Forecasting Changes of Economic Inequality: A Boosting Approach

Christian Pierdzioch[a], Rangan Gupta[b], Hossein Hassani[c] and Emmanuel Sirimal Silva[d]

Submission: May 2019 / Resubmission: July 2019 / Final version: August 2ß19

Abstract

We use a boosting algorithm to forecast changes in three income- and three consumption-based inequality measures. Unlike the existing literature, which basically deals with in-sample predictability, we analyze the role of large number of predictors in out-of-sample prediction of inequality growth. Further, deviating from the annual data-based literature on inequality, we study quarterly UK data covering the period from 1975Q1 to 2016Q1. We find that the boosted forecasting models, at forecasting horizons of up to one year, have to differing extents predictive value for changes in the six different inequality measures. Evidence of predictability is stronger on balance when we use information criteria that result in relatively parsimonious forecasting models than information criteria that are more generous in this regard. In addition to lagged inequality measures, stock-market developments and fiscal deficits, and to a lesser extent the real interest rate, economic policy uncertainty, and output growth turn out to be predictors that are often selected by the algorithm.

JEL classification: C53; D63

Keywords: Inequality; Predictability; Boosting; UK data

Address:

[a] Department of Economics, Helmut Schmidt University, Holstenhofweg 85, P.O.B. 700822, 22008 Hamburg, Germany; Email address: c.pierdzioch@hsu-hh.de.

[b] Department of Economics, University of Pretoria, Pretoria, 0002, South Africa; E-mail address: rangan.gupta@up.ac.za.

[c] Research Institute for Energy Management and Planning, University of Tehran, No. 13 Qods St, Tehran, Iran; Email address: hassani.stat@gmail.com.

[d] Centre for Fashion Business and Innovation Research, Fashion Business School, London College of Fashion, University of the Arts London, 272 High Holborn, London, WC1V 7EY; Email address: e.silva@fashion.arts.ac.uk.

Acknowledgements:
We thank the Deutche Forschungsgemeinschaft (German Science Foundation) for financial support (Project: Exploring the experience-expectation nexus in macroeconomic forecasting using computational text analysis and machine learning; Project number: 275693836). In addition, we thank two anonymous reviewers for helpful and constructive comments. The usual disclaimer applies.

1 Introduction

Inequality is not only a problem in itself, but it also has negative economic, social, and health implications (Chang et al., 2019). High levels of income inequality are not only linked to economic instability, financial crisis, debt, and inflation (van Treeck, 2014; Kumhof et al., 2015; Berg & Ostry 2011; Balcilar et al., 2018), but also to lower social mobility and lower scores in maths, reading, and science, leading to weaker human capital development (Corak, 2016). Inequality is also associated with increased property and violent crimes (Rufrancos et al., 2013). From the perspective of health, living in an unequal society causes stress and status anxiety (Layte & Whelan, 2014). In more equal societies people live longer, are less likely to be mentally ill or obese, along with lower rates of infant mortality (Pickett & Wilkinson, 2009). In addition, inequality affects our perception of others around us and our level of happiness (Delhey & Dragolov, 2013), with people in more unequal societies less likely to trust each other (Gustavsson & Jordahl, 2008), less likely to engage in social or civic participation (Uslaner & Brown, 2005), and less likely to confess that they are happy (van Praag & Ferrer-i-Carbonell, 2009).

Given the wide-ranging negative impact of inequality, its predictability is of paramount importance for appropriate policy-making. Naturally, attempts to forecast the future path of inequality in the United States (US) and the United Kingdom (UK) (which are examples of two countries to have witnessed sharp increases in inequality; Mumtaz & Theophilopoulou, 2017; Chang et al., 2018), have been gaining some momentum through the recent works of Piketty (2014), Gindelsky (2016), and Hood & Waters (2017) based on annual data. Gindelsky (2016) finds that while macroeconomic indicators, human capital, and labor-force metrics often enhance forecasting performance of models, measures of skill-biased technological change are robust predictors of inequality trends for the US. Model selection, however, seems to be sensitive to predictor choice and the the number of lags being considered. Hood & Waters (2017) estimate how the incomes of different households in the UK would develop until 2021-2022, given the current tax as well as benefit policy plans and also the macroeconomic forecasts on earnings and employment from the Office for Budget Responsibility (OBR). They also study macroeconomic scenarios that they describe as more and less optimistic than the OBR's central forecast. Based on their analyses, Hood & Waters (2017) project an increase in income inequality over the next years (particularly if they study income less of housing costs).

Against this backdrop, the aim of this paper is to forecast both income- and consumption-based relative and absolute measures of inequality for the UK, using an unique data set at the (highest possible) quarterly frequency over the period of 1975Q1 to 2016Q1, based on a wide array of macroeconomic and financial variables. The choice of the UK as our case study is driven by data availability of inequality at a quarterly frequency, which is important because accurate forecasting of inequality at a higher frequency should be more relevant to policymakers than at the lower annual frequency. Besides data-based reasons, the decision to look at the UK is based on the massive inequality growth figures, with income (consumption) inequality growth between 1975Q1 to 2016Q1 ranging between 10% to 21% (10% to 28%), and the UK being recognized as an outlier of extreme inequality in the European context (Dorling, 2015).

As far as the econometric approach is concerned, we use a machine-learning technique known as boosting to forecast at a quarterly frequency, movements of three income-based and three consumption-based inequality measures. Boosting is particularly suited to forecast movements of inequality measures as we use several potentially important determinants of economic inequality. Boosting is a data-driven algorithmic machine-learning approach to the selection of predictors for model fitting and forecasting in an environment where the number of predictors is large relative to the number of available time-series data. Variants of boosting have been applied in recent research in economics, for example, to forecast movements of exchange rates and commodity prices (Berge, 2013; Pierdzioch et al., 2015), to compute forecasts of recessions (Ng, 2009; Döpke et al., 2017), to model inflation expectations (Berge, 2017), and to test the rationality of survey forecasts (Pierdzioch & Risse, 2018). While in-sample analyses of the trend in UK's inequality (based on factors such as, skill-based education and technological advances, changes in the family structure, employment status and occupation, structural reforms in the labour market, globalization, and increased international trade) have been previously widely undertaken using annual data (for a detailed review, see Belfield et al., 2017), to the best of our knowledge, this is the first attempt of out-of-sample forecasting of inequality at a high data frequency based on boosting. Given the fact that in-sample predictability does not guarantee out-of-sample forecasting gain, and the suggestion in this regard that the ultimate test of any predictive model is its out-of-sample performance (Campbell, 2008), our analysis at a higher data frequency aims to make a major contribution to the sparse literature on forecasting inequality.

In sum, we can outline our contributions as follows: (i) Unlike the few existing studies on forecasting of income inequality of the UK based on annual projections, we provide a comprehensive analysis on forecasting quarterly inequality growth rate. This is important since, given the multi-dimensional negative impact of inequality, accurate forecasting of the same at a higher frequency should be more relevant to policymakers than at the lower annual frequency to design appropriate policies to reduce inequality and its impacts; (ii) In addition, unlike existing papers which basically deals with in-sample determinants of inequality, we conduct real-time out-of-sample forecasting of inequality and determine the factors responsible for driving future inequality growth of the UK, which in turn is likely to help in better design of policies ahead of time, i.e., before the actual realization of the growth of inequality, and; (iii) We combine information on the various predictors used separately in the literature, and at a higher frequency, based on recent methodological advances associated with machine-learning, in particular boosting, which is able to use multiple predictors for model fitting and forecasting in an environment where the number of predictors is large relative to the number of available time-series data. In other words, differently from the existing literature, we use high-frequency data to produce accurate forecasts of inequality growth by incorporating simultaneously large number of predictors based on innovative econometric techniques, which should all play a role in better design of policies to curb inequality growth in the UK.

We organize the remainder of this research as follows. We describe the boosting algorithm in Section 2 and this is followed by a description of the data that we use in our empirical research in Section 3. We summarize the results of our empirical analysis in Section 4, and conclude in Section 5.

2 The Boosting Algorithm

We use a L2-boosting algorithm (Friedman, 2001; Bühlmann & Yu, 2003; Bühlmann & Hothorn, 2007) to forecast annualized changes, $y_{i,t+h}$, of the log of inequality measure i, where the index h denotes the forecast horizon. In order to compute one-quarter-ahead forecasts, we set $h = 1$. We compute multiperiod changes of the inequality measures as $y_{i,t+h} = (y_{i,t+1} + ... + y_{i,t+h})/h$. We compute forecasts by means of a forecasting model of the general format $y_{i,t+h} = F(\beta_{i,h}, x_{t,i,h}) + u_{i,h,t+1}$, where the function $F(\beta_{i,h}, x_{i,h,t}) = \sum_{j=1}^{k} \beta_j x_{i,h,t,j}$ is a so-called strong learner, $\beta_{i,h,j}$ denotes coefficient $j, j = 1,...,k$, estimated for inequality measure i at forecast horizon h, $x_{i,h,t,j}$ denote the j-th predictor, and $u_{i,h,t+1}$ denotes a disturbance term. The predictors carry the index i because we include lagged changes of an inequality measure (but not lagged changes of the other inequality measures) in the vector of predictors. For the L2-boosting algorithm, the strong learner is the solution to

$$\hat{F}(\hat{\beta}_{i,h}, x_{t,i,h}) = \underset{F(\beta_{i,h}, x_{t,i,h})}{\arg\min}\, E\left[\frac{1}{2}(y_{i,t+h} - F(\beta_{i,h}, x_{t,i,h}))^2\right], \tag{1}$$

where E denotes the expectations operator. Equation (1) is a function-approximation problem that we solve by means of the L2-boosting algorithm. To this end, we initialize the vector of coefficients $\beta_{i,h}^{[0]} = 0$ (which initializes the strong learner, F_0), and then iterate over the following steps:

□ For$(m = 1 : M)\{$

1. We compute the negative gradient vector and estimate univariate regressions of the negative of the gradient vector on the j individual predictors. The estimates produce k weak learners, $\hat{f}_j$. As recommended by Bühlmann and Hothorn (2007), we apply boosting on mean centered data.

2. We identify the best weak learner as the solution of the following minimization problem: $\kappa = \arg\min_j \left[\sum_t (u_{i,h,t+1} - \hat{\gamma}_j x_{i,h,t,j})^2\right]$, where $\hat{\gamma}_j$ denote the estimated coefficients of the univariate regressions (Step 2).

3. We use the best weak learner to update the vector of coefficients $\hat{\beta}_{i,h}^{[m]} = \hat{\beta}_{i,h}^{[m-1]} + s\hat{\gamma}_{i,h}^{[\kappa,m]}$, where s is the learning rate (a smaller s produces more iterations) and $\hat{\gamma}_{i,h}^{[\kappa,m]}$ contains as the only non-zero element the coefficient estimated for κ. Updating the coefficients is equivalent to updating the strong learner: $\hat{F}_{m+1} = \hat{F}_m + s\hat{f}_{m,\kappa}$. We use the updated strong learner to compute a new gradient vector and new weak learners.

} ■

We terminate the boosting algorithm either if it reaches the maximum number of iterations, $m = M$, or in iteration $m^* < M$ if an information criterion, $IC(m^*)$, is minimized. To find the minimum of the information criterion, we run the algorithm m_{break} times. We terminate the algorithm if $m^* = \arg\min_m IC(m)$ satisfies $m^* \leq 0.75 \times m_{break}$. Otherwise, we set $m_{break} = m_{break} + 10$, and then check again whether $m^* \leq 0.75 \times m_{break}$ (for a similar approach, see Mayr et al., 2012).

We study four different information criteria. The first one is the Akaike information Criterion, AIC_{trace} (Hurvich et al., 1998, Bühlmann, 2006), which is defined as follows:

$$AIC_{trace}(m,i,h) = \ln(\hat{\sigma}_{i,h,m}^2) + \frac{1 + df(m)/T_t}{1 - (df(m)+2)/T_t}, \tag{2}$$

where $\hat{\sigma}_{i,h,m}^2$ denotes the residual variance of the boosted forecasting model for inequality measure i at forecast horizon h in iteration m, and T_t denotes the number of observations available in period of time t. The degrees of freedom are defined as $df(m) = \text{trace}(B_m)$, where the matrix B_m is updated according to the recursion $B_m = B_{m-1} + H_\kappa(I - B_{m-1})$, where $H_\kappa = \mathbf{x}_{\kappa,i,h,t}(\mathbf{x}_{\kappa,i,h,t})^\intercal / ||\mathbf{x}_{\kappa,i,h,t}||^2$, I denotes a suitable identity matrix, $||.||$ denotes the Euclidian norm, $\mathbf{x}_{\kappa,i,h,t}$ denotes the vector of observations on the κ-th predictor for inequality measure i at forecast horizon h available in period of time t (that is, the weak learner selected in iteration m, see Bühlmann, 2006; Bühlmann & Hothorn, 2007).

The second information criterion is the generalized Minimum Description Length (gMDL, Hansen & Yu, 2001; Bühlmann & Hothorn, 2007):

$$gMDL_{trace}(m,i,h) = \ln(S) + (df(m)/T_t)\ln(Z), \tag{3}$$

where $S = \frac{T_t \hat{\sigma}^2_{i,h,m}}{T_t - df(m)}$, and $Z = \frac{\sum_{j=1}^{t} y_j^2 - T_t \hat{\sigma}^2_{i,h,m}}{df(m)S}$.

The third and fourth information criteria produce relatively parsimonious forecasting models. To this end, we use the number of selected predictors (the "active set") to define $df(m)$ (see, Hastie 2007). Accordingly, we use two information criteria, $AIC_{actset}(m,i,h)$ and $gMDL_{actset}(m,i,h)$, for which the active set defines the degrees of freedom.

Finally, we use the active set of predictors included in the optimal iteration, m^*, to re-estimate the forecasting model on the original data by the ordinary-least squares technique. We use the estimate to produce summary statistics (like the adjusted coefficient of determination) for the forecasting model, and to compute an out-of-sample forecasts of changes in the inequality measure being studied.

3 Data

We analyze quarterly data that range from March (Q1) 1975 to March (Q1) 2016. The seasonally-adjusted inequality data is for income equivalized by dividing by the square root of the number of people in a household and total consumption per capita of a household. We consider three measures of inequality: the Gini coefficient, the standard deviation (of the data in natural logs), and the difference between the 90th and 10th percentile (with the data in natural logs). The inequality measures are computed using survey data on income and consumption from the family expenditure survey (FES)[1]. Mumtaz & Theophilopoulou (2017) provide an extensive documentation of the construction of the data and the survey. Note that, these authors remove any households reporting zero or negative income, when constructing the income-based measures of inequality.[2]. At this, it is important to highlight that studies like that of Foster (1996) and van de Ven (2011) point out issues of representation problems associated with the survey. In particular, the

[1]The data is downloadable from:https://discover.ukdataservice.ac.uk/series/?sn=200016 and https://discover.ukdataservice.ac.uk/series/?sn=2000028.

[2]We would like to thank Professor Haroon Mumtaz for kindly sharing the inequality data with us.

FES, on one hand, tends to over represent mortgage holders, people living in the countryside, older households, and, on the other hand, people living in council flats, institutions (retirement homes, military), no fixed address holders, ethnic minorities, self employed, manual workers, and younger households are under represented (Dayal et al., 2000; Mumtaz & Theophilopoulou, 2017). Clearly then there is bias and underreporting of top earners when using survey data, compared to when administrative data is used.

It should be noted that while the surveys are recorded at an annual frequency, Mumtaz & Theophilopoulou (2017), following Cloyne & Surico (2017) and Cloyne et al. (2019), assign households to different quarters within a year based on the date of the survey interviews, which, in turn, allows them to calculate the measures of inequality at a quarterly frequency.[3] We abbreviate the three income-based inequality measures as YI1, YI2, and YI3, while the three consumption-based inequality measures are denoted as CI1, CI2, and CI3.

– Please include Figure 1 about here. –

Figure 1 plots the inequality measures. Eyeballing the figure shows that the inequality measures exhibit a discernible trend increase until around the early 1990s. Thereafter, the consumption-based inequality measures, and to a lesser extent the income-based inequality measures, show a tendency to decrease again. Standard unit-root tests (not reported, but available upon request) provided strong evidence that the inequality measures are nonstationary, while their first differences are stationary. For this reason, we forecast changes in the inequality measures. Forecasting changes in the inequality measures is also interesting from the perspective of political economics given that policy makers are likely to be interested in the upticks and downticks of inequality at the forecasting horizons that we study in our research. Clearly, in the long run, policy makers are

[3] It is important to point out that, as survey respondents are asked about consumption and income over a period preceding the interview, the time-series corresponding to each interview is indeed meaningful. To elaborate, the consumption of household X in the second-quarter differs from household Y in the third-quarter not only due to differences in X and Y, but also because Y is asked about variables like wages up to the third-quarter, while X is asked about the same up to the second-quarter.

interested in trends in inequality, but at short-term forecast horizons (with the time interval to the next election getting shorter) changes in inequality are often the subject of controversial debates in the policy arena.

As far as the predictors are concerned, besides lagged inequality measures, we rely on the recent literature to motivate the choice of predictors at the quarterly frequency. For instance, Mumtaz & Theophilopoulou (2017) indicate the role of real Gross Domestic Product (GDP), monetary policy (real interest rate, RIR, i.e., nominal three-month Treasury bill rate less the Consumer Price Index, CPI, based inflation rate), and real effective exchange rate (EER) to be important predictors. As an alternative measure of economic activity, we also consider the unemployment rate (UR). The importance of fiscal policy over and above the monetary stance, as captured by the budget-deficit as percentage of GDP (BUD), has been emphasized by Coibon et al. (2017). We also used decomposed version of the fiscal policy variable by looking at direct (TAX) and indirect (ITAX) taxes, and social benefits (BEN) and public investment (GINV) spendings, with all these disaggregated revenues and expenditures expressed as percentage of the GDP. Asset price (financial market) movements as captured by real stock price (RSP) and real house price (RHP), both of which are computed by deflating the nominal prices with the CPI, have been shown to play a role in driving inequality, as discussed in detail by de Haan & Sturm (2017). Further, uncertainty related to policy decisions, as captured by the news-based economic policy uncertainty (EPU) of Baker et al. (2016),[4] has been shown to be driver of inequality by Balcilar et al. (2019). In addition to policy-related uncertainty, we also use realized volatility (as captured by the sum of daily squared stock log-returns (RV), following Andersen & Bollerslev (1998), as a measure of general economic and financial market uncertainty (Gupta et al., 2018). Besides domestic variables, the role of financial stress in the United States (FSI), has also been pointed out to be important by Mumtaz & Theodoridis (2017). The data on output, unemployment rate, interest rate, and CPI are derived from the Main Economic Indicators (MEI) of the Organisation for Economic Co-operation and Development (OECD), while data on housing prices is derived

[4]The data is available for download from: `http://www.policyuncertainty.com/uk_historical.html` till 2008, and from `http://www.policyuncertainty.com/europe_monthly.html` thereafter.

from the Housing Prices Database of the OECD. The fiscal policy variables are obtained from the public finance and fiscal policy segment of the Economic Outlook databased of the OECD. The real effective exchange rate is sourced from the Effective Exchange Rates Database of the Bank of International Settlements (BIS). Stock market data (both at quarterly and daily frequencies) comes from Datastream. The financial stress data for the United States is based on the research by Püttmann (2018)[5]. Even though other financial stress data is available for the United States from the various Federal Reserves, the length of the data coverage of this index motivates us to prefer this measure over the other alternatives. Barring all the fiscal-policy variables, real interest rate, unemployment rate, RV, EPU and FSI (with the latter two being in their natural logarithmic form) all variables are in their respective growth rates (log-returns) to ensure mean-reversion.

Table A1 at the end of the paper (Appendix) summarizes the variables we use in our empirical analysis, along with their source and the transformations used. Note that our forecasting exercise is not based on real-time data, with the data-vintage being that of 2016Q1 to correspond to the end-date of the inequality measures.

4 Empirical Analysis

4.1 Calibration Issues

We compute out-of-sample forecasts for three different forecast horizons by recursively re-estimating the forecasting model. We present results for one-quarter-ahead forecasts ($h = 1$), and in In Section 4.4 also for two-quarters ahead and one-year ahead forecasts ($h = 2,4$).[6] In order to account for the possibility that the boosting algorithm may include different predictors

[5]Further details regarding the data and the data itself for download can be found at: `http://www.policyuncertainty.com/financial_stress.html`.

[6]As for the timing of events, our forecasting approach is based on the assumption that a forecaster who (i) estimates at the beginning of period $t+2$ a model that predicts period-$t+1$ inequality (in case of $h = 1$) using data on inequality and the other predictors from period t or earlier, and, (ii) uses this model at the end of period $t+2$ to form a forecast of period-$t+2$ inequality using period-$t+1$ (or earlier) data on inequality and the other predictors.

in the forecasting model in different periods of time, we use a recursively expanding estimation window to implement the L2-boosting algorithm. The first estimation window uses the first 10 years of data to train the algorithm, but we also report results for a longer training period. We set the maximum number of iterations to M=250, but the algorithm typically stops much earlier. As for the learning rate, we set $s = 0.25$. As an extension, we also report results for a smaller learning rate and a larger maximum number of iterations (Section 4.4). Choosing a smaller value for the learning rate leads to more iterations. We use the R programming environment for statistical computing for our empirical analysis (R Development Core Team, 2017). Finally, we account for publication lags to avoid a look-ahead bias. Specifically, we assume a publication lag of one quarter for the financial-market-based predictors, and two quarters for real GDP growth, the unemployment rate, and the government-related predictors. We also include lagged predictors (dated $t, t-1, t-2$, and $t-3$) to account for the possibility that lags contain information useful for predicting changes in the inequality measures. Accounting for all data transformations, we estimate the boosting algorithm on data starting in 1976(Q3). In total, the boosting algorithm can include up to 60 predictors (including all lags) in the boosted forecasting models.

4.2 Properties of the Boosted Forecasting Models

Table 1 summarizes key in-sample properties of the boosted forecasting models. The forecasting horizon is one quarter. Panel A shows the average number of iterations for the four information criteria. The two information criteria that use the active set produce fewer iterations than the other two information criteria. For example, the mean number of iterations for the trace-based information criteria is approximately 101 and 77 for income-based inequality measure DYI1, while the mean number of iterations for the other two information criteria is only about 15 iterations.

– Please include Table 1 about here. –

The larger number of iterations that we observe for the trace-based information criteria results on average in boosted forecasting models that include more predictors than under the active-set information criteria, as shown in Panel B. For the consumption-based inequality measures, the boosting algorithm includes on average fewer predictors in the forecasting models than for the income-based inequality measures. Moreover, as one would have expected, the larger number of predictors under the trace-based information criteria results on average in a larger in-sample adjusted R^2 than under the more restrictive information criteria, as plotted in Panel C. The average in-sample adjusted R^2 is largest (about 0.58) for the income-based inequality measure DYI3 for the two trace-based information criteria, and lowest (about 0.25 to 0.26) in case of the consumption-based inequality measure DCI1 for the two active-set-based information criteria

– Please include Figure 2 about here. –

Figure 2 plots for the gMDL-based information criteria the evolution of the in-sample estimation error defined in terms of the square root of the estimated variance of the disturbance term of the boosted forecasting models. The estimation error is larger for the DYI3 and DCI3 inequality measures than for the other four inequality measures. The estimation error is relatively stable over time and shows no abrupt large and disruptive changes. Hence, while the dimension of boosted forecasting models as measured in terms of the number of predictors as well as the composition of the vector of selected predictors change over time, structural breaks in the process generating the inequality measures do not beleaguer the forecasting models. The in-sample estimation error is on average somewhat larger for the active-set-based information criteria than for the trace-based information criteria.

4.3 Importance of Predictors

Table 2 reports results for predictor inclusion in the boosted forecasting models, and Table 3 informs about the relative importance of the predictors. The forecasting horizon is one quarter. Predictor inclusion is a metric of absolute predictor importance. Predictor inclusion is defined as the number of times a predictor (including all its lags) is included in the boosted forecasting models divided by the total number of estimated forecasting models (times the number of lags of the predictors). Relative importance, in turn, is defined as the number of times a predictor (including all its lags) is included in the boosted forecasting models divided by the total number of predictors included in the boosted forecasting models, computed across all recursive estimation windows and predictors. Hence, the numbers in the rows of Table 3 sum up to 100%. Panels A−C report the results for the income-based inequality measures, and Panels D−F summarize the results for the consumption-based inequality measures.

– Please include Table 2 and 3 about here. –

Several results emerge. The lagged changes in the inequality measures are always among the top predictors in terms of predictor inclusion and relative importance. While the inclusion of the lagged changes in the inequality measures in the boosted forecasting models is larger under the trace-based information criteria, their relative importance is larger when we study the active-set-based information criteria rather than for the trace-based information criteria. This result is not surprising given the well-known persistence of inequality (Arestis et al., 2011), and is in line with the earlier observations for the UK by Gindelsky (2016). For the first income-based inequality measure, DYI1, two other important predictors are returns of real stock prices and fiscal policy, as measured in terms of the budget deficit. Returns of real stock prices, the budget deficit, and the real interest rate are also important for the second and third income-based inequality measures. In addition, we find that returns of real stock prices are important for the consumption-based inequality measures, especially under the trace-based information criteria. The budget deficit is also among the important predictors (for the second consumption-based inequality measure in particular), but in Table 2 only when we consider the trace-based information criteria. For all inequality measures, economic policy uncertainty mainly plays a role when we consider the trace-based information criteria. Similarly, output growth is relatively more important when we use the trace-based information criteria to compute the boosted forecasting models. While the role of the stock market has been emphasized by Haan and Sturm (2017) as within-sample predictors of inequality in general, Gindelsky (2016), and Hood & Waters (2017) did not emphasize on the equity market in their forecasting analysis of the UK. Though our findings of the importance of the fiscal policy variable and economic growth is somewhat in line with Hood & Waters (2017). Finally, the real interest rate plays a role in terms of absolute and relative importance for the inequality measures DIY2, DIY3, and DIC3, which in turn supports the important role of monetary policy in affecting inequality, as suggested in Coibion et al. (2017) and, in particular for the U.K., by Mumtaz & Theophilopoulou (2017), based on within-sample analyses. However, the importance of monetary policy decisions does not appear in the out-of-sample analyses of Gindelsky (2016), and Hood & Waters (2017). Further, unlike existing studies on inequality projections of the UK, we show that not only does policy decision matter, but more importantly, uncertainty around policymaking is also very important for predicting the future path of inequality. This final result clearly hints at suggestion of transparency in policy decisions, and reduction of uncertainty is required to enhance investment (as theoretically discussed in Bernanke (1983), Dixit & Pindyck (1994), and recently by Bloom (2009)), and hence reduce inequality growth in the UK.

4.4 Out-of-Sample Forecasting

Intuitively, a forecasting model that includes several predictors and that, thereby, produces a larger in-sample adjusted R^2 and a smaller in-sample estimation error is more likely to overfit the data and, according to the standard bias-variance trade-off (see, for example, Hastie et al., 2009), to produce less favourable out-of-sample results than a more parsimonious forecasting model. The results on the out-of-sample performance of the boosted forecasting models that we summarize in Table 4 confirm this intuition. The forecasting horizon is one quarter. We measure out-of-sample performance in terms of the root-mean-squared forecasting error (RMSFE) and the out-of-sample R^2 proposed by Campbell & Thompson (2008). For implementing the out-of-sample R^2, we use the recursively computed historical mean of changes in the inequality measures as a benchmark forecast. A positive out-of-sample R^2 shows that the boosted forecasting models perform better than the benchmark forecast.

– Please include Table 4 about here. –

In terms of the RMSFE, the boosted forecasting models perform best for the DYI2 and DCI2 inequality measures, followed by the DYI1 and DC1 measures (Panel A). Results further show that the RMSFE is smaller when we study an active-set-based information criterion than for the corresponding trace-based information criterion. Hence, the more restrictive information criteria, which result in relatively parsimonious forecasting models, produce in general more accurate out-of-sample forecasts (in terms of the RMSFE statistic) than the relatively generous (in terms of the number of predictors) information criteria.

The results for the out-of-sample R^2 confirm this result (Panel B). For a given inequality measure, the out-of-sample R^2 is larger for the active-set-based information criterion then for the corresponding trace-based information criterion. This result, however, does not rule out that a trace-based information criterion yields a better out-of-sample R^2 than an active-set-based information criterion. For example, the out-of-sample R^2 is somewhat larger for $gMDL_{trace}$ than for $AIC_{activeset}$ information criterian when we consider the DCI2 inequality measure. Again, the boosted forecasting models perform best (relatively to the historical-mean benchmark) for the DIC2 consumption-based inequality measure (with the exception of the AIC_{trace} information criterion). The boosted forecasting models also perform well for the DYI2 and DYI3 income-based inequality measures when we consider the active-set-based information criteria. The boosted forecasting models realize the worst forecasting performance in terms of the out-of-sample R^2 for the DYI1 inequality measures, where the out-of-sample R^2 takes on negative values for all information criteria.

In addition, we estimate Fair-Shiller regressions (Fair & Shiller, 1990) to analyze the predictive value of the forecasts implied by the boosted forecasting models relative to the forecasts implied by the recursively computed historical mean of the dependent variable. Specifically, we estimate regression equations of the format $y_{i,t+h} = \alpha + \beta_1 \hat{y}^{boost}_{i,t+h} + \beta_2 \hat{y}^{m}_{i,t+h} + u_{i,t+h}$, where a hat denotes a forecast, u denotes the disturbance terms, and the index m denotes the recursively computed historical-mean benchmark forecasts. If the forecasts implied by the boosted forecasting models contain information over and above the information that the benchmark forecasts contain, then the coefficient β_1 should be significantly different from zero, while the coefficient β_2 should be zero. If, in contrast, the predictive value of the boosted forecasts is completely encapsulated in the benchmark forecasts and the latter contains additional information, then the coefficient β_1 should be zero while the coefficient β_2 should be significantly different from zero. If the boosted model and the benchmark forecasts contain exactly the same information then both coefficients, β_1 and β_2, are not separately identified. If the forecasts of both models do not have predictive value, then both coefficients should be zero. Finally, both coefficients should be significantly different from zero if the forecasts implied by the boosted forecasting model and the benchmark forecasts contain independent information. Our interest is primarily in the significance of the coefficient β_1, and so we shall only report results for this coefficient. We use Newey-West standard errors to compute robust standard errors. Panel C of Table 4 reports the p-values of the coefficient β_1. We reject the null hypothesis that this coefficient is zero for all inequality measures.[7]

– Please include Table 5 about here. –

While the L2-boosting algorithm is not tailored to forecast binary-coded dependent variables (in contrast to specialized boosting algorithms like logit boost, see Freund & Schapire, 1997; Friedman, 2001; Bühlmann & Yu, 2003), an alternative way to assess the out-of-sample performance of the boosted forecasting models is to study their directional accuracy. We use the Pesaran & Timmermann (1992) to test the directional accuracy of forecasts. Table 5 summarizes the results (Panel A). The forecasting horizon is one quarter. The results of the Pesaran-Timmermann test show that the boosted forecasting models perform well in terms of directional accuracy. With few exceptions, we can reject the null hypothesis of no predictive directional accuracy. The corresponding sucess rates (Panel B) vary from a mimimum of approximately 0.52 (trace-based information criteria) to a maximum of about 0.62 (active-set-based information criteria) for the DY1 inequality measure. The sucess rates are largest across information criteria for the DIY2,

[7]Results of Fair-Shiller regressions estimated on two-quarter-ahead and four-quarter-ahead forecasts also show that model that the boosting algorithm builds from the lagged inequality measures along with the various economic predictors contains significant additional incremental predictive value relative to the simple benchmark model. Results are not reported, but available from the authors upon request.

DIY3, and DIC2 inequality measures, a result that is consistent with the results of the Pesaran-Timmermann test.

– Please include Table 6 about here. –

Next, we study the Diebold & Mariano (1995) test to compare forecast accuracy. As in the case of the out-of-sample R^2, we consider the recursively computed historical mean of changes in the inequality measures as a benchmark forecast. We report results that we derive using the modified Diebold-Mariano test proposed by Harvey et al. (1997), where we report the p-values for both tests computed using the R package "forecast" (Hyndman, 2017; Hyndman & Khandakar, 2008). We report results for two forecast horizons (one quarter and four quarters), two alternative loss functions (squared-error loss and absolute-error loss), and two learning rates ($s = 0.25$ and $s = 0.1$).[8] Again, the forecasts that we compute using the active-set-based information criteria that give rise to relatively parsimonious forecasting models perform better than the trace-based information criteria. For one-quarter ahead forecasts, the tests yield several significant results for DYI3 and DCI2. For four-quarter ahead forecasts, in turn, we observe significant test results for all three income-based inequality measures under both squared-error loss and absolute-error loss (again for the active-set-based information criteria). Assuming a smaller learning rate strengthens the significance of the test results under the $AIC_{activeset}$ and $gMDL_{activeset}$ information criteria. More precisely, the test results are significant for the DIY3, DCI1, and DCI2 inequality measures under squared-error loss, and for the DIY2 (for the $AIC_{activeset}$ information criterion), DIY3, DCI1 (for the $gMDL_{activeset}$ information criterion), and DCI2 inequality measures under absolute-error loss.[9]

---

[8] We set $M = 500$ for the smaller learning rate. A smaller learning rate makes computations slower, but typically increases accuracy (see, for example, Friedman, 2001).

[9] Based on the suggestion of an anonymous referee, we re-conducted the analysis reported in Panels E and F of 6 with a random-walk, i.e., constant-mean model as a benchmark, and found qualitatively similar results. The analysis has been suppressed to save space, but complete details are available upon request from the authors.

5 Concluding Remarks

The dynamics of economic inequality has been at the center of many controversial policy debates in recent years. Our contribution to these debates is that we have used a relatively simple boosting algorithm to inspect which macroeconomic and financial variables along with lagged inequality measures help to forecast the dynamics of six different income- and consumption-based inequality measures of the UK at forecasting horizons of up to one year. The uniqueness of our analysis is that unlike existing studies, which forecast inequality at an annual frequency, we do so at a higher (quarterly) frequency based on a data set covering the period from 1975Q1 to 2016Q1. This is important, since given the wide-ranging negative impact of inequality, its predictability at higher frequency should be important for appropriate and timely policy-making than when inequality forecasts are available only at an annual (or even lower) frequency. Results show that, especially when we consider information criteria that give rise to relatively parsimonious forecasting models, the forecasts implied by the boosted forecasting models have predictive value (though the trace-based information criteria also work well in some of our forecasting exercises), where the strength of the evidence of predictability varies across inequality measures. We also have documented evidence of directional accuracy. Among the predictors that stand out in terms of absolute and/or relative importance are, in addition to lagged inequality measures, stock-market developments and fiscal deficits and, to a lesser extent, economic policy uncertainty, output growth, and the real interest rate.

From a policy perspective, our results thus highlight the importance of both monetary and fiscal policy variables in affecting the future path of inequality growth. Hence, the policymakers have an important role in reducing inequality by undertaking distributional measures, which in turn would also reduce the persistence in the process of inequality. Given that these policies are also likely to affect economic growth, inequality will also be affected indirectly via these policy decisions, given the role of economic growth in forecasting inequality growth obtained by our analysis. In addition, policymakers would also need to be transparent in their decision making, and reduce policy uncertainty to have positive impact on the reduction of inequality. Finally, our results are in line with the existing literature of stock market development enhancing inequality, and in this regard, the general prescription is that investment in equity market should be made as inclusive as possible for the entire population, which can however, only happen in the long-run as the redistributive policies reduce the disparity across income levels.

As part of future research, given the importance of forecasting inequality at a higher frequency, and also simultaneously realizing the need to model variables such as human-capital attainment and labor-force structure variables, which are only available at annual frequency, it would be interesting to apply the reverse unrestricted mixed data sampling (RU-MIDAS) model of Foroni et al. (2018) to gain further insights into the dynamics and the predictability of economic inequality.

References

Andersen T.G., & Bollerslev T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review, 39*, 885−905. http://dx.doi.org/10.2307/2527343

Arestis, P., Martin, R., & Tyler, P. (2011). The Persistence of Inequality? *Cambridge Journal of Regions Economy and Society, 4*, 3−11. http://dx.doi.org/10.1093/cjres/rsr001

Baker, S., Bloom, N., & Davis, S. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics, 131*, 1593−1636. https://doi.org/10.1093/qje/qjw024

Balcilar, M., Akadiri, S.S., Gupta, R., & Miller, S.M. (2019). Partisan conflict and income inequality in the United States: A nonparametric causality-in-quantiles approach. *Social Indicators Research, 142*, 65−82. https://doi.org/10.1007/s11205-018-1906-3

Balcilar, M., Chang, S., Gupta, R., & Miller, S.M. (2018). The relationship between the inflation rate and inequality across U.S. states: a semiparametric approach. *Quality & Quantity, 52*, 2413-2425. https://doi.org/10.1007/s11135-017-0676-3

Belfield, C., Blundell, R., Cribb, J., Hood, A., & Joyce, R. (2017). Two decades of income inequality in Britain: The role of wages, household earnings and redistribution. *Economica, 84*, 157−179. https://doi.org/10.1111/ecca.12220

Berg, A.G., & Ostry, J.D. (2017). Inequality and unsustainable growth: Two sides of the same coin? *IMF Economic Review, 65*, 792−815. https://doi.org/10.1057/s41308-017-0030-8

Berge, T.J. (2014). Forecasting disconnected exchange rates. *Journal of Applied Econometrics, 29*, 713−4735. https://doi.org/10.1002/jae.2350

Berge, T.J. (2017). Understanding survey based inflation expectations. *International Journal of Forecasting, 34*, 788−801. https://doi.org/10.1016/j.ijforecast.2018.07.003

Bernanke, B. (1983). Irreversibility, Uncertainty, and Cyclical Investment. *Quarterly Journal of Economics, 98*, 85−106. https://doi.org/10.2307/1885568

Bloom, N. (2009). The Impact of Uncertainty Shocks. *Econometrica, 77*, 623−685. https://doi.org/10.3982/E

Bühlmann, P. (2006). Boosting for high-dimensional linear models. *The Annals of Statistics, 34*, 559−583. https://doi.org/10.1214/009053606000000092

Bühlmann, P., & Hothorn, T. (2007). Boosting algorithms: regularization, prediction and model fitting. *Statistical Science, 22*, 477−505. https://doi.org/10.1214/07-STS242

Bühlmann, P., & Yu, B. (2003). Boosting with the L2 Loss: Regression and classification. *Journal of the American Statistical Association, 98*, 324−339. https://doi.org/10.1198/016214503000125

Campbell J.Y., & Thompson, S.B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies, 21*, 1509−1531. https://doi.org/10.1093/rfs/hhm055

Chang, S., Gupta, R., & Miller, S.M. (2018). Causality between per capita real GDP and income inequality in the U.S.: Evidence from a wavelet analysis. *Social Indicators Research, 135*, 269−289. https://doi.org/10.1007/s11205-016-1485-0

Chang, S., Gupta, R., Miller, S.M., & Wohar, M.E. (2019). Growth volatility and inequality in the U.S.: A wavelet analysis. *Physica A: Statistical Mechanics and its Applications, 521*, 48−73. https://doi.org/10.1016/j.physa.2019.01.024

Cloyne, J., & Surico, P. (2017). Household debt and the dynamic effects of income tax changes. *Review of Economic Studies, 84*, 45−81. https://doi.org/10.1093/restud/rdw021

Cloyne, J., Ferreira, C., & Surico, P. (2019). Monetary Policy when Households have Debt: New Evidence on the Transmission Mechanism. *Review of Economic Studies*, https://doi.org/10.1093/restud/rdy074.

Coibion, O., Gorodnichenko, Y., Kueng, L., & Silvia, J. (2017). Innocent bystanders? Monetary policy and inequality in the U.S. *Journal of Monetary Economics, 88*, 70−89. https://doi.org/10.1016/j.jmoneco.2017.05.005

Corak, M. (2016). Inequality from generation to generation: The United States in comparison. Institute for the Study of Labor (IZA), IZA Discussion Papers 9929. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2786013

Dayal, N., Gomulka, J., Mitton, L., Sutherland, H., & Taylor, R. (2000). Enhancing family resources survey income data with expenditure data from the Family Expenditure Survey: Data comparisons. Department of Applied Economics, University of Cambridge, Microsimulation Unit Research Note No. MU/RN/40. https://ideas.repec.org/p/ese/msimrn/mu-rn-40.html

de Haan J., & Sturm, J-E. (2017). Finance and income inequality: A review and new evidence. *European Journal of Political Economy, 50*, 171−195. https://doi.org/10.1016/j.ejpoleco.2017.04.007

Delhey, J., & Dragolov, G. (2013). Why inequality makes Europeans less happy: The role of distrust, status anxiety, and perceived conflict. *European Sociological Review, 30*, 151−165. https://doi.org/10.1093/esr/jct033

Diebold, F. X., & Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics, 13*, 253−263. https://doi.org/10.1080/07350015.1995.10524599

Dixit, A.K., & Pindyck, R.S. (1994). *Investment under uncertainty*. Princeton, New Jersey: Princeton University Press.

Dorling, D. (2015). Income inequality in the UK: Comparisons with five large Western European countries and the USA. *Applied Geography, 61*, 24−34. https://doi.org/10.1016/j.apgeog.2015.02.004

Döpke, J. Fritsche, U., & Pierdzioch, C. (2017). Predicting recessions with boosted regression trees. *International Journal of Forecasting, 33*, 745−759. https://doi.org/10.1016/j.ijforecast.2017.02.003

Fair, R.C., & Shiller, R.J. (1990). Comparing information in forecasts from econometric models. *American Economic Review, 80*, 375−389. https://www.jstor.org/stable/2006672

Foroni, C., Guerin, P., & Marcellino, M. (2018). Using low frequency information for predicting high frequency variables. *International Journal of Forecasting, 34*, 774−787. https://doi.org/10.1016/j.ijforecast.2018.06.004

Foster, K. (1996). A comparison of the census characteristics of respondents and non-respondents to the 1991 Family Expenditure Survey (FES). Office for National Statistics Survey Methodology Bulletin 38. Office for National Statistics Survey Methodology Bulletin 38: 38−40.

Freund, Y. & Schapire. R.E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences, 55*, 119−139. https://doi.org/10.1006/jcss.1997.1504

Friedman, J. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*, 1189−1232. https://doi.org/10.1214/aos/1013203451

Gindelsky, M. (2016). Will inequality continue to rise? Forecasting income inequality in the United States. *In:* The 34th IARIW General Conference, August 21-27, Dresden, Germany. http://www.iariw.org/dresden/gindelsky.pdf

Gupta, R., Ma, J., Risse, M., & Wohar, M.E. (2018). Common business cycles and volatilities in US States and MSAs: The role of economic uncertainty. *Journal of Macroeconomics, 57*, 317−337. https://doi.org/10.1016/j.jmacro.2018.06.009

Gustavsson, M., & Jordahl, H. (2008). Inequality and trust in Sweden: some inequalities are more harmful than others. *Journal of Public Economics, 92*, 348−365. https://doi.

org/10.1016/j.jpubeco.2007.06.010

Hansen, M., & Yu, B. (2001). Model selection and the principle of minimum description length. *Journal of the American Statistical Association, 96*, 746−774. https://doi.org/10.1198/016214501753168398

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting, 13*, 281−291. https://doi.org/10.1016/S0169-2070(96)00719-4

Hastie, T. (2007). Comment: Boosting algorithms: Regularization, prediction and model fitting. *Statistical Science, 22*, 513−515. https://doi.org/10.1214/07-STS242A

Hastie, T. Tibshirani, R., & Friedman, J.(2009). *The elements of statistical learning. Data mining, inference, and prediction*. New York: Springer-Verlag New York. https://doi.org/10.1007/978-0-387-84858-7

Hood, A., & Waters, T. (2017). Living standards, poverty and inequality in the UK: 2016−17 to 2021−22. Report, The Institute for Fiscal Studies. https://www.ifs.org.uk/publications/8957

Hurvich, C., Simonoff, J., & Tsai, C.-L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society Series B, 60*, 271−293. https://doi.org/10.1111/1467-9868.00125

Hyndman, R.J. (2017). forecast: Forecasting functions for time series and linear models. R package version 8.0. http://github.com/robjhyndman/forecast

Hyndman, R.J., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software, 26*, 1−22. https://doi.org/10.18637/jss.v027.i03

Kumof, M. Rancière, R., & Winant, P. (2015). Inequality, Leverage, and Crises. *American Economic Review, 105*, 1217−1245. https://doi.org/10.1257/aer.20110683

Layte, R., & Whelan, C.T. (2014). Who feels inferior? A test of the status anxiety hypothesis of social inequalities in health. *European Sociological Review, 30*, 525−535. https://doi.org/10.1093/esr/jcu057

Mayr, A., Hofner, B., & Schmid, M. (2012). The importance of knowing when to stop. A sequential stopping rule for component-wise gradient boosting. *Methods of Information in Medicine, 51*, 178−186. https://doi.org/10.3414/ME11-02-0030

Mumtaz, H., & Theodoridis, K. (2017). US financial shocks and the distribution of income and consumption in the UK. Working Papers 845, Queen Mary University of London, School of Economics and Finance.

Mumtaz, H., & Theophilopoulou, A. (2017). The impact of monetary policy on inequality in the UK. An empirical analysis. *European Economic Review, 98*, 410−423. https://doi.org/10.1016/j.euroecorev.2017.07.008

Ng, S. (2014). Viewpoint: Boosting recessions. *Canadian Journal of Economics, 47*, 1−34. https://doi.org/10.1111/caje.12070

Pesaran, M.H., & Timmermann, A. (1992). A simple nonparametric test of predictive performance. *Journal of Business and Economic Statistics, 10*, 461−465. https://doi.org/10.1080/07350015.1992.10509922

Pickett, K.E, & Wilkinson, R. (2009). *The Spirit Level: Why Equality is Better for Everyone.* UK: Penguin Books.

Pierdzioch, C., Risse, M., & Rohloff, S. (2015). A boosting approach to forecasting gold and silver returns: economic and statistical forecast evaluation. *Applied Economics Letters, 23*, 347−352. https://doi.org/10.1080/13504851.2015.1073835

Pierdzioch, C., & Risse, M. (2018). A machine-learning analysis of the rationality of aggregate stock market forecasts. *International Journal of Finance and Economics, 23*, 642−654. https://doi.org/10.1002/ijfe.1641

Piketty, T. (2014), *Capital in the twenty-first century.* Cambridge: Harvard University Press.

Püttmann, L. (2018). Patterns of panic: Financial crisis language in historical newspapers. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3156287.

R Development Core Team (2017). R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/. R version 3.3.3.

Rufrancos, H.G., Power, M., Pickett, K.E., & Wilkinson, R. (2013). Income inequality and crime: A review and explanation of the time-series evidence. *Sociology and Criminology-Open Access, 1*, 103. https://doi.org/10.4172/2375-4435.1000103

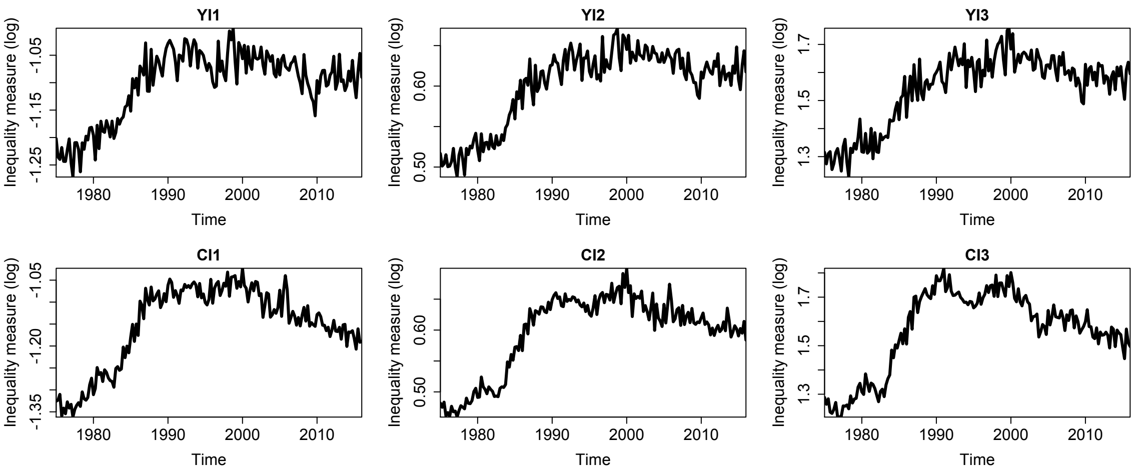
Uslaner, E.M., & Brown, M. (2005). Inequality, trust, and civic engagement. *American Politics Research, 33*, 868−894. https://doi.org/10.1177/1532673X04271903

van de Ven, J. (2011). Expenditure and disposable income trends of UK households: evidence from micro-data. *National Institute Economic Review, 218*, R44R57. https://doi.org/10.1177/002795011121800105

van Praag, B., & Ferrer-i-Carbonell, A. (2009). Inequality and happiness. In: The Oxford Handbook of Economic Inequality, Edited by Wiemer Salverda, Brian Nolan, and Timothy M. Smeeding, Chapter 17, Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199606061.001.0001

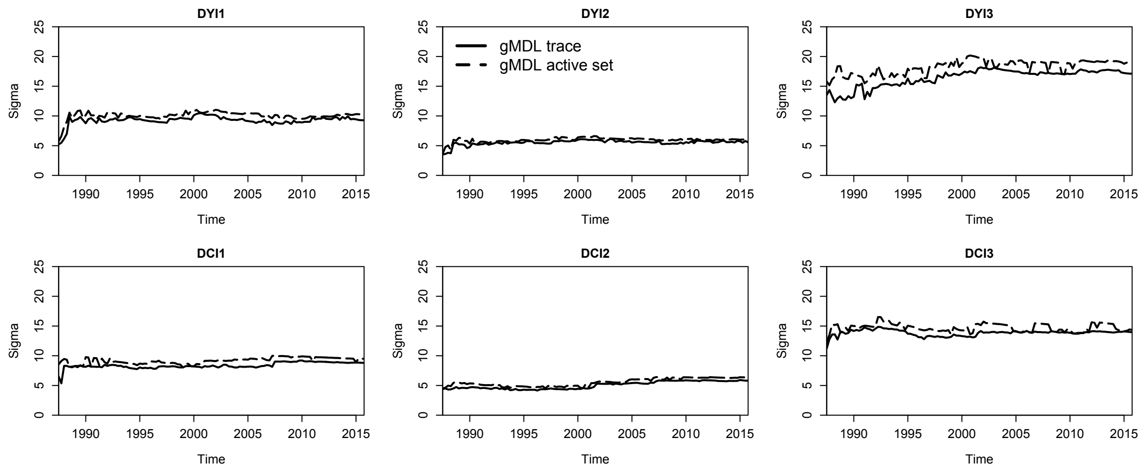
van Treeck, T. (2014). Did inequality cause the U.S. financial crisis? *Journal of Economic Surveys, 28*, 421−448. https://doi.org/10.1111/joes.12028

Figure 1: Inequality Measures



Note: YI denotes the three income-based inequality measures. CI denotes the three consumption-bassed inequality measures. For details on the data, see Section 3.

Figure 2: Evolution of the In-Sample Estimation Error



Note: The in-sample estimation error is defined in terms of the square root of the estimated variance of the disturbance term as computed by estimating by the ordinary-least-squares technique the boosted models using using the predictors selected by the boosting algorithm. A recursively expanding estimation window is used. Training period: 10 years. Forecast horizon: one quarter.

Table 1: Properties of the Boosted Forecasting Models

Panel A: Number of iterations

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	100.9825	96.7281	86.8509	64.1667	61.9035	50.7982
AIC active set	15.3070	16.4123	25.5702	8.1316	7.6140	11.0877
gMDL trace	76.5000	57.0439	66.5965	56.5965	40.4211	47.2895
gMDL active set	15.5439	15.5088	17.4211	8.2719	6.9649	12.1053

Panel B: Number of predictors

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	20.2456	19.7719	19.3947	14.9649	15.4737	15.9123
AIC active set	5.2982	5.8246	7.7807	2.4386	2.6667	4.5526
gMDL trace	17.3070	14.7105	16.1491	13.6316	12.4825	15.1667
gMDL active set	5.3772	5.4825	5.8070	2.4123	2.3158	4.9649

Panel C: In-sample R^2

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	0.4942	0.5138	0.5894	0.3875	0.4199	0.3655
AIC active set	0.3861	0.4126	0.5113	0.2565	0.2939	0.2754
gMDL trace	0.4806	0.4906	0.5808	0.3867	0.4128	0.3661
gMDL active set	0.3885	0.4075	0.4853	0.2623	0.2896	0.2825

Note: Number of iterations is the average number of iterations it takes to minimize an information criterion. Number of predictors is the average number of predictors included in the boosted forecsating models. In-sample R^2 is the average adjusted coefficient of determination. A recursively expanding estimation window is used to compute the boosted forecasting models. Training period: 10 years. Forecast horizon: one quarter. Learning rate: $s = 0.25$.

Table 2: Inclusion of Predictors in the Boosted Forecasting Models

Panel A: DYI1

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GINV	UR
AIC trace	74.34	43.86	52.19	62.72	32.46	11.40	43.64	26.75	69.52	34.65	32.24	3.51	12.06	6.80	0.00
AIC active set	48.46	0.66	11.84	26.32	0.00	0.00	6.58	6.14	32.46	0.00	0.00	0.00	0.00	0.00	0.00
gMDL trace	70.39	38.38	48.46	57.89	21.71	7.46	39.25	23.46	60.75	28.73	23.25	1.97	7.24	3.73	0.00
gMDL active set	48.46	0.88	12.28	26.54	0.00	0.00	6.58	6.58	33.11	0.00	0.00	0.00	0.00	0.00	0.00

Panel B: DYI2

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GNV	UR
AIC trace	78.29	23.68	48.25	54.61	49.12	13.60	33.11	26.97	71.27	33.77	25.44	1.75	20.18	14.25	0.00
AIC active set	43.86	0.44	6.14	28.07	21.27	0.00	0.22	2.41	42.32	0.88	0.00	0.00	0.00	0.00	0.00
gMDL trace	66.89	14.69	41.01	50.66	44.96	3.95	17.76	23.25	64.91	16.67	8.11	0.44	10.75	3.73	0.00
gMDL active set	42.32	0.44	5.48	27.85	16.67	0.00	0.22	1.32	41.89	0.88	0.00	0.00	0.00	0.00	0.00

Panel C: DYI3

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GNV	UR
AIC trace	85.75	33.11	38.16	85.09	48.90	25.88	11.84	30.04	83.77	11.40	11.40	7.68	5.04	6.58	0.22
AIC active set	66.01	1.32	11.62	37.28	21.05	1.97	0.00	7.24	47.59	0.22	0.00	0.00	0.00	0.22	0.00
gMDL trace	82.24	22.37	30.70	73.25	45.83	20.39	5.92	25.00	82.68	5.04	3.29	2.41	1.97	2.41	0.22
gMDL active set	58.77	0.44	5.70	32.46	8.55	1.54	0.00	1.54	36.18	0.00	0.00	0.00	0.00	0.00	0.00

Panel D: DCI1

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GNV	UR
AIC trace	68.42	26.32	34.21	70.83	9.21	3.07	16.89	42.11	46.05	10.53	13.60	9.65	4.17	8.11	10.96
AIC active set	42.11	1.32	3.51	5.04	0.00	0.00	2.41	1.32	3.07	0.44	0.44	0.00	0.00	0.88	0.44
gMDL trace	65.79	23.46	33.33	62.06	6.58	3.07	14.69	39.91	43.20	8.33	11.18	6.80	4.17	7.89	10.31
gMDL active set	42.11	2.63	3.51	5.26	0.00	0.22	1.75	2.19	1.32	0.00	0.44	0.00	0.00	0.44	0.44

Panel E: DCI2

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GNV	UR
AIC trace	75.66	25.44	50.66	57.02	13.82	2.63	11.62	27.41	62.06	10.96	11.62	4.39	8.11	8.99	16.45
AIC active set	41.67	0.66	0.22	1.54	0.44	0.00	2.85	1.75	14.47	1.32	0.00	0.00	0.66	0.22	0.88
gMDL trace	74.56	17.98	39.91	42.54	8.99	1.10	8.11	22.15	51.10	10.09	6.58	0.44	6.58	7.24	14.69
gMDL active set	39.25	1.10	0.44	1.10	0.22	0.00	0.88	1.97	10.09	1.10	0.00	0.00	1.54	0.00	0.22

Panel F: DCI3

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GNV	UR
AIC trace	76.10	25.88	52.85	47.59	39.47	14.25	6.36	20.61	59.43	14.47	16.45	5.04	5.04	0.44	13.82
AIC active set	55.04	1.75	2.19	15.13	10.09	0.00	0.00	7.24	9.87	3.73	0.00	0.00	0.00	0.00	8.77
gMDL trace	75.66	24.78	49.56	47.59	33.99	12.72	6.36	18.20	59.43	14.47	14.25	4.39	3.95	0.00	13.82
gMDL active set	54.17	2.85	4.39	15.35	12.28	0.22	0.00	9.21	14.25	3.73	0.00	0.00	0.66	0.00	7.02

Note: Predictor inclusion is defined as the number of times a predictor (including all its lags) is included in the boosted forecasting models divided by the maximum times a predictor can be included in the boosted forecasting models. Predictor inclusion is expressed in percent. Predictor inclusion is computed across all forecasting periods. A recursively expanding estimation window is used to compute the boosted forecasting models. Training period: 10 years. Forecast horizon: one quarter. Learning rate: $s = 0.25$.

Table 3: Relative Importance of Predictors

Panel A: DYI1

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GINV	UR
AIC trace	14.69	8.67	10.31	12.39	6.41	2.25	8.62	5.29	13.73	6.85	6.37	0.69	2.38	1.34	0.00
AIC active set	36.59	0.50	8.94	19.87	0.00	0.00	4.97	4.64	24.50	0.00	0.00	0.00	0.00	0.00	0.00
gMDL trace	16.27	8.87	11.20	13.38	5.02	1.72	9.07	5.42	14.04	6.64	5.37	0.46	1.67	0.86	0.00
gMDL active set	36.05	0.65	9.14	19.74	0.00	0.00	4.89	4.89	24.63	0.00	0.00	0.00	0.00	0.00	0.00

Panel B: DYI2

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GNV	UR
AIC trace	15.84	4.79	9.76	11.05	9.94	2.75	6.70	5.46	14.42	6.83	5.15	0.35	4.08	2.88	0.00
AIC active set	30.12	0.30	4.22	19.28	14.61	0.00	0.15	1.66	29.07	0.60	0.00	0.00	0.00	0.00	0.00
gMDL trace	18.19	4.00	11.15	13.77	12.22	1.07	4.83	6.32	17.65	4.53	2.21	0.12	2.92	1.01	0.00
gMDL active set	30.88	0.32	4.00	20.32	12.16	0.00	0.16	0.96	30.56	0.64	0.00	0.00	0.00	0.00	0.00

Panel C: DYI3

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GNV	UR
AIC trace	17.68	6.83	7.87	17.55	10.09	5.34	2.44	6.20	17.28	2.35	2.35	1.58	1.04	1.36	0.05
AIC active set	33.93	0.68	5.98	19.17	10.82	1.01	0.00	3.72	24.46	0.11	0.00	0.00	0.00	0.11	0.00
gMDL trace	20.37	5.54	7.60	18.14	11.35	5.05	1.47	6.19	20.48	1.25	0.81	0.60	0.49	0.60	0.05
gMDL active set	40.48	0.30	3.93	22.36	5.89	1.06	0.00	1.06	24.92	0.00	0.00	0.00	0.00	0.00	0.00

Panel D: DCI1

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GNV	UR
AIC trace	18.29	7.03	9.14	18.93	2.46	0.82	4.51	11.25	12.31	2.81	3.63	2.58	1.11	2.17	2.93
AIC active set	69.06	2.16	5.76	8.27	0.00	0.00	3.96	2.16	5.04	0.72	0.72	0.00	0.00	1.44	0.72
gMDL trace	19.31	6.89	9.78	18.21	1.93	0.90	4.31	11.71	12.68	2.45	3.28	1.99	1.22	2.32	3.02
gMDL active set	69.82	4.36	5.82	8.73	0.00	0.36	2.91	3.64	2.18	0.00	0.73	0.00	0.00	0.73	0.73

Panel E: DCI2

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GNV	UR
AIC trace	19.56	6.58	13.10	14.74	3.57	0.68	3.00	7.09	16.04	2.83	3.00	1.13	2.10	2.32	4.25
AIC active set	62.50	0.99	0.33	2.30	0.66	0.00	4.28	2.63	21.71	1.97	0.00	0.00	0.99	0.33	1.32
gMDL trace	23.89	5.76	12.79	13.63	2.88	0.35	2.60	7.10	16.37	3.23	2.11	0.14	2.11	2.32	4.71
gMDL active set	67.80	1.89	0.76	1.89	0.38	0.00	1.52	3.41	17.42	1.89	0.00	0.00	2.65	0.00	0.38

Panel F: DCI3

Criterion	DYI1	RHP	GDP	RSP	RIR	EER	RV	EPU	BUD	FSI	TAX	ITAX	BEN	GNV	UR
AIC trace	19.13	6.50	13.29	11.96	9.92	3.58	1.60	5.18	14.94	3.64	4.13	1.27	1.27	0.11	3.47
AIC active set	48.36	1.54	1.93	13.29	8.86	0.00	0.00	6.36	8.67	3.28	0.00	0.00	0.00	0.00	7.71
gMDL trace	19.95	6.54	13.07	12.55	8.96	3.35	1.68	4.80	15.67	3.82	3.76	1.16	1.04	0.00	3.64
gMDL active set	43.64	2.30	3.53	12.37	9.89	0.18	0.00	7.42	11.48	3.00	0.00	0.00	0.53	0.00	5.65

Note: Relative importance is defined as the number of times a predictor (including all its lags) is included in the boosted forecasting models divided by the total number of predictors included in the boosted forecasting models. Relative importance is expressed in percent. Relative importance is computed across all forecasting periods. A recursively expanding estimation window is used to compute the boosted forecasting models. Training period: 10 years. Forecast horizon: one quarter. Learning rate: $s = 0.25$.

Table 4: Out-of-Sample Performance of the Boosted Forecasting Models

Panel A: Root-mean-squared forecasting error

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	13.5949	8.0121	23.9157	14.1719	9.7574	17.5791
AIC active set	12.9137	7.3146	22.0480	10.7621	7.2680	17.2595
gMDL trace	13.7459	7.9818	22.9855	13.7843	7.2598	17.6378
gMDL active set	12.8182	7.4450	21.7861	10.9554	7.1086	17.4575

Panel B: Out-of-sample R^2

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	-0.1494	-0.0658	0.0146	-0.6891	-0.5051	0.0217
AIC active set	-0.0371	0.1117	0.1625	0.0259	0.1649	0.0570
gMDL trace	-0.1751	-0.0577	0.0898	-0.5979	0.1668	0.0152
gMDL active set	-0.0218	0.0798	0.1823	-0.0094	0.2012	0.0352

Panel C: Fair-Shiller regressions

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	0.0014	0.0006	0.0001°	0.0584	0.1000	0.0001°
AIC active set	0.0001	0.0001°	0.0001°	0.0001°	0.0001°	0.0001°
gMDL trace	0.0032	0.0002	0.0001°	0.0684	0.0001°	0.0001°
gMDL active set	0.0000	0.0002	0.0001°	0.0001	0.0001°	0.0001°

Note: The RMSE is computed as $RMSE = [(1/T)(y_{i,t+h} - \hat{y}^{boost}_{i,t+h})^2]^{0.5}$, where $y_{i,t+h}$ denotes the actual data and T denotes the number of out-of-sample forecasting periods. The out-of-sample R^2 is computed using the recursively estimated historical mean, $\hat{y}^m_{i,t+h}$, of the dependent variable as a benchmark as follows: $R^2 = 1 - [\sum_T (y_{i,t+h} - \hat{y}^{boost}_{i,t+h})^2]/[\sum_T (y_{i,t+h} - \hat{y}^m_{i,t+h})^2]$, where the summation is over the out-of-sample forecasting periods. The p-values of the coefficient β_1 in the Fair-Shiller regression, $y_{i,t+h} = \alpha + \beta_1 \hat{y}^{boost}_{i,t+h} + \beta_2 \hat{y}^m_{i,t+h} + u_{i,t+h}$, are computed using Newey-West standard errors. Training period: 10 years. Forecast horizon: one quarter. Learning rate for the boosted model: $s = 0.25$. A $^\circ$ denotes a p-value samller than 0.0001.

Table 5: Directional Accuracy of Out-of-Sample Forecasts

Panel A: Pesaran-Timmermann test

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	2.6406	2.3550	2.4886	1.4955	2.3364	2.1434
AIC active set	0.8523	2.3550	2.6777	1.8675	2.5044	1.5024
gMDL trace	2.5933	2.0682	2.2653	1.8748	2.7464	2.3549
gMDL active set	0.8523	2.3550	2.2812	1.6671	2.6263	1.8875

Panel B: Sucess rates

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	0.6228	0.6053	0.6140	0.5702	0.6053	0.5965
AIC active set	0.5263	0.6053	0.6228	0.5877	0.6053	0.5614
gMDL trace	0.6140	0.5877	0.6053	0.5877	0.6228	0.6053
gMDL active set	0.5263	0.6053	0.6053	0.5789	0.6140	0.5789

Note: Directional accuracy is analyzed by means of the Pesaran-Timmermann test. The null hypothesis is that the forecasts have no predictive value with regard to the direction of change of the inequality measures. The Pesaran-Timmermann test has a standard normal distribution (the 10% and 5% one-sided critical values are 1.64 and 1.95). In order to set up the test, a contingency table with the marginal events $y_{i,t+h} > 0$, $\hat{y}^{boost}_{i,t+h} < 0$, $y_{i,t+h} \geq 0$, and $\hat{y}^{boost}_{i,t+h} \geq 0$ is constructed. The events in category ij of this contingency table have frequency $\hat{P}_{ij} = T_{ij}/T$. The estimated frequencies of the marginal events are given by $\hat{P}_{0j} = T_{0j}/T$ and $\hat{P}_{i0} = T_{i0}/T$, where T_{0j} (T_{i0}) denotes the sum of the rows (columns). Using the notation also used by Pesaran and Timmermann (1992), the following quantities are defined: $\mathbf{P} = (P_{11}, P_{12}, P_{21}, P_{22})$ and $V = \left(\frac{\partial f(\mathbf{P})}{\partial \mathbf{P}}\right)' (\Psi - \mathbf{P}\mathbf{P}') \left(\frac{\partial f(\mathbf{P})}{\partial \mathbf{P}}\right)$, with $\frac{\partial f(\mathbf{P})}{\partial P_{ij}} = 1 - P_{i0} - P_{0i}$ for $i = j$, and $\frac{\partial f(\mathbf{P})}{\partial P_{ij}} = -P_{i0} - P_{0j}$ otherwise, where $\Psi =$ diagonal matrix with the elements of $\mathbf{P}$ on its diagonal. The Pesaran-Timmermann test is computed as $\sqrt{n}V^{-1/2}S \to N(0,1)$, where $S = \sum_{i=1}^{2} \hat{P}_{ii} - \hat{P}_{i0}\hat{P}_{0i}$ and V and S are evaluated at $\mathbf{P} = \hat{\mathbf{P}}$. In order to test the null hypothesis, positive values of the test (correctly predicted direction of change) are of interest. The success rate is computed be defining $x_1 = \sum_T [\mathbf{1}(y_{i,t+h} > 0) \times (\mathbf{1}\hat{y}^{boost}_{i,t+h} > 0)]$ and $x_2 = \sum_T [\mathbf{1}(y_{i,t+h} < 0) \times (\mathbf{1}\hat{y}^{boost}_{i,t+h} < 0)]$, and then computing success rate $= (x_1 + x_2)/T$, where $\mathbf{1}$ denotes the indicator function and the summation is over the out-of-sample forecasting periods, T. A recursively expanding estimation window is used to compute the boosted forecasting models. Training period: 10 years. Forecast horizon: one quarter. Learning rate: $s = 0.25$. Success rates: correct directional forecasts divided by the total number of forecasts.

Table 6: Diebold-Mariano Tests

Panel A: One-quarter ahead (squared-error loss)

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	0.7818	0.6361	0.4596	0.8618	0.7665	0.4261
AIC active set	0.6086	0.2292	0.0880	0.3896	0.0580	0.2527
gMDL trace	0.8111	0.6253	0.2498	0.8285	0.0932	0.4492
gMDL active set	0.5634	0.3028	0.0643	0.5363	0.0319	0.3408

Panel B: One-quarter ahead (absolute-error loss)

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	0.5750	0.5247	0.5586	0.8546	0.5618	0.6624
AIC active set	0.7115	0.1825	0.1883	0.4018	0.0772	0.6803
gMDL trace	0.5950	0.4922	0.3745	0.7864	0.1554	0.6533
gMDL active set	0.6675	0.2118	0.1586	0.3911	0.0303	0.6713

Panel C: Four-quarters ahead (squared-error loss)

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	0.1452	0.9721	0.5633	0.6894	0.7675	0.8954
AIC active set	0.0000	0.0058	0.1059	0.5216	0.6540	0.4263
gMDL trace	0.1859	0.9200	0.5266	0.6558	0.6557	0.8832
gMDL active set	0.0001	0.0105	0.0674	0.4247	0.6370	0.4022

Panel D: Four-quarters ahead (absolute-error loss)

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	0.1459	0.9658	0.3525	0.8348	0.6895	0.6970
AIC active set	0.0005	0.0105	0.0241	0.6091	0.5021	0.3069
gMDL trace	0.1484	0.8480	0.3146	0.8144	0.5817	0.7149
gMDL active set	0.0005	0.0147	0.0071	0.5071	0.4611	0.2536

Panel E: One-quarter ahead (squared-error loss, smaller learning rate)

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	0.8761	0.6762	0.3983	0.8197	0.7475	0.5777
AIC active set	0.3834	0.1461	0.0712	0.0758	0.0265	0.1379
gMDL trace	0.8293	0.5898	0.3288	0.8207	0.1527	0.5582
gMDL active set	0.3880	0.1852	0.0337	0.0667	0.0242	0.1323

Panel F: One-quarter ahead (absolute-error loss, smaller learning rate)

Criterion	DYI1	DYI2	DYI3	DCI1	DCI2	DCI3
AIC trace	0.7622	0.6155	0.5806	0.7803	0.4640	0.7842
AIC active set	0.4897	0.0907	0.0464	0.1046	0.0367	0.3354
gMDL trace	0.6779	0.4472	0.4849	0.7545	0.1303	0.7543
gMDL active set	0.4838	0.1406	0.0272	0.0757	0.0263	0.2837

Note: p-values of the Diebold-Mariano test. A recursively expanding estimation window is used to compute the boosted forecasting models. Training period: 10 years. Learning rate (Panels A to D): $s = 0.25$. Learning rate (Panels E and F): $s = 0.1$. The recursively computed historical mean is the benchmark forecast. In order to set up the Diebold-Mariano test, the forecast errors for the boosted forecasts, $fe^{boost}_{i,t+h} = y_{i,t+h} - \hat{y}^{boost}_{i,t+h}$, and the recursively-estimated-mean forecasts, $fe^{m}_{i,t+h} = y_{i,t+h} - \hat{y}^{m}_{i,t+h}$, are computed. The loss differential is computed as $d_{i,t+h} = |fe^{m}_{i,t+h}|^p - |fe^{boost}_{i,t+h}|^p$, where $p = 1$ for absolute loss and $p = 2$ for squared-error loss. The Diebold-mariano test is computed as kDM, where $k =$ Harvey-Leybourne-Newbold adjustment factor, $DM = \bar{d}/(\hat{\Omega}(\bar{d})^{0.5})$, where $\bar{d} =$ average loss differential, and $\hat{\Omega}(\bar{d}) =$ estimate of the variance of $\bar{d}$.

Appendix

Table A1: Data

Abbreviation	Economic interpretation and transformations	Publication lag (in quarters)	Data source
RHP	log of real house price, annualized quarter-on-quarter growth rate	1	OECD
GDP	log of real GDP, annualized quarter-on-quarter growth rate	2	OECD
RSP	log of real stock price, annualized quarter-on-quarter growth rate	1	Datastream and OECD
RIR	real interest rate, 3 month treasury bill rate minus CPI inflation	1	OECD
EER	log of the real effective exchange rate, annualized quarter-on-quarter growth rate	1	Bank of International Settlements (BIS)
RV	realized stock-market volatility	1	Datstream
EPU	economic policy uncertainty	1	`http://policyuncertainty.com/uk_monthly.html`
BUD	budget deficit, as percentage of GDP	2	OECD
FSI	U.S. financial stress	1	`http://policyuncertainty.com/financial_stress.html`
TAX	direct taxes, as percentage of GDP	2	OECD
ITAX	indirect taxes, as percentage of GDP	2	OECD
BEN	social benefits, as percentage of GDP	2	OECD
GNV	productive public expenditure (investment), as percentage of GDP	2	OECD
UR	unemployment Rate	2	OECD