HARVARD | BUSINESS | SCHOOL

Is Accounting Useful for Forecasting GDP Growth? A Machine Learning Perspective

Srikant Datar
Apurv Jain
Charles C.Y. Wang
Siyu Zhang

Working Paper 21-113

Is Accounting Useful for Forecasting GDP Growth? A Machine Learning Perspective

Srikant Datar
Harvard Business School

Apurv Jain
MacroXStudio

Charles C.Y. Wang
Harvard Business School

Siyu Zhang
Harvard Business School

Working Paper 21-113

Funding for this research was provided in part by Harvard Business School.

Is Accounting Useful for Forecasting GDP Growth? A Machine Learning Perspective*

Srikant Datar
Harvard Business School

Apurv Jain
MacroXStudio

Charles C.Y. Wang
Harvard Business School

Siyu Zhang
Harvard Business School

December 1, 2020

Abstract

We provide a comprehensive examination of whether, to what extent, and which accounting variables are useful for improving the predictive accuracy of GDP growth forecasts. We leverage statistical models that accommodate a broad set of (341) variables—outnumbering the total time-series observations—and apply machine learning techniques to train, validate, and test the prediction models. For near-term (current and next-quarter) GDP growth, accounting does not improve the out-of-sample accuracy of predictions because the professional forecasters' predictions are relatively efficient. Accounting's predictive usefulness increases for more-distant-term (three- and four-quarters-ahead) GDP growth forecasts: they contribute more to the model's predictions; moreover, their inclusion increases the model's out-of-sample predictive accuracy by 13 to 46%. Overall, four categories of accounting variables—relating to profits, accrual estimates (e.g., loan loss provisions or write-offs), capital raises or distributions, and capital allocation decisions (e.g., investments)—are most informative of the longer-term outlook of the economy.

JEL: E01, E32; E37; E60; M41
Keywords: Accounting; Big Data; Elastic Net; GDP Growth; Machine Learning; Macro Forecasting; Short Fat Data

*Datar (sdatar@hbs.edu) is the Arthur Lowes Dickinson Professor of Business Administration. Jain (apurvj@macroxstudio.com) is the CEO and co-founder of MacroXStudio and a former visiting researcher at Harvard Business School. Wang (charles.cy.wang@hbs.edu) is the Glenn and Mary Jane Creamer Associate Professor of Business Administration at Harvard Business School. Zhang (szhang@hbs.edu) is a Research Associate at Harvard Business School. For helpful comments and suggestions, we are grateful to Anupam Datta, David Kurokawa, and Will Uppington from Truera and Jay Shuran from DataRobot. We also thank Paul Hamilton for excellent research assistance. Comments are welcome, and all errors remain ours.

1 Introduction

A significant amount of economic activity and resource planning depends on expectations about the future states of the economy. Managers rely on economic forecasts to plan future operating activities, such as supply chain management, and financing activities, such as capital budgeting. Accurate macroeconomic forecasts are also important for policymakers, for example, in determining monetary and fiscal policy. Thus, improving the accuracy of the estimates of both prior and future macroeconomic conditions carry far-reaching implications.

Recent literature suggests that publicly listed firms' financial statements contain relevant information for estimating key macroeconomic indicators (Arif and Lee, 2014; Konchitchki and Patatoukas, 2014a,b; Nallareddy and Ogneva, 2017). This is not surprising: public companies account for 50% of all business profits and employ about 30% of private-sector employees. The accrual accounting system reflects firms' historical performance. It also enables managers to exercise discretion in reporting performance based on expectations of *future* business conditions (e.g., in the estimation allowances for doubtful accounts, allowances for loan losses, or write-off of assets). Therefore, in theory, public firms' financial statements could contain information useful for improving the precision of forecasts of future macroeconomic conditions.

Notably, prior research suggests that information in publicly available financial statements could help explain future GDP growth. For example, Konchitchki and Patatoukas (2014a) and Konchitchki and Patatoukas (2014b) show that the consensus forecasts of GDP growth from the Survey of Professional Forecasters (SPF), the oldest and most well-regarded publicly available quarterly survey of U.S. macroeconomic forecasts, do not efficiently incorporate the macroeconomic information content in public firms' earnings and profitability ratios.

This paper examines whether, to what extent, and which type of accounting information is useful for improving the *predictive accuracy* of GDP growth forecasts. It is worth noting that the prior literature does not focus on predictive accuracy as a primary outcome of interest. Instead, it focuses on *explanation*, by assessing whether accounting variables are significant in an in-sample regression of future GDP growth. Moreover, while the prior literature's findings are suggestive, accounting information's predictive relevance remains an open research question for two reasons. First, the prior literature considers the usefulness of a small subset of the accounting system's outputs, such as aggregate net income growth (Konchitchki and Patatoukas, 2014a) or the growth in accounting profitability (Konchitchki and Patatoukas, 2014b). The focus on specific accounting variables is in part necessitated by the nature of the research problem and the solutions (ordinary least squares or "OLS") applied: the time-series observations on macroeconomic performance (e.g., quarterly GDP growth rates) are far smaller than the number of potentially relevant accounting variables, and OLS is ill-suited for such high dimensional data (or "short fat data") problems. Given the richness of the information produced by public firms, the focus on specific variables could understate the overall importance of accounting information for forecasting GDP growth. Second, prior literature's evidence is based on in-sample regression fits, which could lead to data over-fitting. To the extent so, the existing evidence would overstate the usefulness of accounting information for macroeconomic forecasting.

To address these two problems, we apply machine learning techniques and principles to better understand accounting information's usefulness for forecasting GDP growth rates. We train algorithms that can accommodate high dimensional data. In particular, we use the elastic net (Zou and Hastie, 2005), a penalized regression model. The elastic net model is attractive because it is a linear model and thus preserves OLS's interpretability. Moreover, it can estimate linear coefficients even when the number of predictors (or model "features") exceed the number of observations. Finally, it can simultaneously perform regularization, mitigating the over-fitting problem, and variable selection, which allows us to identify the most critical forecast features.

To further tackle the over-fitting problem and estimate the models' out-of-sample prediction error rates, we split the data into training, validation, and holdout sets (Hastie et al., 2009).

We develop our prediction model using the training and validation data. Once finalized, we estimate the predictive performance of the model on the holdout set.

We apply these machine learning techniques to the problem of forecasting quarterly GDP growth estimates (the target variable of interest), specifically the Bureau of Economic Analysis' (BEA) third ("final") estimate of a quarter's GDP growth rates. These estimates are released at the end of the third month after the quarter ends, and are more authoritative because they are based on more complete data (Landefeld et al., 2008) than the initial ("advance") estimate, released in the first month after the quarter ends. We focus on GDP because it is perhaps the most important of the macroeconomic indicators and is consistent with the prior literature examining the relation between accounting information and the economy (Arif and Lee, 2014; Konchitchki and Patatoukas, 2014a,b; Nallareddy and Ogneva, 2017).[1] In particular, we focus on forecasting GDP growth for the current and next four quarters. Obtaining more accurate estimates for all of them is useful for better understanding the economy's trajectory and could have significant implications for managers, policymakers, and investors.

To train the models for forecasting quarterly GDP growth, we consider a large set of 341 features, spanning three categories of predictors constructed using: survey estimates of future GDP growth and BEA estimates for current quarter GDP growth (21 features), information about market returns and stock prices (28 features), and accounting information (292 features). The prior literature also considers these three categories of predictors in forecasting GDP growth. Our analysis differs in that we consider a much broader set of predictors within each category. For example, we consider a broad set of firm-level accounting variables and various ways for summarizing them in the cross-section (e.g., taking the mean, median, and variance). By considering a significantly greater number of variables than prior studies, we can provide a more comprehensive assessment of accounting information's

[1] Baumohl (2012) refers to GDP as the "mother of all economic indicators." The Bloomberg Economic Calendar for the United States indicates the importance of various economic releases and ranks GDP in the most important tier, suggesting the measure's importance to the investment community.

usefulness in forecasting quarterly GDP growth and pinpoint which predictor types are most useful.

In examining the usefulness of accounting information in this forecasting exercise, our empirical analyses ask the following questions. Which elastic net model would the forecaster obtain if she did not have accounting data? Which model would the forecasters obtain if accounting information were available to her? Finally, does the model trained with accounting data produce more accurate predictions in terms of its prediction errors on the holdout sample?

To answer these questions, we train variants of the elastic net with and without accounting information, and compare their holdout predictive accuracy. The most basic version of the model uses as features the survey and BEA estimates. The next benchmark variant adds to the survey and BEA estimates those features relating to market prices and returns. Then, we train two model variants that use accounting information: the first combines accounting information with survey and BEA estimates, and the second uses all available features. For each target variable (e.g., quarter Q, $Q+1$,$Q+2$, $Q+3$, or $Q+4$ GDP growth), we compare the mean-squared prediction errors (MSE) on the holdout sample (the "test MSE"), a standard metric for assessing the general accuracy of a predictive model (Hastie et al., 2009).

We find that accounting information is not particularly useful in improving the accuracy of current quarter or next quarter GDP growth forecasts. This is partly because the consensus estimates of professional forecasters do a good job of forecasting near-horizon GDP growth.

However, we find that accounting information is useful for forecasting more distant quarters' GDP growth rates. For example, we find that, relative to the models trained on survey and BEA estimates and market information, the introduction of accounting features to model training lowers the holdout sample MSE by about 2%, 13%, and 46% for forecasts of $Q+2$, $Q+3$, and $Q+4$ GDP growth.

We also analyze which features are most "important" in the trained elastic net models. In particular, we focus on the "full" elastic net model—trained using the full set of features—and ask the following question: how important are the features based on accounting information in the model's predictions? We consider two methods for evaluating the importance of accounting features in our trained elastic net models, based on the feature coefficients and their Shapley values.

In both sets of analyses, we find that the importance of survey forecasts and BEA estimates are declining as we forecast more-distant-quarter GDP growth. For example, the survey forecasts account for more than 90% of the feature importance in forecasting current quarter GDP growth, but this percentage declines monotonically to 0% in the model that forecasts $Q+4$ GDP growth. In contrast, the importance of accounting features increases as we forecast more-distant-quarter GDP growth. For example, whereas accounting features account for less than 10% of the feature importance for forecasting current quarter GDP growth, this percentage increases monotonically to 90% in the model that forecasts $Q + 4$ GDP growth.

Our analyses of feature importance also shed light on the type of accounting variables useful for GDP growth forecasting. Like the prior literature, we find that accounting features relating to profits are valuable. Moreover, we find that accrual accounting estimates (such as loan loss provisions and write-offs), capital raises or distributions (such as dividends or equity issuances), and capital allocation decisions (such as investments and the growth in PP&E) are also important.

That these additional categories provide signals about future states of the economy may not be entirely surprising. In making decisions about investments, payouts, or share issuances, managers need to consider the current and expected future state of the economy. Moreover, the accrual accounting system allows managers to express their forecasts about their businesses' state and the expected future states of the economy, such as determining reserves for loan losses or uncollectible accounts or determining write-offs. However, the value of these signals may be more subtle, and therefore could provide greater incremental value for longer-horizon GDP forecasts.

Our findings contribute to the literature in several ways. First, we provide an analytical structure and methodologies for evaluating accounting information's predictive relevance for macroeconomic forecasting. In contrast to the prior literature, which focuses on explaining in-sample variation, our work focuses on forecasting models' predictive accuracy, which is more relevant to policymakers, managers, and investors.

We also contribute to the literature by providing a comprehensive assessment of accounting information's usefulness for forecasting quarterly GDP growth. Our results differ from the prior literature in that we do not find accounting information relevant for near-term GDP growth forecasts. To the extent accounting information is relevant, we find that they are most likely useful for more distant quarterly GDP growth forecasts. Also, whereas the prior literature focuses on profitability measures, we document the importance of several other types of accounting information.

Finally, we contribute methodologically to the literature by demonstrating how machine learning techniques can be applied to prediction problems in accounting research. For example, we illustrate the use of regularization techniques for mitigating in-sample over-fitting and introduce some models that can accommodate "short fat" data problems. We also introduce concepts of feature importance and the application of Shapley values for understanding the impact of a feature (or a group of features) to a forecasting model's predictions.

2 Data Construction and Pre-Processing

We take the perspective of a manager or policymaker interested in forecasting the current-quarter or future-quarter state of the macroeconomy. Specifically, she uses information available as of the end of the second month in the quarter (See Appendix A for an illustration of the timing of the forecast and feature measurement). At this point, she observes an advanced estimate of GDP growth for the prior quarter from the BEA, available at the end of the first month of the quarter, and consensus estimates of GDP growth for the current and future quarters from professional forecasters, available in the middle of the second month of the quarter. She wants to forecasts BEA's "final" estimates of nominal GDP growth for the current and four future quarters. These "final" estimates are available three months after the close of the relevant quarter.[2] We focus on this forecasting problem over the 1990 to 2019 period, similar to Konchitchki and Patatoukas (2014a), during which macroeconomic growth has been more relatively stationary. Moreover, we can ascertain when accounting information is made available for the cross-section of firms in this period. Together, our time series contains 120 quarters of data.

We collect consensus forecasts of nominal GDP growth from the Survey of Professional Forecasters (SPF), compiled by the Federal Reserve Bank of Philadelphia.[3] SPF is the longest and most well-respected survey of macroeconomic forecasters. The survey provides estimates of GDP growth (both nominal and real) for the current quarter as well as for the next four quarters. It reports the consensus of professional forecasters in two ways, based on the mean or the median of their forecasts. We include all variants of the consensus forecast in our baseline set of predictors.

We add to the forecaster's information set the reported financials of public companies. Conservatively, we consider the quarterly accounting information available as of the end of the first month of each quarter, based on Compustat's earnings announcement report date ($RDQQ$). We also add to the information set stock market returns in the one, three, six, twelve, and twenty-four months before the end of the first month of the quarter, using the CRSP value-weighted index returns.

To construct accounting-based predictors, we obtain data from the CRSP/Compustat-Quarterly Merged database with a CRSP share code ("shrcd") of 10 or 11, removing ADRs,

[2] Our focus on BEA's "final" or third estimate follows Konchitchki and Patatoukas (2014a). This estimate is more authoritative because it is based on more accurate source data. The advance or first estimate, on the other hand, is based on incomplete data and can contain substantial measurement error. Konchitchki and Patatoukas (2014a) suggests that the final estimate is preferred by professionals when evaluating the accuracy of the SPF consensus forecasts. The Philadelphia Fed compiles BEA's GDP growth estimates: `https://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files/noutput`.

[3] This dataset can be obtained in the Philadelphia Fed's research and data repository: `https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters`.

certificates, and closed-end funds. We then create a master set of accounting variables, summarized at each year-quarter level, that are candidates for our analysis.

We first convert all cumulative quarterly accounting data (e.g., cumulative cash flow from operations from the beginning of the fiscal year to the current quarter) into quarterly flow variables (e.g., cash flow from operations in the current quarter). We also generate additional variables following prior literature, such as profitability ratios (Konchitchki and Patatoukas, 2014b)—net operating profit after taxes, net operating assets, return on net operating assets, asset turnover, profit margin, operating margin before depreciation, depreciation intensity—and net investment (Gutiérrez and Philippon, 2017). Then, for each accounting variable (X_q below), we derive six measures of its growth: quarter-over-quarter growth (i.e., current quarter value minus last quarter value or Eq. (1)), year-over-year same-quarter growth (i.e., current quarter value minus value from four quarters ago or Eq. (2)), growth scaled by total assets (Eq. (3) and (4)), and growth scaled by total revenues (Eq. (5) and (6)).

$$
\begin{aligned}
dif\_X_q &= X_q - X_{q-1} && (1)\\
yoy\_X_q &= X_q - X_{q-4} && (2)\\
sc1\_dif\_X_q &= [X_q - X_{q-1}]/atq_q && (3)\\
sc1\_yoy\_X_q &= [X_q - X_{q-4}]/atq_q && (4)\\
sc2\_dif\_X_q &= [X_q - X_{q-1}]/saleq_q && (5)\\
sc2\_yoy\_X_q &= [X_q - X_{q-4}]/saleq_q && (6)
\end{aligned}
$$

Finally, we summarize the accounting information available to the forecaster each quarter by taking its cross-sectional equal-weighted mean, market-value-weighted mean, median, and standard deviation. Together, the combination of the BEA advance estimate, SPF estimates, accounting information, and market returns yield a set of 69,993 predictors (model "features") and 120 quarters of observations.

Before training predictive models, we pre-process the data to hone in on the most informative features. We remove all features in which more than 15% of the data are missing or exhibit near-zero variance, then winsorize the features at the top and bottom 2% of the distribution to mitigate the influence of outliers. Based on the remaining features, we create a correlation matrix and, in cases where two features exhibit a correlation greater than 90%, we only keep the first feature. These pre-processing steps result in a set of 341 features. These features can be classified into three different types: professional estimates (21 variables that include BEA's advanced estimate of prior quarter GDP growth and SPF consensus forecasts), market-based values (28 variables that include market returns from CRSP and cross-sectional summaries of stock-price-related variables constructed from Compustat), and accounting-based values (292 cross-sectional summaries of accounting-related variables constructed from Compustat).

3 Methodology

This section describes the problem facing the forecaster and explains why it is difficult to fully assess the usefulness of accounting data for predicting future GDP growth. We then explain the statistical models and machine learning techniques we apply to tackle these challenges.

3.1 The Forecaster's Problem and the Bias-Variance Trade-off

To better understand the problem facing the forecaster, suppose that the data comes from the following process:

$$y_t = g(\boldsymbol{x}_t) + \epsilon_t \tag{7}$$

for some unknown function g and ϵ_t is a white noise process with finite variance σ^2. The forecaster observes a sample of observations $\{y_t, x_t\}_{t=1}^T$ but not $g(.)$. She seeks to approximate $g(.)$ with $\hat{g}(.; S)$—a model obtained using a training sample $S \subset \{y_t, x_t\}_{t=1}^T$—to make predictions for y_τ (the "target") that are outside the training sample. Finally, she wishes to maximize her model's *out-of-sample* predictive accuracy.

Using the "test" mean-squared error (MSE) as the evaluative criterion, the average accuracy of an estimator trained on new (unseen) samples can be decompose into three terms (Hastie et al., 2009):[4]

$$
\begin{aligned}
\mathbb{E}_S[(y_t - \hat{g}(\boldsymbol{x}_t;S))^2 \,|\boldsymbol{x}_t = \mathbf{x}] &= \mathbb{E}_S[(g(\mathbf{x}) + \epsilon_t - \hat{g}(\mathbf{x};S))^2] \\
&= \mathbb{E}_S[(g(\mathbf{x}) + \epsilon_t - \hat{g}(\mathbf{x};S) \pm \mathbb{E}_S[\hat{g}(\mathbf{x};S)])^2] \\
&= \sigma^2 + (\mathbb{E}_S[\hat{g}(\mathbf{x};S)] - g(\mathbf{x}))^2 \\
&\quad + \mathbb{E}_S\left(\hat{g}(\mathbf{x};S) - \mathbb{E}_S[\hat{g}(\mathbf{x};S)]\right)^2. \qquad (8)
\end{aligned}
$$

The first term of Eq. (8), σ^2, is the "irreducible error" in this forecasting problem. It is the variance of the target (y_t) around its true mean ($g(\mathbf{x})$), which cannot be avoided no matter how well the forecaster approximates g. The second term is the squared bias of the model predictions, or the amount by which the forecaster's model deviates from the target's true mean. The third term is the variance of the model predictions, or the variability of the model's prediction around its true mean. This terms expresses how much the model's prediction would change if it were trained using a different training sample.

The latter two terms are reducible errors. Ideally, the forecaster would like to obtain a model that generalizes to other samples or whose predictions would not vary too much based on different training samples (i.e., low variance). An estimator with high variance would yield very different predictions from small changes in the training sample. The forecaster would also like a model that produces predictions that approximate the true mean well (i.e., low bias). An estimator with high bias produces predictions that are very different from the actual value.

[4]In concept, the test MSE captures how accurate the statistical method's predictions would be when using unseen samples. For simplicity, here we evaluate the test MSE at a point $\mathbf{x}$. Expectations are taken over different samples, expressed as $\mathbb{E}_S$, so that $\mathbb{E}_S[(y_t - \hat{g}(\boldsymbol{x}_t;S))^2 \,|\boldsymbol{x}_t = \mathbf{x}]$ refers to the average MSE the forecaster obtains from repeatedly estimating g using a large number of training samples and testing each at $\mathbf{x}$. Note that, conditional on $\mathbf{x}$, y_t is deterministic and is unaffected by the sample S.

However, there is generally a trade-off between bias and variance (James, 2013). On one extreme, fitting the data with a highly flexible model (e.g., with a highly non-linear $\hat{g}$ or many predictors) can lead to lower bias and better predictions of each observation. However, the resultant model will be more sensitive to changes in the training sample (high variance) and thus less generalizable. On the other extreme, an overly simplistic model that makes a single prediction regardless of the data yields a high degree of bias but low variance since its predictions are not sensitive to differences in the training sample. Put differently, models that fit a sample's observation "too" well (low bias) would not generalize well to unseen data (high variance). The forecaster must find a balance between bias and variance in selecting a statistical model and a method for training its parameters.

3.2 Challenges with Forecasting GDP Growth Using Accounting Data

Characterizing the usefulness of accounting information for forecasting GDP growth faces two significant challenges. One is the general problem of in-sample over-fitting, which leads to high variance predictions in new data. Another problem is the high dimensionality of accounting data: the accrual accounting system produces much more information than observations of the economy. Certain classes of estimators, such as OLS, cannot accommodate these "short fat data" (or "$p >> N$") structures. Among the models that can, the inclusion of many predictors could help reduce the model's bias but at the expense of its variance.

These twin problems imply that, while the prior literature has provided suggestive evidence, whether and the extent to which accounting information can improve the accuracy of quarterly GDP growth forecasts remains an open question. On the one hand, the evidence from prior research is based on in-sample linear regression fits (i.e., from minimizing training MSE) of GDP growth on accounting information, thus could overstate accounting information's general predictive usefulness. On the other hand, prior research examines tiny subsets of the accounting system's outputs, thus could understate accounting information's importance.

We build on the prior work by applying machine learning techniques and principles designed to mitigate over-fitting, generalize accounting information's predictive usefulness, and accommodate accounting information's high dimensionality. Below, we describe a powerful and flexible model that preserves OLS's linearity and interpretability and can accommodate high dimensional data and mitigate over-fitting. We also describe how training and validation sample splits are done in the time-series forecasting context to calibrate these models and estimate their test MSEs.

3.3 Elastic Net Regression

One method we apply that can deal with the "short fat" data problem and mitigate over-fitting is the elastic net regression model (Zou and Hastie, 2005). Like OLS, the elastic net regression fits a linear model (i.e., $\hat{g}(\boldsymbol{x}_t) = \boldsymbol{x}_t'\boldsymbol{\beta}$). Unlike OLS, which obtains the slope coefficients by minimizing the sum of squared residuals, the elastic net regression does so by minimizing the sum of the squared residuals plus a penalty function:

$$\beta_{enet} = argmin_{\beta} \frac{1}{2T} \left[\sum_{t=1}^{T} (y_t - \beta_0 - \sum_{k=1}^{p} x_{tk}\beta_k) + \lambda_1 \sum_{k=1}^{p} |\beta_k| + \lambda_2 \sum_{k=1}^{p} (\beta_k)^2 \right], \tag{9}$$

or equivalently

$$\beta_{enet} = argmin_{\beta} \frac{1}{2T} \left[\sum_{t=1}^{T} (y_t - \beta_0 - \sum_{k=1}^{p} x_{tk}\beta_k) + \lambda \left(\alpha \sum_{k=1}^{p} |\beta_k| + (1-\alpha) \sum_{k=1}^{p} (\beta_k)^2 \right) \right] \tag{10}$$

for $\alpha = \lambda_1/(\lambda_1 + \lambda_2)$, $\lambda = \lambda_1 + \lambda_2$, and $\lambda_1, \lambda_2 \geq 0$. In this set up, all predictor variables are standardized and the target is normalized to have a mean of zero.

The elastic-net regression belongs to the class of penalized linear regression models, which were developed in response to the poor predictive properties of OLS (e.g., Hoerl and Kennard, 1980). By imposing penalties on the coefficients' magnitudes, these models tend to shrink the coefficient estimates towards zero (e.g., compared to the OLS estimate). In effect, such "regularization" techniques desensitize the predictions of the model to changes in the input, thereby mitigating in-sample over-fitting. At the limit, when λ is large, optimal slope coefficients approach 0, and the model predictions do not vary by predictor values. Thus, the penalty term adds bias to the model's predictions to reduce its variance in unseen samples.

The elastic-net penalty function combines the ridge regression penalty and the lasso penalty: two popularly used regularization methods. The ridge regression (Hoerl and Kennard, 1980) imposes an $L2$ penalty (or squared penalty) on the slope coefficients' magnitudes to the sum of squared residuals from a linear model. In contrast, the lasso regression (Tibshirani, 1996) imposes an $L1$ penalty (or the absolute value penalty).

$$\beta_{ridge} = argmin_{\beta} \frac{1}{2T} \left[\sum_{t=1}^{T} (y_t - \beta_0 - \sum_{k=1}^{p} x_{tk}\beta_k) + \lambda_2 \sum_{k=1}^{p} (\beta_k)^2 \right] \tag{11}$$

$$\beta_{lasso} = argmin_{\beta} \frac{1}{2T} \left[\sum_{t=1}^{T} (y_t - \beta_0 - \sum_{k=1}^{p} x_{tk}\beta_k) + \lambda_1 \sum_{k=1}^{p} |\beta_k| \right] \tag{12}$$

Equivalently, the ridge and lasso solve constrained versions of the OLS optimization problem. Whereas OLS chooses the linear slope coefficients that minimize the sum or squared residuals, ridge does so subject to the constraint that $\sum_{k=1}^{p}(\beta_k)^2 \leq c$ for some constant c, and lasso does so subject to the constraint that $\sum_{k=1}^{p}|\beta_k| \leq l$ for some constant l. Equations (11) and (12) are the Lagrangian forms of the constrained optimization problems.

In a ridge regression, increasing the weight on the penalty term (λ_2 of Eq. (11)) shrinks the magnitude of the optimal slope coefficient estimate *toward* zero (relative to the OLS estimate). Ridge has the advantage of being able to estimate the linear coefficients even when the number of coefficients is greater than the number of observations. However, with a large number of features, a ridge regression model will be complex: although the ridge penalty shrinks the weights on less important variables *towards* 0, it always keeps *all* the variables (Tibshirani, 1996).

In a lasso regression, increasing the weight on the penalty term (λ_1 of Eq. (12)) shrinks the slope all the way *to* 0, yielding more parsimonious models with fewer features. Thus an advantage of the lasso is its ability to perform regularization and feature selection simultaneously. However, when the number of predictors (p) is larger than the number of observations (N), it selects at most N features, shrinking the coefficients on the remaining features to zero. Intuitively, ridge regressions would be more appropriate if most of the variables are useful for forecasting. On the other hand, lasso would be better if the model contains many useless variables for forecasting, leading to a more parsimonious and interpretable model.[5]

By combining the ridge and the lasso penalty functions, the elastic-net regression can offer the ridge and lasso regressions' respective advantages while mitigating their respective weaknesses. In particular, the elastic-net is a flexible penalized linear regression model that can accommodate a "short fat" dataset structure while performing regularization and feature selection. To see how the elastic-net regression relates to the ridge and the lasso, note that the parameter α in Eq. (10) determines the mix of the penalties: the elastic-net regression is equivalent to the ridge regression when $\alpha = 0$ and the lasso regression when $\alpha = 1$, and it is a combination of the two when $\alpha \in (0, 1)$. The parameter λ in Eq. (10) determines how much weight to apply to the penalty and thus how much shrinkage in the linear coefficients there will be. The optimal values of α and λ, the elastic net model's hyperparameters, are chosen using the "validation" portion of the sample, described in the following section.

3.4 Training, Validation, and Testing

To train the elastic net model (i.e., choose the appropriate values of λ and α and the slope coefficients) and assess its predictive accuracy, we divide our 120 quarters of data into training, validation, and testing sets as follows. This approach is standard in machine learning for developing models and testing their predictive performance Hastie et al. (2009). The first 99 quarters (about 80%) of our data are used for training and validation: the beginning 79

[5]Principal components regressions is another linear regression method for dealing with the high-dimensionality problem. However, PCR is very similar to the ridge regression in that both utilize the principal components of the input matrix. In particular, ridge regressions shrink the coefficients related to the principal components with a small variance Hastie et al. (2009), which provide less information in the estimation process.

quarters (about 67%) for training and the remaining 20 quarters (about 20%) for validation. The last 21 quarters (about 18%) of our data are used for testing.[6]

The training set is used to learn the model parameters (i.e., the linear model's slope coefficients); the validation set, or the development sample, is used to select the hyperparameters that determine the elastic net's learning process (i.e., the λ and α). Together, the training and validation sets are used to develop the final forecasting model. Specifically, we consider multiple values of λ and α. For each (λ, α) combination, we estimate the linear coefficients following Eq. (10) using the training sample, and evaluate its performance by making forecasts on the validation set and computing the prediction MSE. We choose (λ^*, α^*) that minimizes the prediction MSE on the validation sample. Put differently, we select the model variant that seems to "generalize" the best from the training set to the validation set. It is possible to overfit the model to the validation set in this model tuning process. To mitigate this possibility, we perform a coarse search over the space of possible (λ, α): specifically, we consider $\lambda \in [0.01, 0.03, 0.1, 0.3, 1, 3, 10]$ and $\alpha \in [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1]$. In the last step of model development, we re-train the model, using the full training and validation samples and setting (λ, α) to (λ^*, α^*), to obtain final slope coefficients.

The test set, or holdout sample, is used to provide an estimate of the error rate of the final model on unseen data (i.e., the test MSE). This sample is "new" in the sense that it is not used during model development. Moreover, it preserves the temporal ordering of the data. This "out-of-sample" approach for dividing the sample into training, validation, and testing sets is specific to, and standard in, time-series forecasting contexts in which time dependency among the observations may exist. In other forecasting contexts where the data can be assumed to be independently and identically distributed or where there is no time-series dependence in the data, approaches that split the sample "in time," such as k-fold cross-validation that randomly partitions the sample upfront into partitions (or

[6] For current quarter GDP growth forecasts, the testing sample contains 21 observations of the target. For Q+1, Q+2, Q+3, and Q+4 quarter GDP growth forecasts, the testing sample contains 20, 19, 18, and 17 observations of the targets.

"folds"), are commonly used for model development. Cerqueira et al. (2017) shows that in real-world settings, the out-of-sample approach provides more accurate test error estimates of time-series forecasting models.

It is also worth noting that using a holdout sample to estimate the final model's out-of-sample error rate is new in the accounting literature that examines the role of accounting information for macroeconomic forecasting. Prior work in this literature uses the full sample for training an OLS and does not engage in testing. The best measure of these models' predictive performance is their in-sample MSEs, which likely overstate how well the models, and their use of accounting data, will perform in predicting new data. Thus, our use of a holdout sample allows for a more realistic view of the forecasting model's test error rates and a better reflection of accounting information's predictive usefulness.

4 Model Performance: Predictive Accuracy

In examining the usefulness of accounting information for forecasting future GDP growth (e.g., next quarter GDP growth), our empirical analyses essentially ask the following questions. Which elastic model would the forecaster obtain if she did not have accounting data? Which elastic net model would the forecasters obtain if accounting information were made available to her? Finally, does the elastic net model trained with accounting data produce more accurate predictions in terms of its prediction errors on the holdout sample?

To answer these questions, we assume the forecaster's position and train elastic net models using four different feature sets. The most basic version of the model uses as features the most recent estimate of GDP growth for the most recent quarter $(Q-1)$ and all variants of consensus forecasts of GDP growth for Q, $Q+1$, $Q+2$, $Q+3$, and $Q+4$, including the mean and median of forecasts as well as forecasts of real and nominal GDP growth. The baseline models considers a total of 21 features, listed under "Feature Set 1" in Appendix B, and its predictions are denoted $\hat{g}_t[Survey + Estimates]$. The next variant augments Feature

Set 1 with 28 features, listed under "Feature Set 2" in Appendix B, relating to market values obtained from CRSP and Compustat: market-level stock returns over the prior 3, 6, 12, and 24 months, and cross-sectional distributional statistics on the quarter-close, quarter-high, and quarter-low stock price. The resultant predictions ($\hat{g}_t[Survey + Estimates + Market]$) are based on a model trained using 49 features. As an alternative, we augment Feature Set 1 with 292 features relating to accounting values, listed under "Feature Set 3" in Appendix B and computed from Compustat. The resultant predictions ($\hat{g}_t[Survey+Estimates+Accounting]$) are based on a model trained using 313 features. Finally, we train elastic net models using the full set of 341 features. Its predictions are denoted $\hat{g}_t[All\ Data]$.

For each set of features considered, we determine the best elastic net model—selecting values for λ, α and training the linear slope coefficients—following the procedure described in Section 3.4. We choose the (λ^*, α^*) that minimizes the prediction MSE on the validation sample. Using these hyper-parameter values, we re-train the linear model using the full training and validation samples. This final model is then used to make predictions on the holdout set. We perform this exercise and develop elastic net models for forecasting the GDP growth for the current quarter (Q) and the GDP growth in each of the four quarters ahead ($Q + 1$, $Q + 2$, $Q + 3$, $Q + 4$). In total, we train 20 elastic net models, four models for each of the five targets.

This section presents summary statistics on these elastic net model forecasts. It also reports the models' prediction accuracy (based on the MSE) in the holdout sample.

We note that our choice of baseline features (Feature Set 1 and 2), with which we train benchmark elastic net models and against which we compare the performance of elastic net models trained using accounting data, follows the prior accounting literature. In particular, Konchitchki and Patatoukas (2014a), Konchitchki and Patatoukas (2014b), and Nallareddy and Ogneva (2017) all use the SPF, estimates of prior quarter GDP growth, and market returns as control variables. Our analysis thus preserves the same structure of comparisons as the prior work, in that we consider the incrementality of accounting information relative to the SPF, prior quarter estimates, and market returns.[7]

However, our analysis is distinctive in two critical ways. First, for each category of features, we consider many more measures than the prior literature. For example, in considering the incrementality of accounting information relative to the professional forecasters' consensus, prior work includes one measure from the SPF as a control (e.g., the mean of the forecast for the quarter of interest). Our models include a much larger number of predictors in each feature category in order for our models to fully capture the totality of information that might be relevant for predicting future GDP growth. Second, our primary focus, and thus the basis for performance evaluation, is predictive accuracy. For this purpose, the statistical significance of a feature's slope coefficient—the primary focus of the prior literature—is unimportant. Instead, our empirical results focus on comparing models' test MSEs, or their prediction errors on the holdout sample. This is a standard metric for assessing the accuracy of a predictive model Hastie et al. (2009).

4.1 Summary Statistics

Table ?? summarizes the target values and the predictions of elastic net model variants on the training and validation sample (i.e., the 99 quarters used for the development of the model). We also report the final model type, based on the hyper-parameter values (α, λ) chosen. For the current and intermediate horizon forecasts (Q and $Q+1$), the inclusion of market and accounting variables results in elastic net models closer to the lasso (α closer to 1), which tend to use a smaller set of variables compared to the ridge. For longer-horizon forecasts ($Q+2$ to $Q+4$), the inclusion of market and accounting variables results in elastic net models closer to the ridge (α closer to 0), which tend to use a broader set of variables compared to the lasso. The differences in the models selected suggest that, relative to the

[7] We do not consider, for example, the forecasts from time-series forecasting models (e.g., ARIMA models). This choice mirrors the prior accounting literature studying the usefulness of accounting for explaining GDP growth rates, which also does not consider ARIMA model forecasts as controls. One reason is that prior work in economics (Ang et al., 2007) has found that the survey forecasts outperform ARIMA models, among others, in macroeconomic forecasting.

survey forecasts and BEA estimates, the extra features may be more informative for more distant GDP growth.

We note that in 6 of the 20 model variants, the final elastic net model is the sample mean. These models produce a single prediction and have zero standard deviation. Notably, three such models are longer-horizon forecasts that use only survey and BEA estimates, suggesting that these features are not useful for forecasting more distant GDP growth. The other three trivial models resulted from the use of market values in short- and intermediate-horizon forecasts, suggesting that, relative to accounting information, historical market returns and prices are of limited value in forecasting future GDP growth.

We also note that similar to OLS, the in-sample mean values of the elastic-net forecasts are equal to the target variable's mean. However, the models' predictions in the holdout sample, reported in Table 2, have varying means. In particular, those models with greater punishment weights λ also tend to have mean prediction values that differ more significantly from the mean of the realizations (g_t). On the other hand, these models' predictions tend to exhibit lower variability (with lower standard deviation or "SD").

4.2 Holdout MSE

We now turn to assess the predictive performance of these models. In principle, if accounting information is useful for forecasting a given target, we should observe that those models trained with accounting data produce more accurate out-of-sample predictions than those trained without accounting data. That is, the inclusion of accounting information in training elastic models should lower the resultant model's test error rate or the MSE of the model's predictions on the holdout sample.

Table ?? reports the performance of the four elastic net model variants for forecasting GDP growth in quarter Q (Panel A) and $Q+1$ (Panel B). In each panel, we report the "Holdout MSE" from the trained model's predictions on the testing data. Columns (1)-(4) report the MSEs for $\hat{g}_t[Survey+Estimates]$, $\hat{g}_t[Survey+Estimates+Market]$, $\hat{g}_t[Survey+$

$Estimates + Accounting]$, $\hat{g}_t[All\ Data]$. Column (5) reports the difference in MSE between $\hat{g}_t[Survey + Estimates + Accounting]$ and $\hat{g}_t[Survey + Estimates]$. Column (6) reports the difference in MSE between $\hat{g}_t[All\ Data]$ and $\hat{g}_t[Survey + Estimates]$. Finally, column (7) reports the difference in MSE between $\hat{g}_t[All\ Data]$ and $\hat{g}_t[Survey + Estimates + Market]$.

Table ?? shows little evidence that the inclusion of accounting information lowers the trained model's test error rates in near-horizon GDP growth forecasts. In both Panels A and B, the elastic net trained using only survey data and BEA estimates, $\hat{g}_t[Survey + Estimates]$, produces the lowest Holdout MSE.

Table ?? reports the performance of the four elastic net model variants for forecasting GDP growth in quarter $Q + 2$ (Panel A), $Q + 3$ (Panel B), and $Q + 4$ (Panel C). Unlike the short-horizon forecast results, we find evidence that the use of accounting data improves the trained model's test error rates for more distant GDP growth forecasts. In each case, $\hat{g}_t[All\ Data]$ produces the lowest MSE.

For the forecasting of $Q + 2$ GDP growth, the Holdout MSE of $\hat{g}_t[All\ Data]$ is lower than that of $\hat{g}_t[Survey + Estimates]$ and $\hat{g}_t[Survey + Estimates + Market]$ by about 2%. For the forecasting of $Q + 3$ GDP growth, the Holdout MSE of $\hat{g}_t[All\ Data]$ is lower than that of $\hat{g}_t[Survey + Estimates]$ and $\hat{g}_t[Survey + Estimates + Market]$ by about 13%. In both of these forecasting exercises, $\hat{g}_t[Survey + Estimates]$ and $\hat{g}_t[Survey + Estimates + Market]$ are based on the trivial model (see Panel D of Table ??), thus the inclusion of accounting data produces $Q + 2$ and $Q + 3$ GDP growth estimates that are more accurate than the in-sample mean of $Q + 2$ and $Q + 3$ GDP growth respectively.

For the forecasting of $Q + 4$ GDP growth, we again find the Holdout MSE of $\hat{g}_t[All\ Data]$ to be the lowest among the four elastic net variants. It is lower than that of $\hat{g}_t[Survey + Estimates]$ by 19% and that of $\hat{g}_t[Survey + Estimates + Market]$ (based on the trivial model) by 46%. Thus, accounting information's usefulness seems to magnify when we forecast GDP growth rates for quarters that are farther away.

5 Feature Importance

Next, we analyze the importance of accounting features for the trained elastic net models. In particular, we focus on the "full" elastic net model—$\hat{g}_t[All\ Data]$—and ask the following question: how important are the features based on accounting information in the model's predictions? We consider two methods for evaluating the importance of accounting features in our trained elastic net models, based on the feature coefficients and their Shapley values.

5.1 Model Coefficients

One way to evaluate feature importance in linear models is to scrutinize the coefficients on the features. Since each elastic net model is a weighted sum of standardized variables, the magnitude of the weights on each variable (or the feature coefficient magnitude) can provide some sense of the "importance" of an individual feature in the model.

For ease of reporting and analysis, we report the sum of the absolute coefficient values by feature category: survey forecasts, market values, or accounting values. We further divide account features into 6 sub-categories: "Capital," "Investments," "Liability," "Profits," "Shares," and "Write-off."

"Capital" refers to those features that provide information about trends in the distribution or raising of capital. For example, the value-weighted average year-over-year growth preferred dividends, scaled by total assets ($vwmean\_yoy\_dvpq$) and the standard deviation of year-over-year growth in total shareholders' equity ($sd\_sc1\_yoy\_teqq$) both belong to the "Capital" category.

"Investments" refers to those features that provide information about what happened to long-term or capital investments. For example, the value-weighted average quarter-over-quarter growth in plant, property, and equipment, scaled by total assets ($vwmean\_sc1\_dif\_ppentq$) and the value-weighted average quarter-over-quarter growth in total non-controlling interests, scaled by total assets ($vwmean\_sc1\_dif\_mibtq$) both belong to the "Investments" category.

"Liability" refers to those features that provide information about what happened to total liabilities. For example, the equal-weighted average quarter-over-quarter growth in accounts payable, scaled by total assets ($mean\_sc1\_dif\_apq$) and the equal-weighted average quarter-over-quarter growth in other liabilities, scaled by total assets ($mean\_sc1\_dif\_loq$) both belong to the "Liability" category.

"Profits" refers to those features that provide information about what happened to earnings, their components, or rates of profitability. For example, the equal-weighted average quarter-over-quarter growth in net margins ($mean\_dif\_pm$) and the standard deviation of quarter-over-quarter growth in non-operating income scaled by current quarter assets ($sd\_sc1\_dif\_nopiq$) both belong to the "Profits" category.

"Shares" refers to those features that provide information about what happened to shares outstanding. For example, the equal-weighted average quarter-over-quarter growth in shares outstanding ($mean\_dif\_cshoq$) and the standard deviation in year-over-year growth in the share count used to compute EPS (sd_yoy_csh12q) both belong to the "Share" category.

"Write-off" refers to those features that provide information about what happened to short-term assets. For example, the value-weighted average year-over-year growth in provision for loan or asset losses, scaled by total assets ($vwmean\_sc1\_yoy\_pllq$), the value-weighted average year-over-year growth in reserve for loan losses, scaled by total assets ($vwmean\_sc1\_yoy\_rllq$), and the equal-weighted average year-over-year growth in extraordinary items and discontinued operations, scaled by total assets ($mean\_sc1\_yoy\_xidoq$) all belong to the "Write-off" category.

Table 5 reports, by each feature category, the sum of absolute feature coefficients and the proportion of total absolute weights each category's features account for in the model. The table also enumerates the number of features in each category used in the model. Panel A shows that the elastic net model selected five features for forecasting current quarter GDP growth. Four of these are survey forecast features, and their coefficients account for 93% of the total absolute weights across features. Only one accounting feature was selected in this model, $sd\_sc1_dif_nopiq$, whose coefficient accounts for 7% of the total. This model did not select any features relating to market values.

Panel B shows that, in forecasting $Q + 1$ GDP growth, the trained elastic net model selected 11 features. Two of these features relate to survey forecasts. For example, the feature with the largest coefficient magnitude is the median of professional forecasts for $Q + 1$ real GDP growth ($drgdp3\_median$), accounting for 31% of the total absolute weights across the 11 features. Together, the survey features account for 36% of the total. The elastic net also selected seven features relating to accounting values, accounting for 37% of the total absolute weights, and 2 features relating to market values, accounting for 27% of the total absolute weights. Among the accounting features, the ones relating to profits are the most important, accounting for 25% of the total absolute weights or 70% of the total absolute weights on accounting features.

Table 6 provides a similar analysis but for the elastic net models trained for forecasting GDP growth in more distant quarters: $Q + 2$ (Panel A), $Q + 3$ (Panel B), and $Q + 4$ (Panel C). Continuing the patterns of Table 5, we find that the importance of survey forecasts, as measured by their absolute weights, are declining as we forecast more-distant-quarter GDP growth. For example, the weights on survey forecasts account for 20% of the total absolute weights across features in the model that forecasts Q+2 GDP growth and 0% in the model that forecasts Q+4 GDP growth.

In contrast, the importance of accounting variables increases as we forecast more-distant-quarter GDP growth. For example, the weights on accounting features account for 65% of the total absolute weights across features in the model that forecasts Q+2 GDP growth and 91% in the model that forecasts Q+4 GDP growth.

Also, whereas the total absolute weights are highly concentrated in a small number of features in the Q and $Q + 1$ forecasts, coefficient magnitudes are more evenly distributed and less concentrated for more-distant GDP growth forecasts. This means that there is a small number of highly informative predictors for near-term GDP growth forecasts. For more-distant GDP growth forecasts, there is a larger number of weakly informative predictors.

It is also instructive to analyze the types of accounting features that are important for forecasting future GDP growth. Focusing on the models for forecasting $Q + 3$ and $Q + 4$ GDP growth (Panel B and C of Table 6), which show the most significant role for accounting information, we find that accounting features relating to profits, capital, write-offs, and investments are the most important. Thus, whereas prior literature has focused on accounting information relating to profitability (Konchitchki and Patatoukas, 2014a,b; Nallareddy and Ogneva, 2017), our findings suggest that accrual accounting estimates (such as provisions and write-offs), capital raises or distributions (such as dividends or equity issuances), and capital allocation decisions (such as investments and the growth in PP&E) are also informative of the longer-term outlook of the economy.

5.2 Shapley Value

As an alternative, we consider an alternative and state-of-the-art method for evaluating the importance of particular features in the predictions of a model, based on the game-theoretic concept of Shapley value (Shapley, 1953). In cooperative game theory, the Shapley value is a way to distribute the payout from the game among the players, who collaborate in a coalition to achieve a payout. A player's Shapley value is her weighted-average marginal contribution ($\nu(M \cup \{i\}) - \nu(M)$ below) to each possible coalition (M below) of players.

Specifically, in an f-player game, player i's Shapley value is defined as

$$\phi_i(\nu) = \frac{1}{|F|} \sum_{M \subseteq F \setminus \{i\}} \binom{|F| - 1}{|M|}^{-1} (\nu(M \cup \{i\}) - \nu(M)), \tag{13}$$

where $F = 1, 2, ..., f$ is the set of f players, M is a subset of the players, and $v(M)$ is a set function specifying the total payout from a game with a subset of players F. To compute Shapley values for each player, consider all possible permutations of each subset of players (for which there are $|F|!$ possibilities). For each permutation, add the players to the coalition in that order and compute the marginal contribution of each player i to the players that come before it (i.e., $\nu(M \cup \{i\}) - \nu(M)$), then average each player's marginal contributions across all permutations.

Consider the following example. Morris and Susan are two employees in a company. Together, they can produce a profit of \$100 ($v(Susan, Morris) = 100$). Morris can produce \$20 by himself ($v(Morris) = 20$). Susan can produce \$50 by herself ($v(Susan) = 50$). The company does not produce any profit without either of them ($v(\{\}) = 0$).

Permutation	Marginal Value for Susan	Marginal Value of Morris
(Susan, Morris)	\$50	\$50
(Morris, Susan)	\$80	\$20
Shapley Value	\$65	\$35

As the computations above show, Susan and Morris' average marginal contribution, and thus their Shapley values, are \$65 and \$35. Shapley (1953) shows that the Shapley value is the *unique* way for distributing the total payout that satisfies the following four axioms: The individual allocations add up to $v(N)$, so that all the grand coalition surplus is distributed ("efficiency"); two players that contribute the same to each coalition will receive the same allocation ("symmetry"); the allocation to a player for two games is the sum of her allocation in each game ("additivity"); and if a player contributes nothing to each coalition, he receives zero allocation ("dummy"). In this sense, the Shapley value is attractive in that it provides a "fair" attribution of the total payout to players. However, calculating these values can be computationally intensive: the computation is exponential in the number of players.

Recently, Shapley values have been applied to explain machine-learning model predictions (Štrumbelj and Kononenko, 2010; Datta et al., 2016; Lundberg and Lee, 2017). In this context, the "game" is the task of predicting a particular instance of the data (e.g., a specific quarter's GDP growth). The feature values for that instance are the "players" that cooperate to make the prediction. The "payout" is the difference between the model's prediction for a given instance and the average prediction across instances. The intuition behind the Shapley value for a feature in this game is similar to the above example. Consider all possible permutations of feature values, which participate in a game to contribute to a prediction. The Shapley value of a feature value is the average change in the prediction that the preceding features receive when the feature value is added to the model. Another way to think about the Shapley value in the prediction context is to imagine that the feature values enter a room in random order and that all feature values in the room participate in the game of making the prediction. The Shapley value of a feature value is the average change in the prediction that the coalition already in the room receives when the feature value joins them.

The estimation of Shapley values for feature attribution, particularly for models with many features, can be very computationally intensive, and substantial work has been devoted to approximation them using Monte Carlo sampling (see, e.g., Merrick and Taly, 2020, for a review). Intuitively, these Monte Carlo methods simulate the idea of feature values "entering a room in random order," using some reference distribution.

While the computational details of the Shapely values of features are beyond the scope of this paper, we emphasize their interpretation and properties and explain the advantages of using them for feature attribution. The Shapely value of a feature value is the average contribution (to the difference between the specific prediction for an instance of the data and the average prediction of the dataset) of a feature value across different coalitions of feature values. (Note that it is not the difference in the prediction when the feature is removed from the model).

Feature Shapley values exhibit four desirable properties. First, the difference between the prediction for an instance and the average prediction of the dataset is entirely attributed across features (i.e., the efficiency axiom). Second, if a feature value does not change the model's performance when it is added to the training data, it receives zero Shapley value (i.e., the dummy axiom). Third, two features are given the same Shapley value if they produce the same change in the model's predictions when they are individually added to the training data (the symmetry axiom). Finally, when the overall prediction score is the sum over all predictions, the value of a feature is the sum of its value for each prediction (the additivity axiom). The last property suggests that the Shapley value for a feature across a set of predictions (e.g., the training sample, which can be thought of as a collection of games) is the sum of its Shapley values for each prediction (individual game).

Because of these properties, one advantage of Shapley values is that it is a theoretically founded and unique way for attributing the predictions of a model to features that satisfy the efficiency, symmetry, dummy, additivity axioms. Shapley values are desirable in that they "fairly" attribute the difference between the predictions for an instance of the data and the average prediction. Another advantage is that Shapley values can be computed not only for individual features but also for groups of features (e.g., accounting-related features) to estimate the impact of groups of variables to model predictions.

Tables 7 and 8 report the Shapley values of the features in the "full" elastic-net models ($\hat{g}_t[\mathit{All\ Data}]$). Following the format of 5 and 6, we report the sum of absolute Shapley values by each feature category and the percent of the total Shapley values each category accounts for. We also report the results for sub-categories of accounting features.[8]

The results of our Shapley value analysis are similar to those of Tables 5 and 6. For current quarter GDP growth forecasts, survey forecasts are the most important features. As we forecast more distant-quarter GDP growth, survey forecasts decline in importance while accounting-based features increase. In the elastic net models forecasting $Q + 3$ and $Q + 4$ GDP growth, accounting values account for 89% and 90% of the total Shapley values. Moreover, in these models, where accounting features are most important, we again find that features relating to profits, investments, capital, and write-off are the most important.

For robustness, Table 9 reports Shapley values for the three groups of features: based on survey forecasts, market values, and accounting values. One way to think about the Shapley value for a feature group is to imagine that the feature groups enter a room in random order and that all feature groups in the room participate in the game of making the prediction. The Shapley value of a feature group is the average change in the prediction that the coalition already in the room receives when the feature group joins them.

[8]Our computations follow the sampling methods of Datta et al. (2016). We thank David Kurokawa, Anupam Datta, and the Truera team for their assistance with the implementation of these computations.

The findings from feature-group Shapley values are similar to those of individual-feature Shapley values. For the quarter-Q GDP growth forecast, survey forecast features account for 89% of the prediction based on Shapley values, while features based on accounting values only account for 11%. For more distant GDP growth forecasts, accounting features play an increasingly important role, contributing 40%, 52%, 79%, and 81% to the predictions of quarter-$Q+1$, $Q+2$, $Q+3$, and $Q+4$ GDP growth forecasts. Market-value-based features exhibit the greatest importance for the $Q+1$ forecasts (29%) but subsequently decline (to 18% for the $Q+4$ forecast).

6 Discussion and Conclusion

Our empirical analyses suggest that accounting information, to the extent that it could be incrementally useful for forecasting future GDP growth rates, is most useful for more distant periods, particularly for quarters that are one-year or more ahead. Unlike the prior literature in accounting, we do not find compelling evidence that accounting information improves the accuracy of immediate-term (i.e., current and next quarter) GDP growth forecasts.

Our findings are broadly consistent with the prior literature in economics that studies the usefulness of the SPF. For example, studies found that professional forecasters' consensus is efficient (Ang et al., 2007), particularly their immediate-horizon forecasts (Clements, 2015). This literature in economics has not examined the efficiency of the SPF relative to accounting information.

Our interpretation of the findings is that professional forecasters tend to have excellent information relating to more immediate horizon macroeconomic performance. Also, forecasters are likely to invest more time and effort on the immediate horizon forecasts, which garner the most attention from policymakers and investors. Thus, to the extent opportunities exist for improving upon professional forecasts, they are more likely to exist for forecasts relating to more distant periods.

It is worth noting that, although accounting statements reflect the results of historical business activity, accounting reports can provide signals relating to the future state of the economy. One reason is that certain decisions managers make, such as investments, shareholder payouts, or equity issuances, depend on managerial assessments about the economy's current and future states. For example, when managers perceive a high degree of economic uncertainty or expect future declines, they are more likely to delay investments, particularly those that are harder to reverse (e.g., Dixit and Pindyck, 1994). Another reason is that certain accrual accounting estimates reflect managers' forecasts about the state of their businesses and the expected future states of the economy. For example, a manager writes off an asset when she determines the asset to be no longer a "probable future economic benefit," or unlikely to yield future cash flows. A manager provisions for doubtful accounts or loan losses when she estimates that some of the receivables are unlikely to be realized in future cash. Thus, patterns in the growth in certain accounting attributes of public firms could embed information about future macroeconomic growth.

Our findings differ from the prior literature that analyzes the role of accounting information for macroeconomic forecasting, both in terms of the context in which accounting is likely to be more useful and the specific accounting measures that are most important. The differences in our findings stem from two primary sources. The first is our use of a holdout sample, a standard approach in machine learning for estimating the test error rates for our trained models. Our findings point to the possibility that the prior literature's findings, based on how well accounting information explains GDP growth in-sample, may be partly attributable to over-fitting. Another source of difference is the model used. Prior literature uses OLS, which is prone to be sensitive to extreme values that can accentuate the over-fitting problem, particularly in small samples. Also, OLS is ill-suited for short fat data problems, requiring researchers to pre-select a set of predictors substantially smaller in number than the total observations in the dataset. Our use of generalized penalized regression models has the advantages of preserving linear models' interpretability while accommodating short fat data structures, allowing the data to ascertain which of the many accounting-related features provide the most useful information for macroeconomic forecasting. Our finding that a variety of other accounting features outside of earnings or profitability are useful for longer horizon macroeconomic forecasting—such as accrual estimates (e.g., growth in provisions or write-offs), and capital raising, distribution, or allocation decisions (e.g., growth in equity, dividend payout, and investments)—is consistent with theory and intuition.

Our results are based on the elastic net model. However, in untabulated analyses, we obtain similar conclusions using tree-based methods, such as random forest and gradient boosted trees, that allow for non-linearities among the features in making forecasts. That is, $\hat{g}()$ of Eq. (8) is allowed to be highly non-linear in the inputs.

Our work provides an analytical structure and methodologies for evaluating the role of accounting information in macroeconomic forecasting. A caveat to our work, and the interpretation of our findings, is that although we have considered many features, they are by no means exhaustive. The approach we have taken, which uses survey forecasts from the SPF and market returns to build benchmark forecasts, is motivated by the prior literature in accounting Konchitchki and Patatoukas (2014a,b); Nallareddy and Ogneva (2017). This literature uses the SPF forecasts and market returns as control variables in linear regression analyses. However, future research can consider additional sources of non-accounting signals in evaluating the relative usefulness of accounting information for forecasting GDP growth. In this sense, a modest interpretation of this paper's results is that, to the degree accounting information is useful for forecasting future GDP growth, it is more likely to be so for longer-horizon forecasts.

Future research may also consider alternative ways for aggregating accounting performance in the cross-section, such as by industry, or examine a broader set of more sophisticated machine learning algorithms. These applications of big data and machine learning algorithms will be essential in deepening our understanding of the connection between accounting information and future macroeconomic performance.

References

Ang, A., G. Bekaert, and M. Wei (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics 54*(4), 1163–1212. 18, 28

Arif, S. and C. M. Lee (2014). Aggregate investment and investor sentiment. *Review of Financial Studies 27*(11), 3241–3279. 1, 3

Baumohl, B. (2012). *The secrets of economic indicators: Hidden clues to future economic trends and investment opportunities*. FT Press. 3

Cerqueira, V., L. Torgo, J. Smailović, and I. Mozetič (2017). A comparative study of performance estimation methods for time series forecasting. *2017 International Conference on Data Science and Advanced Analytics (DSAA)*, 529–538. 16

Clements, M. P. (2015). Are professional macroeconomic forecasters able to do better than forecasting trends? *Journal of Money, Credit and Banking 47*(2-3), 349–382. 28

Datta, A., S. Sen, and Y. Zick (2016). Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems. In *Proceedings of 37th IEEE Symposium on Security and Privacy*. 25, 27

Dixit, A. K. and R. S. Pindyck (1994). *Investment Under Uncertainty*. Princeton University Press. 29

Gutiérrez, G. and T. Philippon (2017). Investmentless Growth: An Empirical Investigation. *Brookings Papers on Economic Activity* (Fall), 89–169. 8

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Second ed.). Springer Science & Business Media. 2, 4, 10, 14, 18

Hoerl, A. E. and R. W. Kennard (1980). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics 12*(1), 69–82. 12, 13

James, G., W. D. H. T. T. R. (2013). *An Introduction to Statistical Learning with Applications in R* (1 ed.). New York: Springer. 11

Konchitchki, Y. and P. N. Patatoukas (2014a). Accounting earnings and gross domestic product. *Journal of Accounting and Economics 57*(1), 76–88. 1, 2, 3, 7, 17, 24, 30

Konchitchki, Y. and P. N. Patatoukas (2014b). Taking the pulse of the real economy using financial statement analysis: Implications for macro forecasting and stock valuation. *Accounting Review 89*(2), 669–694. 1, 2, 3, 8, 17, 24, 30

Landefeld, J. S., E. P. Seskin, and B. M. Fraumeni (2008). Taking the pulse of the economy: Measuring GDP. *Journal of Economic Perspectives 22*(2), 193–216. 3

Lundberg, S. M. and S. I. Lee (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems 2017-Decem*(Section 2), 4766–4775. 25

Merrick, L. and A. Taly (2020, 9). The Explanation Game: Explaining Machine Learning Models Using Shapley Values. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Volume 12279 LNCS, pp. 17–38. 26

Nallareddy, S. and M. Ogneva (2017). Predicting restatements in macroeconomic indicators using accounting information. *Accounting Review 92*(2), 151–182. 1, 3, 17, 24, 30

Shapley, L. S. (1953). A Value for n-Person Games. In A. W. Tucker and R. D. Luce (Eds.), *Contributions to the Theory of Games II*, Volume II, pp. 307–317. Princeton University Press. 24, 25

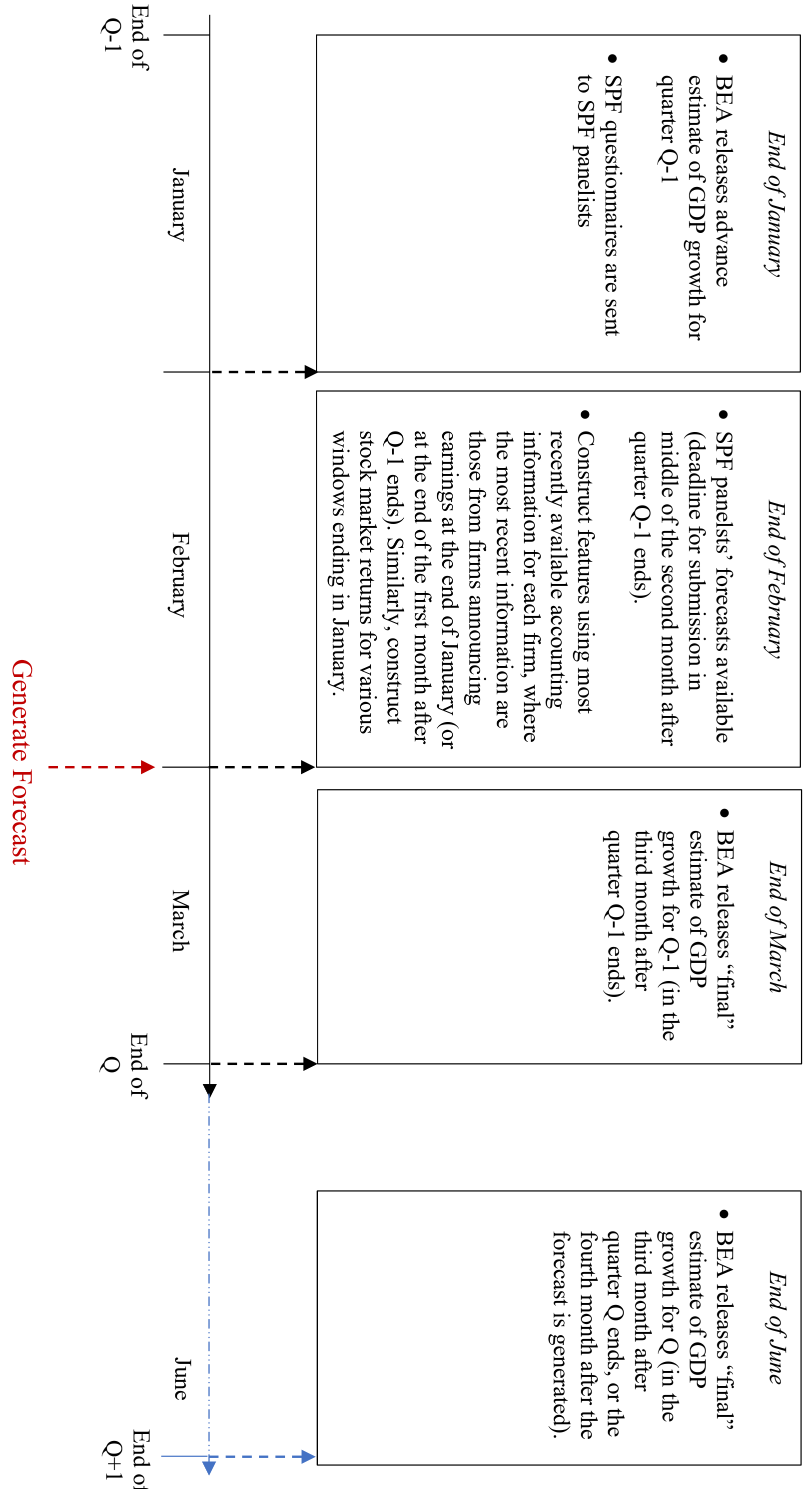
Štrumbelj, E. and I. Kononenko (2010). An efficient explanation of individual classifications using game theory. *Journal of Machine Learning Research 11*, 1–18. 25

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society 58*(1), 267–288. 13

Zou, H. and T. Hastie (2005). Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society 67*(2), 301–320. 2, 12

A Timing of Forecast and Data Availability

This figure illustrates the timing of our forecast, the timing of the measurement of features, and the timing of the target value realizations.



B Description of Features

This table enumerates the features used to train elastic net models in our analysis. The first set of features ("Feature Set 1") are obtained from the Survey of Professional Forecasters (SPF) and the Bureau of Economic Analysis (BEA). We consider both the mean and the median of professional forecasts as measures of consensus. The second set of features ("Feature Set 2") are obtained from CRSP and Compustat. We obtain monthly, prior 3-month, 6-month, 12-month, and 24-month CRSP value-weighted market returns. We also compute market capitalization from CRSP and firm-level price information from Compustat. For these variables, we compute the year-over-year ("yoy") and quarter-over-quarter ("dif") growth, and summarize their cross-sectional equal-weighted average ("mean"), value-weighted average ("vw"), and standard deviation ("sd"). The third set of features ("Feature Set 3") are accounting variables obtained from Compustat. We compute the same variants of these variables as the non-market-return features in Feature Set 2. For each growth variable variant, we also consider a version that is scaled by total assets and a version that is scaled by total revenues.

Variable	Description	Variants
Feature Set 1. Survey+Estimates		
dngdp2	Survey of Nominal GDP Growth of current quarter	median, mean
dngdp3	Survey of Nominal GDP Growth of next quarter	median, mean
dngdp4	Survey of Nominal GDP Growth of 2nd quarter forward	median, mean
dngdp5	Survey of Nominal GDP Growth of 3rd quarter forward	median, mean
dngdp6	Survey of Nominal GDP Growth of 4th quarter forward	median, mean
drgdp2	Survey of Real GDP Growth of current quarter	median, mean
drgdp3	Survey of Real GDP Growth of next quarter	median, mean
drgdp4	Survey of Real GDP Growth of 2nd quarter forward	median, mean
drgdp5	Survey of Real GDP Growth of 3rd quarter forward	median, mean
drgdp6	Survey of Real GDP Growth of 4th quarter forward	median, mean
lag_first	BEA's First Estimate of Last Quarter's Nominal GDP Growth	
Feature Set 2. Market Values		
crspmkt	Stock Market Return	*ret*, *ret*3*m*, *ret*6*m*, *ret*12*m*, *ret*24*m*
mcap_fqend_crsp	Calender Quarter End Market Capitalization, $\lvert prc \rvert \times shrout$	vwmean_dif, sd_yoy, mean_dif, mean_yoy,vwmean_yoy, sd_dif
PRCCQ	Price Close - Quarter	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
PRCHQ	Price High - Quarter	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
PRCLQ	Price Low - Quarter	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
Feature Set 3. Accounting Values		
APQ	Account Payable	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
CAPSQ	Capital Surplus/Share Premium Reserve	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
CHEQ	Cash and Short-Term Investments	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
CSH12Q	Common Shares Used to Calculate Earnings Per Share (12 Months Moving Average)	mean_dif, sd_yoy, sd_dif, mean_yoy, vwmean_dif
CSHFDQ	Com Shares for Diluted EPS	sd_dif, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
CSHOQ	Common Shares Outstanding	sd_dif, vwmean_dif, mean_dif, mean_yoy, sd_yoy
CSHPRQ	Common Shares Used to Calculate Earnings Per Share - Basic	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
CSHTRQ	Common Shares Traded - Quarter	sd_dif, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
CSTKQ	Common/Ordinary Stock (Capital)	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
DEP	Depreciation Intensity dep = dpq / saleq	vwmean_yoy, mean_yoy
DPQ	Depreciation and Amortization - Total	vwmean_yoy, mean_yoy
DVPQ	Dividends - Preferred/Preference	sd_yoy, vwmean_yoy, mean_yoy

Appendix B Continued.

Compustat Variable	Description	Variants
DVPSXQ	Div per Share - Exdate - Quarter	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
EPSF12	Earnings Per Share (Diluted) - Excluding Extraordinary Items - 12 Months Mo	sd_dif, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
EPSFIQ	Earnings Per Share (Diluted) - Including Extraordinary Items	sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
EPSFXQ	Earnings Per Share (Diluted) - Excluding Extraordinary items	sd_dif, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
EPSPIQ	Earnings Per Share (Basic) - Including Extraordinary Items	sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
EPSPXQ	Earnings Per Share (Basic) - Excluding Extraordinary Items	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
EPSX12	Earnings Per Share (Basic) - Excluding Extraordinary Items - 12 Months Moving	sd_dif, mean_dif, vwmean_yoy, sd_yoy, vwmean_dif, mean_yoy
IBQ	Income Before Extraordinary Items	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
ICAPTQ	Invested Capital - Total - Quarterly	sd_dif, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
LOQ	Liabilities - Other	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
LSEQ	Liabilities and Stockholders Equity - Total	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
LTQ	Liabilities - Total	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
MIBQ	Noncontrolling Interest - Redeemable - Balance Sheet	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
MIBTQ	Noncontrolling Interests - Total - Balance Sheet	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
MIIQ	Noncontrolling Interest - Income Account	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
NCOQ	Net Charge-Offs	vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
ni_gr_scaled	Year over Year Earnings growth scaled by beginning period shareholders' equity	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
NIQ	Net Income (Loss)	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
NOPATQ	operating income after depreciation * (1-0.35)	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
NOPIQ	Non-Operating Income (Expense) - Total	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
OIBDPQ	Operating Income Before Depreciation - Quarterly	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
PIQ	Pretax Income	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
PLLQ	Provision for Loan/Asset Losses	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
PM	Operating Margin	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
PPENTQ	Property Plant and Equipment - Total (Net)	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
PSTKQ	Preferred/Preference Stock (Capital) - Total	sd_yoy, vwmean_yoy, mean_yoy
qCSHPRY	Common Shares Used to Calculate Earnings Per Share - Basic	sd_yoy, vwmean_yoy, mean_yoy
qEPSFIY	Earnings Per Share (Diluted) - Including Extraordinary Items	vwmean_yoy, mean_yoy
qEPSPXY	Earnings Per Share (Basic) - Excluding Extraordinary Items	vwmean_yoy, mean_yoy
qIBCOMY	Income Before Extraordinary Items - Available for Common	vwmean_yoy, mean_yoy
qIBY	Income Before Extraordinary Items	vwmean_yoy, mean_yoy
qNIY	Net Income (Loss)	vwmean_yoy, mean_yoy
qNOPIY	Non-Operating Income (Expense) - Total	mean_yoy, vwmean_yoy
qOIADPY	Operating Income After Depreciation - Year-to-Date	sd_yoy, vwmean_yoy, mean_yoy
qSALEY	Sales/Turnover (Net)	vwmean_yoy, mean_yoy

Appendix B Continued.

Compustat Variable	Description	Variants
RECTQ	Receivables - Total	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
REQ	Retained Earnings	sd_dif, vwmean_yoy, mean_yoy, sd_yoy, vwmean_dif, mean_dif
RLLQ	Reserve for Loan/Asset Losses	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
SALEQ	Sales/Turnover	mean_yoy, sd_dif, mean_dif, vwmean_yoy, vwmean_dif, sd_yoy
SEQQ	Stockholders Equity	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
TEQQ	Stockholders Equity	vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
TIEQ	Interest Expense (Financial Services)	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
TXTQ	Income Taxes	vwmean_yoy, sd_yoy, sd_dif, mean_dif, vwmean_dif, mean_yoy
XIDOQ	Extraordinary Items and Discontinued Operations	sd_yoy, vwmean_yoy, sd_dif, mean_dif, vwmean_dif, mean_yoy
XINTQ	Interest and Related Expense	vwmean_yoy, mean_yoy
XOPRQ	Operating Expense	sd_dif, sd_yoy, vwmean_dif, mean_dif, vwmean_yoy, mean_yoy
XSGAQ	Selling, General and Administrative Expenses	mean_dif, vwmean_dif, vwmean_yoy, sd_yoy, mean_yoy, sd_dif

Figure 1. Training-Validation-Testing Splits

This figure illustrates how our sample is split into training, validation, and testing sets for the forecasting of current-quarter GDP growth. Our data contains 120 quarters of information on features that can be used for our forecasting exercise, spanning 1990Q1 to 2019Q4.The first 99 quarters (about 80%) of our data are used for training and validation: the beginning 79 quarters (about 67%) for training and the remaining 20 quarters (about 20%) for validation. The last 21 quarters (about 18%) of our data are used for testing. For Q+1, Q+2, Q+3, and Q+4 quarter GDP growth forecasts, the testing sample contains 20, 19, 18, and 17 observations of the targets.

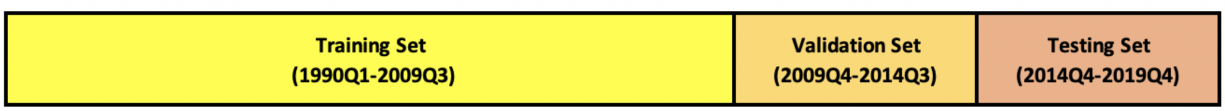


Figure 2. Model Tuning using the Validation Set

This figure illustrates how the hyperparameters of the elastic-net models (α, λ) are selected using the validation data. We consider multiple values of λ and α. For each (λ, α) combination, we estimate the linear coefficients following Eq. (10) using the training sample, and evaluate its performance by making forecasts on the validation set and computing the prediction MSE. The "optimal parameters" are those that minimizes the prediction MSE on the validation sample, or the maximum point in the below (Log(Lambda), Alpha, -MSE) surface. We plot the surface using Log(Lambda) and -MSE to ease visual discernment..

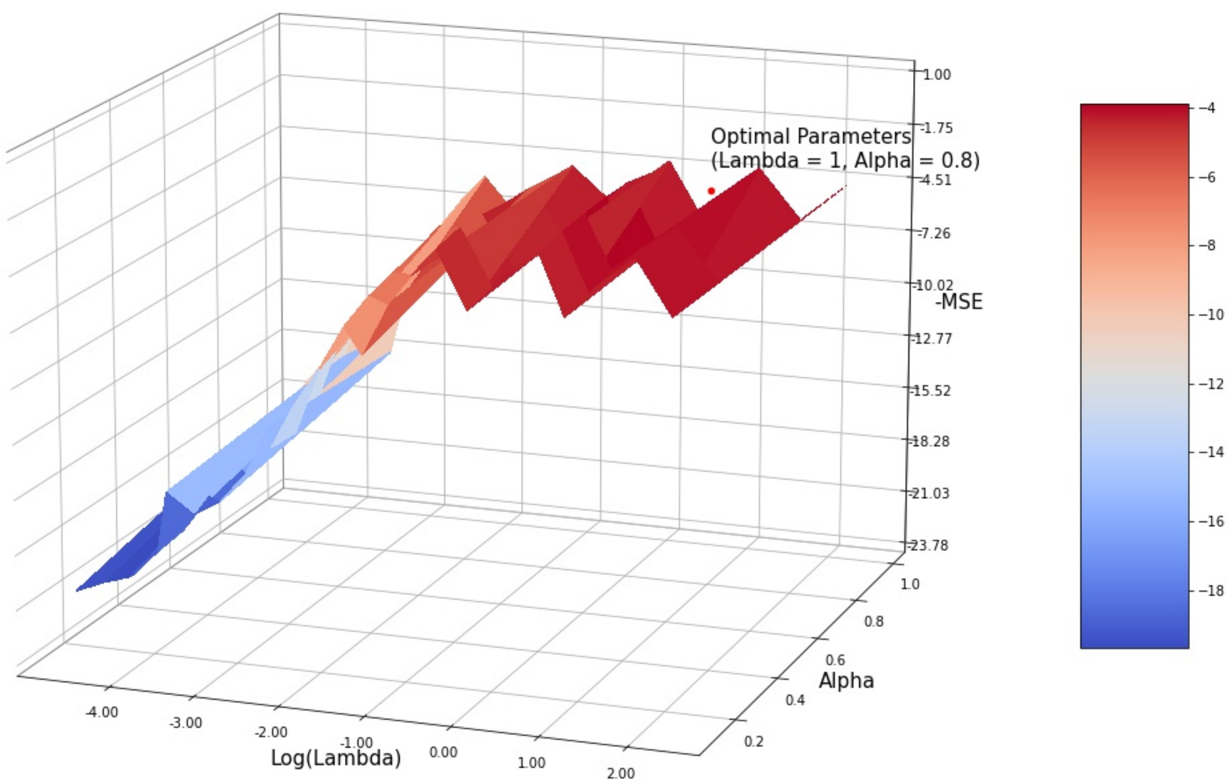


Table 1.
Summary Statistics: Model Prediction on Training and Validation Sample

This table reports descriptive statistics for the target (g_t) and the trained models' predictions ($\hat{g}_t$) on the development sample (training sample plus validation sample) for current quarter (Panel A), Q+1 (Panel B), Q+2 (Panel C), Q+3 (Panel D), and Q+4 (Panel E) nominal GDP growth. For each variable, we report its observation count (Obs), minimum (Min), 5^{th} percentile ($p5$), 25^{th} percentile ($p25$), average ($Mean$), median ($p50$), 75^{th} percentile ($p75$), 95^{th} percentile ($p95$), maximum (Max), and standard deviation (SD). The baseline models is trained using Feature Set 1, and its predictions are denoted $\hat{g}_t[Survey + Estimates]$. The next variant is trained using both Feature Set 1 and 2, and resultant predictions are denoted $\hat{g}_t[Survey + Estimates + Market]$. The final two models are trained with accounting data. $\hat{g}_t[Survey + Estimates + Accounting]$ refers to the predictions from a model trained with Feature Set 1 and 3, and $\hat{g}_t[All\ Data]$ refers to the predictions from a model trained with all features. We also report the hyper-parameters (α and λ) of each trained model. Feature sets are enumerated in Appendix B.

Panel A: Elastic Net

	α	λ	Obs	Min	$p5$	$p25$	$Mean$	$p50$	$p75$	$p95$	Max	SD
A. Current Quarter GDP Growth												
g_t			20	-1.712	-0.181	3.273	4.006	3.945	5.278	6.628	6.839	1.927
$\hat{g}_t[Survey{+}Estimates]$	0.30	0.30	20	3.640	3.642	3.972	4.425	4.244	4.818	5.747	5.804	0.643
$\hat{g}_t[Survey{+}Estimates{+}Market]$	0.60	1.00	20	4.074	4.077	4.238	4.573	4.472	4.813	5.393	5.444	0.416
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	0.80	1.00	20	4.156	4.174	4.301	4.568	4.527	4.746	5.158	5.223	0.317
$\hat{g}_t[All\ Data]$	0.80	1.00	20	4.156	4.174	4.301	4.568	4.527	4.746	5.158	5.223	0.317
B. Q+1 GDP Growth												
g_t			20	-1.712	-0.181	3.068	3.821	3.911	4.533	6.628	6.839	1.895
$\hat{g}_t[Survey{+}Estimates]$	0.10	3.00	20	3.982	3.986	4.364	4.506	4.409	4.682	5.122	5.125	0.311
$\hat{g}_t[Survey{+}Estimates{+}Market]$	0.50	3.00	20	4.589	4.589	4.589	4.589	4.589	4.589	4.589	4.589	0.000
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	0.90	1.00	20	4.328	4.347	4.478	4.539	4.526	4.594	4.736	4.755	0.109
$\hat{g}_t[All\ Data]$	1.00	0.30	20	3.540	3.598	3.954	4.174	4.221	4.393	4.806	4.988	0.350
C. Q+2 GDP Growth												
g_t			20	-1.712	-0.973	2.931	3.615	3.881	4.533	6.628	6.839	2.100
$\hat{g}_t[Survey{+}Estimates]$	0.40	3.00	20	4.535	4.535	4.535	4.535	4.535	4.535	4.535	4.535	0.000
$\hat{g}_t[Survey{+}Estimates{+}Market]$	0.40	3.00	20	4.535	4.535	4.535	4.535	4.535	4.535	4.535	4.535	0.000
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	0.10	3.00	20	3.683	3.774	3.986	4.193	4.155	4.334	4.797	4.872	0.293
$\hat{g}_t[All\ Data]$	0.10	3.00	20	3.626	3.736	4.037	4.217	4.162	4.345	4.914	5.001	0.323
D. Q+3 GDP Growth												
g_t			20	-1.712	-0.973	2.931	3.736	3.945	5.278	6.628	6.839	2.174
$\hat{g}_t[Survey{+}Estimates]$	0.80	1.00	20	4.543	4.543	4.543	4.543	4.543	4.543	4.543	4.543	0.000
$\hat{g}_t[Survey{+}Estimates{+}Market]$	0.90	1.00	20	4.543	4.543	4.543	4.543	4.543	4.543	4.543	4.543	0.000
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	0.50	1.00	20	4.200	4.208	4.316	4.391	4.391	4.463	4.587	4.618	0.113
$\hat{g}_t[All\ Data]$	0.10	3.00	20	3.878	3.910	4.064	4.193	4.142	4.322	4.551	4.648	0.190
E. Q+4 GDP Growth												
g_t			20	-1.712	-0.973	2.931	3.669	3.881	5.167	6.628	6.839	2.165
$\hat{g}_t[Survey{+}Estimates]$	0.70	1.00	20	4.567	4.567	4.567	4.567	4.567	4.567	4.567	4.567	0.000
$\hat{g}_t[Survey{+}Estimates{+}Market]$	0.30	0.10	20	2.604	2.786	3.406	3.863	3.899	4.295	4.757	4.924	0.588
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	0.20	1.00	20	1.248	1.952	3.634	3.874	3.964	4.453	4.866	4.930	0.822
$\hat{g}_t[All\ Data]$	0.20	1.00	20	1.260	1.936	3.758	3.898	3.974	4.439	4.862	4.868	0.838

Summary Statistics: Model Prediction on Training and Validation Sample

Panel B: Random Forest

	Max Depth	*Max Features*	*N Estimators*	*Obs*	*Min*	*p5*	*p25*	*Mean*	*p50*	*p75*	*p95*	*Max*	*SD*
A. Current Quarter GDP Growth													
g_t				20	-1.712	-0.181	3.273	4.006	3.945	5.278	6.628	6.839	1.927
$\hat{g}_t$[*Survey+Estimates*]	12	1	300	20	0.129	1.163	3.660	4.133	4.132	5.002	6.023	6.161	1.334
$\hat{g}_t$[*Survey+Estimates+Market*]	10	1	100	20	-0.074	1.211	3.577	4.031	4.054	4.714	5.931	5.974	1.297
$\hat{g}_t$[*Survey+Estimates+Accounting*]	6	1	300	20	0.520	1.624	3.871	4.092	4.273	4.605	5.375	5.487	1.039
$\hat{g}_t$[*All Data*]	4	2	300	20	2.166	2.812	4.021	4.194	4.294	4.443	5.007	5.126	0.609
B. Q+1 GDP Growth													
g_t				20	-1.712	-0.181	3.068	3.821	3.911	4.533	6.628	6.839	1.895
$\hat{g}_t$[*Survey+Estimates*]	6	1	100	20	1.488	2.023	3.807	4.092	3.996	4.969	5.320	5.428	0.939
$\hat{g}_t$[*Survey+Estimates+Market*]	10	1	100	20	0.537	1.432	3.447	3.887	3.763	4.695	5.650	5.797	1.158
$\hat{g}_t$[*Survey+Estimates+Accounting*]	4	2	100	20	0.253	1.124	3.225	3.897	3.936	4.696	5.711	6.107	1.287
$\hat{g}_t$[*All Data*]	6	3	500	20	0.263	1.021	3.143	3.855	3.938	4.637	5.807	6.134	1.312
C. Q+2 GDP Growth													
g_t				20	-1.712	-0.973	2.931	3.615	3.881	4.533	6.628	6.839	2.100
$\hat{g}_t$[*Survey+Estimates*]	2	1	100	20	3.691	3.914	4.299	4.443	4.478	4.622	4.910	4.962	0.287
$\hat{g}_t$[*Survey+Estimates+Market*]	8	1	100	20	0.634	0.905	3.100	3.682	3.835	4.217	5.836	5.891	1.304
$\hat{g}_t$[*Survey+Estimates+Accounting*]	8	2	200	20	0.379	0.751	2.894	3.657	3.872	4.577	5.587	5.854	1.369
$\hat{g}_t$[*All Data*]	8	2	200	20	0.511	0.870	3.053	3.657	3.787	4.588	5.585	5.695	1.301
D. Q+3 GDP Growth													
g_t				20	-1.712	-0.973	2.931	3.736	3.945	5.278	6.628	6.839	2.174
$\hat{g}_t$[*Survey+Estimates*]	12	1	300	20	0.539	1.013	3.249	3.851	3.952	4.994	5.464	5.490	1.329
$\hat{g}_t$[*Survey+Estimates+Market*]	4	2	100	20	0.567	1.081	2.846	3.663	3.887	4.479	5.393	5.440	1.323
$\hat{g}_t$[*Survey+Estimates+Accounting*]	4	3	200	20	1.233	1.758	3.809	3.955	4.314	4.471	4.871	5.058	0.934
$\hat{g}_t$[*All Data*]	4	2	300	20	0.902	1.218	3.474	3.818	3.975	4.637	5.430	5.431	1.220
E. Q+4 GDP Growth													
g_t				20	-1.712	-0.973	2.931	3.669	3.881	5.167	6.628	6.839	2.165
$\hat{g}_t$[*Survey+Estimates*]	12	1	200	20	0.693	0.866	3.369	3.726	3.850	4.625	5.271	5.316	1.260
$\hat{g}_t$[*Survey+Estimates+Market*]	2	1	500	20	3.616	3.665	3.963	4.065	4.094	4.135	4.465	4.499	0.218
$\hat{g}_t$[*Survey+Estimates+Accounting*]	10	1	200	20	3.445	3.546	3.928	4.090	4.178	4.300	4.463	4.534	0.288
$\hat{g}_t$[*All Data*]	8	1	200	20	0.561	0.779	3.275	3.762	3.963	4.600	5.476	5.592	1.359

Summary Statistics: Model Prediction on Training and Validation Sample

Panel C: Neural Network

	Hidden Layer Sizes	α	*Obs*	*Min*	*p5*	*p25*	*Mean*	*p50*	*p75*	*p95*	*Max*	*SD*
A. Current Quarter GDP Growth												
g_t			20	-1.712	-0.181	3.273	4.006	3.945	5.278	6.628	6.839	1.927
$\hat{g}_t[Survey+Estimates]$	3	0.1000	20	3.811	3.869	4.089	4.539	4.440	4.797	5.650	5.788	0.566
$\hat{g}_t[Survey+Estimates+Market]$	2	1.0000	20	3.544	3.585	3.878	4.320	4.289	4.598	5.383	5.506	0.547
$\hat{g}_t[Survey+Estimates+Accounting]$	4	0.0001	20	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.000
$\hat{g}_t[All\ Data]$	2	0.0010	20	2.036	2.506	3.749	4.268	4.281	5.135	5.774	5.901	0.999
B. Q+1 GDP Growth												
g_t			20	-1.712	-0.181	3.068	3.821	3.911	4.533	6.628	6.839	1.895
$\hat{g}_t[Survey+Estimates]$	5	0.0001	20	3.861	3.899	4.255	4.452	4.333	4.585	5.239	5.239	0.395
$\hat{g}_t[Survey+Estimates+Market]$	3	1.0000	20	2.880	3.072	3.527	3.972	3.926	4.279	5.095	5.340	0.590
$\hat{g}_t[Survey+Estimates+Accounting]$	5	0.0010	20	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.000
$\hat{g}_t[All\ Data]$	3	1.0000	20	0.806	1.081	2.703	3.215	3.324	3.843	4.931	5.017	1.015
C. Q+2 GDP Growth												
g_t			20	-1.712	-0.973	2.931	3.615	3.881	4.533	6.628	6.839	2.100
$\hat{g}_t[Survey+Estimates]$	5	0.1000	20	3.416	3.613	4.097	4.404	4.319	4.731	5.355	5.387	0.523
$\hat{g}_t[Survey+Estimates+Market]$	3	1.0000	20	2.620	2.891	4.034	4.453	4.643	4.880	5.558	5.899	0.771
$\hat{g}_t[Survey+Estimates+Accounting]$	5	0.0010	20	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.000
$\hat{g}_t[All\ Data]$	3	1.0000	20	0.803	1.304	3.134	4.074	3.746	4.875	7.827	8.033	1.735
D. Q+3 GDP Growth												
g_t			20	-1.712	-0.973	2.931	3.736	3.945	5.278	6.628	6.839	2.174
$\hat{g}_t[Survey+Estimates]$	3	0.0100	20	3.756	3.781	4.046	4.196	4.117	4.292	4.779	4.818	0.286
$\hat{g}_t[Survey+Estimates+Market]$	3	1.0000	20	3.536	3.637	4.070	4.300	4.265	4.457	5.144	5.404	0.440
$\hat{g}_t[Survey+Estimates+Accounting]$	5	0.0010	20	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.000
$\hat{g}_t[All\ Data]$	3	1.0000	20	4.067	4.352	5.310	7.010	6.430	8.823	10.703	10.731	2.168
E. Q+4 GDP Growth												
g_t			20	-1.712	-0.973	2.931	3.669	3.881	5.167	6.628	6.839	2.165
$\hat{g}_t[Survey+Estimates]$	3	0.1000	20	3.796	3.874	4.052	4.217	4.114	4.274	4.809	4.871	0.295
$\hat{g}_t[Survey+Estimates+Market]$	3	1.0000	20	3.114	3.160	3.557	3.964	3.949	4.404	4.911	5.187	0.551
$\hat{g}_t[Survey+Estimates+Accounting]$	5	0.0100	20	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.000
$\hat{g}_t[All\ Data]$	3	1.0000	20	1.886	2.070	3.116	3.726	3.962	4.293	5.069	5.354	0.902

Summary Statistics: Model Prediction on Training and Validation Sample

Panel D: Model Blender (Elastic Net and Random Forest, 1:1)

	Obs	*Min*	*p5*	*p25*	*Mean*	*p50*	*p75*	*p95*	*Max*	*SD*
A. Current Quarter GDP Growth										
g_t	20	-1.712	-0.181	3.273	4.006	3.945	5.278	6.628	6.839	1.927
$\hat{g}_t[Survey{+}Estimates]$	20	2.025	2.486	3.890	4.279	4.290	4.826	5.566	5.727	0.868
$\hat{g}_t[Survey{+}Estimates{+}Market]$	20	2.163	2.735	4.002	4.293	4.404	4.673	5.459	5.509	0.729
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	20	3.834	3.995	4.298	4.498	4.509	4.716	4.920	4.942	0.281
$\hat{g}_t[All\ Data]$	20	2.080	2.651	4.016	4.281	4.444	4.611	5.374	5.375	0.721
B. Q+1 GDP Growth										
g_t	20	-1.712	-0.181	3.068	3.821	3.911	4.533	6.628	6.839	1.895
$\hat{g}_t[Survey{+}Estimates]$	20	2.945	3.110	4.098	4.299	4.180	4.771	5.142	5.155	0.581
$\hat{g}_t[Survey{+}Estimates{+}Market]$	20	2.685	3.153	3.984	4.279	4.232	4.613	5.170	5.300	0.563
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	20	2.350	2.798	3.903	4.195	4.214	4.662	5.169	5.293	0.671
$\hat{g}_t[All\ Data]$	20	2.211	2.492	3.714	4.012	4.052	4.478	4.964	5.101	0.706
C. Q+2 GDP Growth										
g_t	20	-1.712	-0.973	2.931	3.615	3.881	4.533	6.628	6.839	2.100
$\hat{g}_t[Survey{+}Estimates]$	20	4.113	4.224	4.417	4.489	4.506	4.579	4.723	4.748	0.144
$\hat{g}_t[Survey{+}Estimates{+}Market]$	20	2.744	2.888	3.829	4.138	4.214	4.407	5.160	5.162	0.622
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	20	2.517	2.607	3.581	3.997	4.065	4.650	4.936	4.943	0.711
$\hat{g}_t[All\ Data]$	20	2.549	2.606	3.533	3.981	4.155	4.414	4.847	4.855	0.665
D. Q+3 GDP Growth										
g_t	20	-1.712	-0.973	2.931	3.736	3.945	5.278	6.628	6.839	2.174
$\hat{g}_t[Survey{+}Estimates]$	20	2.541	2.778	3.896	4.197	4.247	4.768	5.003	5.017	0.665
$\hat{g}_t[Survey{+}Estimates{+}Market]$	20	3.411	3.617	3.956	4.197	4.260	4.422	4.598	4.622	0.309
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	20	3.617	3.626	4.128	4.243	4.357	4.455	4.577	4.651	0.296
$\hat{g}_t[All\ Data]$	20	3.328	3.379	3.955	4.119	4.265	4.344	4.470	4.546	0.324
E. Q+4 GDP Growth										
g_t	20	-1.712	-0.973	2.931	3.669	3.881	5.167	6.628	6.839	2.165
$\hat{g}_t[Survey{+}Estimates]$	20	2.630	2.717	3.968	4.146	4.209	4.596	4.919	4.942	0.630
$\hat{g}_t[Survey{+}Estimates{+}Market]$	20	3.283	3.329	3.714	3.951	4.034	4.189	4.501	4.665	0.353
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	20	0.886	1.396	3.450	3.786	3.844	4.400	5.048	5.096	1.025
$\hat{g}_t[All\ Data]$	20	0.903	1.436	3.476	3.784	3.925	4.432	4.868	4.876	0.963

Table 2.
Summary Statistics: Model Prediction on Holdout Sample

This table reports descriptive statistics for the target (g_t) and the trained models' predictions ($\hat{g}_t$) on the the holdout sample for current quarter (Panel A), Q+1 (Panel B), Q+2 (Panel C), Q+3 (Panel D), and Q+4 (Panel E) nominal GDP growth. For each variable, we report its observation count (Obs), minimum (Min), 5^{th} percentile ($p5$), 25^{th} percentile ($p25$), average ($Mean$), median ($p50$), 75^{th} percentile ($p75$), 95^{th} percentile ($p95$), maximum (Max), and standard deviation (SD). The baseline models is trained using Feature Set 1, and its predictions are denoted $\hat{g}_t[Survey + Estimates]$. The next variant is trained using both Feature Set 1 and 2, and resultant predictions are denoted $\hat{g}_t[Survey + Estimates + Market]$. The final two models are trained with accounting data. $\hat{g}_t[Survey + Estimates + Accounting]$ refers to the predictions from a model trained with Feature Set 1 and 3, and $\hat{g}_t[All\ Data]$ refers to the predictions from a model trained with all features. We also report the hyper-parameters (α and λ) of each trained model. Feature sets are enumerated in Appendix B.

Panel A: Elastic Net

	α	λ	Obs	Min	$p5$	$p25$	$Mean$	$p50$	$p75$	$p95$	Max	SD
A. Current Quarter GDP Growth												
g_t			21	-0.235	1.446	3.385	3.954	4.067	4.915	6.128	7.611	1.648
$\hat{g}_t[Survey+Estimates]$	0.30	0.30	21	3.260	3.282	3.848	4.340	4.411	4.719	5.243	5.286	0.617
$\hat{g}_t[Survey+Estimates+Market]$	0.60	1.00	21	3.925	4.050	4.332	4.593	4.644	4.801	5.129	5.170	0.360
$\hat{g}_t[Survey+Estimates+Accounting]$	0.80	1.00	21	3.884	4.106	4.335	4.526	4.561	4.734	4.925	4.928	0.295
$\hat{g}_t[All\ Data]$	0.80	1.00	21	3.884	4.106	4.335	4.526	4.561	4.734	4.925	4.928	0.295
B. Q+1 GDP Growth												
g_t			20	-0.235	0.605	3.448	4.032	4.082	4.944	6.870	7.611	1.650
$\hat{g}_t[Survey+Estimates]$	0.10	3.00	20	3.985	4.066	4.237	4.442	4.445	4.652	4.782	4.802	0.234
$\hat{g}_t[Survey+Estimates+Market]$	0.50	3.00	20	4.589	4.589	4.589	4.589	4.589	4.589	4.589	4.589	0.000
$\hat{g}_t[Survey+Estimates+Accounting]$	0.90	1.00	20	4.224	4.257	4.325	4.434	4.438	4.528	4.636	4.653	0.127
$\hat{g}_t[All\ Data]$	1.00	0.30	20	2.715	2.801	3.134	3.499	3.411	3.943	4.184	4.211	0.448
C. Q+2 GDP Growth												
g_t			19	1.446	1.446	3.510	4.257	4.098	4.972	7.611	7.611	1.345
$\hat{g}_t[Survey+Estimates]$	0.40	3.00	19	4.535	4.535	4.535	4.535	4.535	4.535	4.535	4.535	0.000
$\hat{g}_t[Survey+Estimates+Market]$	0.40	3.00	19	4.535	4.535	4.535	4.535	4.535	4.535	4.535	4.535	0.000
$\hat{g}_t[Survey+Estimates+Accounting]$	0.10	3.00	19	3.755	3.755	3.957	4.223	4.297	4.432	4.632	4.632	0.263
$\hat{g}_t[All\ Data]$	0.10	3.00	19	3.795	3.795	3.927	4.226	4.357	4.419	4.554	4.554	0.253
D. Q+3 GDP Growth												
g_t			18	1.446	1.446	3.510	4.153	4.082	4.915	7.611	7.611	1.303
$\hat{g}_t[Survey+Estimates]$	0.80	1.00	18	4.543	4.543	4.543	4.543	4.543	4.543	4.543	4.543	0.000
$\hat{g}_t[Survey+Estimates+Market]$	0.90	1.00	18	4.543	4.543	4.543	4.543	4.543	4.543	4.543	4.543	0.000
$\hat{g}_t[Survey+Estimates+Accounting]$	0.50	1.00	18	4.202	4.202	4.291	4.371	4.396	4.416	4.559	4.559	0.096
$\hat{g}_t[All\ Data]$	0.10	3.00	18	3.974	3.974	4.217	4.288	4.309	4.372	4.597	4.597	0.145
E. Q+4 GDP Growth												
g_t			17	1.446	1.446	3.738	4.203	4.098	4.915	7.611	7.611	1.326
$\hat{g}_t[Survey+Estimates]$	0.70	1.00	17	4.567	4.567	4.567	4.567	4.567	4.567	4.567	4.567	0.000
$\hat{g}_t[Survey+Estimates+Market]$	0.30	0.10	17	2.469	2.469	3.034	3.435	3.477	3.869	4.164	4.164	0.499
$\hat{g}_t[Survey+Estimates+Accounting]$	0.20	1.00	17	3.440	3.440	3.874	4.111	4.109	4.453	4.772	4.772	0.405
$\hat{g}_t[All\ Data]$	0.20	1.00	17	3.086	3.086	3.766	4.060	4.050	4.395	4.722	4.722	0.437

Summary Statistics: Model Prediction on Holdout Sample

Panel B: Random Forest

	Max Depth	*Max Features*	*NEstimators*	*Obs*	*Min*	*p5*	*p25*	*Mean*	*p50*	*p75*	*p95*	*Max*	*SD*
A. Current Quarter GDP Growth													
g_t				21	-0.235	1.446	3.385	3.954	4.067	4.915	6.128	7.611	1.648
$\hat{g}_t[Survey{+}Estimates]$	12	1	300	21	3.906	3.917	4.153	4.440	4.357	4.851	5.133	5.168	0.398
$\hat{g}_t[Survey{+}Estimates{+}Market]$	10	1	100	21	3.818	4.039	4.184	4.374	4.363	4.538	4.846	4.880	0.283
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	6	1	300	21	2.414	3.398	3.640	3.989	3.975	4.440	4.673	4.887	0.559
$\hat{g}_t[All\ Data]$	4	2	300	21	3.281	3.722	4.023	4.286	4.313	4.613	4.777	4.811	0.417
B. Q+1 GDP Growth													
g_t				20	-0.235	0.605	3.448	4.032	4.082	4.944	6.870	7.611	1.650
$\hat{g}_t[Survey{+}Estimates]$	6	1	100	20	3.854	3.870	4.309	4.628	4.563	4.972	5.403	5.454	0.458
$\hat{g}_t[Survey{+}Estimates{+}Market]$	10	1	100	20	3.427	3.503	3.969	4.325	4.279	4.787	5.209	5.242	0.540
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	4	2	100	20	2.832	3.068	3.709	4.084	4.159	4.429	5.050	5.247	0.582
$\hat{g}_t[All\ Data]$	6	3	500	20	2.988	3.152	3.659	4.167	4.232	4.634	5.046	5.098	0.607
C. Q+2 GDP Growth													
g_t				19	1.446	1.446	3.510	4.257	4.098	4.972	7.611	7.611	1.345
$\hat{g}_t[Survey{+}Estimates]$	2	1	100	19	4.068	4.068	4.252	4.469	4.507	4.678	4.807	4.807	0.229
$\hat{g}_t[Survey{+}Estimates{+}Market]$	8	1	100	19	3.170	3.170	3.811	4.110	4.058	4.540	4.936	4.936	0.477
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	8	2	200	19	3.202	3.202	3.590	4.043	4.158	4.336	4.962	4.962	0.476
$\hat{g}_t[All\ Data]$	8	2	200	19	3.265	3.265	3.550	3.990	4.219	4.312	4.693	4.693	0.432
D. Q+3 GDP Growth													
g_t				18	1.446	1.446	3.510	4.153	4.082	4.915	7.611	7.611	1.303
$\hat{g}_t[Survey{+}Estimates]$	12	1	300	18	3.857	3.857	4.204	4.471	4.471	4.645	5.444	5.444	0.452
$\hat{g}_t[Survey{+}Estimates{+}Market]$	4	2	100	18	2.860	2.860	3.547	3.886	3.944	4.285	4.665	4.665	0.512
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	4	3	200	18	2.676	2.676	3.559	4.054	4.308	4.496	4.863	4.863	0.631
$\hat{g}_t[All\ Data]$	4	2	300	18	2.953	2.953	3.673	4.225	4.402	4.695	4.951	4.951	0.647
E. Q+4 GDP Growth													
g_t				17	1.446	1.446	3.738	4.203	4.098	4.915	7.611	7.611	1.326
$\hat{g}_t[Survey{+}Estimates]$	12	1	200	17	3.769	3.769	4.285	4.569	4.533	4.844	5.304	5.304	0.419
$\hat{g}_t[Survey{+}Estimates{+}Market]$	2	1	500	17	3.747	3.747	4.084	4.160	4.164	4.263	4.402	4.402	0.184
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	10	1	200	17	3.745	3.745	4.116	4.223	4.297	4.381	4.473	4.473	0.209
$\hat{g}_t[All\ Data]$	8	1	200	17	2.970	2.970	3.755	3.940	4.041	4.158	4.331	4.331	0.354

Summary Statistics: Model Prediction on Holdout Sample

Panel C: Neural Network

	Hidden Layer Sizes	α	*Obs*	*Min*	*p5*	*p25*	*Mean*	*p50*	*p75*	*p95*	*Max*	*SD*
A. Current Quarter GDP Growth												
g_t			21	-0.235	1.446	3.385	3.954	4.067	4.915	6.128	7.611	1.648
$\hat{g}_t$[*Survey+Estimates*]	3	0.1000	21	3.816	4.089	4.458	4.678	4.696	4.893	5.343	5.435	0.408
$\hat{g}_t$[*Survey+Estimates+Market*]	2	1.0000	21	3.055	3.575	4.086	4.529	4.487	5.027	5.403	5.929	0.704
$\hat{g}_t$[*Survey+Estimates+Accounting*]	4	0.0001	21	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.000
$\hat{g}_t$[*All Data*]	2	0.0010	21	2.850	2.867	3.187	3.990	3.910	4.375	5.208	5.620	0.812
B. Q+1 GDP Growth												
g_t			20	-0.235	0.605	3.448	4.032	4.082	4.944	6.870	7.611	1.650
$\hat{g}_t$[*Survey+Estimates*]	5	0.0001	20	3.666	3.792	4.184	4.360	4.408	4.595	4.824	4.857	0.314
$\hat{g}_t$[*Survey+Estimates+Market*]	3	1.0000	20	2.772	2.841	3.368	3.746	3.713	4.160	4.783	4.998	0.559
$\hat{g}_t$[*Survey+Estimates+Accounting*]	5	0.0010	20	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.000
$\hat{g}_t$[*All Data*]	3	1.0000	20	-1.107	-0.747	1.185	1.678	1.678	2.175	3.879	3.996	1.249
C. Q+2 GDP Growth												
g_t			19	1.446	1.446	3.510	4.257	4.098	4.972	7.611	7.611	1.345
$\hat{g}_t$[*Survey+Estimates*]	5	0.1000	19	3.788	3.788	4.275	4.423	4.389	4.678	4.952	4.952	0.324
$\hat{g}_t$[*Survey+Estimates+Market*]	3	1.0000	19	1.482	1.482	3.884	4.157	4.247	4.803	5.416	5.416	0.866
$\hat{g}_t$[*Survey+Estimates+Accounting*]	5	0.0010	19	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.000
$\hat{g}_t$[*All Data*]	3	1.0000	19	2.717	2.717	3.101	5.381	5.388	7.706	8.888	8.888	2.208
D. Q+3 GDP Growth												
g_t			18	1.446	1.446	3.510	4.153	4.082	4.915	7.611	7.611	1.303
$\hat{g}_t$[*Survey+Estimates*]	3	0.0100	18	3.904	3.904	4.216	4.305	4.297	4.407	4.718	4.718	0.212
$\hat{g}_t$[*Survey+Estimates+Market*]	3	1.0000	18	3.784	3.784	4.095	4.270	4.228	4.456	4.816	4.816	0.288
$\hat{g}_t$[*Survey+Estimates+Accounting*]	5	0.0010	18	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.000
$\hat{g}_t$[*All Data*]	3	1.0000	18	3.705	3.705	5.677	8.574	7.923	12.166	14.141	14.141	3.397
E. Q+4 GDP Growth												
g_t			17	1.446	1.446	3.738	4.203	4.098	4.915	7.611	7.611	1.326
$\hat{g}_t$[*Survey+Estimates*]	3	0.1000	17	3.889	3.889	4.201	4.305	4.305	4.417	4.643	4.643	0.205
$\hat{g}_t$[*Survey+Estimates+Market*]	3	1.0000	17	3.514	3.514	3.795	4.080	4.023	4.329	4.738	4.738	0.368
$\hat{g}_t$[*Survey+Estimates+Accounting*]	5	0.0100	17	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.787	0.000
$\hat{g}_t$[*All Data*]	3	1.0000	17	1.246	1.246	2.467	3.211	2.959	3.741	5.593	5.593	1.153

Summary Statistics: Model Prediction on Holdout Sample

Panel D: Model Blender (Elastic Net and Random Forest, 1:1)

	Obs	*Min*	*p5*	*p25*	*Mean*	*p50*	*p75*	*p95*	*Max*	*SD*
A. Current Quarter GDP Growth										
g_t	21	-0.235	1.446	3.385	3.954	4.067	4.915	6.128	7.611	1.648
$\hat{g}_t[Survey{+}Estimates]$	21	3.723	3.765	4.127	4.390	4.326	4.635	5.206	5.209	0.436
$\hat{g}_t[Survey{+}Estimates{+}Market]$	21	3.896	3.982	4.228	4.492	4.489	4.560	5.104	5.226	0.374
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	21	3.824	3.940	4.224	4.407	4.371	4.639	4.839	4.859	0.289
$\hat{g}_t[All\ Data]$	21	3.253	3.734	4.000	4.236	4.280	4.577	4.685	4.798	0.399
B. Q+1 GDP Growth										
g_t	20	-0.235	0.605	3.448	4.032	4.082	4.944	6.870	7.611	1.650
$\hat{g}_t[Survey{+}Estimates]$	20	3.920	3.990	4.261	4.535	4.538	4.800	5.014	5.034	0.324
$\hat{g}_t[Survey{+}Estimates{+}Market]$	20	3.847	3.979	4.190	4.408	4.437	4.627	4.785	4.845	0.254
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	20	3.498	3.634	3.869	4.101	4.106	4.342	4.517	4.612	0.281
$\hat{g}_t[All\ Data]$	20	2.890	3.108	3.528	3.771	3.856	4.028	4.286	4.474	0.355
C. Q+2 GDP Growth										
g_t	19	1.446	1.446	3.510	4.257	4.098	4.972	7.611	7.611	1.345
$\hat{g}_t[Survey{+}Estimates]$	19	4.302	4.302	4.394	4.502	4.521	4.606	4.671	4.671	0.115
$\hat{g}_t[Survey{+}Estimates{+}Market]$	19	3.944	3.944	4.144	4.293	4.238	4.515	4.611	4.611	0.201
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	19	3.536	3.536	3.798	4.123	4.298	4.425	4.589	4.589	0.338
$\hat{g}_t[All\ Data]$	19	3.586	3.586	3.819	4.116	4.326	4.432	4.620	4.620	0.353
D. Q+3 GDP Growth										
g_t	18	1.446	1.446	3.510	4.153	4.082	4.915	7.611	7.611	1.303
$\hat{g}_t[Survey{+}Estimates]$	18	4.200	4.200	4.373	4.507	4.507	4.594	4.994	4.994	0.226
$\hat{g}_t[Survey{+}Estimates{+}Market]$	18	3.726	3.726	4.103	4.226	4.240	4.400	4.604	4.604	0.245
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	18	3.955	3.955	4.057	4.243	4.300	4.383	4.473	4.473	0.173
$\hat{g}_t[All\ Data]$	18	3.873	3.873	4.030	4.260	4.319	4.447	4.582	4.582	0.216
E. Q+4 GDP Growth										
g_t	17	1.446	1.446	3.738	4.203	4.098	4.915	7.611	7.611	1.326
$\hat{g}_t[Survey{+}Estimates]$	17	4.168	4.168	4.426	4.568	4.550	4.706	4.936	4.936	0.209
$\hat{g}_t[Survey{+}Estimates{+}Market]$	17	3.257	3.257	3.582	3.790	3.823	3.935	4.210	4.210	0.273
$\hat{g}_t[Survey{+}Estimates{+}Accounting]$	17	3.224	3.224	3.935	4.020	4.087	4.190	4.464	4.464	0.331
$\hat{g}_t[All\ Data]$	17	3.113	3.113	3.817	3.960	3.989	4.240	4.452	4.452	0.406

Table 3.
Model Performance: Holdout MSE (Q and Q+1)

This table reports the trained models' mean-squared prediction errors on the holdout sample (*Holdout MSE*) for forecasting quarter Q (current quarter, Panel A) and Q+1 (Panel B) nominal GDP growth. Column (1)–(4) report the Holdout MSEs for $\hat{g}_t[Survey + Estimates]$ (trained using Feature Set 1), $\hat{g}_t[Survey + Estimates + Market]$ (trained using Features Sets 1 and 2), $\hat{g}_t[Survey + Estimates + Accounting]$ (trained using Feature Set 1 and 3), and $\hat{g}_t[All\ Data]$ (trained using Feature Set 1, 2, and 3) respectively. The last two rows of each panel report the hyper-parameters (α and λ) of each trained model, and feature sets are enumerated in Appendix B. Column (5) reports the difference in holdout MSE between $\hat{g}_t[Survey + Estimates + Accounting]$ and $\hat{g}_t[Survey + Estimates]$. Column (6) reports the difference in holdout MSE between $\hat{g}_t[All\ Data]$ and $\hat{g}_t[Survey + Estimates]$. Column (7) reports the difference in holdout MSE between $\hat{g}_t[All\ Data]$ and $\hat{g}_t[Survey + Estimates + Market]$.

Panel A: Elastic Net

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{g}_t[Survey+ Estimates]$	$\hat{g}_t[Survey+ Estimates+ Market]$	$\hat{g}_t[Survey + Estimates+ Accounting]$	$\hat{g}_t[All\ Data]$	(3)–(1)	(4)–(1)	(4)–(2)
Panel A: Forecasting Current Quarter (Quarter Q) GDP Growth							
Holdout MSE	2.4680	2.6539	2.7177	2.7177	0.2497	0.2497	0.0639
α	0.3	0.6	0.8	0.8	–	–	–
λ	0.3	1.0	1.0	1.0	–	–	–
Panel B: Forecasting Q+1 GDP Growth							
Holdout MSE	2.7946	2.8974	2.8212	3.0998	0.0266	0.3052	0.2023
α	0.1	0.5	0.9	1.0	–	–	–
λ	3.0	3.0	1.0	0.3	–	–	–

Model Performance: Holdout MSE (Q and Q+1)

Panel B: Random Forest

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{g}_t$[*Survey+ Estimates*]	$\hat{g}_t$[*Survey+ Estimates+ Market*]	$\hat{g}_t$[*Survey + Estimates+ Accounting*]	$\hat{g}_t$[*All Data*]	(3)–(1)	(4)–(1)	(4)–(2)
Panel A: Forecasting Current Quarter (Quarter Q) GDP Growth							
Holdout MSE	2.1714	2.2880	1.8524	2.1941	-0.3190	0.0226	-0.0940
max_depth	12	10	6	4	–	–	–
max_features	1	1	1	2	–	–	–
n_estimators	300	100	300	300	–	–	–
Panel B: Forecasting Q+1 GDP Growth							
Holdout MSE	2.6278	2.5963	2.1146	1.8228	-0.5132	-0.8051	-0.7736
max_depth	6	10	4	6	–	–	–
max_features	1	1	2	3	–	–	–
n_estimators	100	100	100	500	–	–	–

Panel C: Neural Network

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{g}_t$[*Survey+ Estimates*]	$\hat{g}_t$[*Survey+ Estimates+ Market*]	$\hat{g}_t$[*Survey + Estimates+ Accounting*]	$\hat{g}_t$[*All Data*]	(3)–(1)	(4)–(1)	(4)–(2)
Panel A: Forecasting Current Quarter (Quarter Q) GDP Growth							
Holdout MSE	2.8669	3.2659	12.6182	2.2418	9.7512	-0.6252	-1.0241
Hidden_layer_sizes	3	2	4	2	–	–	–
α	0.1	1	0.0001	0.001	–	–	–
Panel B: Forecasting Q+1 GDP Growth							
Holdout MSE	2.7507	2.9759	13.1218	9.8876	10.3711	7.1369	6.9116
Hidden_layer_sizes	5	3	5	3	–	–	–
α	0.0001	1	0.001	1	–	–	–

Panel D: Model Blender (Elastic Net and Random Forest, 1:1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{g}_t$[*Survey+ Estimates*]	$\hat{g}_t$[*Survey+ Estimates+ Market*]	$\hat{g}_t$[*Survey + Estimates+ Accounting*]	$\hat{g}_t$[*All Data*]	(3)–(1)	(4)–(1)	(4)–(2)
Panel A: Forecasting Current Quarter (Quarter Q) GDP Growth							
Holdout MSE	2.2412	2.4607	2.4121	2.1923	0.1709	-0.0489	-0.2684
Panel B: Forecasting Q+1 GDP Growth							
Holdout MSE	2.6768	2.6056	2.2640	2.3205	-0.4128	-0.3563	-0.2852

Table 4.
Model Performance: Holdout MSE (Q+2 to Q+4)

This table reports the trained models' mean-squared prediction errors on the holdout sample (*Holdout MSE*) for forecasting quarter Q+2 (Panel A), Q+3 (Panel B), and Q+4 (Panel C) nominal GDP growth. Column (1)–(4) report the Holdout MSEs for $\hat{g}_t[Survey + Estimates]$ (trained using Feature Set 1), $\hat{g}_t[Survey + Estimates + Market]$ (trained using Features Sets 1 and 2), $\hat{g}_t[Survey + Estimates + Accounting]$ (trained using Feature Set 1 and 3), and $\hat{g}_t[All\ Data]$ (trained using Feature Set 1, 2, and 3) respectively. The last two rows of each panel report the hyper-parameters (α and λ) of each trained model, and feature sets are enumerated in Appendix B. Column (5) reports the difference in holdout MSE between $\hat{g}_t[Survey + Estimates + Accounting]$ and $\hat{g}_t[Survey + Estimates]$. Column (6) reports the difference in holdout MSE between $\hat{g}_t[All\ Data]$ and $\hat{g}_t[Survey + Estimates]$. Column (7) reports the difference in holdout MSE between $\hat{g}_t[All\ Data]$ and $\hat{g}_t[Survey + Estimates + Market]$.

Panel A: Elastic Net

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{g}_t[Survey+$ $Estimates]$	$\hat{g}_t[Survey+$ $Estimates+$ $Market]$	$\hat{g}_t[Survey\ +$ $Estimates+$ $Accounting]$	$\hat{g}_t[All\ Data]$	(3)–(1)	(4)–(1)	(4)–(2)
Panel A: Forecasting Q+2 GDP Growth							
Holdout MSE	1.7919	1.7919	1.8612	1.7646	0.0693	-0.0273	-0.0273
α	0.4	0.4	0.1	0.1	–	–	–
λ	3.0	3.0	3.0	3.0	–	–	–
Panel B: Forecasting Q+3 GDP Growth							
Holdout MSE	1.7565	1.7565	1.7029	1.5365	-0.0536	-0.2200	-0.2200
α	0.8	0.9	0.5	0.1	–	–	–
λ	1.0	1.0	1.0	3.0	–	–	–
Panel C: Forecasting Q+4 GDP Growth							
Holdout MSE	1.7877	2.6555	1.5004	1.4406	-0.2872	-0.3471	-1.2149
α	0.7	0.3	0.2	0.2	–	–	–
λ	1.0	0.1	1.0	1.0	–	–	–

Model Performance: Holdout MSE (Q and Q+1)

Panel B: Random Forest

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{g}_t$[*Survey+ Estimates*]	$\hat{g}_t$[*Survey+ Estimates+ Market*]	$\hat{g}_t$[*Survey + Estimates+ Accounting*]	$\hat{g}_t$[*All Data*]	(3)–(1)	(4)–(1)	(4)–(2)
Panel A: Forecasting Q+2 GDP Growth							
Holdout MSE	1.7834	1.6751	1.5287	1.4934	-0.2547	-0.2900	-0.1817
max_depth	2	8	8	8	–	–	–
max_features	1	1	2	2	–	–	–
n_estimators	100	100	200	200	–	–	–
Panel B: Forecasting Q+3 GDP Growth							
Holdout MSE	1.3623	1.6181	1.1277	1.2521	-0.2346	-0.1101	-0.3660
max_depth	12	4	4	4	–	–	–
max_features	1	2	3	2	–	–	–
n_estimators	300	100	200	300	–	–	–
Panel C: Forecasting Q+4 GDP Growth							
Holdout MSE	1.4797	1.6803	1.3724	1.3325	-0.1072	-0.1472	-0.3477
max_depth	12	2	10	8	–	–	–
max_features	1	1	1	1	–	–	–
n_estimators	200	500	200	200	–	–	–

Model Performance: Holdout MSE (Q and Q+1)

Panel C: Neural Network

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{g}_t$[*Survey+ Estimates*]	$\hat{g}_t$[*Survey+ Estimates+ Market*]	$\hat{g}_t$[*Survey + Estimates+ Accounting*]	$\hat{g}_t$[*All Data*]	(3)–(1)	(4)–(1)	(4)–(2)
Panel A: Forecasting Q+2 GDP Growth							
Holdout MSE	1.9029	2.2688	13.7575	7.1989	11.8546	5.2959	4.9300
Hidden_layer_sizes	5	3	5	3	–	–	–
α	0.1	1	0.001	1	–	–	–
Panel B: Forecasting Q+3 GDP Growth							
Holdout MSE	1.7381	1.7288	12.9367	27.4416	11.1986	25.7035	25.7127
Hidden_layer_sizes	3	3	5	3	–	–	–
α	0.01	1	0.001	1	–	–	–
Panel C: Forecasting Q+4 GDP Growth							
Holdout MSE	1.7969	2.1152	13.3226	3.2494	11.5258	1.4525	1.1342
Hidden_layer_sizes	3	3	5	3	–	–	–
α	0.1	1	0.01	1	–	–	–

Panel D: Model Blender (Elastic Net and Random Forest, 1:1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{g}_t$[*Survey+ Estimates*]	$\hat{g}_t$[*Survey+ Estimates+ Market*]	$\hat{g}_t$[*Survey + Estimates+ Accounting*]	$\hat{g}_t$[*All Data*]	(3)–(1)	(4)–(1)	(4)–(2)
Panel A: Forecasting Q+2 GDP Growth							
Holdout MSE	1.7741	1.6670	1.5858	1.5751	-0.1884	-0.1990	-0.0919
Panel B: Forecasting Q+3 GDP Growth							
Holdout MSE	1.5099	1.6859	1.4290	1.3710	-0.0810	-0.1389	-0.3149
Panel C: Forecasting Q+4 GDP Growth							
Holdout MSE	1.5924	1.9772	1.3645	1.2949	-0.2280	-0.2975	-0.6823

Table 5.
Feature Importance: Q and Q+1

This table summarize the importance of features for the full elastic net model ($\hat{g}_t[All\ Data]$ trained using all 341 features) by each feature type: survey forecasts, market values, and accounting values. Within accounting features, we further summarize feature importance by the following six sub-categories: "Capital," "Investments," "Liability," "Profits," "Shares," and "Write-off." For each feature type, we report the total number of features in column (1), the sum of the features' absolute coefficients in column (2), and the percentage of the total absolute feature coefficients accounted for by the feature category in column (3). Panel A (B) reports feature importance for the elastic net model trained to forecast current quarter (Q+1) GDP growth.

	(1) Feature Numbers	(2) Feature |Coeff|	(3) %
Panel A: Forecasting Current Quarter GDP Growth			
$\sum$Survey Forecasts	4	0.81	93.16
$\sum$Market Values	0	0.00	0.00
$\sum$Accounting Values	1	0.06	6.84
$\hookrightarrow$ Profits	1	0.06	6.84
Panel B: Forecasting Q+1 GDP Growth			
$\sum$Survey Forecasts	2	0.72	36.06
$\sum$Market Values	2	0.54	27.10
$\sum$Accounting Values	7	0.74	36.84
$\hookrightarrow$ Profits	3	0.51	25.49
$\hookrightarrow$ Shares	2	0.19	9.43
$\hookrightarrow$ Write-off	1	0.04	1.89
$\hookrightarrow$ Liability	1	0.00	0.02

Table 6.
Feature Importance: Q+2 to Q+4

This table summarize the importance of features for the full elastic net model ($\hat{g}_t[All\ Data]$ trained using all 341 features) by each feature type: survey forecasts, market values, and accounting values. Within accounting features, we further summarize feature importance by the following six sub-categories: "Capital," "Investments," "Liability," "Profits," "Shares," and "Write-off." For each feature type, we report the total number of features in column (1), the sum of the features' absolute coefficients in column (2), and the percentage of the total absolute feature coefficients accounted for by the feature category in column (3). Panel A (B) [C] reports feature importance for the elastic net model trained to forecast quarter Q+2 (Q+3) [Q+4] GDP growth.

	(1) Total Features	(2) Feature |Coeff|	(3) %
Panel A: Forecasting Q+2 GDP Growth			
$\sum$Survey Forecasts	11	0.31	19.50
$\sum$Market Values	7	0.25	15.68
$\sum$Accounting Values	49	1.03	64.82
$\hookrightarrow$ Profits	16	0.23	14.62
$\hookrightarrow$ Write-off	6	0.20	12.68
$\hookrightarrow$ Liability	5	0.18	11.21
$\hookrightarrow$ Investments	8	0.18	11.13
$\hookrightarrow$ Shares	6	0.12	7.77
$\hookrightarrow$ Capital	8	0.12	7.41
Panel B: Forecasting Q+3 GDP Growth			
$\sum$Survey Forecasts	3	0.01	0.88
$\sum$Market Values	5	0.15	10.89
$\sum$Accounting Values	46	1.18	88.23
$\hookrightarrow$ Write-off	10	0.31	23.48
$\hookrightarrow$ Capital	7	0.28	20.71
$\hookrightarrow$ Investments	7	0.25	18.74
$\hookrightarrow$ Profits	15	0.19	14.20
$\hookrightarrow$ Liability	6	0.14	10.55
$\hookrightarrow$ Shares	1	0.01	0.54
Panel C: Forecasting Q+4 GDP Growth			
$\sum$Survey Forecasts	0	0.00	0.00
$\sum$Market Values	5	0.27	9.46
$\sum$Accounting Values	46	2.56	90.54
$\hookrightarrow$ Profits	19	0.80	28.35
$\hookrightarrow$ Capital	8	0.74	26.12
$\hookrightarrow$ Investments	8	0.52	18.56
$\hookrightarrow$ Write-off	8	0.32	11.18
$\hookrightarrow$ Shares	2	0.14	4.96
$\hookrightarrow$ Liability	1	0.04	1.38

Table 7.
Shapley Values of Individual Features: Q and Q+1

This table summarize the importance of features for the full elastic net model ($\hat{g}_t[All\ Data]$ trained using all 341 features) by each feature type: survey forecasts, market values, and accounting values. Within accounting features, we further summarize feature importance by the following six sub-categories: "Capital," "Investments," "Liability," "Profits," "Shares," and "Write-off." For each feature type, we report the total number of features in column (1), the sum of the features' absolute Shapley values in column (2), and the percentage of the total absolute feature Shapley values accounted for by the feature category in column (3). Panel A (B) reports feature importance for the elastic net model trained to forecast current quarter (Q+1) GDP growth.

	(1) Total Features	(2) Shapley Value	(3) %
Panel A: Forecasting Third Estimate of Current Quarter GDP Growth			
$\sum$Survey Forecasts	4	0.52	89.51
$\sum$Market Values	0	0.00	0.00
$\sum$Accounting Values	1	0.06	10.49
$\hookrightarrow$ Profits	1	0.06	10.49
Panel B: Forecasting Q+1 GDP Growth			
$\sum$Survey Forecasts	2	0.44	29.14
$\sum$Market Values	2	0.41	26.81
$\sum$Accounting Values	7	0.67	44.05
$\hookrightarrow$ Profits	3	0.50	32.91
$\hookrightarrow$ Shares	2	0.15	9.59
$\hookrightarrow$ Write-off	1	0.02	1.53
$\hookrightarrow$ Liability	1	0.00	0.02

Table 8.
Shapley Values of Individual Features: Q+2 and Q+4

This table summarize the importance of features for the full elastic net model ($\hat{g}_t[\textit{All Data}]$ trained using all 341 features) by each feature type: survey forecasts, market values, and accounting values. Within accounting features, we further summarize feature importance by the following six sub-categories: "Capital," "Investments," "Liability," "Profits," "Shares," and "Write-off." For each feature type, we report the total number of features in column (1), the sum of the features' absolute Shapley values in column (2), and the percentage of the total absolute feature Shapley values accounted for by the feature category in column (3). Panel A (B) [C] reports feature importance for the elastic net model trained to forecast quarter Q+2 (Q+3) [Q+4] GDP growth.

	(1) Total Features	(2) Shapley Value	(3) %
Panel A: Forecasting Q+2 GDP Growth			
$\sum$Survey Forecasts	11	0.20	17.12
$\sum$Market Values	7	0.18	15.46
$\sum$Accounting Values	49	0.80	67.42
$\hookrightarrow$ Profits	16	0.19	15.73
$\hookrightarrow$ Investments	8	0.15	12.81
$\hookrightarrow$ Liability	5	0.14	11.68
$\hookrightarrow$ Write-off	6	0.13	11.08
$\hookrightarrow$ Capital	8	0.10	8.27
$\hookrightarrow$ Shares	6	0.09	7.84
Panel B: Forecasting Q+3 GDP Growth			
$\sum$Survey Forecasts	3	0.01	0.72
$\sum$Market Values	5	0.11	10.67
$\sum$Accounting Values	46	0.91	88.61
$\hookrightarrow$ Investments	7	0.23	22.18
$\hookrightarrow$ Write-off	10	0.21	20.28
$\hookrightarrow$ Capital	7	0.20	19.22
$\hookrightarrow$ Profits	15	0.16	15.67
$\hookrightarrow$ Liability	6	0.11	10.77
$\hookrightarrow$ Shares	1	0.01	0.50
Panel C: Forecasting Q+4 GDP Growth			
$\sum$Market Values	0	0.00	0.00
$\sum$Market Values	5	0.20	9.66
$\sum$Accounting Values	46	1.91	90.34
$\hookrightarrow$ Profits	19	0.66	31.19
$\hookrightarrow$ Investments	8	0.45	21.24
$\hookrightarrow$ Capital	8	0.45	21.10
$\hookrightarrow$ Write-off	8	0.21	10.03
$\hookrightarrow$ Shares	2	0.11	5.14
$\hookrightarrow$ Liability	1	0.03	1.64

Table 9.
Shapley Values of Feature Groups

This table summarize the importance of each feature group—survey forecasts, market values, and accounting values—for the full elastic net model ($\hat{g}_t[\mathit{All\ Data}]$ trained using all 341 features). Columns (1) and (2) report results for the quarter Q forecasts, columns (3) and (4) report results for quarter $Q+1$ forecasts, columns (5) and (6) report results for quarter $Q+2$ forecasts, columns (7) and (8) report results for quarter $Q+3$ forecasts, and columns (9) and (10) report results for quarter $Q+1$ forecasts. For each feature group, we report the sum of absolute Shapley values in odd columns and the percentage of the total absolute Shapley values accounted for by the feature group in even columns.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Q		Q+1		Q+2		Q+3		Q+4	
	Shapely Value	%	Shapely Value	%	Shapely Value	%	Shapely Value	%	Shapely Value	%
Survey Forecasts	0.49	88.92	0.42	31.10	0.17	27.76	0.01	1.62	0.00	0.00
Market Values	0.00	0.00	0.38	28.80	0.12	19.79	0.09	19.87	0.15	18.86
Accounting Values	0.06	11.08	0.54	40.10	0.32	52.44	0.35	78.51	0.63	81.14