Belief Distortions and Macroeconomic Fluctuations

Francesco Bianchi, Sydney C. Ludvigson, and Sai Ma

Duke University, New York University, Board of Governors

How Important are Belief Distortions?

- How important are **belief distortions** in economic decision making and what role do they play in **macroeconomic fluctuations**?

How Important are Belief Distortions?

- How important are **belief distortions** in economic decision making and what role do they play in **macroeconomic fluctuations**?
- **Large theoretical literatures**: emerged to argue that **systematic expectational errors embedded in beliefs** have important dynamic effects on economy. Agents make systematic errors for many reasons:

How Important are Belief Distortions?

- How important are **belief distortions** in economic decision making and what role do they play in **macroeconomic fluctuations**?
- **Large theoretical literatures**: emerged to argue that **systematic expectational errors embedded in beliefs** have important dynamic effects on economy. Agents make systematic errors for many reasons:
  1. Face limits on their ability to acquire and process information (e.g., Sims (2003), Reis (2006a,b), Woodford (2013), Coibion and Gorodnichenko (2015))
  2. Use simple extrapolative rules (De Long, Shleifer, Summers, and Waldmann (1990); Barberis, Shleifer, and Vishny (1998); Barberis, Greenwood, Jin, and Shleifer (2015)).
  3. Intentionally adopt conservatively pessimistic beliefs due to aversion to ambiguity (e.g., Hansen and Sargent (2008); Epstein and Schneider (2010); Ilut and Schneider (2015)).
  4. Over-weight personal experiences (e.g., Malmendier and Nagel (2011); Malmendier and Nagel (2015)) or relevance of incoming data (e.g, Bordalo, Gennaioli, and Shleifer (2018); Gennaioli and Shleifer (2018))
  5. Engage in Bayesian learning about a mean with a prior that assumes skewness (e.g., Afrouzi, Veldkamp, et al. (2019)).

How Important are Belief Distortions?

- All previous theories presume agents exhibit some form of **belief distortion**, which we **define here** in general terms as:
  - *Systematic mis-weighting of available information demonstrably pertinent to the accuracy of the belief.*

How Important are Belief Distortions?

- All previous theories presume agents exhibit some form of **belief distortion**, which we **define here** in general terms as:
  - *Systematic mis-weighting of available information demonstrably pertinent to the accuracy of the belief.*
- **A first challenge** in empirically assessing role of belief distortions in economic outcomes is that there is **no widely accepted measure of belief distortions**.

How Important are Belief Distortions?

- All previous theories presume agents exhibit some form of **belief distortion**, which we **define here** in general terms as:
  - *Systematic mis-weighting of available information demonstrably pertinent to the accuracy of the belief.*
- **A first challenge** in empirically assessing role of belief distortions in economic outcomes is that there is **no widely accepted measure of belief distortions**.
- Use of surveys seems promising to measure agents' beliefs. But existing studies vary by:
  - specific survey data used
  - segment of the population surveyed
  - time frame to which the survey responses pertain
  - survey questions analyzed
  - extraneous econometric methodology

How Important are Belief Distortions?

- All previous theories presume agents exhibit some form of **belief distortion**, which we **define here** in general terms as:
  - *Systematic mis-weighting of available information demonstrably pertinent to the accuracy of the belief.*
- **A first challenge** in empirically assessing role of belief distortions in economic outcomes is that there is **no widely accepted measure of belief distortions**.
- Use of surveys seems promising to measure agents' beliefs. But existing studies vary by:
  - specific survey data used
  - segment of the population surveyed
  - time frame to which the survey responses pertain
  - survey questions analyzed
  - extraneous econometric methodology
- **A second challenge:** Given the range of theories and large amount of information that is in fact **ex-ante available and possibly pertinent** to decision making, there is **no widely accepted benchmark model of belief formation** to measure distortions in survey responses.

Our Work

Goals:

1. Provide a measure of distortions that is as wide-ranging as possible.
2. Assess the role of belief distortions in macroeconomic fluctuations.

- Measure beliefs across a **range of surveys, respondents, and questions about future economic outcomes**.
- Adopt the **perspective of survey respondents**:
  1. Out-of-sample nature of the forecast
  2. Real-time data-rich environment
  3. Heterogeneity in beliefs
- Construct and study a **broad measure of belief distortion in economic decision making**.

Operationalizing Measurement of Belief Distortions

- **Three key ingredients.**

Operationalizing Measurement of Belief Distortions

- **Three key ingredients.**
  1. **Direct evidence on what economic decision makers actually believe.**
     - Use data from different surveys, different types of decision makers, different survey questions, and broad cross-sections of survey respondents with different beliefs.

Operationalizing Measurement of Belief Distortions

- **Three key ingredients**.
  1. **Direct evidence on what economic decision makers actually believe**.
     - Use data from different surveys, different types of decision makers, different survey questions, and broad cross-sections of survey respondents with different beliefs.
  2. **Algorithm to acquire and process the theoretically vast quantities of real-time information**.
     - Failure to take into account the **data-rich environment** or the **the out-of-sample nature of real-time decisions**, can lead to erroneous conclusions about whether beliefs are distorted.
     - Use **machine learning tools** to **combine** information from the **data-rich** environment with the **survey forecasts**. **Machine learning benchmark** in principle free of human biases, to process hundreds of pieces of information available in real-time at mixed sampling intervals.

Operationalizing Measurement of Belief Distortions

- **Three key ingredients.**
  1. **Direct evidence on what economic decision makers actually believe.**
     - Use data from different surveys, different types of decision makers, different survey questions, and broad cross-sections of survey respondents with different beliefs.
  2. **Algorithm to acquire and process the theoretically vast quantities of real-time information.**
     - Failure to take into account the **data-rich environment** or the **the out-of-sample nature of real-time decisions**, can lead to erroneous conclusions about whether beliefs are distorted.
     - Use **machine learning tools** to **combine** information from the **data-rich** environment with the **survey forecasts**. **Machine learning benchmark** in principle free of human biases, to process hundreds of pieces of information available in real-time at mixed sampling intervals.
  3. **A sufficiently long time series on the first two.**
     - **Reduce sampling noise**, as necessary to distinguish bad luck in a random environment (i.e., pure random error) from a systematic mis-weighting of information.
     - Assess the role of any distortions in **dynamic macroeconomic fluctuations**.

     $\Rightarrow$ Study the bias of **respondent type** (mean, median, percentiles) as opposed to single respondent.

Revisiting Well Known Empirical Examples

- Importance of minding challenges of **real-world decision making** can be illustrated by revisiting some well-known empirical findings.

- Two key aspects that we need to model:

  1. **Real-time adaptation to new information & out-of-sample nature of decisions**. As an example, revisit some findings from Coibion and Gorodnichenko (2015) (CB).

  2. **The data-rich environment in which survey respondents operate**. As an example, revisit findings from Chauvet and Potter (2013) (CP).

Real-Time, Out-of-Sample Decision Making

Coibion and Gorodnichenko (2015) in-sample regressions

CG In-Sample Regressions of Forecast Errors on Forecast Revisions

Regression: $\pi_{t+3,t} - \mathbb{F}_t\left[\pi_{t+3,t}\right] = \alpha + \beta\left(\mathbb{F}_t\left[\pi_{t+3,t}\right] - \mathbb{F}_{t-1}\left[\pi_{t+3,t}\right]\right) + \delta\pi_{t+2,t-1} + \epsilon_t$

	(1)	(2)	(3)	(4)
	Panel A: Sample: 1969:Q1 - 2014:Q4		**Panel B**: Sample: 1969:Q1 - 2018:Q2	
Constant	0.001	-0.077	-0.022	-0.116
t-stat	(0.005)	(-0.442)	(-0.167)	(-0.758)
$\mathbb{F}_t\left[\pi_{t+3,t}\right] - \mathbb{F}_{t-1}\left[\pi_{t+3,t}\right]$	1.194**	1.141**	1.186**	1.116**
t-stat	(2.496)	(2.560)	(2.478)	(2.532)
$\pi_{t+2,t-1}$		0.021		0.027
t-stat		(0.435)		(0.574)
$\bar{R}^2$	0.195	0.197	0.193	0.195

Notes: annual inflation is defined as $\pi_{t+3,t} = \frac{P_t}{P_{t-1}} \times \frac{P_{t+1}}{P_t} \times \frac{P_{t+2}}{P_{t+1}} \times \frac{P_{t+3}}{P_{t+2}}$, $\mathbb{F}_t\left[\pi_{t+3,t}\right]$ is the mean Survey of Professional Forecasters (SPF) forecast of annual inflation as of time t. Panel A presents the sample in Coibion and Gorodnichenko (2015) and Panel B updates the sample to 2017:Q4. Following CG, regressions are run and forecast errors computed using forecasts of real-time inflation data available four quarters after the period being forecast. Newey-West corrected (t-statistics) with lags = 4. Newey-West HAC: *sig. at 10%. **sig. at 5%. ***sig. at 1%.

Real-Time, Out-of-Sample Decision Making

Coibion and Gorodnichenko (2015) in-sample regressions

- Forecast errors predictable by forecast revisions.

CG In-Sample Regressions of Forecast Errors on Forecast Revisions

Regression: $\pi_{t+3,t} - \mathbb{F}_t\left[\pi_{t+3,t}\right] = \alpha + \beta\left(\mathbb{F}_t\left[\pi_{t+3,t}\right] - \mathbb{F}_{t-1}\left[\pi_{t+3,t}\right]\right) + \delta\pi_{t+2,t-1} + \epsilon_t$

	(1)	(2)	(3)	(4)
	Panel A: Sample: 1969:Q1 - 2014:Q4		**Panel B**: Sample: 1969:Q1 - 2018:Q2	
Constant	0.001	-0.077	-0.022	-0.116
t-stat	(0.005)	(-0.442)	(-0.167)	(-0.758)
$\mathbb{F}_t\left[\pi_{t+3,t}\right] - \mathbb{F}_{t-1}\left[\pi_{t+3,t}\right]$	1.194^{**}	1.141^{**}	1.186^{**}	1.116^{**}
t-stat	(2.496)	(2.560)	(2.478)	(2.532)
$\pi_{t+2,t-1}$		0.021		0.027
t-stat		(0.435)		(0.574)
$\bar{R}^2$	0.195	0.197	0.193	0.195

Notes: annual inflation is defined as $\pi_{t+3,t} = \frac{P_t}{P_{t-1}} \times \frac{P_{t+1}}{P_t} \times \frac{P_{t+2}}{P_{t+1}} \times \frac{P_{t+3}}{P_{t+2}}$, $\mathbb{F}_t\left[\pi_{t+3,t}\right]$ is the mean Survey of Professional Forecasters (SPF) forecast of annual inflation as of time t. Panel A presents the sample in Coibion and Gorodnichenko (2015) and Panel B updates the sample to 2017:Q4. Following CG, regressions are run and forecast errors computed using forecasts of real-time inflation data available four quarters after the period being forecast. Newey-West corrected (t-statistics) with lags = 4. Newey-West HAC: *sig. at 10%. **sig. at 5%. ***sig. at 1%.

Real-Time, Out-of-Sample Decision Making

Coibion and Gorodnichenko (2015) in-sample regressions

- Forecast errors predictable by forecast revisions.
- Other information (e.g., lagged $\pi_{t+2,t-1}$) insignificant once forecast revisions included.

CG In-Sample Regressions of Forecast Errors on Forecast Revisions

Regression: $\pi_{t+3,t} - \mathbb{F}_t\left[\pi_{t+3,t}\right] = \alpha + \beta\left(\mathbb{F}_t\left[\pi_{t+3,t}\right] - \mathbb{F}_{t-1}\left[\pi_{t+3,t}\right]\right) + \delta\pi_{t+2,t-1} + \epsilon_t$

	(1)	(2)	(3)	(4)
	Panel A: Sample: 1969:Q1 - 2014:Q4		**Panel B**: Sample: 1969:Q1 - 2018:Q2	
Constant	0.001	-0.077	-0.022	-0.116
t-stat	(0.005)	(-0.442)	(-0.167)	(-0.758)
$\mathbb{F}_t\left[\pi_{t+3,t}\right] - \mathbb{F}_{t-1}\left[\pi_{t+3,t}\right]$	1.194**	1.141**	1.186**	1.116**
t-stat	(2.496)	(2.560)	(2.478)	(2.532)
$\pi_{t+2,t-1}$		0.021		0.027
t-stat		(0.435)		(0.574)
$\bar{R}^2$	0.195	0.197	0.193	0.195

Notes: annual inflation is defined as $\pi_{t+3,t} = \frac{P_t}{P_{t-1}} \times \frac{P_{t+1}}{P_t} \times \frac{P_{t+2}}{P_{t+1}} \times \frac{P_{t+3}}{P_{t+2}}$, $\mathbb{F}_t\left[\pi_{t+3,t}\right]$ is the mean Survey of Professional Forecasters (SPF) forecast of annual inflation as of time t. Panel A presents the sample in Coibion and Gorodnichenko (2015) and Panel B updates the sample to 2017:Q4. Following CG, regressions are run and forecast errors computed using forecasts of real-time inflation data available four quarters after the period being forecast. Newey-West corrected (t-statistics) with lags = 4. Newey-West HAC: *sig. at 10%. **sig. at 5%. ***sig. at 1%.

Out-of-Sample GG Regressions

- Forecast model:

$$\pi_{t+3,t} - \mathbb{F}_t \left[\pi_{t+3,t}\right] = \alpha + \beta \left(\mathbb{F}_t \left[\pi_{t+3,t}\right] - \mathbb{F}_{t-1} \left[\pi_{t+3,t}\right]\right) + \epsilon_t$$

- Rolling (or recursive) regressions used to estimate coefficients at t and predict $\pi_{t+3,t}$ in *subsequent periods*. CG model forecasts:

$$\widehat{\pi}_{t+3,t} = \widehat{\alpha}^{(t)} + \left(1 + \widehat{\beta}^{(t)}\right) \mathbb{F}_t \left[\pi_{t+3,t}\right] - \widehat{\beta}^{(t)} \mathbb{F}_{t-1} \left[\pi_{t+3,t}\right]$$

$$\text{survey error}_t = \mathbb{F}_t \left[\pi_{t+3,t}\right] - \pi_{t+3,t}$$
$$\text{CG model error}_t = \widehat{\pi}_{t+3,t} - \pi_{t+3,t}$$

- mean-square-forecast errors computed over samples of size T^F as

$$\text{MSE}_{\mathbb{F}} = \left(T^F\right)^{-1} \sum_{s=1}^{T^F} \left(\text{survey error}_{t+s}\right)^2$$
$$\text{MSE}_{\text{CG}} = \left(T^F\right)^{-1} \sum_{s=1}^{T^F} \left(\text{CG model error}_{t+s}\right)^2$$

Out-of-Sample CG Regressions

- Compare ratio of mean OOS forecast errors.

Mean square errors (MSE): CG model and SPF

Forecast model: $\widehat{\pi}_{t+3,t} - \mathbb{F}_t[\pi_{t+3,t}] = \widehat{\alpha}^{(t)} + \widehat{\beta}^{(t)}\mathbb{F}_t[\pi_{t+3,t}] - \widehat{\beta}^{(t)}\mathbb{F}_{t-1}[\pi_{t+3,t}]$		
Method	Forecast Sample	$MSE_{CG}/MSE_{\mathbb{F}}$
Rolling 5 years	1975:Q4 - 2018:Q2	1.4
Rolling 10 years	1980:Q4 - 2018:Q2	1.3
Rolling 20 years	1990:Q4 - 2018:Q2	1.3
Recursive 5 years	1975:Q4 - 2018:Q2	1.7
Recursive 10 years	1980:Q4 - 2018:Q2	1.6
Recursive 20 years	1990:Q4 - 2018:Q2	1.3

Notes: The table reports the ratio of MSEs of the CG model forecast over the survey forecast. $\pi_{t+3,t} = \frac{P_t}{P_{t-1}} \times \frac{P_{t+1}}{P_t} \times \frac{P_{t+2}}{P_{t+1}} \times \frac{P_{t+3}}{P_{t+2}}$. The regression estimation uses the latest vintage of inflation in real time and, following CG, computes forecast errors real-time data available four quarters after the period being forecast. The sample spans the period 1969:Q1 - 2018:Q2.

Out-of-Sample CG Regressions

- Compare ratio of mean OOS forecast errors.
- Across range of rolling or recursive windows, CG model performs much worse than survey forecast.

Mean square errors (MSE): CG model and SPF

Forecast model: $\widehat{\pi}_{t+3,t} - \mathbb{F}_t[\pi_{t+3,t}] = \widehat{\alpha}^{(t)} + \widehat{\beta}^{(t)}\mathbb{F}_t[\pi_{t+3,t}] - \widehat{\beta}^{(t)}\mathbb{F}_{t-1}[\pi_{t+3,t}]$		
Method	Forecast Sample	$MSE_{CG}/MSE_{\mathbb{F}}$
Rolling 5 years	1975:Q4 - 2018:Q2	1.4
Rolling 10 years	1980:Q4 - 2018:Q2	1.3
Rolling 20 years	1990:Q4 - 2018:Q2	1.3
Recursive 5 years	1975:Q4 - 2018:Q2	1.7
Recursive 10 years	1980:Q4 - 2018:Q2	1.6
Recursive 20 years	1990:Q4 - 2018:Q2	1.3

Notes: The table reports the ratio of MSEs of the CG model forecast over the survey forecast. $\pi_{t+3,t} = \frac{P_t}{P_{t-1}} \times \frac{P_{t+1}}{P_t} \times \frac{P_{t+2}}{P_{t+1}} \times \frac{P_{t+3}}{P_{t+2}}$. The regression estimation uses the latest vintage of inflation in real time and, following CG, computes forecast errors real-time data available four quarters after the period being forecast. The sample spans the period 1969:Q1 - 2018:Q2.

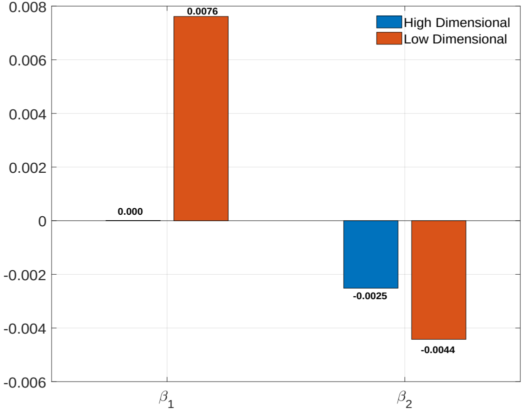
Out-of-sample nature of real-time decision making

- CG argue that results for mean forecasts support **information rigidities**, not departures from rationality (e.g., adaptive expectations).
- If professional forecasters are at the "frontier", we could have expected the apparent relation in-sample to be eliminated if it were exploitable.
- **Implication:** Even agents (such as our machine) who face no substantive information processing limitations will optimally downweight information that might appear relevant ex post if it fails to improve ex ante forecasts.

High- v.s. Low-Dimensional Decision Making

Chauvet and Potter (2013) (CP) out-of-sample example

- Equally important to take into account the **high-dimensional nature of decision making**.

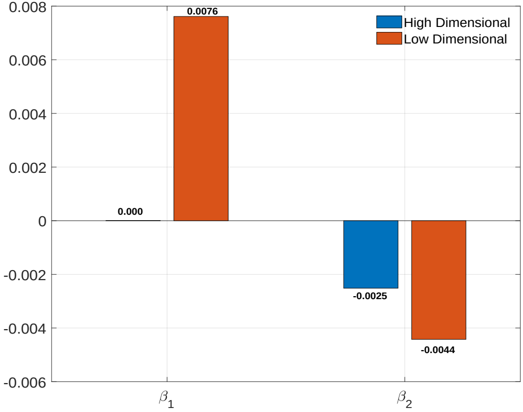


Autoregressive Coefficients in high- v.s. low-dimensional out-of-sample forecasts. Average autoregressive coefficients from one-quarter-ahead rolling regressions of real GDP growth on predictors. β_1 is the average coefficient on the first AR lag; β_2 is the average coefficient on the second. The high dimension estimation entertains very large numbers of potential predictors, in addition to the autoregressive lags, while the low dimension setting uses only two additional predictors. The sample spans 1995:Q1-2018:Q2.

High- v.s. Low-Dimensional Decision Making

Chauvet and Potter (2013) (CP) out-of-sample example

- CP: AR(2) works best for out-of-sample GDP growth forecasts.

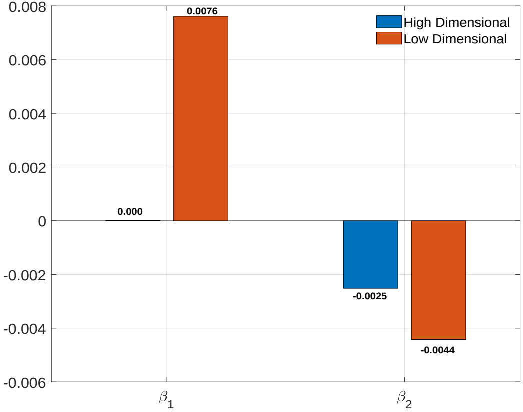


Autoregressive Coefficients in high- v.s. low-dimensional out-of-sample forecasts. Average autoregressive coefficients from one-quarter-ahead rolling regressions of real GDP growth on predictors. β_1 is the average coefficient on the first AR lag; β_2 is the average coefficient on the second. The high dimension estimation entertains very large numbers of potential predictors, in addition to the autoregressive lags, while the low dimension setting uses only two additional predictors. The sample spans 1995:Q1-2018:Q2.

High- v.s. Low-Dimensional Decision Making

Chauvet and Potter (2013) (CP) out-of-sample example

- Information in AR lags **less important** in a **high-dimensional setting**.



Autoregressive Coefficients in high- v.s. low-dimensional out-of-sample forecasts. Average autoregressive coefficients from one-quarter-ahead rolling regressions of real GDP growth on predictors. β_1 is the average coefficient on the first AR lag; β_2 is the average coefficient on the second. The high dimension estimation entertains very large numbers of potential predictors, in addition to the autoregressive lags, while the low dimension setting uses only two additional predictors. The sample spans 1995:Q1-2018:Q2.

Machine Learning and Econometric Model

Beliefs and Biases: Outline

- Do survey forecasts **systematically** mis-weight information? We use large amount of information **("big data")** and **real-time machine learning benchmark model of beliefs**.
- Our belief distortion (**"bias"**) measure for respondent-type i is defined as the **difference** between her survey forecast and the machine benchmark.

$$\underbrace{bias^{(i)}_{j,t+h}}_{\text{Type } i \text{ bias}} = \underbrace{\mathbb{F}^{(i)}_t \left[y_{j,t+h} \right]}_{\text{Her survey forecast}} - \underbrace{\mathbb{E}^{(i)}_t \left[y_{j,t+h} \right]}_{\text{Machine benchmark}}$$

- Roadmap:
  1. Describe **machine benchmark**.
  2. Describe **data**.
  3. Describe **evidence**. *Are* there biases? Do they vary over time and across agents? Do they matter? Are they correlated with anything in aggregate economy? (VAR analysis.)

Notation:

- Let $y_{j,t+h}$ denote a series indexed by j whose value in period $h \geq 1$ a survey forecaster is asked to predict at time t.
- Let $\mathbb{F}_t^{(i)}$ denote a survey forecast made at time t and let superscript (i) refer to the ith **respondent-type**, where i denotes either a respondent-type with the mean belief, "$i = mean$" or at the ith percentile of the forecast distribution, i.e., "$i = 65$" refers to a belief at the 65th percentile.
- e.g., $\mathbb{F}_t^{(65)}\left[y_{j,t+h}\right]$ denotes the survey forecast of $y_{j,t+h}$ that is formed at time t by a survey respondent in the 65th percentile of the survey distribution.

Machine Learning and Econometric Model

Access and process relevant information with machine learning algorithm:

- It is imperative that the **data used in the algorithm** be as **rich as possible**, so measure of belief distortion does not *miss pertinent information*.
- *Relevant information not considered by the benchmark can lead to spurious estimates of belief distortions and their dynamics*.
- To address this problem we take a **two pronged approach**:
    1. **Diffusion index estimation**: relatively small number of dynamic **factors** are estimated from hundreds (or potentially thousands) of economic time-series.
    2. **Machine Learning**: regularized estimation, optimally trades off costs of downweighting information for benefits of reduced parameter estimation error.

Machine Learning Model of Beliefs

- Consider the **machine learning model of expectation formation**:

$$y_{j,t+h} = \alpha_j^{(i)} + \beta_{j\mathbb{F}}^{(i)} \mathbb{F}_t^{(i)} \left[y_{j,t+h} \right] + \mathbf{B}_{j\mathcal{Z}}^{(i)} \mathcal{Z}_{jt} + \epsilon_{jt+h}, \quad (1)$$

where $\mathcal{Z}_{jt}$ is a vector of variables and factors available at the time of the forecast.

- Define the *machine efficient benchmark* for type i as a set of parameter restrictions that would imply the survey forecast efficiently processes all available information at time t:

$$\beta_{j\mathbb{F}}^{(i)} = 1; \ \mathbf{B}_{j\mathcal{Z}}^{(i)} = \mathbf{0}; \ \alpha_j^{(i)} = 0. \quad (2)$$

Machine Learning Model of Beliefs

- Consider the **machine learning model of expectation formation**:

$$y_{j,t+h} = \alpha_j^{(i)} + \beta_{j\mathbb{F}}^{(i)} \mathbb{F}_t^{(i)} \left[y_{j,t+h} \right] + \mathbf{B}_{j\mathcal{Z}}^{(i)} \mathcal{Z}_{jt} + \epsilon_{jt+h}, \quad (1)$$

  where $\mathcal{Z}_{jt}$ is a vector of variables and factors available at the time of the forecast.

- Define the *machine efficient benchmark* for type i as a set of parameter restrictions that would imply the survey forecast efficiently processes all available information at time t:

$$\beta_{j\mathbb{F}}^{(i)} = 1; \; \mathbf{B}_{j\mathcal{Z}}^{(i)} = \mathbf{0}; \, \alpha_j^{(i)} = 0. \quad (2)$$

- Deviations from above benchmark reveal **systematic expectational errors**, as measured by the **mis-weighting of information** contained in $\mathcal{Z}_{jt}$ or "1" ($\mathbf{B}_{j\mathcal{Z}}^{(i)} \neq \mathbf{0}$ or $\alpha_j^{(i)} \neq 0$) and/or the respondent-type's own forecast, $\mathbb{F}_t^{(i)} \left[y_{j,t+h} \right]$ ($\beta_{j\mathbb{F}}^{(i)} \neq 1$).

Machine Learning Model of Beliefs

$$y_{j,t+h} = \alpha_j^{(i)} + \beta_{j\mathbb{F}}^{(i)} \mathbb{F}_t^{(i)} \left[y_{j,t+h} \right] + \mathbf{B}_{j\mathcal{Z}}^{(i)} \mathcal{Z}_{jt} + \epsilon_{jt+h},$$

- Several points about the machine learning benchmark bear noting.

Machine Learning Model of Beliefs

$$y_{j,t+h} = \alpha_j^{(i)} + \beta_{j\mathbb{F}}^{(i)} \mathbb{F}_t^{(i)} \left[y_{j,t+h} \right] + \mathbf{B}_{j\mathcal{Z}}^{(i)} \mathcal{Z}_{jt} + \epsilon_{jt+h},$$

- Several points about the machine learning benchmark bear noting.
  1. Measure is *individual-specific benchmark* that adopts the perspective of a forecaster at the ith percentile.
  2. Machine given any information that the respondent-type could have known at t, which includes her own forecast $\mathbb{F}_t^{(i)} \left[y_{j,t+h} \right]$ and other information in $\mathcal{Z}_t$.
  3. $\beta_{j\mathbb{F}}^{(i)} \neq 1$ means $\mathbb{F}_t^{(i)} \left[y_{j,t+h} \right]$ could have been improved by re-weighting her own forecast against the other information in $\mathcal{Z}_t$ which potentially includes $\mathbb{F}_{t-1}^{(s \neq i)} \left[y_{j,t+h} \right]$.

Machine Learning Model of Beliefs

$$y_{j,t+h} = \alpha_j^{(i)} + \beta_{j\mathbb{F}}^{(i)} \mathbb{F}_t^{(i)} \left[y_{j,t+h} \right] + \mathbf{B}_{j\mathcal{Z}}^{(i)} \mathcal{Z}_{jt} + \epsilon_{jt+h},$$

- Several points about the machine learning benchmark bear noting.
  1. Measure is *individual-specific benchmark* that adopts the perspective of a forecaster at the ith percentile.
  2. Machine given any information that the respondent-type could have known at t, which includes her own forecast $\mathbb{F}_t^{(i)} \left[y_{j,t+h} \right]$ and other information in $\mathcal{Z}_t$.
  3. $\beta_{j\mathbb{F}}^{(i)} \neq 1$ means $\mathbb{F}_t^{(i)} \left[y_{j,t+h} \right]$ could have been improved by re-weighting her own forecast against the other information in $\mathcal{Z}_t$ which potentially includes $\mathbb{F}_{t-1}^{(s \neq i)} \left[y_{j,t+h} \right]$.
- Even with factors, the **number of possible predictors in benchmark model** can be large or even exceed number of observations. We turn to machine learning algorithm.

Machine Learning Algorithm

Simplify notation:

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

where $\mathcal{X}_t = (1, \mathcal{X}_{1t,...,}\mathcal{X}_{Kt})'$ collects the variables $\left(\mathbb{F}_t^{(i)}\left[y_{j,t+h}\right], \mathcal{Z}_{jt}\right)$ into a vector with "1" and $\boldsymbol{\beta}_j^{(i)} \equiv \left(\alpha_j^{(i)}, \beta_{j\mathbb{F}}^{(i)}, \mathbf{B}_{j\mathcal{Z}}^{(i)}\right)' \equiv \left(\beta_0^{(i)}, \beta_1^{(i)}, ...\beta_K^{(i)}\right)'$ collects all the coefficients. Let $X_t = \left(y_{j,t+h}, \mathcal{X}_t'\right)'$.

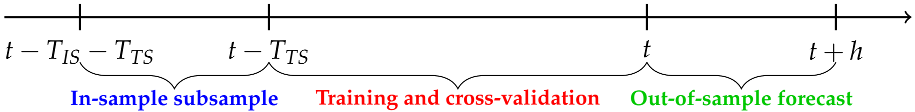
- Consider estimators of $\boldsymbol{\beta}_j^{(i)}$,

$$\hat{\boldsymbol{\beta}}_j^{(i)} = m\left(X_t, \boldsymbol{\lambda}\right),$$

- $m\left(X_t, \boldsymbol{\lambda}\right)$ defines an estimator of $\boldsymbol{\beta}_j^{(i)}$ as a function of X_t and a non-negative regularization parameter vector $\boldsymbol{\lambda}$
- $\boldsymbol{\lambda}$ estimated with **real-time training sample**. Denote combined estimator $\hat{\boldsymbol{\beta}}_j^{(i)}\left(X_t, \hat{\boldsymbol{\lambda}}\right)$.
- Possible estimators: Least Absolute Shrinkage and Selection Operator (LASSO), Random Forest, Ridge, Elastic Net (EN) (combines LASSO and Ridge).

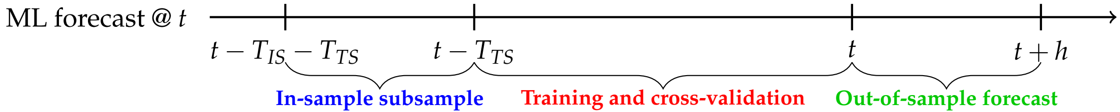
Machine Learning Algorithm Summary

ML forecast @ t



1. **Sample partitioning**: At time t, the prior sample is partitioned into an **in-sample** subsample with T_{IS} quarters and a **hold-out** **training** subsample with T_{TS} quarters.

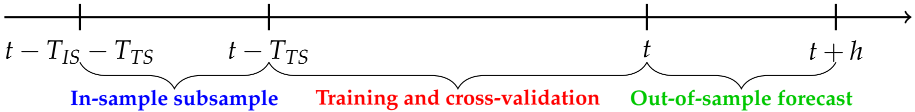
Machine Learning Algorithm Summary



1. **Sample partitioning**: At time t, the prior sample is partitioned into an **in-sample** subsample with T_{IS} quarters and a **hold-out training** subsample with T_{TS} quarters.
2. **In-sample estimation**: Given a pair (λ_1, λ_2), calculate EN coefficients $\hat{\boldsymbol{\beta}}_j^{(i)}(X_t, \boldsymbol{\lambda})$

Machine Learning Algorithm Summary

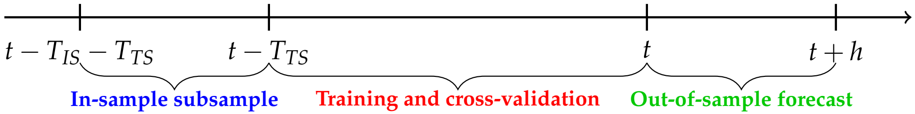
ML forecast @ t



1. **Sample partitioning**: At time t, the prior sample is partitioned into an **in-sample** subsample with T_{IS} quarters and a **hold-out** **training** subsample with T_{TS} quarters.
2. **In-sample estimation**: Given a pair (λ_1, λ_2), calculate EN coefficients $\hat{\boldsymbol{\beta}}_j^{(i)}(X_t, \boldsymbol{\lambda})$
3. **Optimal regularization**: Optimal $\hat{\boldsymbol{\lambda}}$ delivers $\hat{\boldsymbol{\beta}}_j^{(i)}(X_t, \boldsymbol{\lambda})$ that minimizes mean-square forecast errors from rolling regressions over the training sample.

Machine Learning Algorithm Summary

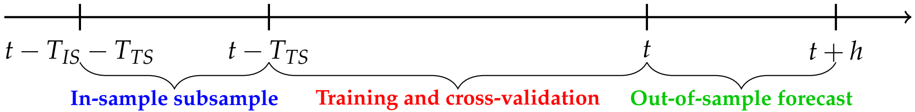
ML forecast @ t



1. **Sample partitioning**: At time t, the prior sample is partitioned into an **in-sample** subsample with T_{IS} quarters and a **hold-out** **training** subsample with T_{TS} quarters.
2. **In-sample estimation**: Given a pair (λ_1, λ_2), calculate EN coefficients $\hat{\boldsymbol{\beta}}_j^{(i)}(X_t, \boldsymbol{\lambda})$
3. **Optimal regularization**: Optimal $\hat{\boldsymbol{\lambda}}$ delivers $\hat{\boldsymbol{\beta}}_j^{(i)}(X_t, \boldsymbol{\lambda})$ that minimizes mean-square forecast errors from rolling regressions over the training sample.
4. **Optimal sample size**: Steps 1-3 are repeated over a grid of window sizes T_{IS} and T_{TS}. Optimal window sizes $\hat{T}_{IS}$ and $\hat{T}_{TS}$ minimizes the MSE over the **training** sample

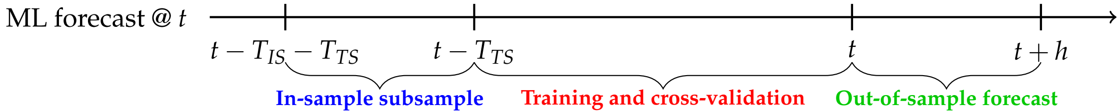
Machine Learning Algorithm Summary

ML forecast @ t



1. **Sample partitioning**: At time t, the prior sample is partitioned into an **in-sample** subsample with T_{IS} quarters and a **hold-out** **training** subsample with T_{TS} quarters.
2. **In-sample estimation**: Given a pair (λ_1, λ_2), calculate EN coefficients $\hat{\boldsymbol{\beta}}_j^{(i)}(X_t, \boldsymbol{\lambda})$
3. **Optimal regularization**: Optimal $\hat{\boldsymbol{\lambda}}$ delivers $\hat{\boldsymbol{\beta}}_j^{(i)}(X_t, \boldsymbol{\lambda})$ that minimizes mean-square forecast errors from rolling regressions over the training sample.
4. **Optimal sample size**: Steps 1-3 are repeated over a grid of window sizes T_{IS} and T_{TS}. Optimal window sizes $\hat{T}_{IS}$ and $\hat{T}_{TS}$ minimizes the MSE over the **training** sample
5. **Out-of-sample prediction**: Values of regressors at t used to make a true out-of-sample prediction of y_{t+h}, using $\hat{\boldsymbol{\beta}}_{j,t}^{(i)}\left(X_{t,\hat{T}_{IS}}, \hat{\boldsymbol{\lambda}}_t^{(i)}\right)$.

Machine Learning Algorithm Summary



1. **Sample partitioning**: At time t, the prior sample is partitioned into an **in-sample** subsample with T_{IS} quarters and a **hold-out training** subsample with T_{TS} quarters.
2. **In-sample estimation**: Given a pair (λ_1, λ_2), calculate EN coefficients $\hat{\boldsymbol{\beta}}_j^{(i)}(X_t, \boldsymbol{\lambda})$
3. **Optimal regularization**: Optimal $\hat{\boldsymbol{\lambda}}$ delivers $\hat{\boldsymbol{\beta}}_j^{(i)}(X_t, \boldsymbol{\lambda})$ that minimizes mean-square forecast errors from rolling regressions over the training sample.
4. **Optimal sample size**: Steps 1-3 are repeated over a grid of window sizes T_{IS} and T_{TS}. Optimal window sizes $\hat{T}_{IS}$ and $\hat{T}_{TS}$ minimizes the MSE over the **training** sample
5. **Out-of-sample prediction**: Values of regressors at t used to make a true out-of-sample prediction of y_{t+h}, using $\hat{\boldsymbol{\beta}}_{j,t}^{(i)}\left(X_{t,\hat{T}_{IS}}, \hat{\boldsymbol{\lambda}}_t^{(i)}\right)$.
6. **Roll forward and repeat**: Roll forward t to $t+1$ and repeat steps 1-5.

Machine Learning Algorithm

- The **benchmark machine learning belief** at time t:

$$\mathbb{E}_t^{(i)}\left[y_{j,t+h}\right] = \mathcal{X}_t' \hat{\beta}_j^{(i)}\left(X_t, \hat{\lambda}\right).$$

- Out-of-Sample Forecast Errors:

$$\text{survey error}_{t+h}^{(i)} = \mathbb{F}_t^{(i)}\left[y_{j,t+h}\right] - y_{j,t+h} \quad (3)$$

$$\text{machine error}_{t+h}^{(i)} = \mathbb{E}_t^{(i)}\left[y_{j,t+h}\right] - y_{j,t+h} \quad (4)$$

- Out-of-sample mean-square-forecast errors (MSE):

$$\text{survey MSE} \equiv MSE_{\mathbb{F}} = \frac{1}{P}\sum_{t=1}^{P}(\text{survey error}_{t+h})^2 \quad (5)$$

$$\text{machine MSE} \equiv MSE_{\mathbb{E}} = \frac{1}{P}\sum_{t=1}^{P}(\text{machine error}_{t+h})^2 \quad (6)$$

Machine Learning Algorithm

- To assess belief distortions we need to **compare forecast accuracy** across the survey and machine learning benchmark.
- Need a sufficiently large number of obs. on relative accuracy to **distinguish bad luck** in a random environment from **systematic error**.
- *If* machine benchmark *consistently* produces more reliable forecasts over an **extended sample**, we take as evidence of systematic expectational errors and quantify their magnitude by looking at the ratio of MSEs. **Otherwise, conclude no evidence of systematic error**.

Machine Learning Algorithm

- To assess belief distortions we need to **compare forecast accuracy** across the survey and machine learning benchmark.
- Need a sufficiently large number of obs. on relative accuracy to **distinguish bad luck** in a random environment from **systematic error**.
- *If* machine benchmark *consistently* produces more reliable forecasts over an **extended sample**, we take as evidence of systematic expectational errors and quantify their magnitude by looking at the ratio of MSEs. **Otherwise, conclude no evidence of systematic error**.
- Obtain a *dynamic* measure of belief distortions by taking the difference between the survey and the machine forecast.

$$bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t \left[y_{j,t+h} \right] - \mathbb{E}^{(i)}_t \left[y_{j,t+h} \right]. \quad (7)$$

- **Ex-ante perspective** conceptually distinct from *pure random forecast error*, since it measures **systematic expectational errors**, not ex-post *mistakes*.

Data

Surveys

- We consider three surveys for real GDP growth and inflation
  1. Survey of Professional Forecasters (SPF)
  2. Michigan Survey of Consumers (SOC)
  3. Blue Chip (BC)
- For each survey/variable we consider mean, median, and several percentiles
  - Exception: SOC's GDP only forecast is constructed from a qualitative balance score

Real Time Data-Rich Environment

At each forecast date, we construct a **rich dataset** of variables observed on or before the day of the survey deadline.

- **Macro Factors** use **92 real-time macro variables** (Philadelphia Fed)
- **Energy prices** (BLS)
- **Monthly financial factors** $\mathcal{D}^F$ use **147 monthly financial series**
- **Daily financial factors** $\mathcal{D}^D$ cover **87 daily indicators**
- **Lagged moments of SPF responses**
- **Additional variables** such as detrended output, trend inflation, term-structure slope,...

Results: How Distorted are Beliefs?

Forecast Comparison: Machine v.s. SPF Inflation

Machine learning v.s. SPF forecasts of inflation

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

Percentile	Median	Mean	5th	10th	20th
$MSE_{\mathbb{E}}$	0.38	0.42	0.51	0.43	0.40
$MSE_{\mathbb{F}}$	0.45	0.44	0.90	0.58	0.48
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.85	0.95	0.56	0.74	0.83
	25th	30th	40th	60th	70th
$MSE_{\mathbb{E}}$	0.41	0.40	0.39	0.36	0.39
$MSE_{\mathbb{F}}$	0.45	0.45	0.44	0.49	0.56
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.90	0.88	0.89	0.74	0.70
	75th	80th	90th	95th	
$MSE_{\mathbb{E}}$	0.41	0.41	0.53	0.65	
$MSE_{\mathbb{F}}$	0.61	0.70	0.96	1.36	
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.67	0.59	0.55	0.47	

Relative Mean-square-forecast-errors. $MSE_{\mathbb{E}}$ and $MSE_{\mathbb{F}}$ are the machine learning benchmark and survey mean-squared-forecast-errors, respectively. Forecast errors are for a 4-Quarter ahead forecast and averaged over the evaluation period. The sample is 1969:Q3 to 2018:Q3.

Forecast Comparison: Machine v.s. SPF Inflation

- Machine model performs much better suggesting belief distortion, **even in heralded consensus π forecasts**.

Machine learning v.s. SPF forecasts of inflation

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

Percentile	Median	Mean	5th	10th	20th
$MSE_{\mathbb{E}}$	0.38	0.42	0.51	0.43	0.40
$MSE_{\mathbb{F}}$	0.45	0.44	0.90	0.58	0.48
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.85	0.95	0.56	0.74	0.83
	25th	30th	40th	60th	70th
$MSE_{\mathbb{E}}$	0.41	0.40	0.39	0.36	0.39
$MSE_{\mathbb{F}}$	0.45	0.45	0.44	0.49	0.56
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.90	0.88	0.89	0.74	0.70
	75th	80th	90th	95th	
$MSE_{\mathbb{E}}$	0.41	0.41	0.53	0.65	
$MSE_{\mathbb{F}}$	0.61	0.70	0.96	1.36	
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.67	0.59	0.55	0.47	

Relative Mean-square-forecast-errors. $MSE_{\mathbb{E}}$ and $MSE_{\mathbb{F}}$ are the machine learning benchmark and survey mean-squared-forecast-errors, respectively. Forecast errors are for a 4-Quarter ahead forecast and averaged over the evaluation period. The sample is 1969:Q3 to 2018:Q3.

Forecast Comparison: Machine v.s. SPF GDP Growth

- For GDP growth, machine model always more accurate than SPF respondents, **no matter what their beliefs**.

Machine learning v.s. SPF forecasts of GDP growth

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

Percentile	Median	Mean	5th	10th	20th
$MSE_{\mathbb{E}}$	2.35	2.41	2.22	2.27	2.12
$MSE_{\mathbb{F}}$	2.63	2.60	3.10	2.73	2.59
$MSE_{\mathbb{E}} / MSE_{\mathbb{F}}$	0.89	0.93	0.72	0.83	0.82
	25th	30th	40th	60th	70th
$MSE_{\mathbb{E}}$	2.21	2.28	2.34	2.34	2.31
$MSE_{\mathbb{F}}$	2.57	2.57	2.60	2.70	2.81
$MSE_{\mathbb{E}} / MSE_{\mathbb{F}}$	0.86	0.89	0.90	0.87	0.82
	75th	80th	90th	95th	
$MSE_{\mathbb{E}}$	2.32	2.43	2.39	2.54	
$MSE_{\mathbb{F}}$	2.87	2.96	3.38	3.90	
$MSE_{\mathbb{E}} / MSE_{\mathbb{F}}$	0.81	0.82	0.71	0.65	

Relative Mean-square-forecast-errors. $MSE_{\mathbb{E}}$ and $MSE_{\mathbb{F}}$ are the machine learning benchmark and survey mean-squared-forecast-errors, respectively. Forecast errors are for a 4-Quarter ahead forecast and averaged over the evaluation period. The estimation sample is 1969:Q3 to 2018:Q3.

Forecast Comparison: Machine v.s. SOC Inflation

- **Much larger average belief distortions for households**

Machine learning v.s. SOC forecasts of inflation

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

Percentile	Median	Mean	5th	10th	20th
$MSE_{\mathbb{E}}$	1.64	1.97	3.25	2.28	1.76
$MSE_{\mathbb{F}}$	2.84	4.65	15.11	8.27	3.87
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.58	0.42	0.22	0.28	0.46
	25th	30th	40th	60th	70th
$MSE_{\mathbb{E}}$	1.83	1.76	1.50	1.75	1.70
$MSE_{\mathbb{F}}$	3.16	2.62	2.30	4.69	8.26
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.58	0.67	0.65	0.37	0.21
	75th	80th	90th	95th	
$MSE_{\mathbb{E}}$	1.65	1.83	2.51	2.72	
$MSE_{\mathbb{F}}$	10.62	15.05	47.03	84.92	
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.16	0.12	0.05	0.03	

Relative Mean-square-forecast-errors. $MSE_{\mathbb{E}}$ and $MSE_{\mathbb{F}}$ are the machine learning benchmark and survey mean-squared-forecast-errors, respectively. Forecast errors are for a 4-Quarter ahead forecast and averaged over the evaluation period. The estimation sample is 1981:Q3 to 2018:Q3.

Forecast Comparison: Machine v.s. SOC GDP Growth

- **Similarly large distortions for SOC forecasts of economic growth**

Machine learning v.s. SOC forecasts of GDP Growth

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

Percentile	Median
$\text{MSE}_{\mathbb{E}}$	2.41
$\text{MSE}_{\mathbb{F}}$	3.24
$\text{MSE}_{\mathbb{E}} / \text{MSE}_{\mathbb{F}}$	0.74

Relative Mean-square-forecast-errors. $MSE_{\mathbb{E}}$ and $MSE_{\mathbb{F}}$ are the machine learning benchmark and survey mean-squared-forecast-errors, respectively. Forecast errors are for a 4-Quarter ahead forecast and averaged over the evaluation period. The estimation sample is 1978:Q1 to 2018:Q3.

Forecast Comparison: Machine v.s. BC Inflation

- **Machine model also improves over BC survey forecasts**

Machine learning v.s. BC forecasts of inflation

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

Percentile	Median	Mean	5th	10th	20th
$MSE_{\mathbb{E}}$	0.41	0.40	0.48	0.40	0.43
$MSE_{\mathbb{F}}$	0.49	0.48	0.83	0.67	0.51
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.84	0.84	0.58	0.60	0.85
	25th	30th	40th	60th	70th
$MSE_{\mathbb{E}}$	0.42	0.41	0.43	0.40	0.39
$MSE_{\mathbb{F}}$	0.49	0.47	0.47	0.51	0.56
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.85	0.86	0.91	0.78	0.69
	75th	80th	90th	95th	
$MSE_{\mathbb{E}}$	0.39	0.38	0.42	0.44	
$MSE_{\mathbb{F}}$	0.59	0.65	0.87	1.14	
$MSE_{\mathbb{E}}/MSE_{\mathbb{F}}$	0.65	0.59	0.48	0.38	

Relative Mean-square-forecast-errors. $MSE_{\mathbb{E}}$ and $MSE_{\mathbb{F}}$ are the machine learning benchmark and survey mean-squared-forecast-errors, respectively. Forecast errors are for a 4-Quarter ahead forecast and averaged over the evaluation period. The estimation sample is 1981:Q3 to 2018:Q3.

Forecast Comparison: Machine v.s. BC GDP Growth

- **Similarly large distortions for BC forecasts of economic growth**

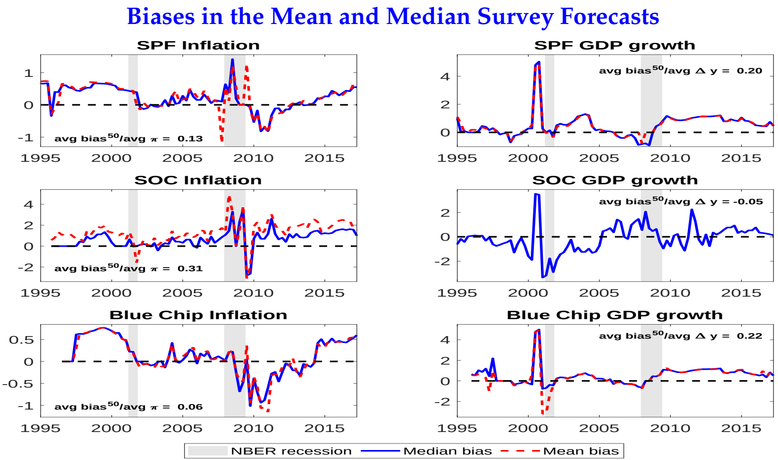
Machine learning v.s. BC forecasts of GDP Growth

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

Percentile	Median	Mean	5th	10th	20th
$MSE_{\mathbb{E}}$	2.14	2.29	2.29	2.08	2.37
$MSE_{\mathbb{F}}$	2.81	2.77	2.97	2.76	2.67
$MSE_{\mathbb{E}} / MSE_{\mathbb{F}}$	0.76	0.83	0.77	0.75	0.89
	25th	30th	40th	60th	70th
$MSE_{\mathbb{E}}$	2.19	2.20	2.11	2.20	2.29
$MSE_{\mathbb{F}}$	2.68	2.71	2.76	2.89	2.99
$MSE_{\mathbb{E}} / MSE_{\mathbb{F}}$	0.82	0.81	0.77	0.76	0.73
	75th	80th	90th	95th	
$MSE_{\mathbb{E}}$	2.23	2.23	2.28	2.54	
$MSE_{\mathbb{F}}$	3.06	3.18	3.50	3.82	
$MSE_{\mathbb{E}} / MSE_{\mathbb{F}}$	0.70	0.65	0.67	0.66	

Relative Mean-square-forecast-errors. $MSE_{\mathbb{E}}$ and $MSE_{\mathbb{F}}$ are the machine learning benchmark and survey mean-squared-forecast-errors, respectively. Forecast errors are for a 4-Quarter ahead forecast and averaged over the evaluation period. The estimation sample is 1978:Q1 to 2018:Q3.

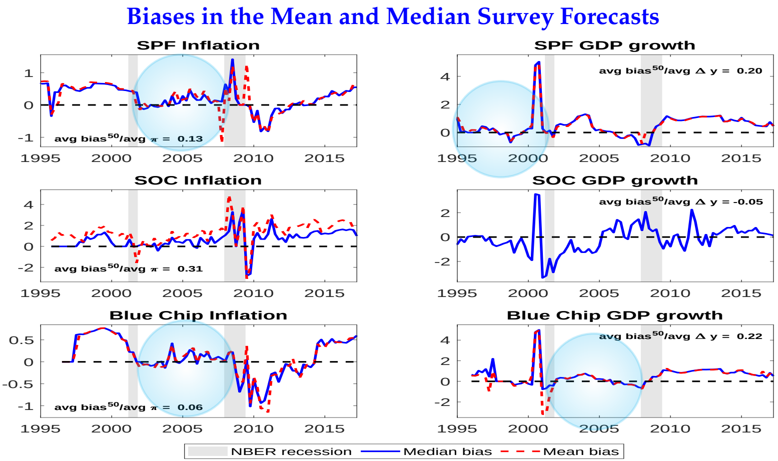
Dynamics of Belief Distortions



Biases in the consensus forecasts. The figure reports the time series $bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t \left[y_{j,t+h}\right] - \mathbb{E}^{(i)}_t \left[y_{j,t+h}\right]$ for $i = 50, mean$. The sample spans the period 1995:Q1-2018:Q2.

Dynamics of Belief Distortions

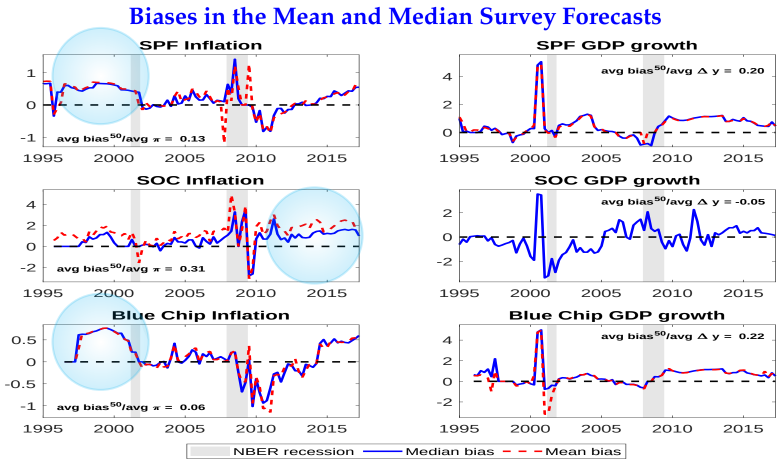
- For SPF and BC, several periods with no or small biases.



Biases in the consensus forecasts. The figure reports the time series $bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t\left[y_{j,t+h}\right] - \mathbb{E}^{(i)}_t\left[y_{j,t+h}\right]$ for $i = 50, mean$. The sample spans the period 1995:Q1-2018:Q2.

Dynamics of Belief Distortions

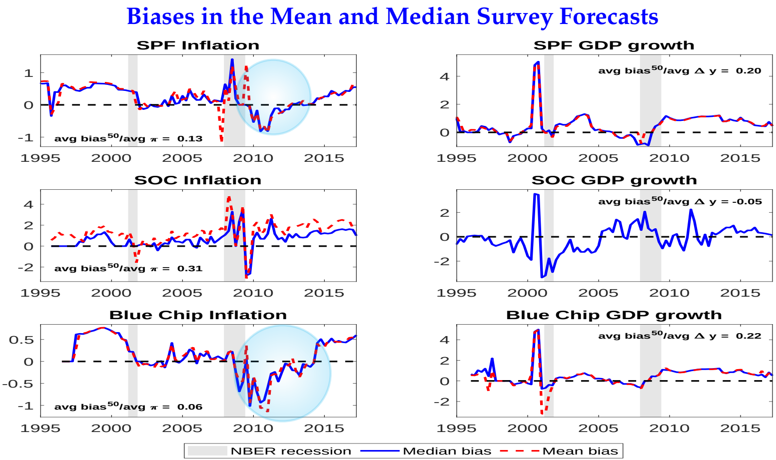
- In all surveys, periods where π expectations were systematically biased upward...



Biases in the consensus forecasts. The figure reports the time series $bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t\left[y_{j,t+h}\right] - \mathbb{E}^{(i)}_t\left[y_{j,t+h}\right]$ for $i = 50, mean$. The sample spans the period 1995:Q1-2018:Q2.

Dynamics of Belief Distortions

- ...and systematically biased downward.

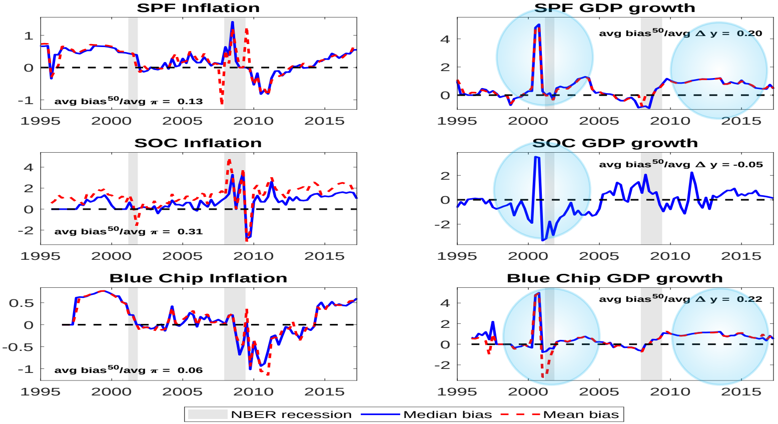


Biases in the consensus forecasts. The figure reports the time series $bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t \left[y_{j,t+h}\right] - \mathbb{E}^{(i)}_t \left[y_{j,t+h}\right]$ for $i = 50, mean$. The sample spans the period 1995:Q1-2018:Q2.

Dynamics of Belief Distortions

- For GDP growth, all surveys show extended periods of **over-optimism**, especially post Great Recession, and right before the 2001 recession.

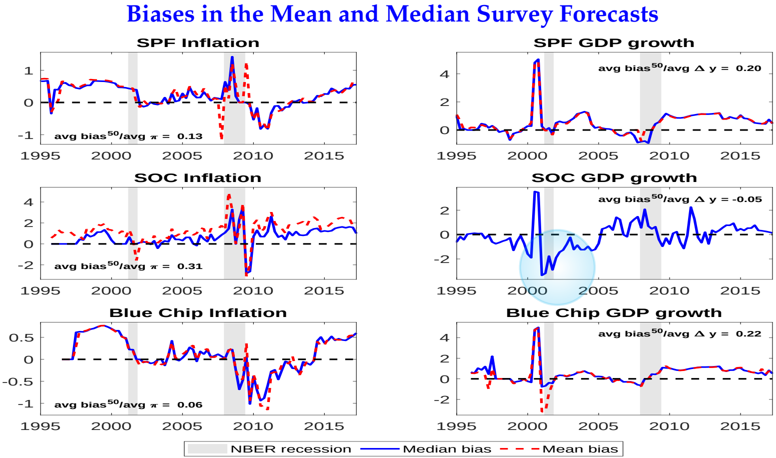
Biases in the Mean and Median Survey Forecasts



Biases in the consensus forecasts. The figure reports the time series $bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t\left[y_{j,t+h}\right] - \mathbb{E}^{(i)}_t\left[y_{j,t+h}\right]$ for $i = 50, mean$. The sample spans the period 1995:Q1-2018:Q2.

Dynamics of Belief Distortions

- HHs' SOC: **over-pessimism** right after 2001 recession

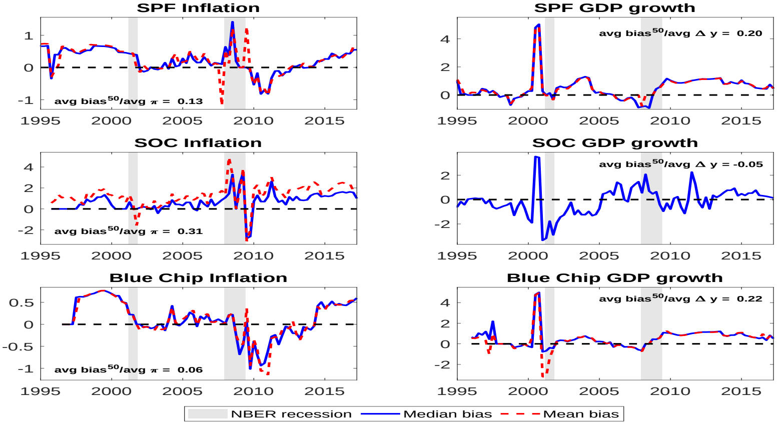


Biases in the consensus forecasts. The figure reports the time series $bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t \left[y_{j,t+h}\right] - \mathbb{E}^{(i)}_t \left[y_{j,t+h}\right]$ for $i = 50, mean$. The sample spans the period 1995:Q1-2018:Q2.

Dynamics of Belief Distortions

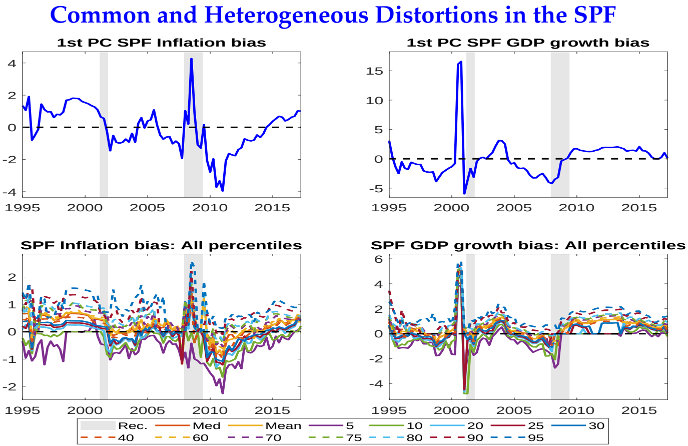
- **Biases large!** (Especially in SOC): inflation biases => systematic errors on order of 0.5-2%; GDP growth by 1-4%.

Biases in the Mean and Median Survey Forecasts



Biases in the consensus forecasts. The figure reports the time series $bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t \left[y_{j,t+h}\right] - \mathbb{E}^{(i)}_t \left[y_{j,t+h}\right]$ for $i = 50, mean$. The sample spans the period 1995:Q1-2018:Q2.

Heterogeneity: SPF

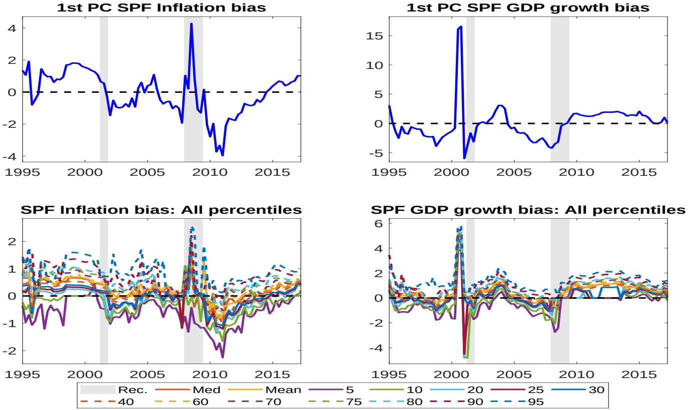


Biases in the SCF. The figure reports the time series $bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t \left[y_{j,t+h}\right] - \mathbb{E}^{(i)}_t \left[y_{j,t+h}\right]$. The sample is 1995:Q1-2018:Q2.

Heterogeneity: SPF

- First principal component of bias shows *common* distortions

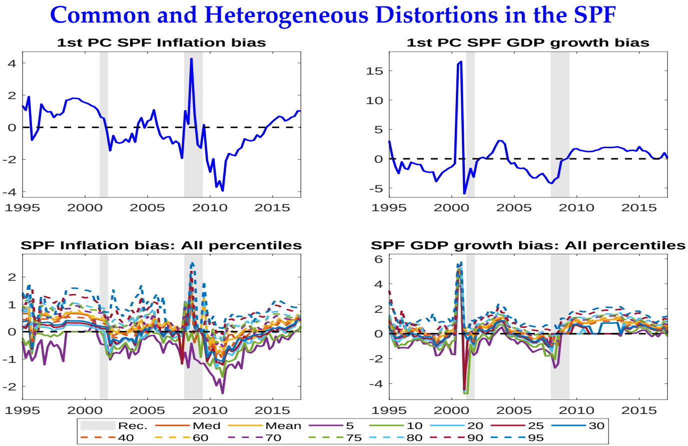
Common and Heterogeneous Distortions in the SPF



Biases in the SCF. The figure reports the time series $bias_{j,t+h}^{(i)} = \mathbb{F}_t^{(i)}\left[y_{j,t+h}\right] - \mathbb{E}_t^{(i)}\left[y_{j,t+h}\right]$. The sample is 1995:Q1-2018:Q2.

Heterogeneity: SPF

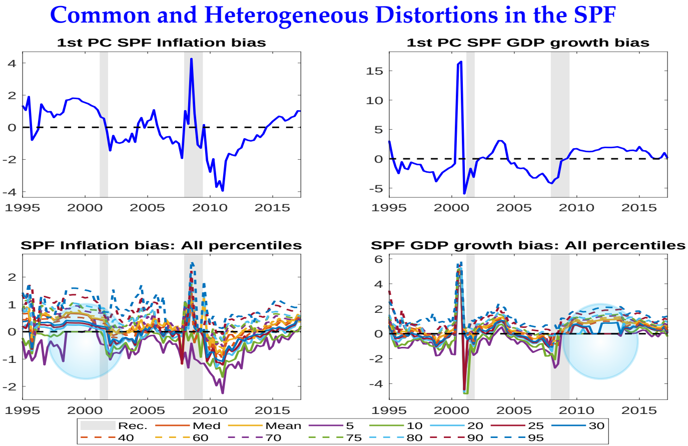
- Substantial **heterogeneity in beliefs**.



Biases in the SCF. The figure reports the time series $bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t \left[y_{j,t+h}\right] - \mathbb{E}^{(i)}_t \left[y_{j,t+h}\right]$. The sample is 1995:Q1-2018:Q2.

Heterogeneity: SPF

- Ex-ante perspective implies that more than one forecaster can show no bias.

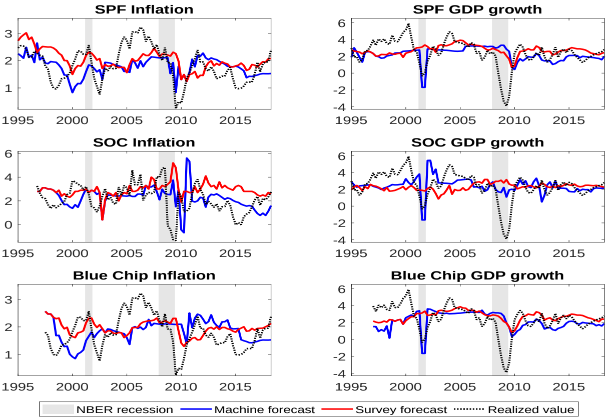


Biases in the SCF. The figure reports the time series $bias^{(i)}_{j,t+h} = \mathbb{F}^{(i)}_t \left[y_{j,t+h}\right] - \mathbb{E}^{(i)}_t \left[y_{j,t+h}\right]$. The sample is 1995:Q1-2018:Q2.

Forecasts v.s. Actual

- *Ex-post forecast mistakes* distinct from ex ante expectational error.

Forecasted versus Actual Inflation, GDP Growth

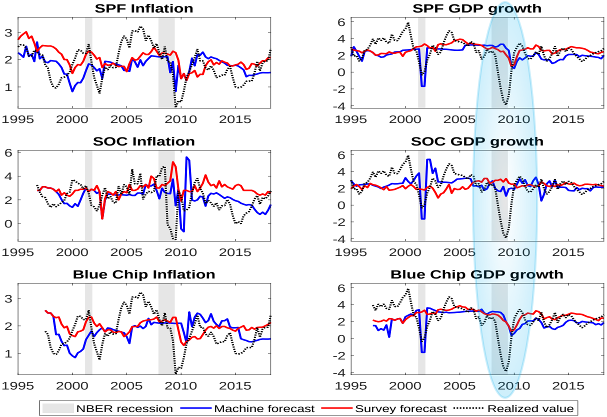


Forecasted and Actual variables. For each variable and survey, the figure reports the median survey forecast of inflation or GDP growth over the next 4 quarters, the corresponding the machine forecast, and the actual inflation or GDP growth during this period. The sample is 1995:Q1-2018:Q2.

Forecasts v.s. Actual

- Underscores the distinction between *luck* and systematic bias. **Large forecast errors *per se* not evidence of bias (e.g. GR)**.

Forecasted versus Actual Inflation, GDP Growth

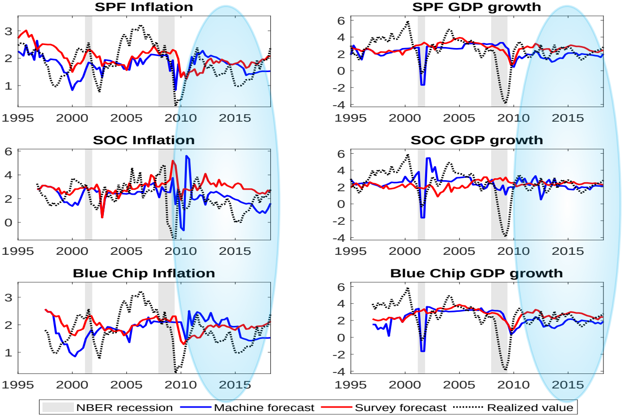


Forecasted and Actual variables. For each variable and survey, the figure reports the median survey forecast of inflation or GDP growth over the next 4 quarters, the corresponding the machine forecast, and the actual inflation or GDP growth during this period. The sample is 1995:Q1-2018:Q2.

Forecasts v.s. Actual

- **SPF: machine more accurate in last 5 years**, while surveys overpredicted π and GDP growth.

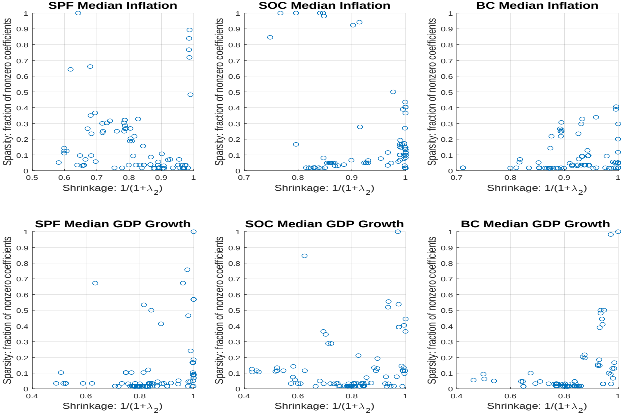
Forecasted versus Actual Inflation, GDP Growth



Forecasted and Actual variables. For each variable and survey, the figure reports the median survey forecast of inflation or GDP growth over the next 4 quarters, the corresponding the machine forecast, and the actual inflation or GDP growth during this period. The sample is 1995:Q1-2018:Q2.

Degree of Sparsity and Shrinkage

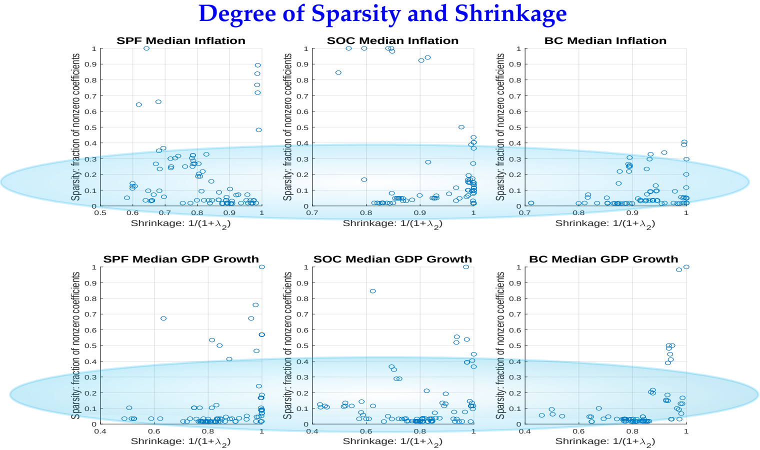
Degree of Sparsity and Shrinkage



Scatterplots of ridge and LASSO penalties. For each observation in the evaluation sample from 1995:1-2018:Q2, the y-axis displays the degree of sparsity implied by the estimated Lasso penalty, λ_1, in units of the fraction of non-zero regression coefficients, and the x-axis displays the degree of shrinkage implied by the estimated L^2 penalty, λ_2 in units of $1/\left(1+\lambda_2\right)$.

Degree of Sparsity and Shrinkage

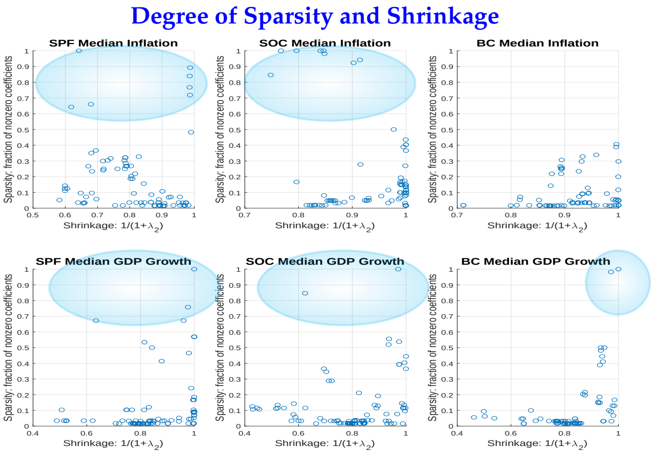
- Machine often selects **only a few variables**, but...



Scatterplots of ridge and LASSO penalties. For each observation in the evaluation sample from 1995:1-2018:Q2, the y-axis displays the degree of sparsity implied by the estimated Lasso penalty, λ_1, in units of the fraction of non-zero regression coefficients, and the x-axis displays the degree of shrinkage implied by the estimated L^2 penalty, λ_2 in units of $1/\left(1+\lambda_2\right)$.

Degree of Sparsity and Shrinkage

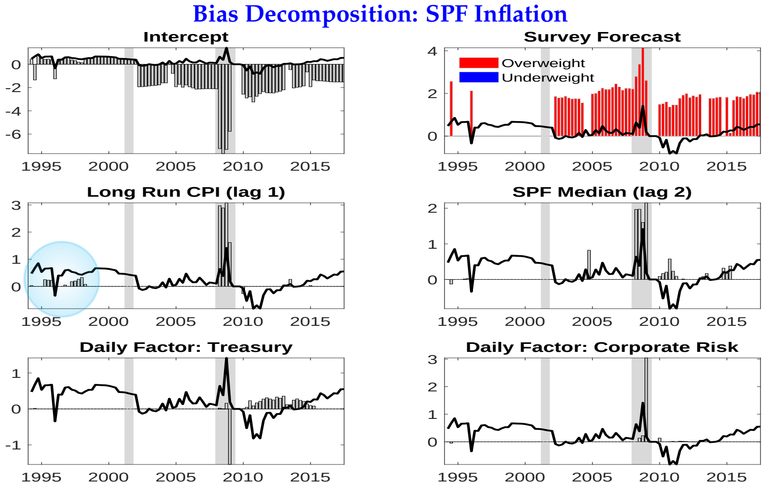
- ...there are **several exceptions**



Scatterplots of ridge and LASSO penalties. For each observation in the evaluation sample from 1995:1-2018:Q2, the y-axis displays the degree of sparsity implied by the estimated Lasso penalty, λ_1, in units of the fraction of non-zero regression coefficients, and the x-axis displays the degree of shrinkage implied by the estimated L^2 penalty, λ_2 in units of $1/\left(1+\lambda_2\right)$.

Bias Decomposition: SPF Inflation

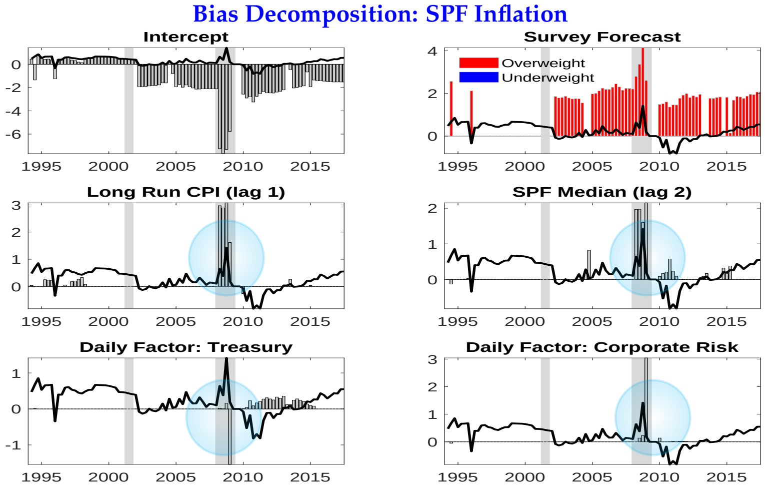
- The selected variables **change over time**



The figures plots contributors to the median bias $\mathbb{F}_t^{(50)}\left[y_{j,t+h}\right] - \mathbb{E}_t^{(50)}\left[y_{j,t+h}\right] = -\hat{\alpha}_j^{(50)} + \left(1 - \hat{\beta}_{j\mathbb{F}}^{(50)}\right)\mathbb{F}_t^{(50)}\left[y_{j,t+h}\right] - \hat{\mathrm{B}}_{j\mathcal{Z}}^{(50)\prime}\mathcal{Z}_{jt}$ at each time t.

Bias Decomposition: SPF Inflation

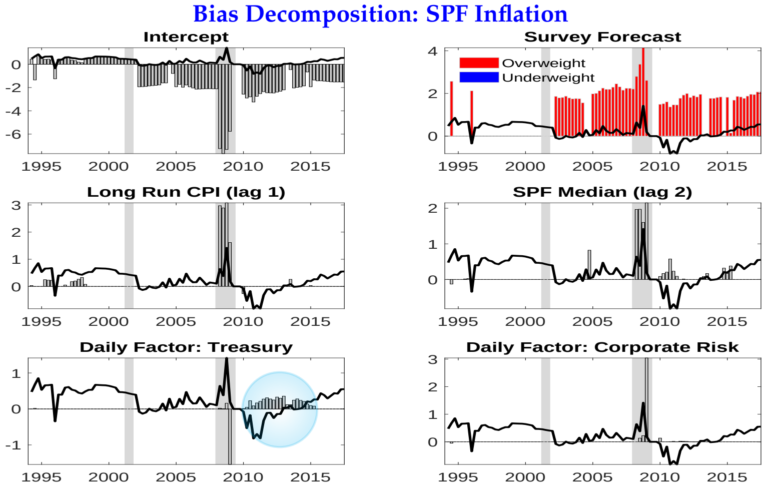
- The selected variables **change over time**



The figures plots contributors to the median bias $\mathbb{F}_t^{(50)}\left[y_{j,t+h}\right] - \mathbb{E}_t^{(50)}\left[y_{j,t+h}\right] = -\hat{\alpha}_j^{(50)} + \left(1 - \hat{\beta}_{j\mathbb{F}}^{(50)}\right)\mathbb{F}_t^{(50)}\left[y_{j,t+h}\right] - \hat{\mathrm{B}}_{j\mathcal{Z}}^{(50)\prime}\mathcal{Z}_{jt}$ at each time t.

Bias Decomposition: SPF Inflation

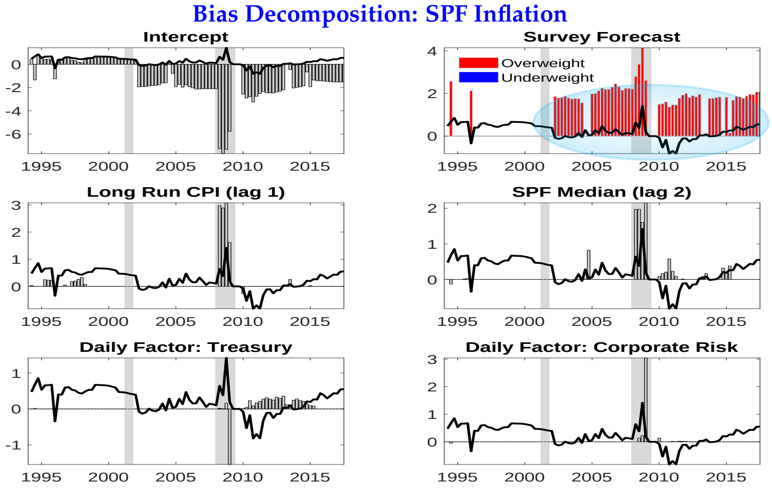
- The selected variables **change over time**



The figures plots contributors to the median bias $\mathbb{F}_t^{(50)}\left[y_{j,t+h}\right] - \mathbb{E}_t^{(50)}\left[y_{j,t+h}\right] = -\hat{\alpha}_j^{(50)} + \left(1 - \hat{\beta}_{j\mathbb{F}}^{(50)}\right)\mathbb{F}_t^{(50)}\left[y_{j,t+h}\right] - \hat{B}_{j\mathcal{Z}}^{(50)\prime}\mathcal{Z}_{jt}$ at each time t.

Bias Decomposition: SPF Inflation

- Forecasters **over-confident** about their own forecasts



The figures plots contributors to the median bias $\mathbb{F}_t^{(50)}\left[y_{j,t+h}\right] - \mathbb{E}_t^{(50)}\left[y_{j,t+h}\right] = -\hat{\alpha}_j^{(50)} + \left(1 - \hat{\beta}_{j\mathbb{F}}^{(50)}\right)\mathbb{F}_t^{(50)}\left[y_{j,t+h}\right] - \hat{\mathrm{B}}_{j\mathcal{Z}}^{(50)\prime}\mathcal{Z}_{jt}$ at each time t.

Results: Do Belief Distortions Matter for Macro Fluctuations?

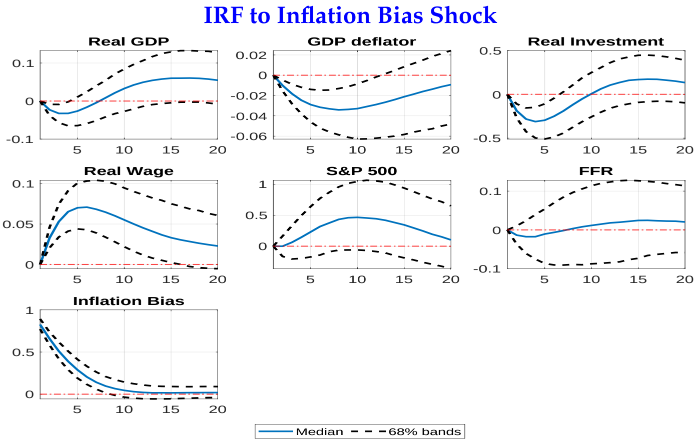
Do Belief Distortions Matter for Macro Fluctuations?

- Illustrate using VAR analysis, with 1st PC of belief distortion, $\overline{bias}_{j,t+h}$ across all surveys and percentiles
- Separate VARs for inflation and Real GDP biases
- VAR variables: real GDP, GDP deflator, real investment, real wage, S&P 500, FFR, $\overline{bias}_{j,t+h}$.

Do Belief Distortions Matter for Macro Fluctuations?

- Illustrate using VAR analysis, with 1st PC of belief distortion, $\overline{bias}_{j,t+h}$ across all surveys and percentiles
- Separate VARs for inflation and Real GDP biases
- VAR variables: real GDP, GDP deflator, real investment, real wage, S&P 500, FFR, $\overline{bias}_{j,t+h}$.
- COV matrix of VAR residuals orthogonalized using a Cholesky decomposition with variables ordered as listed above.
  - Conservative approach: Bias can affect other variables only with a lag.
- A **bias shock** is a movement in belief distortions that is orthogonal to the aggregate economic state.
- We use a Bayesian approach with flat priors.

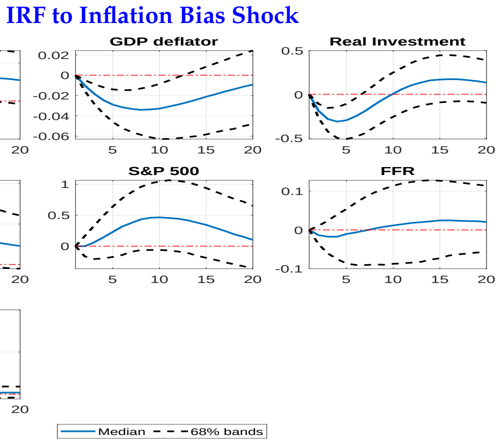
Responses to an Inflation Bias Shock, SPF and BC



Impulse responses to a one standard deviation inflation bias shock. Estimates from quarterly VAR with one lag. An increase in the bias means that forecasters systematically over-predict future inflation. Units are in percentage points. The sample is 1995:Q1-2018:Q2.

Responses to an Inflation Bias Shock, SPF and BC

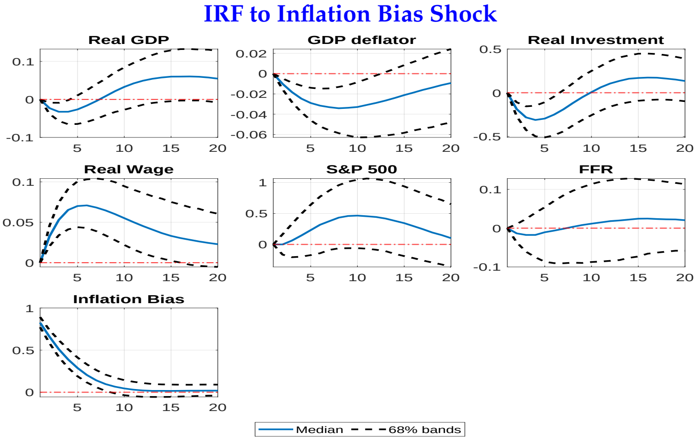
- Systematic over-prediction of π that is *orthogonal* to economic state **is associated with a recession**. GDP growth, investment and stock market fall.



Impulse responses to a one standard deviation inflation bias shock. Estimates from quarterly VAR with one lag. An increase in the bias means that forecasters systematically over-predict future inflation. Units are in percentage points. The sample is 1995:Q1-2018:Q2.

Responses to an Inflation Bias Shock, SPF and BC

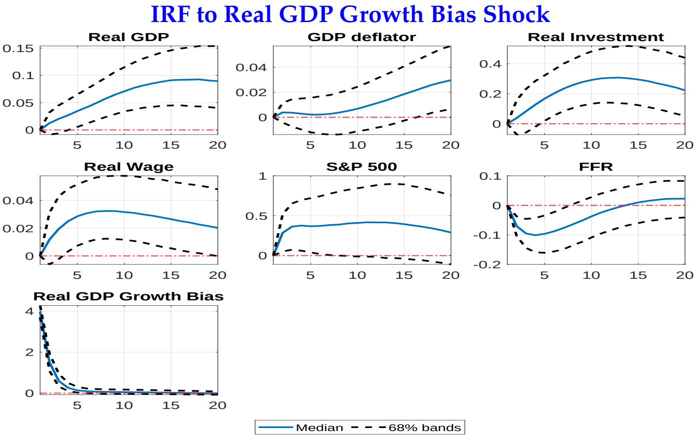
- Why? One story: note real wages rise, so this is tantamount to an **adverse cost-push shock**.



Impulse responses to a one standard deviation inflation bias shock. Estimates from quarterly VAR with one lag. An increase in the bias means that forecasters systematically over-predict future inflation. Units are in percentage points. The sample is 1995:Q1-2018:Q2.

VAR: Real GDP Growth Bias Shock

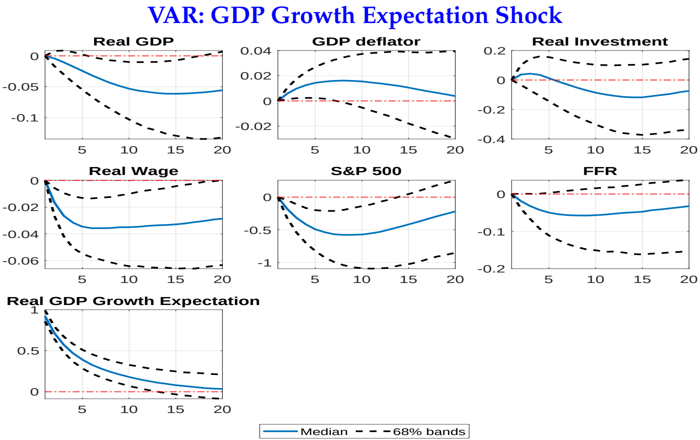
- Systematic over-prediction of GDP growth that is *orthogonal* to economic state **is associated with an expansion**. GDP growth, investment and stock market increase.



Impulse responses to a one standard deviation GDP growth bias shock. Estimates from a quarterly VAR with one lag. An increase in the bias means that forecasters overestimate future Real GDP growth. Units are percentage points. The sample is 1995:Q1-2018:Q2.

VAR: GDP Growth Expectation Shock

- A positive shock to **expected GDP growth** has the opposite effect, causing a **recession**



Impulse responses to a one standard deviation GDP expectation shock. Estimates from a quarterly VAR with one lag. The GDP growth expectation $\overline{\mathbb{F}}_t^{\Delta y}$ is constructed as the first principle component of GDP growth survey forecast across all surveys and percentiles.

Conclusions

We provide new measures of **systematic expectational errors** in survey responses and relates them to macroeconomic activity.

- We confront the two challenges faced by forecasters
  1. **Out-of-sample** nature of the forecast
  2. **Data-rich environment**
  3. **Heterogeneity in beliefs**

  Separate **luck** from **systematic expectational errors**
- Large improvements of the machine over the surveys: **Large biases**
- Across all surveys, respondents appear to be **overconfident**, placing too much weight on their own forecast relative to other information
- Significant **variation in the selection** of variables over time.
- Fluctuations in belief distortions exhibit **important dynamic relations with the macroeconomy**.

Afrouzi, H., L. Veldkamp, et al. (2019): "Biased Inflation Forecasts," in *2019 Meeting Papers*, no. 894. Society for Economic Dynamics.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015): "X-CAPM: An extrapolative capital asset pricing model," *Journal of financial economics*, 115(1), 1–24.

Barberis, N., A. Shleifer, and R. W. Vishny (1998): "A Model of Investor Sentiment," *Journal of Financial Economics*, 49(3), 307–43.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018): "Diagnostic expectations and credit cycles," *The Journal of Finance*, 73(1), 199–227.

Chauvet, M., and S. Potter (2013): "Forecasting output," in *Handbook of Economic Forecasting*, vol. 2, pp. 141–194. Elsevier.

Coibion, O., and Y. Gorodnichenko (2015): "Information rigidity and the expectations formation process: A simple framework and new facts," *American Economic Review*, 105(8), 2644–78.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990): "Positive feedback investment strategies and destabilizing rational speculation," *the Journal of Finance*, 45(2), 379–395.

Epstein, L. G., and M. Schneider (2010): "Ambiguity and asset markets," *Annual Review of Financial Economics*, 2, 315–346.

Gennaioli, N., and A. Shleifer (2018): *A crisis of beliefs: Investor psychology and financial fragility*. Princeton University Press.

HANSEN, L. P., AND T. J. SARGENT (2008): *Robustness*. Princeton university press.

ILUT, C. L., AND M. SCHNEIDER (2015): "Ambiguous Business Cycles," *American Economic Review,* forthcoming.

MALMENDIER, U., AND S. NAGEL (2011): "Depression babies: do macroeconomic experiences affect risk taking?," *The Quarterly Journal of Economics*, 126(1), 373–416.

——— (2015): "Learning from inflation experiences," *The Quarterly Journal of Economics*, 131(1), 53–87.