Forecasting GDP Growth Rates: A Large Panel Micro Data Approach

Yongmiao Hong
U. of Chinese Academy of Sciences

Naijing Huang
Central U. of Finance and Economics

Yicheng Wang
Peking U.

Zixuan Zhao
Central U. of Finance and Economics

Jan 4, 2025
ASSA, San Francisco

Motivation

- Gross Domestic Product (GDP) is one of the key indicators in the macro economy, providing a broad measurement of the overall economic activity.
- GDP forecast is of great importance:
  - *Government and Central Bank:* Accurate GDP forecasts guide economic and monetary policies to ensure stability and growth.
  - *Investors:* GDP forecasts provide critical insights for market trends, enabling investors to make timely adjustments to their decisions.
  - *Households:* Economic growth projections influence employment opportunities, income levels, and consumption decisions.

Literature

- GDP is hard to predict (Granger, 1996, JAE):
    - The economy is unstable: economic structure change, technology progress, monetary policy (e.g., more aggressive against inflation), Covid 19...
    - Model misspecification
    - ...

Literature: Methodology

- Classic GDP forecasting models:
  - **Univariate time series models**: *autoregression (AR)* and *random walk (RW)* are always used as benchmark models (Marcellino et al., 2016; Cepni et al., 2019)
  - **Multivariate time series models**:
    - *VAR* (Sims, 1980; Lutkepohl, 2013; Chan, 2023; Samadi and Herath, 2023)
    - *Bayesian VAR (BVAR)* (Litterman, 1986; Koop, 2013; Carriero et al., 2016)
    - *Global VAR (GVAR)* (Pesaran et al., 2009)
    - *Augmented VAR* (Chudik et al., 2016)
    - ...
  - **Factor models**: e.g., *dynamic factor models (DFM)* (Marcellino et al., 2003; Stock and Watson, 2016; among many others);

Literature: Methodology Cont.

- Classic GDP forecasting models:
  - **Machine learning (ML) models:**
    - Swanson and White (1997, REStats): artificial neural network (*ANN*) performs better in short-term GDP forecast.
    - Mogliani and Simoni (2021, JoE): Bayesian mixed-data sampling (*MIDAS*) penalized regression outperforms *AR* model.
    - Longo et al. (2022, JEDC): combine recurrent neural network (*RNN*) with *DFM* to forecast U.S. GDP.
    - ...

Literature: Commonly Used Predictors

- **FRED-MD and FRED-QD** (McCracken and Ng, 2016, 2020), covering key macroeconomic and financial indicators, are widely used in forecasting GDP (Mogliani and Simoni, 2021; Chan, 2023; Samadi and Herath, 2023).
- **Survey-based predictors**:
  - The Survey of Professional Forecasters (SPF) (Giannone et al., 2008; Swanson and White, 1997)
  - Survey data from Central Bank for Euro Area (Glas and Heinisch, 2023)
- **Accounting predictors**: accounting earnings (Konchitchki and Patatoukas, 2014, JAE), accounting profits (Konchitchki and Patatoukas, 2014, AR), special items (Abdalla and Carabias, 2022, AR) $\cdots$
- ...

Literature: Commonly Used Predictors

- **Common Features of Predictors**: all previous studies use aggregate predictors for GDP forecasting.
  - An aggregate economic variable is a statistic that summarizes the economic activities of an entire area, sector, or population.
  - The formation of many aggregate variables involves summarizing micro-level data to create a single index that represents a larger group or population. For example:
    - Composite stock price index: $S\&P$ 500, weighted average of individual stock prices
    - Professional forecaster predictor: the mean or median of professional forecasters' predictions
    - ...

Drawbacks: Aggregate Predictors

- **Aggregate-level predictors have drawbacks, leading to the loss of information, which smoothes out important features among economic agents:**
  - Heterogeneity
  - Interaction
  - Nonlinearity
  - Jumps
  - Heavy tails
  - Sudden structural breaks
  - ...

From Macro/Aggregate to Micro/Individual Predictors

- We propose a **micro-forecasting framework** for GDP forecasting, generally defined as:

$$y_{t+h} = F_h(\mathbf{X}_t; \gamma) + \varepsilon_{t+h}$$

  - y_{t+h}: GDP growth rate, h is the forecast horizon
  - **Key difference:** $\mathbf{X_t} = (1,\ x_{1t},\ x_{2t},\ \cdots,\ x_{it},\ \cdots,\ x_{Nt})'$ is an $(N+1) \times 1$ predictor vector, a large panel of individual level predictors:
    - Individual level accounting earnings of all listed firms of U.S.
    - 20,173 firms in total

- **Question:** Why accounting earnings?

Why Individual Accounting Earnings Predict Future GDP?

- GDP by definition includes the accounting earning information:
  - **Firm performance:** accounting earnings → corporate profit → persistence of firm performance → future GDP growth (Konchitchki and Patatoukas, 2014, JAE, 2014, AR; Nallareddy and Ogneva, 2017; Gaertner et al., 2020)
  - **Production network:** individual accounting earning $\xrightarrow{\textit{production networks}}$ upstream or downstream firms' earnings → future GDP growth (Gabaix, 2011; Acemoglu et al., 2012; Carvalho and Gabaix, 2013; Baqaee and Farhi, 2019)

Why Individual Accounting Earnings Predict Future GDP?

- Information Channel:
  - **Information on firm operation:** individual accounting earnings $\xrightarrow{\textit{reveal}}$ unobservable information on firms' performance and decisions $\rightarrow$ future GDP growth (Foster et al., 1984; Bernard and Thomas, 1990; Shivakumar and Urcan, 2017)
  - **Information on unobservable policy shocks:** individual accounting earnings $\xrightarrow{\textit{reveal}}$ unobservable heterogeneous information about policy shocks $\rightarrow$ heterogeneous impacts on different firms $\rightarrow$ future GDP growth (Anilowski et al., 2007; Shivakumar, 2007; Patatoukas, 2014; Kim et al., 2016; Konchitchki et al., 2016; Kalay et al., 2018; Rouxelin et al., 2018; Kausar and Park, 2024)

Why Individual Accounting Earnings Predict Future GDP?

- Liquidity and Wealth Effects (Campbell, 1999; Poterba, 2000; Jermann and Quadrini, 2007; Miao and Wang, 2012; Gilchrist et al., 2017; Shivakumar and Urcan, 2017; Gertler and Gilchrist, 2018; Chodorow-Reich et al., 2021)
  - **Firm:** individual accounting earnings → financial constraints → production → future GDP growth
  - **Household:** individual accounting earnings → wealth of households → consumption behavior → future GDP growth
- It is important to note that these connections are not necessarily causal but instead indicate a correlation between individual accounting earnings and future GDP.

Key to the Micro Forecasting Framework: $F_h(\cdot; \gamma)$

- A large model is needed to handle a large panel of individual accounting earnings:
  - Allowing for interactions among individual accounting earnings.
  - Allowing to take into account:
    - possible heterogeneity,
    - nonlinearity, jumps, structural breaks and etc.
- What is a "large model"?
  - Large Language Model (LLM)
    - We are not talking about ChatGPT: currently the best technique to understand textual data

Large Models in Econometrics

- Large models in econometrics contain
  - a large set of predicting variables
  - a large set of unknown parameters
  - interactions among predicting variables
  - possible nonlinear structures
  - ...

Machine Learning Algorithms

- We apply various machine learning algorithms to the large panel of micro-level predictors.
- Why using machine learning models?
    - Directly handling a high-dimensional set of predictors without suffering from the curse of dimensionality
    - Without any functional form assumptions (as in *RF* and *GBRT*)
    - Allowing for heterogeneity and/or interactions among individuals (as in *RF* and *GBRT*)
    - Suitable for out-of-sample forecasting thanks to the use of training data and validating data, as well as the use of regularization

Main Findings: GDP Forecasting

- Our micro-forecasting approach using a large panel of individual accounting earnings substantially **enhances the accuracy** of GDP growth rate forecasting, compared to:
  - Benchmark univariate model: *AR*
    - MSE improvement is almost 40%
  - Aggregate approaches:
    - Aggregate predictors: FRED-QD, aggregate accounting earnings.
    - Aggregate modeling approaches: ML vs. Principle Component Regression Model

Outline

- Data
- Forecasting Models
- Empirical Results
- Concluding Remarks

Data

Data

- **Forecast Target: GDP Growth Rate**
  - We use the quarterly accumulated changing rate of real GDP for the U.S. (Kylian et al., 2014).
  - Source: from *FRED* Database.
- **Predictors: Individual Accounting Earnings**
  - Quarterly accounting earnings of all listed firms of U.S., 20,173 firms in total
  - Sources: from *Compustat Quarterly Preliminary History*
- **Sample Periods**
  - Full-sample period: 1989Q2 - 2023Q4
  - Out-of-sample period: 2000Q1 - 2023Q4
  - Our out-of-sample period covers the significant economic shocks, such as the global financial crisis, recessions, and the COVID-19 pandemic.

Forecasting Models

Forecasting Models

- Our micro-forecasting framework to predict GDP growth rate is:

$$g_{t+h} = F_h(\mathbf{X_t}; \gamma) + \varepsilon_{t+h}$$

  - g_{t+h} represents the value of GDP growth rate at quarter $t + h$.
  - h is the forecast horizon, including 1-, 2-, 4-, 8-quarter forecasts respectively.
  - $\mathbf{X_t} = (1,\ x_{1t},\ x_{2t},\ \cdots,\ x_{it},\ \cdots,\ x_{Nt})'$ is an $(N+1) \times 1$ predictor vector, including the accounting earnings from individual firms and their one-period lags, as well as the GDP growth rate in period t and its four lags.

Forecasting Models

- **ML Algorithms**
  - **Penalized Regression Models**
    - Adaptive LASSO (*adaLASSO*)
    - Ridge Regression (*Ridge*)
    - Elastic Net
  - **Nonlinear Algorithms**
    - Random Forest (*RF*)
    - Gradient Boosting Regression Tree (*GBRT*)
- **Benchmark Models** in GDP forecasting literature
    - Autoregressive (*AR*) Model (Marcellino et al., 2016; Cepni et al., 2019)

Forecast Evaluation Criterion

- The accuracy of the out-of-sample forecasts is measured by the **mean squared errors (MSE)**:

$$MSE = \frac{1}{T_{oos}} \sum_{t \in \mathcal{P}} (g_{t+h} - \hat{g}_{t+h})^2$$

where $\mathcal{P}$ denotes the set of out-of-sample periods, and T_{oos} is the number of observations in $\mathcal{P}$.

Empirical Results

Empirical Results: U.S. GDP, MSE Ratios to *AR* Model

- Using **individual accounting earnings** significantly **improves** the predictive accuracy of GDP growth rate forecasts over **AR benchmark model**.
- The improvement is more remarkable **in short forecast horizon ($h = 1$) and long forecast horizon ($h = 8$)**.

Table: MSE Ratio of Micro-Forecasting Model Relative to *AR* Model: U.S. Real GDP

Micro-Forecasting Model	Forecast Horizon			
	h=1	h=2	h=4	h=8
Adaptive LASSO	0.643	0.870	0.860	0.846
LASSO	0.626	0.857	0.871	0.865
Elastic Net	**0.607**	**0.834**	0.879	0.850
Ridge	1.175	1.124	0.915	0.750
RF	0.623	0.863	0.895	0.658
GBRT	0.679	0.926	**0.780**	**0.604**

Concluding Remarks

Concluding Remarks

(1) We propose a novel micro-forecasting framework in macroeconomic forecast, characterized by:

- A large panel of micro-level predictors;
- Machine learning models, which are utilized to integrate the heterogeneous information in the micro-level predictors.

Concluding Remarks

(2) As an application of our micro-forecasting framework, we forecast GDP growth rates based on a large panel of individual accounting earnings.

- Our micro-forecasting approach substantially **enhances the accuracy** of GDP forecasting, performing better than:
    - Benchmark univariate model *AR*;
    - Aggregate modeling approaches such as *factor model*;
    - Aggregate predictors including FRED-QD and aggregate accounting earnings.

Concluding Remarks

(3) It would be interesting to explore whether micro-forecasting approach also works in forecasting other macro time series:

- exchange rates
- unemployment rates
- inflation rates
- nowcasting
- ...

Thanks!