS and TIS saturation at $\alpha = 0.1\%$ with no lagged dependent variable or conditioning regressors, denoted the sTIS model:

$$\widehat{y}_t = \underset{(0.05)}{0.27} - \underset{(0.087)}{0.23}\, t + \underset{(0.17)}{0.54}\tau_{\{1992\}} + \underset{(0.20)}{0.41}\tau_{\{1995\}} + \underset{(0.15)}{0.17}\tau_{\{2004\}} + \underset{(0.17)}{0.34}\tau_{\{2009\}}, \tag{5}$$

with

$$\widehat{\sigma} = 0.0064, \quad R^2_{adj} = 0.95, \quad F_{ar}(2,34) = 4.97^{*}, \quad F_{arch}(1,40) = 0.12,$$
$$\chi^2_{nd}(2) = 0.07, \quad F_{het}(10,31) = 1.0, \quad F_{reset}(2,34) = 4.72^{*},$$
$$F_{chow}(5,36) = 0.71, \quad T = 1975\text{–}2016.$$

Four trend breaks are retained with no impulse indicators, and the trend at the end of the sample is falling, reversing the results of equation (4). The trend breaks do not coincide with those found in the general model. The trend breaks found are likely to be reflecting breaks captured by the other regressors in the general model equation (3). If we augment equation (5) with the trend breaks detected in equation (3) over the full sample, then the forecasts are extremely accurate, again driven by the trend break in 2010, shown by the dashed green lines in Figure 6, middle row.

The final univariate benchmark includes an autoregressive lag in the set of fixed regressors, where selection over the impulses and trends is applied at $\alpha = 0.1\%$, denoted ArTIS:

$$\widehat{y}_t = \underset{(0.003)}{0.003} + \underset{(0.020)}{0.023}t + \underset{(0.086)}{0.90}\, y_{t-1} - \underset{(0.56)}{2.08}\iota_{\{2010\}}, \tag{6}$$

with

$$\widehat{\sigma} = 0.0054, \quad R^2_{adj} = 0.965, \quad F_{ar}(2,36) = 0.53, \quad F_{arch}(1,40) = 0.006,$$
$$\chi^2_{nd}(2) = 3.09, \quad F_{het}(4,36) = 2.69^{*}, \quad F_{reset}(2,36) = 2.90,$$
$$F_{chow}(5,38) = 1.01, \quad T = 1975\text{–}2016.$$

Just one impulse indicator is retained, so the lagged dependent variable is capturing the changing trends, giving an estimate of $\widehat{\beta}_2$ close to unity. The resulting estimate for the long run of $0.23 = 0.023/(1-0.9)$ for the full sample trend is the same as the sTIS model.

Figure 6 illustrates the in-sample results for equations (4), (5) and (6). Panels (i), (iv) and (vii) record the fitted values and outcomes, with panels (ii), (v) and (viii) recording the residuals and forecast errors. Panels (iii), (vi) and (ix) report the indicator trajectories of the models

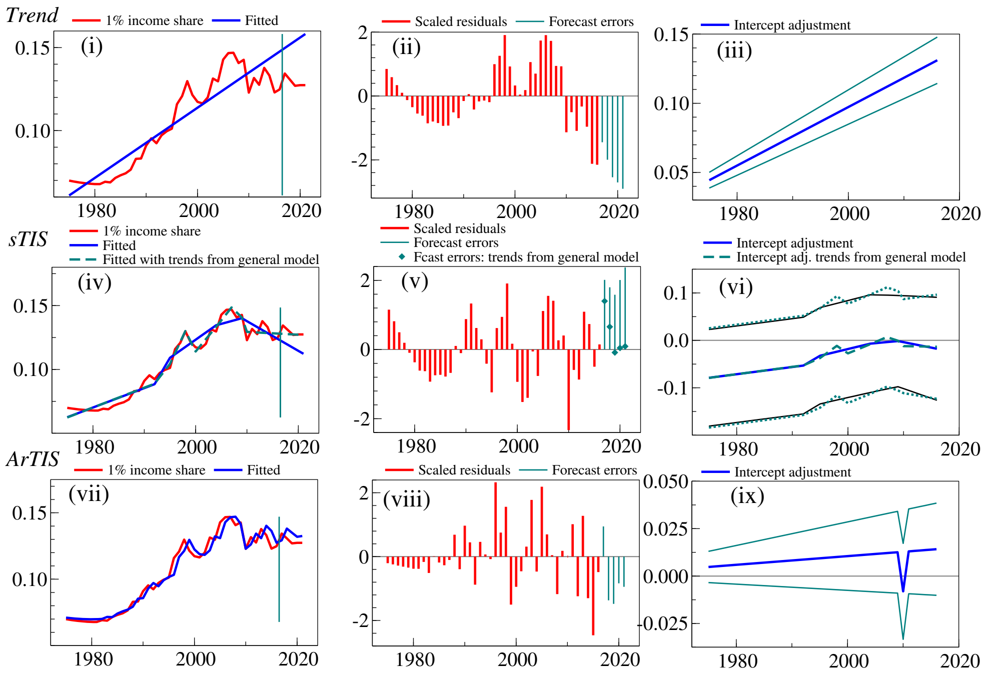

FIGURE 6 In-sample results. *Notes*: Top row for model equation (4), middle row for model equation (5), bottom row for model equation (6), for 1% top income shares for the UK. (i), (iv), (vii) show fitted values and outcomes; (ii), (v), (viii) show residuals scaled by the estimated equation standard error; and (iii), (vi), (ix) show indicator trajectories of the models.

that drive the forecast performance. The models differ significantly in their economic interpretation; not handling trend breaks gives a very different interpretation of the second half of the data. Saturation without the lagged regressand suggests that income inequality has flattened or been falling since 2008, whereas ignoring the breaks or using the lagged regressand to capture the shifts suggests that inequality is still rising over the 2010s. Omitting relevant variables results in trend breaks that proxy shifts in those variables. Including the breaks retained from the general model with conditioning variables results in a more accurate estimate of the underlying trend in the data, as can be seen by the dashed green lines in panels (iv), (v) and (vi).

The in-sample analysis emphasizes that all formulations are characterized by many changing trends. If the objective is forecasting, then all specifications will give substantially different extrapolations of the trend over the forecast horizon. Pinning down the correct trend estimate at the forecast origin is crucial to obtaining helpful forecasts of income inequality, which Section 5 considers.

5 | ALTERNATIVE FORECASTS OF TOP INCOME SHARES

A recursive forecasting exercise is undertaken, where all models are estimated from an in-sample period, $t = 1975, \ldots, T_j$, holding back the final 18 years of observations from the available data. A sequence of 14 multi-step forecasts for $h = 1, \ldots, 5$ years ahead are produced for $T_j = 2002, \ldots, 2015$, using recursive selection and re-estimation at each step from the forecast origin, so there is no knowledge of the data in the out-of-sample period for each recursion.

In the in-sample exercise above, there are significant trend shifts, so given that such breaks occur in-sample, it is highly likely that breaks will also happen in the future, which means that any forecasting models that are not robust after trend breaks out-of-sample will be unlikely to forecast well. Hence we compare standard forecasting models to a suite of devices that are designed to be robust after breaks have occurred. We undertake recursive forecasting to evaluate whether trend breaks can be rapidly detected and allowed for when forecasting.

The general model selected from equation (2) is an open model that includes contemporaneous conditioning regressors. Dynamic *ex ante* forecasts for the general model are produced by a two-step procedure, first producing offline forecasts for the conditioning variables, and then using these forecasts in the dynamic projections of income inequality. Forecasts of the conditioning variables are generated using three vector autoregression (VAR) specifications but allowing for outliers and breaks using IIS:

$$\mathbf{z}_{k,t} = \boldsymbol{\gamma} + \mathbf{A}\mathbf{z}_{k,t-1} + \sum_{j=1}^{T} \boldsymbol{\delta}_j \iota_{\{j\}} + \boldsymbol{\nu}_{k,t}, \tag{7}$$

for $t = 1975, \dots, T_j$, where $\mathbf{z}_{k,t}$ is one of three sets of regressors, including

$$\begin{aligned}
\mathbf{z}_1 &= (\Delta g, \Delta(w-p), \Delta p, Tax)', \\
\mathbf{z}_2 &= \left(\widetilde{\Delta g}^{5yr}, \widetilde{\Delta(w-p)}^{5yr}, \widetilde{\Delta p}^{5yr}\right)', \\
\mathbf{z}_3 &= (\Delta KOFGI, PC_1, PC_2, PC_3, PC_4)'.
\end{aligned}$$

Selection is applied at the 1% significance level (longer lags were considered, but one lag was sufficient to account for serial correlation in all cases). Recursive dynamic 1-step- to 5-step-ahead forecasts are produced for all conditioning regressors, which are then plugged in to the recursively selected and estimated general model to produce *ex ante* forecasts. Alternatively, the contemporaneous regressors could be excluded to close the system for 1-step-ahead forecasts; see Hendry and Mizon (2012) for a discussion of the potential benefits and drawbacks of open models.

A range of forecasting models is considered to provide benchmarks against the general model with the aim of establishing where the gains from addressing non-stationarities in the form of changing trends arise, including the following.

(i) A dynamic single-equation model allowing for explanatory variables, outliers, and breaks in trend (and possible non-linearities) using automatic model selection, producing *ex ante* forecasts using offline forecasts of retained regressors, denoted *general* (see equation (3) for the final recursion).

(ii) An autoregressive model with constant, denoted *AR(1)*, often used as a standard benchmark.

(iii) A deterministic trend model, denoted *Trend* (see equation (4) for the final recursion).

(iv) TIS and IIS with the constant and an overall trend fixed, selecting significant indicators at $\alpha = 0.1\%$, denoted *sTIS* (see equation (5) for the final recursion). The forecasts are computed over $h = 1, \dots, H$ horizons as

$$\hat{y}_{T+h|T} = \hat{\beta}_0 + \hat{\beta}_1(T+h) \tag{8}$$

as the indicators and broken trends end before the forecast horizon. Hence the purpose of TIS is to obtain the most accurate estimates of $\hat{\beta}_0$ and $\hat{\beta}_1$ in-sample for forecast performance. Also, *sTIS smoothed* is computed by estimating sTIS over $t = 1, \dots, T-j$ for $j = 0, \dots, 5$, producing the forecasts over $T+1, \dots, T+H$ for each model, and taking the average of the six resulting forecasts.

TABLE 1 1- to 5-year-ahead RMSFEs for UK top 1% income shares.

	General	AR(1)	Trend	sTIS	ArTIS	Cardt	RW	SRW	SRB	Ave
1-step	*1.58*	0.91	1.56	1.57	1.29	0.91	**0.88**	**0.88**	1.00	0.90
2-step	1.91	1.13	1.77	2.15	*2.20*	1.07	1.02	**1.01**	1.33	1.15
3-step	1.96	1.54	2.03	2.80	*3.57*	1.39	1.27	**1.26**	1.90	1.52
4-step	2.22	1.76	2.24	3.27	*5.09*	1.45	**1.24**	1.25	2.21	1.70
5-step	2.47	2.09	2.46	3.69	*7.33*	1.58	**1.23**	1.29	2.63	2.07
All steps	2.05	1.55	2.04	2.80	*4.45*	1.30	**1.14**	1.15	1.91	1.52

Notes: Forecasts from in-sample period 1975–2002 up to 1975–2016, with RMSFE over all horizons. Bold denotes smallest, and italic denotes largest.

(v) TIS and IIS with the constant, an overall trend and lagged regressand fixed, selecting significant indicators at $\alpha = 0.01\%$, denoted *ArTIS* (see equation (6) for the final recursion). The forecasts are computed as

$$\hat{y}_{T+h|T} = \frac{\hat{\beta}_0(1-\hat{\beta}_2^h)}{1-\hat{\beta}_2}(1+h) + \hat{\beta}_1 \sum_{j=0}^{h-1} \hat{\beta}_2^j (T+h-j) + (\hat{\beta}_2)^h y_T. \tag{9}$$

(vi) Cardt forecasts (see Subsection 3.2.2), denoted *Cardt*.
(vii) A random walk $\hat{y}_{T+h|T} = y_T$, denoted *RW*.
(viii) A smoothed random walk in levels based on a window of $n = 5$ observations, denoted *SRW* (details in Online Appendix I).
(ix) A smoothed robust predictor in levels based on a window of $n = 5$ observations, denoted *SRB* (details in Online Appendix I).
(x) The equally weighted average of all forecasts (i)–(ix), denoted *Ave*.

5.1 | Forecast results

Table 1 records the RMSFEs for the forecasting models over the five forecast horizons, with each cell consisting of an average of fifteen recursive forecasts, along with the average RMSFE across all forecast horizons ('All steps'), and Table 2 records the equivalent mean absolute errors (MAEs) along with maximum absolute errors in parentheses.

There is a clear ranking across all forecast horizons, with the random walk or smoothed random walk producing the most accurate forecasts, and the univariate models with trend saturation producing the poorest forecasts. In almost all cases, the calculated RMSFEs increase as the horizon increases as expected, indicating that producing longer-term forecasts is difficult given the uncertainty over the trend and the cumulating errors. However, the 5-step-ahead forecasts from RW and SRW are smaller than the 1-step from the general model. Figure 7 records the forecast paths for the 1-step-ahead forecasts (Figure 7(a)) and 5-step-ahead forecasts (Figure 7(b)) comparing the RW and general models. There is a systematic upward bias in the general model forecasts; intercept correction was applied to the forecasts from the general model but did not yield a significant forecast improvement.

Table 3 reports the proportion of forecasts in which the predicted sign change is the same as the actual sign change.[11] Model RW is not included in the table as the forecast is a no-change prediction. The sign predictions tend to be worse at longer horizons, but across all models, the sign is predicted correctly only approximately 50% of the time, reflecting the difficulty of inequality forecasting.

TABLE 2 1- to 5-year-ahead MAEs over 15 forecasts.

	General	AR(1)	Trend	sTIS	ArTIS	Cardt	RW	SRW	SRB	Ave
1-step	1.33 (*2.99*)	0.75 (2.22)	*1.34* (2.73)	1.31 (2.98)	1.10 (2.64)	0.75 (2.08)	0.72 (**1.98**)	0.71 (**1.98**)	0.80 (2.24)	**0.68** (2.26)
2-step	1.57 (3.67)	0.97 (2.27)	1.55 (2.90)	*1.92* (3.48)	1.80 (*5.54*)	0.95 (2.02)	0.92 (**1.78**)	**0.90** (**1.78**)	1.25 (2.16)	0.96 (2.10)
3-step	1.70 (3.42)	1.23 (3.59)	1.83 (3.01)	2.54 (4.53)	*2.98* (*9.26*)	1.18 (2.99)	**1.03** (**2.40**)	**1.03** (2.47)	1.63 (4.18)	1.26 (3.20)
4-step	1.97 (3.66)	1.28 (4.31)	2.04 (3.24)	2.86 (5.02)	*4.13* (*13.5*)	1.17 (3.39)	**0.96** (**2.67**)	0.97 (**2.67**)	1.66 (5.27)	1.40 (3.97)
5-step	2.21 (4.55)	1.53 (4.43)	2.26 (3.42)	3.25 (5.53)	*5.81* (*19.9*)	1.22 (3.09)	**1.00** (**2.69**)	1.03 (**2.69**)	1.94 (5.22)	1.72 (4.10)
All steps	1.75 (4.55)	1.15 (4.43)	1.80 (3.42)	2.38 (5.53)	*3.16* (*19.9*)	1.05 (3.39)	**0.93** (**2.69**)	**0.93** (**2.69**)	1.46 (5.27)	1.20 (4.10)

Notes: Maximum absolute forecast errors in parentheses. All steps reports the average and maximum across all forecast horizons for UK top 1% income shares. Bold denotes smallest, and italic denotes largest.

FIGURE 7 Forecast paths. *Notes*: (a) 1-step-ahead and (b) 5-step-ahead forecasts for the general model and the RW model, recorded with the top 1% income share for the UK.

TABLE 3 Proportion of forecasts in which the predicted sign change is the same as the actual sign change for UK top 1% income shares.

	General	AR(1)	Trend	sTIS	ArTIS	Cardt	SRW	SRB	Ave
1-step	0.60	0.60	0.60	0.53	0.40	0.47	0.67	0.53	0.60
2-step	0.60	0.60	0.60	0.47	0.53	0.47	0.80	0.40	0.60
3-step	0.60	0.53	0.67	0.27	0.47	0.33	0.60	0.33	0.47
4-step	0.40	0.40	0.53	0.27	0.33	0.27	0.60	0.47	0.33
5-step	0.47	0.47	0.47	0.33	0.40	0.33	0.67	0.40	0.33
All steps	0.53	0.52	0.57	0.37	0.43	0.37	0.67	0.43	0.47

The general model is significantly worse than the random walk and its variant, so including relevant explanatory variables need not lead to improved forecast performance. There are two explanations for this result.

First, explanatory variables must be forecast, which introduces additional forecast uncertainty. Table 4 records the forecast performance of the explanatory variables based on the three VAR specifications (see Online Appendix II for the full sample estimates). While some variables are forecastable in that their RMSFE is smaller than the in-sample equation standard error ($\hat{\sigma}$), others are particularly difficult to forecast.[12]

Second, estimation of the trend extrapolated into the forecast period can introduce further uncertainty if the trend is estimated imprecisely. The forecast period commences in 2003, so the break in 2007 and the outlier in 2010 identified in equation (3) occur during the out-of-sample forecast period. If we consider the dynamic forecasts between 2007 and 2012, when identifying these breaks is particularly difficult, then the average RMSFE across all horizons is 2.9 compared to 1.5 for the forecasts over the earlier and later periods combined. Figure 8(a) demonstrates the

TABLE 4 1- to 5-year-ahead RMSFE for forecasts for conditioning variables in the general model.

	$\hat{\sigma}$ (1975–2002)	RMSFE 1-step	2-step	3-step	4-step	5-step	Ave
Δg	1.55	2.40	2.03	2.14	3.97	5.00	3.36
Δp	1.47	1.05	1.04	1.05	1.27	1.33	1.17
$\Delta(w-p)$	1.39	1.77	1.73	1.77	2.27	2.83	2.14
$\widetilde{\Delta g}^{5yr}$	0.34	0.53	0.83	1.08	1.24	1.51	1.10
$\widetilde{\Delta p}^{5yr}$	0.62	1.57	2.19	3.01	3.70	4.33	3.16
$\Delta\widetilde{(w-p)}^{5yr}$	0.26	0.38	0.71	1.03	1.18	1.40	1.02
Tax	0.98	0.70	0.60	0.61	0.59	0.65	0.64
$\Delta KOFGI$	0.63	0.75	0.90	0.88	1.01	1.16	0.96
PC_1	0.46	0.86	1.01	0.90	0.89	1.03	0.95
PC_2	0.49	0.77	0.83	0.84	0.98	1.15	0.94
PC_3	0.57	0.74	0.84	1.18	1.27	1.29	1.10
PC_4	0.55	0.85	0.95	1.15	1.65	2.06	1.43

Notes: Forecasts from in-sample period 1975–2002 up to 1975–2016, along with average over all forecast horizons and in-sample equation standard error ($\hat{\sigma}$) for comparison. First six rows are all $\times 100$.

larger forecast errors over 2010–12 for the general model by recording the 1-step-ahead squared forecast errors compared to the random walk. The general model also has larger forecast errors over 2003–4 as the break in trend in 2000 is imprecisely estimated with just a couple of in-sample observations. Trend uncertainty dominates forecast performance, and the random walk essentially takes an agnostic view of the direction of the trend, which is advantageous over periods of uncertainty.

Saturation can lead to wild forecasts by extrapolating imprecisely estimated trends when the forecast origin is close to a turning point in the data. When trends are uncertain, there is a benefit to dampening trends close to the forecast origin. One method of implementing this is recorded in Figure 8(b) for the 5-step-ahead forecasts, where the smoothed sTIS forecasts are computed by applying saturation to the in-sample observations over $t = 1, \dots, T-j$, where $j = 0, \dots, 5$. These models are then used to forecast over $T+1, \dots, T+H$, and the resulting forecasts are averaged using an equal weighting. Such smoothing results in a reduction in RMSFE from 2.80 to 1.97. While robust devices still dominate in forecast performance, such dampening approaches can yield significant benefits over using a model selected and estimated over $t = 1, \dots, T$ for forecasting.

Including an autoregressive term in the trend saturation model leads to more variable forecasts, which can occur for two reasons. First, the ArTIS model can result in a coefficient estimate greater than 1 on the lagged regressand unless constrained optimization is used. In practice, to avoid such outcomes once the indicators are retained, the lagged regressand could be constrained to 1 or less than 1, possibly with a further round of selection over the indicators. Cardt automatically dampens any estimated explosive roots. Second, if impulse indicators are retained at the forecast origin, then they can obscure the correct trend, resulting in systematically biased forecasts. In practical applications, the decision as to whether to retain an impulse indicator at the forecast origin will depend on whether the forecaster thinks the last observation is an outlier due to measurement error (in which case it is a sensible forecast strategy) or due to a structural change that is likely to persist over the forecast horizon (where the indicator should not be present in the forecasting model). A forecast model using just TIS and not IIS would provide a more accurate estimate of the in-sample trend, but is a less parsimonious way to model outliers in the data.

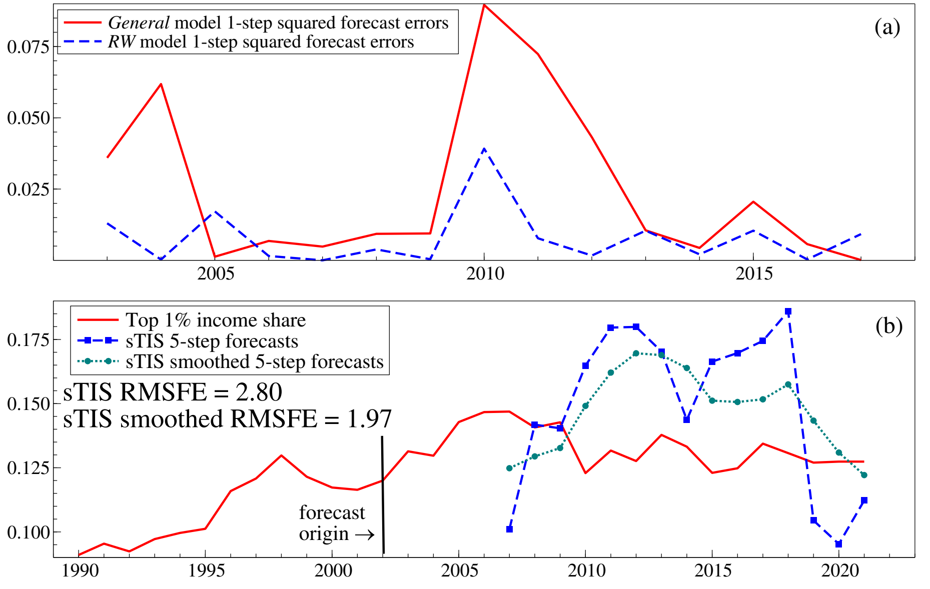

FIGURE 8 Estimation of trends. *Notes*: (a) 1-step-ahead squared forecast errors for the general model and the RW model. (b) 5-step-ahead dynamic forecasts for sTIS (dashed blue line) recorded against smoothed sTIS forecasts (dotted green line).

Many of the forecasts illustrate a typical 'hedgehog' pattern, with forecasts deviating systematically from the outturns in the same direction; Figure 9(a) records the forecast paths for the general model, and Figure 9(b) for the 2-year-ahead consensus forecasts for the sterling effective exchange rate (from Nickell 2009). This feature derives from the forecasting models being equilibrium-correction specifications, where the pre-forecast-origin mean differs from the forecast-period outcomes, yet models keep predicting back to the old equilibrium. Not doing this is why the random walk and its variants dominate in forecasting performance, which matches the empirical findings of Meese and Rogoff (1983). Figure 9(b) reflects this result: the RW would have been no better than the consensus forecast from 1995 to 1997, but after, it is far better as it is able to adjust rapidly to the shift in equilibrium, whereas the consensus does not. Random walks were known to be applicable for speculative markets since Bachelier (1900), and were shown by Nelson (1972) to work empirically in more general settings. Hendry (2006) establishes the explanation for this performance: not only are random walks able to adapt rapidly to shifts in equilibrium, but the lagged dependent variable is the DGP, albeit lagged one period, which contains everything needed for forecasting. The two drawbacks are that it also contains the lagged error term, so it is a noisy measure and does not anticipate breaks, but as the random walk contains all explanatory variables relevant to the top 1% income share without the need to forecast those variables, it is unsurprising that it forecasts better.

TIS is designed for modelling in-sample shifts but does not help when there are trend breaks in the forecast period. Dampening, which is what the robust devices achieve, helps in these settings. The models that capture in-sample trends in the data are not the best forecasting models, reinforcing the result in Castle and Hendry (2011) that models should not be judged on their forecast performance, and highlighting the disconnect between well-specified models in-sample and the forecast performance of models out-of-sample.

As RW dominates at the 5-year horizon, we make projections for 2026 using a random walk and compare to the general model; see Figure 10. The central projection for the random walk is 12.7%. As the residuals from this model are not independent and identically distributed, the

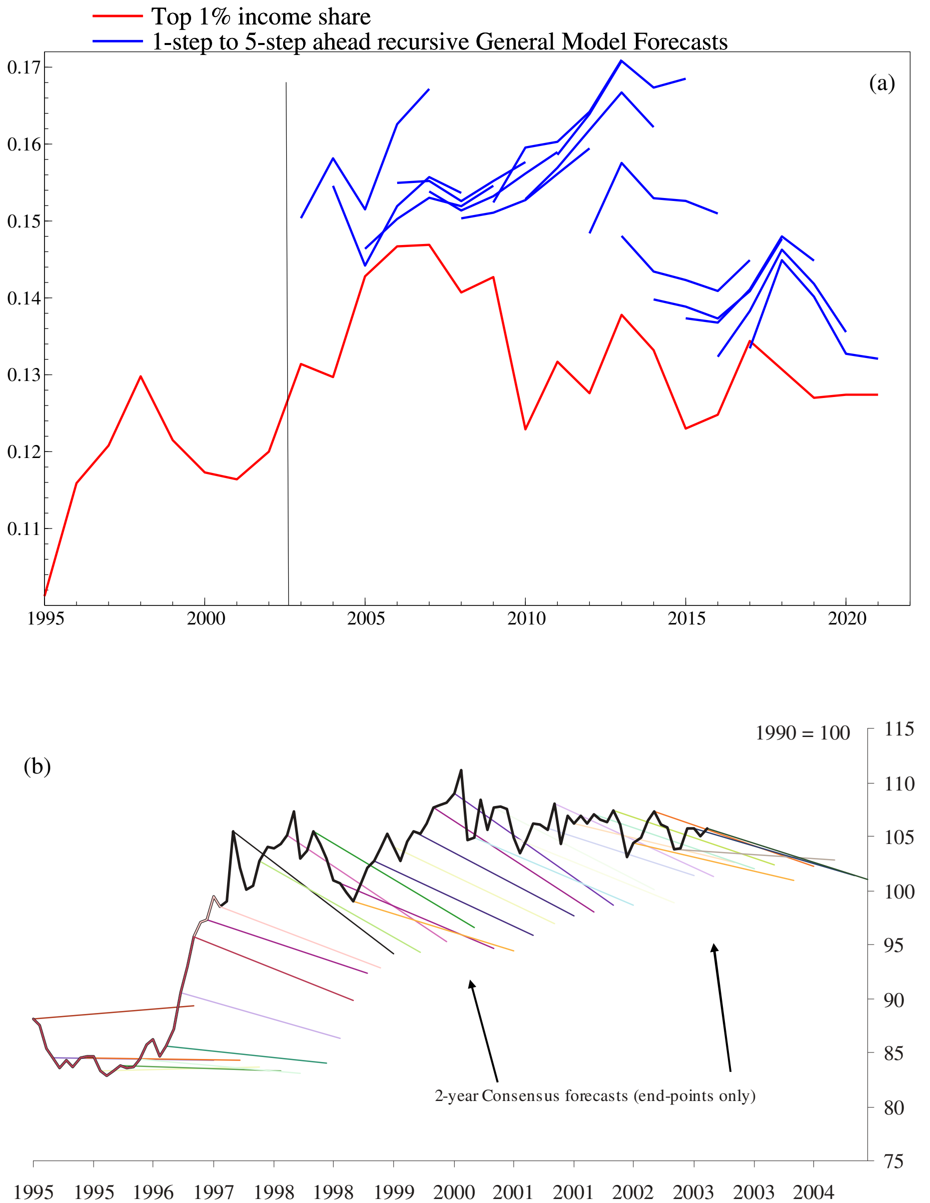

FIGURE 9 More forecast paths. *Notes*: (a) Sequences of 1-step-ahead to 5-step-ahead dynamic forecasts of top 1% income share from the general model. (b) 2-year-ahead consensus forecasts of the sterling effective exchange rate.

forecast standard errors are not informative but are likely larger than the historical range of the data since 1995. Conversely, the general model predicts a small decline in inequality to 11.8% in 2026, with uncertainty bands from 10.8% to 12.8%.

5.2 | Other country results

Table 5 reports average RMSFEs and mean forecast errors (MFEs) for a range of countries using the univariate models. There is no systematic dominance of any forecasting method. RW and

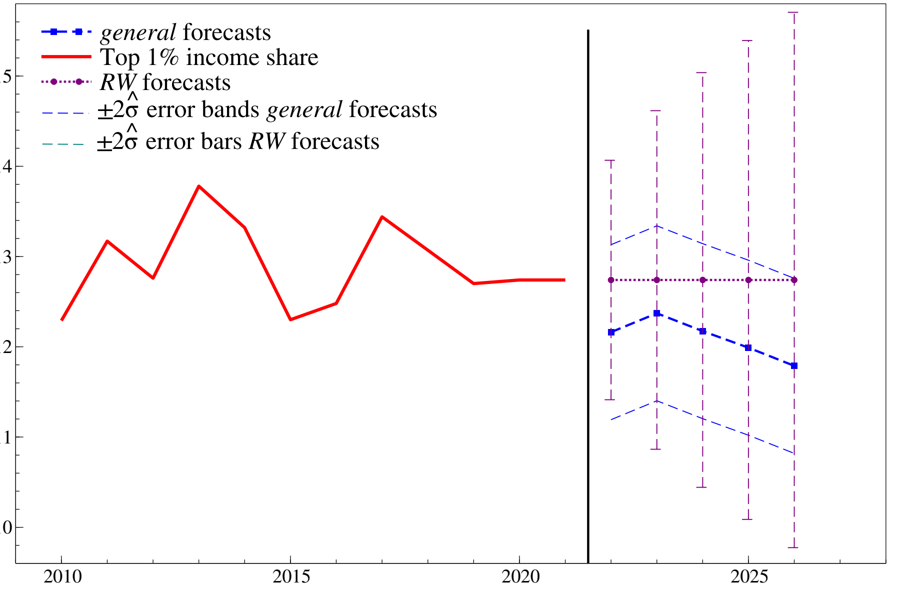

FIGURE 10 Forecasts to 2026 for the general and RW models.

SRW deliver the best forecast results for the majority of countries, but sTIS performs best for the USA. Applying TIS at 0.1% to the USA top 1% income share over the period 1920–2016 results in 10 trend breaks (1923, 1926, 1929, 1931, 1936, 1942, 1945, 1947, 1978, 1998), with the most recent break in 1998. Given the time between the last break date and the forecast origin, the estimate of the trend is precise (with standard error 0.0002). Furthermore, there are no trend breaks over the forecast period. In cases where there are no trend breaks near the forecast origin or over the forecast horizon, we would expect methods that approximate the trend to forecast well.

6 | DISCUSSION

We explain the results of the forecasting exercise for top income shares in the UK using the eight forecasting theorems when stationarity is relaxed in Section 2.

(1a) *A model of the DGP with known parameters in-sample need not dominate in forecasting.* Although the general forecasting model uses estimated rather than known parameters, it is designed to approximate the unknown DGP as a congruent specification. However, it does not dominate in forecasting because all the variables need to be forecast adding to uncertainty, and their distributions need not stay constant between the in-sample and forecast periods. In particular, changes in trend at the forecast origin can lead to poor forecasts. Furthermore, the conditional model is an approximation to the DGP but is not a facsimile, and this can introduce additional uncertainty that affects forecast performance.

(2a) *Adding non-causal variables to the in-sample model can improve forecasts.* The general model does not find the lagged dependent variable to be significant in the model, but instead all of the inertia is explained by exogenous variables. However, RW and SRW, both of which are functions of the lagged dependent variable, dominate in the forecasting exercise. Replacing the end-of-sample trend with the lagged dependent variable in which its coefficient is forced to be unity—that is, adding a random walk component to the general model—worsens in-sample fit but improves the forecast performance.

TABLE 5 Average RMSFEs and MFEs across 1- to 5-year-ahead horizons for top 1% income shares.

	AR(1)	sTIS	ArTIS	RW	SRW	SRB	Ave
Average RMSFE across horizons							
USA	1.22	**1.00**	1.10	1.06	1.02	*1.25*	1.09
Canada	1.52	2.48	*2.87*	1.50	**1.44**	2.05	1.71
Japan	0.68	*1.74*	0.74	0.68	**0.67**	1.05	0.80
France	**0.61**	*1.26*	0.79	0.66	0.65	0.81	0.70
Australia	0.75	*0.94*	0.84	0.72	**0.67**	0.91	0.76
Germany	3.34	2.86	**1.47**	1.66	*4.00*	3.32	3.64
Netherlands	*6.57*	2.05	2.65	**0.60**	0.67	0.61	0.63
New Zealand	0.73	*1.10*	0.96	0.75	0.68	**0.66**	0.66
Switzerland	1.02	1.11	*1.37*	**0.72**	0.86	0.86	0.86
Sweden	1.11	3.95	*5.95*	1.45	**0.84**	0.86	0.85
UK	1.55	2.80	*4.45*	**1.14**	1.15	1.91	1.52
Average MFE across horizons							
USA	1.80	**0.10**	*1.85*	1.57	1.50	1.43	1.46
Canada	*2.02*	**0.12**	1.46	1.67	1.66	0.97	1.31
Japan	1.55	1.25	*1.69*	1.44	1.52	**0.77**	1.14
France	*0.93*	−0.49	**0.05**	0.87	0.83	0.78	0.81
Australia	0.41	*−0.59*	−0.06	0.17	0.20	**−0.04**	0.08
Germany	−3.20	1.25	**0.69**	0.73	*3.21*	2.38	2.80
Netherlands	*−6.55*	1.52	2.62	**0.07**	0.28	0.10	0.19
New Zealand	0.57	*−0.79*	−0.68	−0.19	−0.26	**−0.12**	−0.19
Switzerland	0.16	−0.09	*0.59*	−0.27	**−0.02**	−0.03	−0.03
Sweden	0.81	0.24	*1.69*	−0.17	0.18	**0.15**	0.17
UK	-0.62	*−1.10*	−0.86	**0.01**	−0.04	−0.25	0.28

Notes: Forecasts from in-sample period 1921–2002 up to 1921–2016. Bold denotes smallest RMSFE, italic denotes largest (%).

(3a) *Non-causal models can outperform correct in-sample causal relationships.* This holds as the random walk and its variant produce better forecasts than the general model.

(4a) *Current conditional expectations of future values can be badly biased and not minimum MSFE when outcomes are drawn from different distributions.* Figure 2 shows how the distribution of the top 1% income share has shifted over time. Without modelling these changes, the conditional expectation of the top income share would be biased, as can be seen in the top row of Figure 6, which assumes a constant trend over the in-sample period. Hence if distributions shift in the forecast period, then the conditional expectation calculated over the in-sample period will be biased and will not deliver minimum MSFE forecasts.

(5a) *The correct estimated in-sample model can deliver worse forecasts than the average of several misspecified predictors.* The equally weighted average forecast has a smaller RMSFE and MAE than the general model at all forecast horizons.

(6a) *Long-run interval forecasts are potentially unbounded.* Uncertainty bands increase dramatically, particularly for the ArTIS forecasts. For the forecasts to 2026, the difference in calculated error bands between the general model forecasts and the RW forecasts is stark. The assumption of stochastic unit roots has a significant effect on the interval forecasts, but even absent unit roots, distributional shifts will lead to larger interval forecasts. To correctly reflect future uncertainty, these should be calculated from the in-sample standard error when saturation estimators are excluded.

(7a) *Parameter estimation uncertainty can substantively increase interval forecasts, as can changes in correlations between conditioning variables at or near the forecast origin.* We do not address this directly (for an analysis, see Castle *et al.* 2010), although cases in which explosive roots were estimated caused severe forecast failure.

(8a) *Misspecification testing to develop congruent models need not improve forecast performance.* The automatic model selection algorithm applied to select the general model requires a congruent model specification. Diagnostic testing is undertaken at the outset (if feasible, i.e. fewer variables than observations), and reductions are not applied if a misspecification test is rejected at a specified significance level. Despite ensuring that general is well specified in-sample, it did not dominate in the forecasting exercise, whereas the robust devices all failed some misspecification tests but produced more accurate forecasts.

The forecast exercise established that breaks in trend are pervasive and need to be modelled in-sample to avoid biased parameter estimates; our method for doing so treated break dates and magnitudes as unknown, estimating them using TIS without imposing prior notions about their existence or timing. However, such methods lead to increased uncertainty at turning points in the data, and can lead to erratic forecasts if the trend changes at or close to the forecast origin. During these periods, robust methods that are agnostic about the direction of the trend, namely the random walk and smoothed random work, are preferable.

7 | CONCLUSIONS

A well-specified constant parameter model of the top 1% pre-tax income share in the UK is built in which economic and political variables are significant. Variables including real wage growth, inflation and the average tax burden, along with institutional factors such as the degree of globalization and the state of democracy, are statistically significant in the model. The analysis focuses on aggregate factors rather than distributional changes to help to identify underlying trend breaks in income inequality, which points to rising inequality over the 30 years from the 1980s, with a decline in inequality over the financial crisis period, before flatlining since the 2010s.

Saturation estimators allow for changing trends in the data that are not explained by the conditioning variables. The advantage of such an approach is that a congruent approximation to the DGP can be found while allowing for omitted, possibly latent, variables, measurement errors and definitional changes, and exogenous shocks. Automatic model selection allows for rapid updating as the data accrue and change, resulting in a learning algorithm. That said, the selected model demonstrates remarkable stability over the last 45 years, which means that policy implications can be drawn from the economic and political variables found to be significant.

A pseudo-out-of-sample forecasting exercise finds that the well-specified constant parameter model does not forecast as well as the random walk and its smoothed variant. This result is also borne out in the international comparisons (for univariate models) in which the random walk and the smoothed random walk mostly outperform the other forecasting devices. The counterexample of the USA indicates that the random walk does well when there are breaks in the underlying trend; if the trend is constant over the forecast period, then alternative forecasting devices that have smaller error variances can outperform. This result is due to the facts that (a) the random walk does not have an embedded constant equilibrium so can adapt rapidly to distributional shifts, and (b) the random walk contains the DGP and hence all conditioning variables without needing to estimate and forecast them. The drawback of the additional lagged error term results in more variable but unbiased forecasts.

Our results highlight the difficulties in producing medium-term forecasts of income inequality, and may point towards the dearth of *ex ante* forecasts in the literature. Confirming these results on a range of income deciles is left for future research. There is a disconnect between well-specified models in-sample and the forecast performance of models out-of-sample. Locating breaks and trend shifts is essential for building congruent economic models in order to understand the development of inequality. Forecasting trends in inequality is important to guide future economic policy, but is fraught with difficulties given the non-stationarities that will almost certainly occur over future forecast periods. However, based on our results that robust predictors deliver more accurate forecasts of income inequality, we project that the UK top 1% income share will be approximately 12.7% in 2026, but with wide estimated uncertainty bands.

Forecasting theorems that embed stationarity assumptions provide cold comfort when undertaking practical forecasting for non-stationary data. Stochastic trends (dynamic unit roots) alone would not result in the frequent systematic forecast failure seen in economic forecasting, as growth rates would be stationary. The more problematic form of non-stationarity is distributional shifts, particularly in the form of equilibrium-mean or trend breaks. A future area of research is to develop ways of using TIS robustly out-of-sample. TIS is crucial for developing congruent constant-parameter models in-sample for wide-sense non-stationary data, but can be costly out-of-sample, so allowing for flexible dampened trends over the forecast period could be fruitful.

ACKNOWLEDGMENTS

This paper was prepared for the *Economica 100 challenge*. Financial support from the Robertson Foundation (award 9907422), Nuffield College and the Norwegian Research Council (project number 324472) is gratefully acknowledged, as are helpful comments from Susana Campos-Martins, Tim Besley and four anonymous referees, with thanks to Brian Nolan and Luis Valenzuela for providing the data on top income shares.

ENDNOTES

1 All calculations and graphs use *PcGive* and *OxMetrics*.

2 The theoretical literature on the effects of income inequality on economic growth is inconclusive, but Cingano (2014) finds that income inequality has a statistically significant negative impact on growth, with redistributive policies towards reducing income inequality having no adverse growth consequences.

3 All results are obtained using OxMetrics version 9; see Doornik (2022) and Doornik and Hendry (2021).

4 There are five indicator saturation estimators (ISEs): IIS for modelling outliers; step (SIS) for modelling location shifts (see Castle *et al.* 2015); trend (TIS) for modelling trend shifts, with an application in Walker *et al.* (2019); multiplicative (MIS) for modelling parameter changes (see Castle *et al.* 2020); and 'designed' (DIS) for modelling repeating data shift patterns, such as the impacts of volcanic eruptions on temperatures (see Pretis *et al.* 2016). Combinations of these are also used (see, for example, Ericsson and Reisman 2012). Software to implement ISEs is available in *Autometrics*, part of *PcGive* (see Doornik 2009), as *gets* in R (see Pretis *et al.* 2018), and in EViews (see https://eviews.com/help/helpintro.html#page/content/Regress2-Indicator_Saturation.html, accessed 1 May 2024).

5 See https://covid-19.bsvgateway.org (accessed 1 May 2024).

6 See https://www.healthdata.org (accessed 1 May 2024).

7 See https://kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-globalisation-index.html (accessed 2 May 2024).

8 See https://www.idea.int/gsod/gsod (accessed 2 May 2024).

9 We also considered the logistic transformation of y_t as the income share has support on [0, 1], but the transformation is highly correlated with y_t so the results are very similar.

10 All coefficients and parameter standard errors are multiplied by 100, apart from the constant, real wage growth and the moving averages of real wage growth and inflation. Lower-case represents logs. Coefficient standard errors are shown in parentheses, $\hat{\sigma}$ is the residual standard deviation, F_{ar} tests for residual autocorrelation (see Godfrey 1978), F_{arch} tests for autoregressive conditional heteroscedasticity (see Engle 1982), F_{het} tests for residual heteroscedasticity (see White 1980), $\chi^2_{nd}(2)$ tests for non-normality (see Doornik and Hansen 2008), F_{reset} tests non-linearity (see Ramsey 1969), and F_{chow} tests for parameter constancy (see Chow 1960). For tests, *, ** indicate significance at 5%, 1%, respectively.

11 For example, for the 5-step-ahead forecast from a forecast origin 2002, the forecast was 16.7% for the general model. The actual level of inequality rose from 12% in 2002 to 14.69% in 2007, so the 5-year actual change was positive, as was the forecast.

12 For example, GDP growth per capita has large and increasing forecast errors relative to the in-sample fit, partly but not fully due to the rapid fall in GDP during the Covid-19 pandemic, which rebounded the following year. Forecasts from

smoothed GDP per capita also demonstrate the difficulty in forecasting this variable. As the variable is not retained in equation (3), its role is limited to forecasting the other relevant macroeconomic regressors in $\mathbf{z}_1$ rather than having a direct effect on the top 1% income share.

REFERENCES

Atkinson, A. B. and Piketty, T. (eds) (2007). *Top Incomes Over the Twentieth Century*. Oxford: Oxford University Press.

——— and ——— (eds) (2010). *Top Incomes: A Global Perspective*. Oxford: Oxford University Press.

———, ——— and Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature*, **49**, 3–71. https://www.aeaweb.org/articles?id=10.1257/jel.49.1.3

——— and Voitchovsky, S. (2011). The distribution of top earnings in the UK since the Second World War. *Economica*, **78**, 440–59.

Bachelier, L. (1900). Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, **3**, 21–86; see also M. H. A. Davis, https://www.ma.imperial.ac.uk/ajacquie/IC_AMDP/IC_AMDP_Docs/Literature/Davis_Bachelier.pdf (accessed 2 May 2024).

Blanchet, T., Saez, E. and Zucman, G. (2022). Real-time inequality. NBER Working Paper no. 30229. NBER, Cambridge: Massachusetts. http://www.nber.org/papers/w30229.

Boot, T. and Pick, A. (2018). Optimal forecasts from Markov switching models. *Journal of Business & Economic Statistics*, **36**(4), 628–42.

——— and ——— (2020). Does modeling a structural break improve forecast accuracy? *Journal of Econometrics*, **215**(1), 35–59.

Castle, J. L., Clements M. P., and Hendry D. F. (2015). Robust approaches to forecasting. *International Journal of Forecasting* **31**, 99–112. https://doi.org/10.1016/j.ijforecast.2014.11.002.

———, Doornik, J. A. and Hendry, D. F. (2020). Multiplicative-indicator saturation. Working Paper, Nuffield College, Oxford University.

———, ——— and ——— (2021). Forecasting principles from experience with forecasting competitions. *Forecasting*, **3**(1), 138–65. https://www.mdpi.com/2571-9394/3/1/10.

———, ———, ——— and Pretis, F. (2015). Detecting location shifts during model selection by step-indicator saturation. *Econometrics*, **3**(2), 240–64. https://doi.org/10.3390/econometrics3020240.

———, ———, ——— and ——— (2019). *Trend-indicator saturation*. Working Paper, Nuffield College, Oxford University.

———, Fawcett, N. W. P. and Hendry, D. F. (2010). Forecasting with equilibrium-correction models during structural breaks. *Journal of Econometrics*, **158**, 25–36. https://doi.org/10.1016/j.jeconom.2010.03.004.

———, ——— and ——— (2011). Forecasting breaks and during breaks. In M. P. Clements and D. F. Hendry (eds), *Oxford Handbook of Economic Forecasting*. Oxford: Oxford University Press, pp. 315–53. https://doi.org/10.1093/oxfordhb/9780195398649.013.0012.

——— and Hendry, D. F. (2010). A low-dimension portmanteau test for non-linearity. *Journal of Econometrics*, **158**(2), 231–45.

——— and ——— (2011). On not evaluating economic models by forecast outcomes. *Istanbul University Journal of the School of Business Administration*, **40**, 1–21. https://dergipark.org.tr/en/pub/iuisletme/issue/9251/115734.

——— and ——— (2022). Econometrics for modelling climate change. In J. Hamilton (ed.), *Oxford Research Encyclopedia of Economics and Finance*. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780190625979.013.675.

——— and Shephard N. (eds.) (2009). *The Methodology and Practice of Econometrics*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199237197.001.0001.

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.

Cingano, F. (2014). Trends in income inequality and its impact on economic growth. *OECD Social, Employment and Migration Working Paper* **no. 163**. OECD Publishing, Paris: France. https://doi.org/10.1787/5jxrjncwxv6j-en.

Clements, M. P., Franses, P. H. and Swanson, N. R. (2004). Forecasting economic and financial time series with non-linear models. *International Journal of Forecasting*, **20**, 169–83.

——— and Hendry, D. F. (1998). *Forecasting Economic Time Series*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511599286.

——— and ——— (1999). *Forecasting Non-stationary Economic Time Series*. Cambridge, MA: MIT Press.

——— and Krolzig, H.-M. (1998). A comparison of the forecast performance of Markov-switching and threshold autoregressive models of US GNP. *Econometrics Journal*, **1**, C47–C75.

Dabla-Norris, E., Kochhar, K., Suphaphiphat, N., Ricka, F. and Tsounta, E. (2015). Causes and consequences of income inequality: a global perspective. IMF Staff Discussion Note SND/15/13.

Doornik, J. A. (2009). Autometrics. In J. L. Castle and N. Shephard (eds), *The Methodology and Practice of Econometrics*. Oxford: Oxford University Press, pp. 88–121. https://doi.org/10.1093/acprof:oso/9780199237197.003.0004.
——— (2022). *OxMetrics: An Interface to Empirical Modelling*, 9th edn. London: Timberlake Consultants Press.
———, Castle, J. L. and Hendry, D. F. (2020). Card forecasts for M4. *International Journal of Forecasting*, **36**, 129–34.
———, ——— and ——— (2021). Modeling and forecasting the COVID-19 pandemic time-series data. *Social Science Quarterly*, **102**, 2070–87. https://doi.org/10.1111/ssqu.13008.
——— and Hansen, H. (2008). An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, **70**, 927–39. https://doi.org/10.1111/j.1468-0084.2008.00537.x.
——— and Hendry, D. F. (2021). *Empirical Econometric Modelling using PcGive*, Vol. **I**, 9th edn. London: Timberlake Consultants Press.
Dreher, A., Gaston, N. and Martens, P. (2008). *Measuring Globalization—Gauging Its Consequences*. Berlin: Springer.
Elliott, G. and Müller, U. K. (2014). Pre and post break parameter inference. *Journal of Econometrics*, **180**(2), 141–57. https://doi.org/10.1016/j.jeconom.2014.03.007.
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. (2017). Economic forecasting in theory and practice: an interview with David F. *Hendry. International Journal of Forecasting*, **33**, 523–42. https://doi.org/10.1016/j.ijforecast.2016.10.001.
——— and Reisman, E. L. (2012). Evaluating a global vector autoregression for forecasting. *International Advances in Economic Research*, **18**, 247–58. https://doi.org/10.1007/s11294-012-9357-0.
Gindelsky, M. (2018). Modeling and forecasting income inequality in the United States. Working Paper no. 2018-07, Bureau of Economic Analysis.
Godfrey, L. G. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica*, **46**, 1303–13. https://doi.org/10.2307/1913830.
Hammersley, J. M. and Handscomb, D. C. (1964). *Monte Carlo Methods*. London: Chapman and Hall.
Hendry, D. F. (1974). Stochastic specification in an aggregate demand model of the United Kingdom. *Econometrica*, **42**, 559–78. https://doi.org/10.2307/1911791.
——— (2006). Robustifying forecasts from equilibrium-correction models. *Journal of Econometrics*, **135**, 399–426. https://doi.org/10.1016/j.jeconom.2005.07.029.
——— and Clements, M. P. (2003). Economic forecasting: some lessons from recent research. *Economic Modelling*, **20**(2), 301–29. https://doi.org/10.1016/S0264-9993(02)00055-X.
——— and Johansen, S. (2015). Model discovery and Trygve Haavelmo's legacy. *Econometric Theory*, **31**, 93–114. https://doi.org/10.1017/S0266466614000218.
———, ——— and Santos, C. (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, **23**, 317–35; Erratum, 337–9. https://doi.org/10.1007/s00180-007-0054-z.
——— and Mizon, G. E. (2012). Open-model forecast-error taxonomies. In X. Chen and N. R. Swanson (eds), *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis*. New York: Springer, pp. 219–40.
——— and Trivedi, P. K. (1972). Maximum likelihood estimation of difference equations with moving-average errors: a simulation study. *Review of Economic Studies*, **39**, 117–45. https://doi.org/10.2307/2296866.
Hills, J., Brewer, M., Jenkins, S. P., Lister, R., Lupton, R., Machin, S., Mills, C., Modood, T., Rees, T. and Riddell, S. (2010). An anatomy of economic inequality in the UK: report of the National Equality Panel. LSE Research Online Documents on *Economics* **no. 28344**. London School of Economics and Political Science, LSE Library.
Hood, A. and Waters, T. (2017). Living standards, poverty and inequality in the UK: 2016–17 to 2021–22. *Institute for Fiscal Studies Report* **no. R127**. London: United Kingdom. https://www.ifs.org.uk/publications/8957.
Hubmer, J., Krusell, P. and Smith, A. A. (2021). Sources of US wealth inequality: past, present, and future. *NBER Macroeconomics Annual*, **35**, 391–455. https://doi.org/10.1086/712332.
Johansen, S. and Nielsen, B. (2009). An analysis of the indicator saturation estimator as a robust regression estimator. In J. L. Castle and N. Shephard (eds), *The Methodology and Practice of Econometrics*. Oxford: Oxford University Press, pp. 1–36. https://doi.org/10.1093/acprof:oso/9780199237197.003.0001.
Kapetanios, G. (1999). Threshold models for trended time series. Cambridge Working Papers in *Economics* **no. 9905**. Faculty of Economics, University of Cambridge.
Koopman, S. J., Harvey, A. C., Doornik, J. A. and Shephard, N. (2007). *Structural Time Series Analysis, Modelling, and Prediction using STAMP*, 4th edn. London: Timberlake Consultants Press.
Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2020). The M4 competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, **36**(1), 54–74. https://doi.org/10.1016/j.ijforecast.2019.04.014.
Martinez, A. B., Castle, J. L. and Hendry, D. F. (2022). Smooth robust multi-horizon forecasts. *Advances in Econometrics*, **43**, 143–65. https://doi.org/10.1108/S0731-90532021000043A008.
Meese, R. and Rogoff, K. (1983). Empirical exchange rate models of the seventies: do they fit out of sample? *Journal of International Economics*, **14**(1–2), 3–24.

Miller, P. J. (1978). Forecasting with econometric methods: a comment. *Journal of Business*, **51**, 579–86. https://www.jstor.org/stable/2352648.

Nelson, C. R. (1972). The prediction performance of the FRB-MIT-PENN model of the US economy. *American Economic Review*, **62**, 902–17. https://www.jstor.org/stable/1815208.

Nickell, S. J. (2009). *Monetary policy making in the UK*. Mimeo, Nuffield College, Oxford.

Pesaran, M. H. and Timmermann, A. G. (2005). Small sample properties of forecasts from autoregressive models under structural breaks. *Journal of Econometrics*, **129**(1), 183–217. https://doi.org/10.1016/j.jeconom.2004.09.007.

Pierdzioch, C., Gupta, R., Hassani, H. and Silva, E. S. (2022). Forecasting changes of economic inequality: a boosting approach. *Social Science Journal*, **59**(2), 252–68. https://doi.org/10.1016/j.soscij.2019.09.001.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Cambridge, MA: The Belknap Press of Harvard University Press.

Potrafke, N. (2015). The evidence on globalization. *World Economy*, **38**(3), 509–52. https://doi.org/10.1111/twec.12174.

Pretis, F., Reade, J. J. and Sucarrat, G. (2018). Automated general-to-specific (GETS) regression modeling and indicator saturation for outliers and structural breaks. *Journal of Statistical Software*, **86**(3), 1–44. https://www.jstatsoft.org/article/view/v086i03.

———, Schneider, L., Smerdon, J. E. and Hendry, D. F. (2016). Detecting volcanic eruptions in temperature reconstructions by designed break-indicator saturation. *Journal of Economic Surveys*, **30**, 403–29. https://doi.org/10.1111/joes.12148.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least squares regression analysis. *Journal of the Royal Statistical Society B*, **31**, 350–71. https://www.jstor.org/stable/2984219.

Roine, J. and Waldenström, D. (2015). Long-run trends in the distribution of income and wealth. In A. B. Atkinson and F. Bourguignon (eds), *Handbook of Income Distribution*, Vol. **2**. Amsterdam: Elsevier, pp. 469–592.

Walker, A., Pretis, F., Powell-Smith, A. and Goldacre, B. (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**, I5205. 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**, 817–38. https://doi.org/10.2307/1912934.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Castle, J. L., Doornik, J. A. and Hendry, D. F. (2024). Forecasting the UK top 1% income share in a shifting world. *Economica*, **91**(363), 1047–1074. https://doi.org/10.1111/ecca.12533