Machine Learning in Finance

1.Create indicators via ML

Conceptual X / Y $\longrightarrow$ Input/Output: X/Y

Different data source:
Text, Video, Picture, Music…

2.Utilize the algorithm advantages of ML

Input X $\xrightarrow{\text{Forecasting problems}}$ Output: Y

Including different dimensional data

Choose another important Y

3.Inside ML

Input X $\xrightarrow{\text{Man VS. Machine}}$ Output: Y

1.The bias between Man and ML (Bianchi et al., 2022; van Binsbergen et al., 2023)

2.How machine conducts economic production (Abis and Veldkamp, 2023)

3.How machine changes the way we live (Cao et al., 2023)

Belief Distortions and Macroeconomic Fluctuations

Bianchi Francesco, Ludvigson Sydney C. and Ma Sai

解读：雷印如

2023 年 12 月 5 日

Background

- What is the **belief distortion**?
  - An ex ante expectational error generated by the systematic misweighting of available information demonstrably pertinent to the accuracy of belief
- Why the belief distortion exists?
  - Information capacity constraints (Coibion and Gorodnichenko, 2015)
  - Extrapolative rules (Barberis et al. 2015)
  - Sentiments (Angeletos et al., 2018b)
  - Ambiguity aversion (Bianchi et al., 2018)

Motivation

- Important question without sufficient empirical results
  - Theoretical literature: **systematic expectational errors** embedded in beliefs can have important dynamic effects on the economy
  - Fundamental challenge: **no objective measure** of such distortions
- Literature finds little agreement on how to measure theoretical standard
  - **Benchmark differs** according to the specific survey (population, topic, time period, empirical methodology...)
- This paper combines a data-rich environment with ML to provide new estimates of time-varying systematic belief distortion

Why take ML as benchmark?

- Machine learning is itself a model of belief formation
- Optimized approaches to real world decision always require the efficient processing of large amounts of information
  - Machine algorithms are advantageous in high-dimensional information
- Machine algorithm can systematically adapt to new information
  - the approach does not run the risk of spurious indication because of the existence of structural breaks or the arrival of new information

Research Question

- Whether survey respondents with different beliefs **systematically misweight** pertinent economic information?
  - MSE comparison
- What kind of errors are the respondent-types making and why?
  - Dynamics of belief distortions
  - Bias decomposition
- What we find when comparing to well-known prior empirical studies?
  - In sample VS. Out of sample
  - How belief distortions change over the business cycle

Contribution

- On the empirical side, the paper reports evidence of belief distortions and relates them to economic activity (Bordalo et al. 2019)
- First to find that systematic expectational errors in professional macroeconomic predictions are partly attributable to overreliance
- Our work also connects with a pre-existing econometric forecasting literature in a data-rich context along with ML (Ang et al., 2007)
- Contributing to literature on ML in correcting errors in human judgment and improving predictive accuracy. (Martin and Nagel, 2019)

Data: three different surveys

- the Survey of Professional Forecasters (SPF) quarterly survey
  - professional forecasters in a variety of institutions
  - nowcasts and quarterly forecasts from one to four quarters ahead
- the University of Michigan Survey of Consumers (SOC)
  - household expectations about inflation, and regression for GDP growth
- the Blue Chip Survey (BC)
  - executives and professional forecasters at financial firms
  - quarter-over-quarter percentage change forecast (4 to 5)

Data: Predictor variables

- Real-Time Macro Data (quarterly) 92 real-time macro variables
  - Federal Reserve Bank of Philadelphia's Real-Time Dataset
- Monthly Financial Data (147 monthly financial indicators)
- Daily Financial Data (87)
  - Combine 39 commodities prices, 16 corporate risk variables, 9 equities, 7 foreign exchange, and 16 government securities into diffusion index
- Additional Nonfactor Data
  - i th percentile's own nowcast, lags of forecasts, and high-order moments

Estimation and Machine Learning

- With these data in hand, consider the following machine learning empirical specification for forecasting, $T_{test} \geq 84$ quarters, rolling

$$y_{j,t+h} = \mathcal{X}_t' \boldsymbol{\beta}_{jh}^{(i)} + \epsilon_{jt+h}, \tag{5}$$

where $\mathcal{X}_t = \left(1, \mathbb{F}_t^{(i)}[y_{j,t+h}], \mathcal{Z}_{jt}\right)'$ and $\boldsymbol{\beta}_{jh}^{(i)} \equiv \left(\alpha_{jh}^{(i)}, \beta_{jh\mathbb{F}}^{(i)}, \left(\mathbf{B}_{jh\mathcal{Z}}^{(i)}\right)\right)'$.

Let $\mathbf{X}_T = \left(y_{j,1}, \ldots y_{j,T}, \ldots \mathcal{X}_1', \ldots, \mathcal{X}_T'\right)'$ be the vector containing all observations in a sample of size T. We consider estimators of $\boldsymbol{\beta}_{jh}^{(i)}$ that take the form

$$\hat{\boldsymbol{\beta}}_{jh}^{(i)} = m\left(\mathbf{X}_T, \boldsymbol{\lambda}^{(i)}\right),$$

- We use the ElasticNet estimator to achieve both shrinkage and sparsity
- T_V= [12, 16, 20, 24, 28] and T_E= [24, 28, 32, 36, 40, 44, 48, 52, 56, 60]

Estimation and Machine Learning

- Output growth behaves differently in recessions than in expansions
- Turning points are often anticipated by a flat or inverted yield curve

$$
\begin{aligned}
\Delta gdp_{j,t+h} &= \alpha_{jh}^{(i)} + \beta_{jh\mathbb{F}}^{(i)} \mathbb{F}_t^{(i)}\left[y_{j,t+h}\right] + \mathbf{B}_{j\mathcal{Z}}^{(i)\prime} \mathcal{Z}_{jt} + \epsilon_{jt+h} \quad \text{if } slope_{kt} > \widehat{tr}_{kt}, \\
\Delta gdp_{j,t+h} &= B_{kt} I_{kt} \quad \text{if } slope_{kt} \leq \widehat{tr}_{kt},
\end{aligned}
$$

where B_{kt} is a parameter and I_{kt} is a dummy variable that depends on a yield curve measure at time t.

- $slope_{kt}$ is a yield spread measure at time t and $\hat{tr}_{kt}$ is a threshold
  - term spread (10y3m, 10y1y, or 10y2y)
  - The machine searches in real time for the specific threshold

Estimation and Machine Learning

- Distortions in survey responses are quantified by the ratio MSE_E/MSE_F over an extended forecast evaluation sample of size P

$$\mathbb{E}_t^{(i)}[y_{j,t+h}] \equiv \mathcal{X}_t' \hat{\beta}_{jh,t}^{(i)}\left(\mathbf{X}_T, \hat{\lambda}_t^{(i)}\right).$$

Forecast errors are differentially denoted for the survey and machine

$$\text{survey error}_{t+h}^{(i)} = \mathbb{F}_t^{(i)}[y_{j,t+h}] - y_{j,t+h},$$

$$\text{machine error}_{t+h}^{(i)} = \mathbb{E}_t^{(i)}[y_{j,t+h}] - y_{j,t+h}.$$

Survey and machine mean squared errors (MSEs) are denoted with $\mathbb{F}$ and $\mathbb{E}$ subscripts, i.e.,

$$\text{(6)} \qquad \text{survey MSE} \equiv MSE_{\mathbb{F}} = (1/P)_{i=1}^{P}\left(\text{survey error}_{t+h}^{(i)}\right)^2$$

$$\text{(7)} \qquad \text{machine MSE} \equiv MSE_{\mathbb{E}} = (1/P)_{i=1}^{P}\left(\text{machine error}_{t+h}^{(i)}\right)^2,$$

Estimation and Machine Learning

- Measures of uncertainty about forecast point forecasts are inherently interesting for reasons other than forecaster bias

Specifically, consider a hybrid forecast of $y_{j,t+h}$, denoted $\mathbb{EF}_t^{(i)}[y_{j,t+h}]$, obtained as a weighted average of the machine and the survey forecasts:

$$\mathbb{EF}_t^{(i)}[y_{j,t+h}] \equiv w\mathbb{E}_t^{(i)}[y_{j,t+h}] + (1-w)\mathbb{F}_t^{(i)}[y_{j,t+h}],$$

The optimal weight w^* placed on the machine forecast is defined as the one that minimizes the hybrid MSE over our evaluation samples, i.e.,

$$w^* = \arg\min MSE_{\mathbb{EF}} = \arg\min(1/P)_{i=1}^{P}\left(\text{hybrid error}_{t+h}^{(i)}\right)^2,$$

- If ML is only marginally better than the survey, w^* would be close to 0.5

Empirical Results: MSE comparison

- The overall magnitude by which the machine model improves on the survey forecasts is in most cases sizable

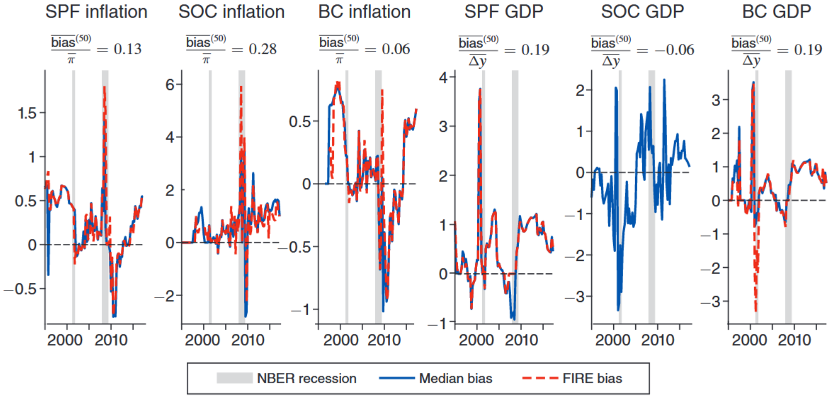
$$\text{ML: } y_{j,t+h} = \alpha_{jh}^{(i)} + \beta_{jh\mathbb{F}}^{(i)} \mathbb{F}_t^{(i)}[y_{j,t+h}] + \mathbf{B}_{jh\mathcal{Z}}^{(i)} \mathcal{Z}_{jt} + \epsilon_{jt+h}$$

Inflation forecasts							
Percentile:	Median	5th	10th	20th	25th	30th	40th
Survey of Professional Forecasters (SPF)							
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.85	0.56	0.74	0.83	0.90	0.88	0.89
OOS R^2	0.15	0.44	0.26	0.17	0.10	0.12	0.11
w^*	0.68	0.82	0.78	0.74	0.63	0.66	0.66
$MSE_{\mathbb{R}}/MSE_{\mathbb{F}}$	0.82	0.41	0.64	0.77	0.81	0.82	0.84
	60th	70th	75th	80th	90th	95th	
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.74	0.70	0.67	0.59	0.55	0.47	
OOS R^2	0.26	0.30	0.33	0.41	0.45	0.53	
w^*	0.78	0.76	0.74	0.80	0.79	0.83	
$MSE_{\mathbb{R}}/MSE_{\mathbb{F}}$	0.76	0.66	0.61	0.53	0.39	0.27	

- Dispersed information is not the most relevant source of belief distortion

Empirical Results: Dynamics of belief distortions

- We turn to investigate the dynamics of systematic expectational error



- GDP “optimism” and CPI “pessimism”

Empirical Results: Dynamics of belief distortions

- The machine does better at the end rather than the beginning

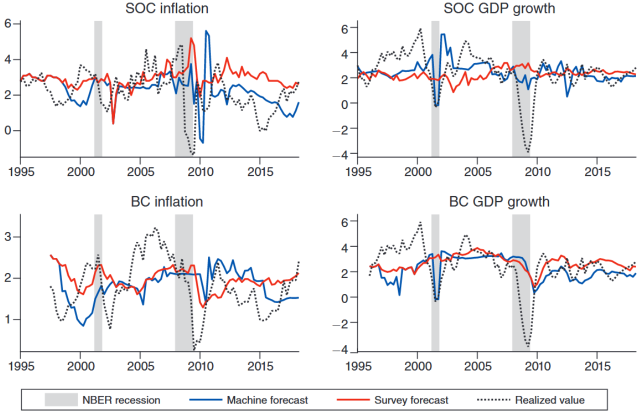
$$\text{ML: } y_{j,t+h} = \alpha_{jh}^{(i)} + \beta_{jh\mathbb{F}}^{(i)} \mathbb{F}_t^{(i)} [y_{j,t+h}] + \mathbf{B}_{jh\mathcal{Z}}^{(i)} \mathcal{Z}_{jt} + \epsilon_{jt+h}$$

Median inflation forecasts, $MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$					
SPF		SOC		BC	
1995:I–2018:II	2013:II–2018:II	1995:I–2018:II	2013:II–2018:II	1997:III–2018:II	2013:II–2018:II
0.85	0.63	0.62	0.37	0.84	0.67
Median GDP forecasts, $\mathrm{MSE}_{\mathbb{E}}/\mathrm{MSE}_{\mathbb{F}}$					
SPF		SOC		BC	
1995:I–2018:II	2013:II–2018:II	1995:I–2018:II	2013:II–2018:II	1996:I–2018:II	2013:II–2018:II
0.88	0.82	0.78	0.71	0.87	0.67

- Bounded rationality in the form of limitations on the human capacity are unlikely to fully explain these findings

Empirical Results: Dynamics of belief distortions

- The machine does not perform well in every time period of our sample



- Large forecast errors were made during Great Recession by both ML and professional forecasters

Empirical Results: Bias decomposition

- What kind of errors in judgment are the respondent-types making?

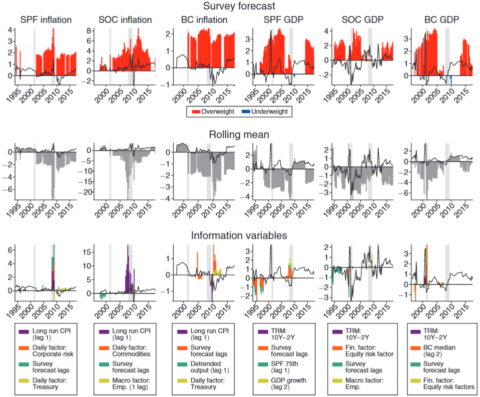
$$bias^{(i)}_{j,t+h} \equiv \mathbb{F}^{(i)}_{j,t+h|t} - \mathbb{E}^{(i)}_{j,t+h|t} = \mathbb{F}^{(i)}_t[y_{j,t+h}] - \hat{\alpha}_{jh} - \hat{\beta}^{(i)}_{jh\mathbb{F}}\mathbb{F}^{(i)}_t[y_{j,t+h}] - \hat{\mathbf{B}}^{(i)\prime}_{jh\mathcal{Z}}\mathcal{Z}_{jt}$$

$$= \underbrace{\left[-\hat{\alpha}^{(i)}_{jh}\right]}_{\text{Intercept}} + \underbrace{\left[\left(1-\hat{\beta}^{(i)}_{j\mathbb{F}}\right)\mathbb{F}^{(i)}_t[y_{j,t+h}]\right]}_{\text{Survey}} + \underbrace{\left[-\hat{\mathbf{B}}^{(i)\prime}_{j\mathcal{Z}}\mathcal{Z}_{jt}\right]}_{\text{Info variables}}.$$

- If $\hat{\beta}^i_{jF,t} < 1$, this implies that the machine improves forecasts by downweighting the survey response (red bar)
- Any estimate of $\hat{\alpha}^i_{jh,t} \neq 0$ and $\hat{B}^i_{jZ,t} \neq 0$ indicates that the machine improved forecasts by giving greater absolute weight to $Z_{j,k,t}$ or $\hat{\alpha}^i_{jh,t}$

Empirical Results: Bias decomposition

- What kind of errors in judgment are the respondent-types making?

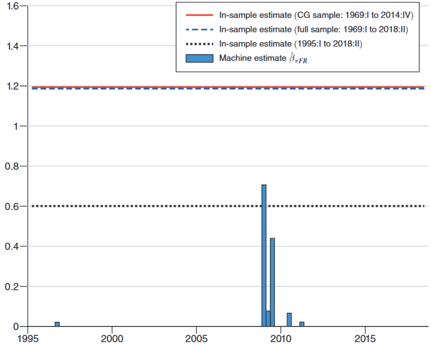


- Substantial **overreliance by professional forecasters** on the private or judgmental component of their predictions

Empirical Results: Comparison to prior studies

- CG(2015) found that mean survey forecast errors are positively predicted by ex ante mean forecast revisions in in-sample regressions

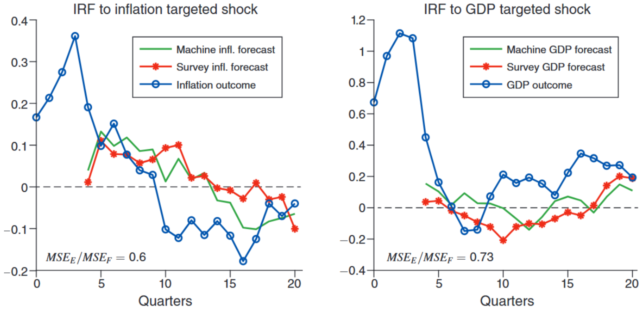
$$\pi_{j,t+3} - \mathbb{F}_t^{(\mu)}[\pi_{j,t+3}] = \alpha_\pi^{(\mu)} + \beta_{\pi\mathrm{FR}}^{(\mu)}\left(\mathbb{F}_t^{(\mu)}[\pi_{t+3}] - \mathbb{F}_{t-1}^{(\mu)}[\pi_{t+3}]\right) + \mathbf{B}_{\pi\mathcal{Z}}^{(\mu)'}\mathcal{Z}_{\pi t} + \epsilon_{\pi t+h}.$$



- No such evidence in ML- OOS

Empirical Results: Over the Business Cycle

- AHS(2020) divide cyclical shocks as the inflation-targeted shock and the GDP-targeted shock



- Cyclical shocks are not well observed in real time, even by a machine
- This may be because shocks are constructed from an in-sample est.

Conclusions

- This paper provides new measures of belief distortions in survey responses based on a novel dynamic machine learning algorithm and relates them to macroeconomic activity
  - Expectations about both inflation and GDP growth are biased upward on average
  - Respondents overweight on the marginal information embedded in their own belief and underweight on other publicly available information
- The results suggest that artificial intelligence algorithms can be productively deployed to correct errors in human judgment and improve predictive accuracy

New Ideas

- Further Analysis
  - Belief Distortion in Macroeconomics and Stock Returns
  - Belief Distortion in Macroeconomics and Systemic Risk
  - Whether Central Bank Communication mitigate the bias
- New Information
  - Textual data (News or Central Bank Communication) for improving the forecasting power when facing with cyclical shocks
  - Authoritative video data for high-frequency trading, obtain indicators to predict stock returns, and compare the predictability between texts and videos, ect.

央行沟通有助于改善宏观经济预测吗？———基于文本数据的高维稀疏建模

林建浩，陈良源，罗子豪，张一帆

2023 年 12 月 5 日

Motivation

- 宏观经济预测不仅是重要的经济学方法论问题更是政策实践的重要挑战
- 央行沟通可以通过创造信息、减少噪声和信息效应的期限结构三个信息渠道影响经济主体的预期和行为
  - 创造新信息 (Peek et al., 2003)
  - 减少噪声（陈良源等，2021）
  - 央行沟通的信息效应具有期限结构
- **利用央行沟通文本能否提升宏观经济预测效果？**

Contribution

- 不同于关注央行沟通对预期变量影响的文献（卞志村和张义，2012），本文从央行沟通是否有助于提高宏观经济预测角度验证了预期管理有效性
- 丰富了高维数据和实时数据在宏观经济预测中的作用文献（刘汉和刘金全，2011；刘涛雄和徐晓飞，2015）
- 丰富了非结构化的文本大数据在金融领域的研究（姜富伟等，2021）

Data

- 宏观经济变量数据的收集和选取（中经网和国家统计局）
  - Tagert：CPI、PPI、GDP、IVA、FAI、M2
  - Input：实际产出类、物流类、实际投资类、房地产类、实际消费类、进口类、期货市场、股市类、外汇类、财政收支类、货币和信贷类、价格类、利率类、债市类以及美国经济指标、国际大宗商品指标共 152 个月度指标
- 央行沟通文本（2003-2019）
  - 书面沟通：《货币政策执行报告》（66 篇）
  - 口头沟通：央行发言人讲话、采访、新闻发布会、时报文档（417 次）

Forecasting Model

- 小型 VAR 模型：根据 BIC 准则确定最优滞后阶数为 2 阶
- 中型 VAR 模型：纳入更多的经济变量（最优滞后阶数为 1 阶）
- 动态因子模型（DFM）：贝叶斯估计从高维经济变量中提取共同因子
- DFM + 历史信息模型（最优）
- FAVAR 模型：主成分分析 +VAR

Main Results

- 样本外模型预测精度，样本外均方根误差 (RMSE) 对预测能力进行衡量

表 4 **基于 DFM 模型和不同信息集下央行沟通测度预测能力分析**

目标变量	预测方法	预测阶数 τ	DFM 模型	DFM + 1 阶历史信息	DFM + 多阶历史信息
			(1)	(2)	(3)
GDP	实时预报		40.41%	18.47%	−12.99%
	预测	1	30.13%	7.08%	−18.18%
		3	32.37%	6.83%	−3.40%
		12	15.38%	3.74%	0.53%
IVA	实时预报		10.91%	7.82%	6.25%
	预测	1	7.40%	0.63%	−7.13%
		3	1.72%	1.57%	1.17%
		12	2.32%	1.82%	6.16%

- 央行沟通能够显著提高宏观变量预测精度，难以提升历史信息 +DFM
  - 中国的央行沟通更多是传达历史和现状信息，可能存在共线性问题
  - 对于长期预测而言，央行沟通信息反而能够发挥其特殊的作用