Subjective Beliefs in Asset Pricing

Stefan Nagel

University of Chicago,
NBER, CEPR, and CESIfo

June 2021

Subjective beliefs in asset pricing

- Asset prices are forward looking: Beliefs about future payoffs crucial
- Basic asset pricing relationship

$$P_t = \tilde{\mathbb{E}}[M_{t+1}(D_{t+1} + P_{t+1})|\boldsymbol{x}_t]$$

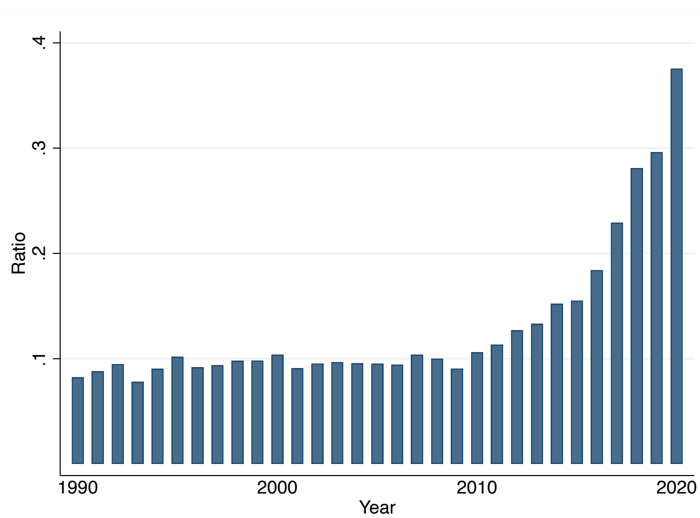
  for asset with payoff D_{t+1} and stochastic discount factor M_{t+1} reflecting investors' preferences.
- Investors' subjective beliefs $\tilde{\mathbb{E}}[.|\boldsymbol{x}_t]$ based on vector of variables $\boldsymbol{x}_t$ they observe
  - Expectations about future fundamentals
  - Expectations about future beliefs of investors pricing the asset at $t+1$
  - Perceptions of risk

Subjective beliefs in asset pricing

- Yet study of investor beliefs was until recently mostly absent in asset pricing
- Still dominant paradigm is rational expectations (RE) = agents know model, parameter values, and forecast rationally
  - Given information $\boldsymbol{x}_t$ they know how to form $\mathbb{E}[.|\boldsymbol{x}_t]$
- RE convenient: Investor beliefs implied by (known) true fundamentals process. No need for separate specification (and empirical study of) investor beliefs.
- But RE potentially overlooks a main driver of asset price variation

Subjective beliefs in asset pricing: Google Scholar

$$\text{Ratio} = \frac{\#\text{ papers with mention of "investor beliefs"}}{\#\text{ papers with mention of "asset prices"}}$$



Agenda

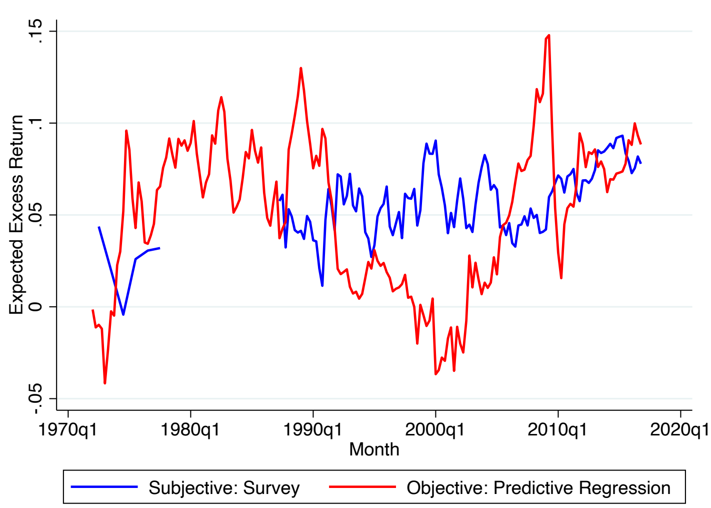
1. Illustration of role of subjective beliefs in asset pricing
2. Belief formation based on historical experience
3. Belief formation when observed data are high-dimensional
4. Conclusion

1. Illustration of role of subjective beliefs in asset pricing

Example: Aggregate stock market valuation

- Return predictability evidence: price/dividend ratio
  - predicts stock market excess returns
  - does not predict future dividend growth
- Standard RE explanations: Countercyclical time-varying expected excess returns due to
  - time-varying risk aversion (Campbell and Cochrane 1999)
  - time-varying risk (Bansal and Yaron 2004)
- How do these RE-implied beliefs compare with survey measures of beliefs about future excess returns?

Disconnect between objective and subjective expected stock market excess returns



Source: Subjective = one-year expected stock market returns in excess of one-year Treasury yield from various individual investor surveys in Nagel and Xu (2021). Objective = Fitted value from predictive regression of stock market excess returns on repurchase-adjusted log price-dividend ratio estimated on quarterly data 1927-2019.

Disconnect between objective and subjective expected bond market excess returns

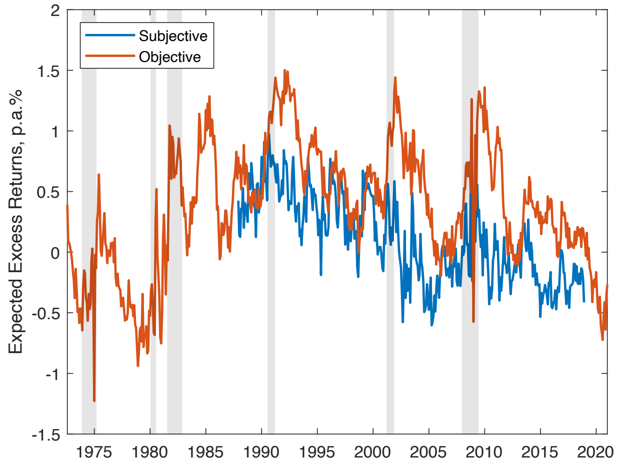


FIGURE 2
Subjective and Objective Bond Risk Premium

The blue line plots the survey-implied subjective bond risk premium from Xu (2020). The red line plots the expected bond risk premium from using the Ludvigson and Ng (2009) factor and the Cieslak and Povala (2015) factor as predictors. We use the maturity-weighted average log excess returns on zero-coupon bonds from two-year to ten-year. The zero-coupon yields are from Liu and Wu (forthcoming) and the survey forecasts are from the Blue Chip Financial Forecasts.

Investor belief formation: Parameter learning

- Once we discard RE, we need model of belief formation
- How do agents get from observed data, $\boldsymbol{x}_t$, to beliefs

$$\tilde{\mathbb{E}}[y_{t+1}|\boldsymbol{x}_t] = f(\boldsymbol{x}_t; \boldsymbol{g})$$

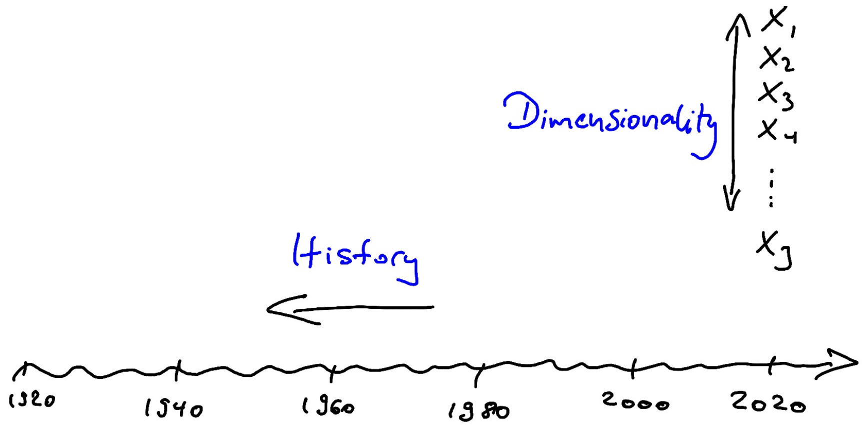
- In this talk, I assume that agents approximate $f(\boldsymbol{x}_t; \boldsymbol{g})$ with linear models

$$y_{t+1} = \boldsymbol{x}_t' \boldsymbol{g} + e_{t+1}$$

- Therefore, focus on formation of beliefs about $\boldsymbol{g}$

What is the information that influences beliefs?

- Typical economic models: Clearly defined, small info set
- Reality for decision makers: Messy, huge amount of potentially, but not necessarily relevant info



2. Belief formation based on historical experience

Macroeconomic experiences and macroeconomic expectations: Example of inflation

- *"An entire generation of young adults has grown up since the mid-1960s knowing* **only inflation** *... In the circumstances, it is hardly surprising that many citizens have begun to wonder whether it is realistic to anticipate a return to general price stability... "* — *Paul Volcker,* ***June 1979***.

- *"There is an entire generation of traders who have grown up investing in the post-global financial crisis world of* **no inflation** *... People shouldn't underestimate how uncertain things will look if we are entering a new paradigm."*
  — *Sonal Desai, CIO, Franklin Templeton,* ***May 2021*** .

Macroeconomic experiences and macroeconomic expectations

- Many empirical properties of macroeconomic expectations can be understood through a model of experience-based expectations formation in which individuals
  - draw on lifetime experiences of macroeconomic data to form expectations
  - put higher weight more recently experienced data
- Experience-based expectations formation can help understand what standard macro models struggle to explain
  - Asset price volatility
  - Credit booms and busts
  - Forward guidance puzzle

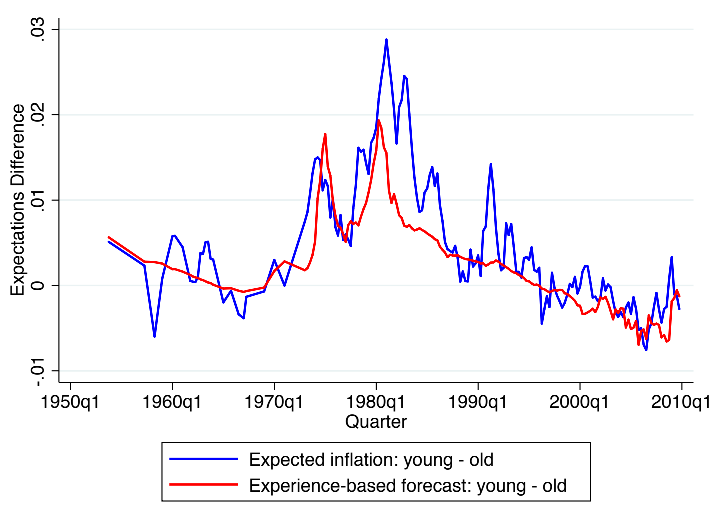
Adaptive parameter learning from life-time experience

- Individuals forecast inflation, π_{t+1}, with AR(1) model

$$\hat{\pi}_{t+1|t} = \tilde{\alpha}_t + \tilde{\phi}_t \pi_t$$

- Parameter estimates $\tilde{\alpha}_t$ and $\tilde{\phi}_t$ obtained at t with
  - least squares regression
  - on individuals' life-time data set up to time t
  - more weight on recent data (half-life $\approx$ 10 years)

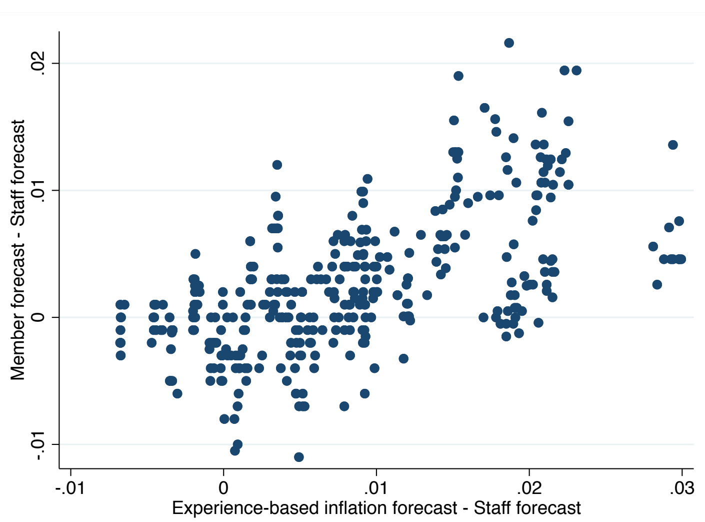
Differences in inflation expectations and experiences



Based on Malmendier and Nagel (2016). **Inflation Expectations**: Michigan Survey of Consumers, one-year expected inflation rate. **Experience-based forecast**: AR(1) model forecast estimated based on weighted life-time inflation data for each survey respondent. Figure shows **differences**: average for individuals of age < 40 minus average for individuals of age > 60.

Out-of-sample evidence from NY Fed Survey

Hawks and doves on the FOMC: Experienced inflation and individual member inflation forecasts



From Malmendier, Nagel, and Yan (2021). **Member forecast**: from semi-annual Monetary Policy Report to Congress, 1992 - 2004. **Staff forecast**: Greenbook forecast. **Experience-based forecast**: AR(1) model forecast estimated based on weighted lifetime inflation data for each FOMC member.

Open question: Stronger persistence of extreme experiences?

- Plausible that extreme experiences stick more strongly in (collective) memory
  - Challenge: difficult to cleanly identify in empirical data
- Example: German Hyperinflation of 1923

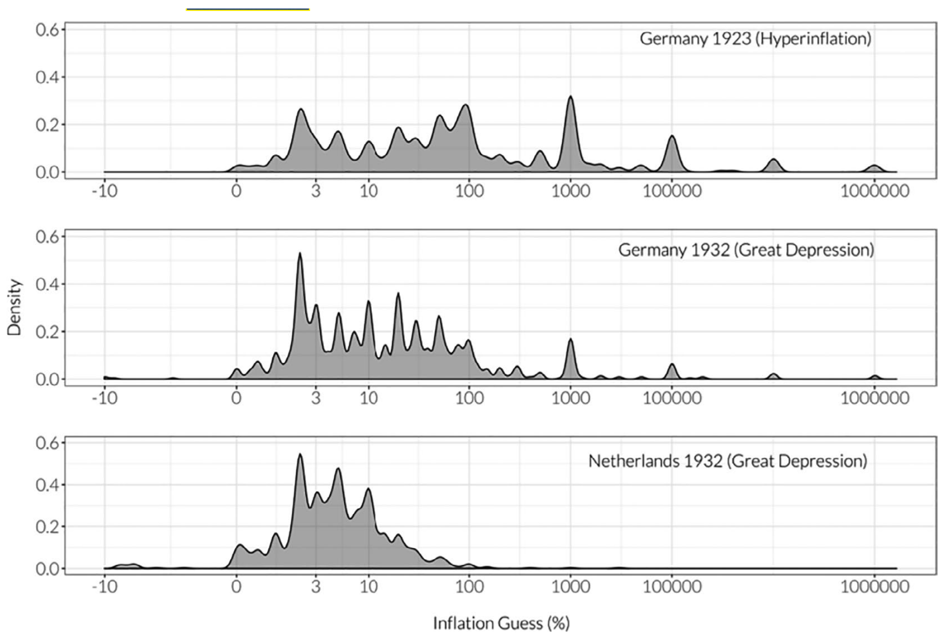
Open question: Stronger persistence of extreme experiences?

C1. Do you agree with the following statement? "The control of inflation is one of the most important missions of US [German, Brazilian] economic policy."

	1 Fully agree	2	3 Undecided	4	5 Completely disagree	
US all	56%	28%	7%	7%	2%	n=123
Born < 1940	69%	13%	11%	4%	2%	n=45
Born > 1950	44%	38%	2%	13%	2%	n=45
Germany all	76%	18%	5%	1%	0%	n=174
Born < 1940	90%	8%	1%	1%	0%	n=77
Born > 1950	51%	40%	7%	2%	0%	n=55
Brazil	56%	32%	2%	4%	7%	n=57

Source: Shiller (1997). See also Ehrmann and Tzamourani (2012) for broader evidence from World Values Survey.

Open question: False (collective) memories of history?



Source: Haffert, Redeker, and Rommel (2021)

Open question: Direct personal experience vs. indirect experience

- Plausible that direct personal experience with an event (e.g., portfolio loss in stock market crash) leaves stronger impression than indirect experience (e.g., reading about stock market crash w/o personal investment loss)
- Some initial evidence from Andersen, Hanspal, Nielsen (2019): Experience with bank collapse in Denmark during financial crisis in 2008 has stronger effect if personally experienced, vs. experienced by relatives, vs. just living in same town as collapsed bank.

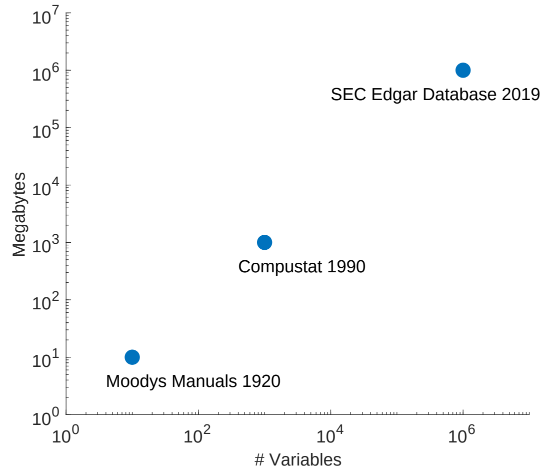
Belief formation based on historical experience: Summary

- Bayesian prescription: use all available relevant data to form posterior—but, in practice, what is "all" and what is "relevant"?
- In practice, individuals seem to rely heavily on slowly fading memory of personal experiences
- But many aspects of how individuals' memory of financial and macroeconomic history is formed, maintained, and recalled remain to be explored

3. Belief formation when observed data are high-dimensional

Investors' Big Data problem

- Investors forecasting cash flows face huge number of potential predictors
- $\mathbb{E}$ [c.f.] = f(predictors) unknown: high-dimensional learning problem
- How do investors deal with this problem?



Investors' Big Data problem

- Machine learning (ML) as inspiration for human forecasting in high-dimensional environments?

  "some common limits on human prediction might be understood as the kinds of errors made by poor implementations of machine learning." — *Camerer (2018).*

- (Supervised) ML methods: Generalized regression techniques to accommodate
  - high-dimensional data: Regularization, shrinkage, sparsity
  - nonlinearity: Trees, neural networks, ...
- Here: Focus on high-dimensionality within linear framework

Example: Learning about parameters of a linear model

- Suppose cash flows for N firms generated as

$$\boldsymbol{y}_t = \boldsymbol{X}_{t-1}\boldsymbol{g} + \boldsymbol{e}_t, \qquad \boldsymbol{e}_t \sim N(0, \boldsymbol{I})$$

  where $\boldsymbol{X}_{t-1}$ is an $N \times J$ predictor matrix.

- Elements of $N \times J$ predictor matrix $\boldsymbol{X}_{t-1}$ drawn IID $N(0,1)$.
  - Fixed regressors case: $\boldsymbol{X}$ drawn once, then fixed
  - Stochastic regressors case: $\boldsymbol{X}_t$ drawn new each period.

- "Nature" draws $\boldsymbol{g} \sim N(0, \sigma_g^2 \boldsymbol{I})$ once at the beginning, where

$$\sigma_g^2 = \frac{\theta}{J}, \qquad \theta = \text{const.}$$

  which keeps total explanatory power of $\boldsymbol{X}_{t-1}\boldsymbol{g}$ constant as we change J.

OLS estimation?

- At $t = 1$, agent uses observed data $\boldsymbol{y}_1$ and $\boldsymbol{X}_0$ to construct estimate $\tilde{\boldsymbol{g}}_1$, to then forecast

$$\hat{\boldsymbol{y}}_2 = \boldsymbol{X}_1 \tilde{\boldsymbol{g}}_1$$

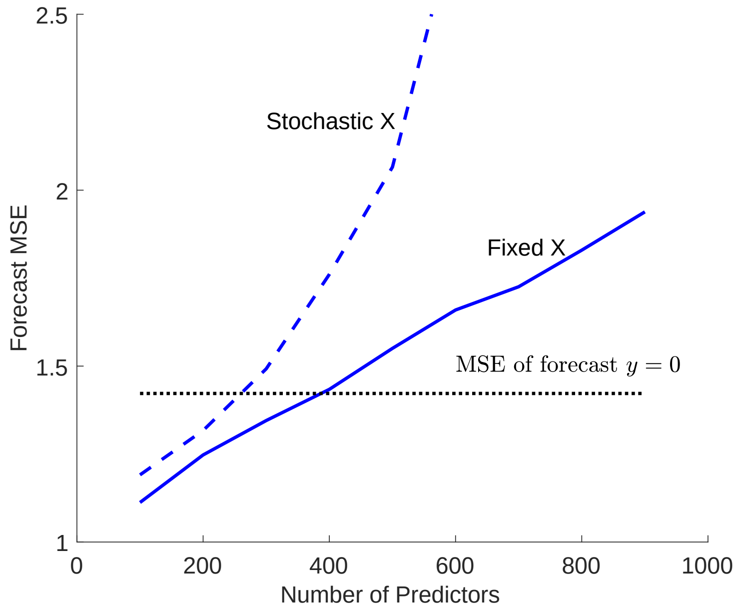
- Consider case where number of predictors (J) may be of close to similar magnitude as number of firms (N)

- Estimate $\tilde{\boldsymbol{g}}_1$ with OLS?

- Examine forecast performance

$$\text{Forecast MSE} = \frac{1}{N}(\boldsymbol{y}_2 - \hat{\boldsymbol{y}}_{2,OLS})'(\boldsymbol{y}_2 - \hat{\boldsymbol{y}}_{2,OLS})$$

Poor forecast performance of OLS in high-dimensional setting



$N = 1000$, $\theta = 0.5$.

Bayesian approach

- OLS = Bayesian posterior mean if prior beliefs about $\boldsymbol{g}$ diffuse
- Diffuse prior implies variance of predictable cash flow components $\boldsymbol{X}_{t-1}\boldsymbol{g} \to \infty$. Not economically sensible
- Economically sensible prior should be an informative prior that recognizes that very large magnitudes of elements of $\boldsymbol{g}$ aren't economically plausible
- Martin and Nagel (2021) explore asset pricing when investors learn with objectively correct prior, i.e., $\boldsymbol{g} \sim N(0, \sigma_g^2 \boldsymbol{I})$, same as distribution that nature draws $\boldsymbol{g}$ from.
  - Stronger assumption than Bayesian rationality

Forecast error predictability

- With informative prior, the posterior mean shrinks OLS coefficient estimates to zero

$$\tilde{\boldsymbol{g}}_t = \boldsymbol{\Gamma}_t \tilde{\boldsymbol{g}}_{OLS,t}$$

via shrinkage matrix $\boldsymbol{\Gamma}_t$ (w/ diffuse prior $\boldsymbol{\Gamma}_t = \boldsymbol{I}$)

- Forecast error

$$\boldsymbol{y}_{t+1} - \hat{\boldsymbol{y}}_{t+1|t} = \underbrace{\boldsymbol{X}_t(\boldsymbol{I} - \boldsymbol{\Gamma}_t)\boldsymbol{g}}_{\text{induced by shrinkage}} + \underbrace{\boldsymbol{X}_t\boldsymbol{\Gamma}_t(\tilde{\boldsymbol{g}}_{OLS,t} - \boldsymbol{g})}_{\text{reduced by shrinkage}} + \boldsymbol{e}_t$$

- First two components correlated with columns of $\boldsymbol{X}_t$: in-sample forecast error predictability

In-sample forecast error predictability

- Consider an econometrician studying forecast error predictability ex-post, regressing $\boldsymbol{y}_{t+1} - \hat{\boldsymbol{y}}_{t+1|t}$ on $\boldsymbol{X}_t$, which yields coefficients $\boldsymbol{b}_{t+1}$

- Econometrician's in-sample prediction of forecast error

$$\boldsymbol{X}_t\boldsymbol{b}_{t+1} = \boldsymbol{X}_t(\boldsymbol{I} - \boldsymbol{\Gamma}_t)\boldsymbol{g} + \boldsymbol{X}_t\boldsymbol{\Gamma}_t(\tilde{\boldsymbol{g}}_{OLS,t} - \boldsymbol{g}) + \boldsymbol{X}_t(\boldsymbol{X}_t'\boldsymbol{X}_t)^{-1}\boldsymbol{X}_t'\boldsymbol{e}_t$$

vs. actual forecast error

$$\boldsymbol{y}_{t+1} - \hat{\boldsymbol{y}}_{t+1|t} = \boldsymbol{X}_t(\boldsymbol{I} - \boldsymbol{\Gamma}_t)\boldsymbol{g} + \boldsymbol{X}_t\boldsymbol{\Gamma}_t(\tilde{\boldsymbol{g}}_{OLS,t} - \boldsymbol{g}) + \boldsymbol{e}_t$$

- Econometrician will detect in-sample predictability due to the two learning-induced components of the forecast error

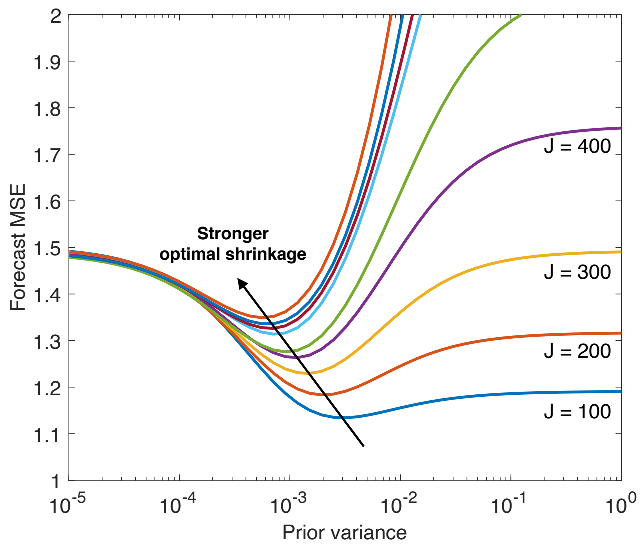
Absence of out-of-sample forecast error predictability

- Econometrician's out-of-sample prediction of $t+2$ forecast error
$$\hat{\boldsymbol{f}}_{t+2|t+1} = \boldsymbol{X}_t \boldsymbol{b}_{t+1}$$
vs. actual forecast error
$$\boldsymbol{y}_{t+2} - \hat{\boldsymbol{y}}_{t+2|t+1}$$
- Martin and Nagel (2021) show that
$$\mathbb{E}[(\boldsymbol{y}_{t+2} - \hat{\boldsymbol{y}}_{t+2|t+1})\hat{\boldsymbol{f}}'_{t+2|t+1}] = 0$$
- The strong assumption of objectively correct prior beliefs is crucial for this no-OOS predictability result

Open question: Which deviation from objective priors benchmark model?

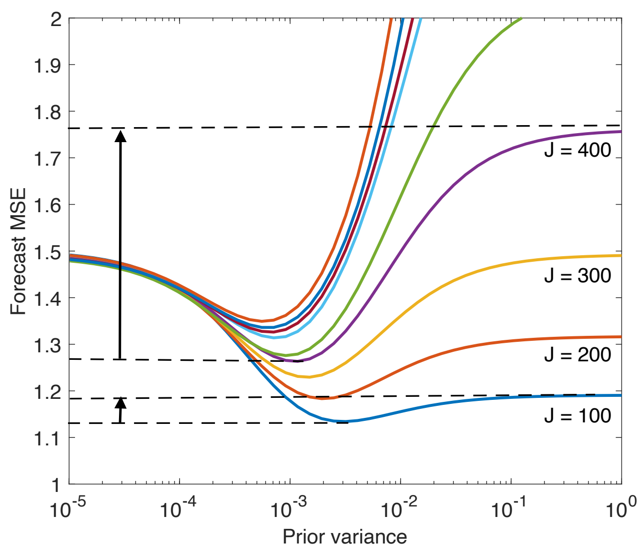
- The assumption of objectively correct priors may be too strong to describe real-world investor belief formation
  - It would not be irrational to have a different prior
  - People may not employ statistically optimal degree of shrinkage
- Ultimately an empirical question how people deal with high dimensionality
  - Tendency to overfit large models (insufficient shrinkage)?
  - Tendency to impose excessive shrinkage or sparsity?
- Within Bayesian model, we can think of insufficient/excessive shrinkage as prior variance above/below objective var($\boldsymbol{g}$)
- Priors matter more when J is large

Optimal shrinkage stronger with higher J



$N = 1000$; $\sim N(0, \sigma_g^2 \boldsymbol{I})$ with $\sigma_g^2 = \frac{\theta}{J}$ and $\theta = 0.5$

Insufficient shrinkage hurts forecast performance more when J is high



$N = 1000$; $\sim N(0, \sigma_g^2 \boldsymbol{I})$ with $\sigma_g^2 = \frac{\theta}{J}$ and $\theta = 0.5$

Open question: Underreaction and overreaction as miscalibrated prior?

- Forecast error

$$\boldsymbol{y}_{t+1} - \hat{\boldsymbol{y}}_{t+1|t} = \underbrace{\boldsymbol{X}_t(\boldsymbol{I} - \boldsymbol{\Gamma}_t)\boldsymbol{g}}_{\substack{\text{underreaction to signal} \\ \text{in } \boldsymbol{X}_t}} + \underbrace{\boldsymbol{X}_t\boldsymbol{\Gamma}_t(\tilde{\boldsymbol{g}}_{OLS,t} - \boldsymbol{g})}_{\substack{\text{overreaction to noise} \\ \text{in } \boldsymbol{e}_t, \boldsymbol{e}_{t-1}, \ldots}} + \boldsymbol{e}_t$$

- Bayesian agent optimally balances UR and OR based on prior beliefs about likely magnitude of $\boldsymbol{g}$ elements
- Empirical underreaction phenomena: can we understand them as excessive shrinkage, sparsity?
- Empirical overreaction phenomena: can we understand them as overfitting, lack of shrinkage & sparsity?

Open question: Heterogeneous priors as source of belief dispersion?

- Consider H individuals forecasting $\boldsymbol{y}_2 = \boldsymbol{X}_1\boldsymbol{g} + \boldsymbol{e}_2$ after observing $\boldsymbol{y}_1$.
- As before, "nature" draws $\boldsymbol{g} \sim N(0, \sigma_g^2\boldsymbol{I})$, where

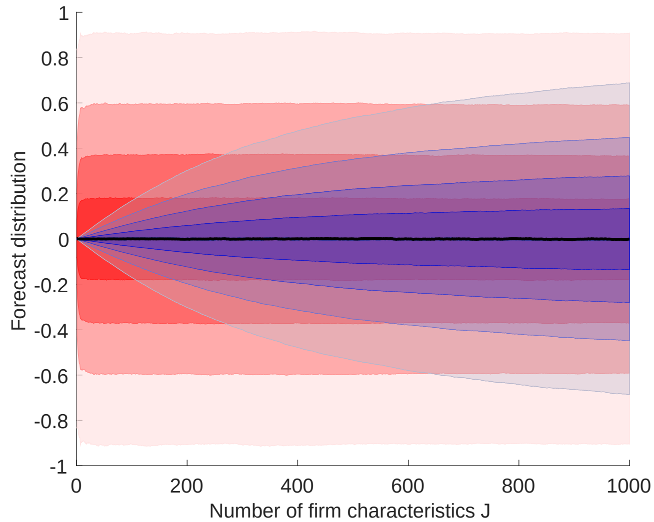
$$\sigma_g^2 = \frac{\theta}{J}, \qquad \theta = \text{const.}$$

- Now prior beliefs heterogeneous: Individuals' prior means are dispersed, individual h's prior mean is drawn from

$$\boldsymbol{\mu}_h \sim N(0, \sigma_g^2\boldsymbol{I})$$

- To what extent does dispersion in priors translate into dispersion in posteriors? How does the extent vary with J?

Dispersion in prior means more strongly affects posterior mean dispersion when J is high



Blue = Quantiles of distribution of elements of posterior means, $\hat{\boldsymbol{y}}_2$, across $H = 100$ individuals; Red = Quantiles of distribution of prior means.

Open question: Heterogeneous priors as source of belief dispersion?

- With different priors (Laplace prior, slab-and-spike prior, ...) instead of normal priors, Bayesian shrinkage can induce sparsity: some variables drop out of the forecasting model
  - Akin to Lasso regression
- Heterogeneity in priors in this case can mean disagreement about which variables are seen as relevant for forecasting

Subjective beliefs in high-dimensional settings: Summary

- Prior beliefs loom large in high-dimensional settings
  - Priors that are miscalibrated relative to objective reality generate large, predictable forecast errors
  - Prior dispersion can be source of belief dispersion

- Many open questions about how humans forecast in high-dimensional environments
  - Excessive shrinkage/sparsity vs. overfitting noise
  - Potential differences between purely human forecasting and machine-aided human forecasting

- Similar questions about nonlinearities (that we ignored here) and role of priors
  - Simpler linear models as approximations?
  - Overfitting of nonlinear patterns?

4. Conclusion

Conclusion

- Understanding subjective beliefs of investors important for understanding asset prices
- Rather than reverse-engineering beliefs from asset prices, assumptions about belief formation in asset pricing models should be grounded in
  - Beliefs data from surveys
  - Data on investor portfolio decisions
  - Experimental data
- Main challenge: Finding pervasive, simple belief formation principles that work in a variety of settings
- In this talk, I have highlighted two areas of key interest for asset pricing
  - Learning from history: Affects low frequency time-series variation in asset prices
  - Learning in high-dimensional settings: Affects cross-sectional variation in asset prices

References

1. Andersen, S., Hanspal, T. and Nielsen, K.M., 2019. Once bitten, twice shy: The power of personal experiences in risk taking. *Journal of Financial Economics*, 132(3), pp.97-117.
2. Malmendier, U., and S. Nagel, 2011. Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking? *Quarterly Journal of Economics* 126, 373–416.
3. Malmendier, U., and S. Nagel, 2016. Learning from Inflation Experiences. *Quarterly Journal of Economics* 131, 53–87.
4. Malmendier, U., Nagel, S., and Yan, Z., 2021. The Making of Hawks and Doves. *Journal of Monetary Economics* 117, 19–42.
5. Martin, I., and S. Nagel, 2021. Market Efficiency in the Age of Big Data. Working paper, University of Chicago and LSE.
6. Nagel, S. and Z. Xu. 2018. Asset Pricing with Fading Memory. *Review of Financial Studies*, forthcoming.
7. Shiller, R., 1997. Why Do People Dislike Inflation? In: Romer, C.D. and D.H. Romer (eds.), *Reducing Inflation: Motivation and Strategy*, University of Chicago Press.
8. Xu, Z., 2020. Expectation Formation in the Treasury Bond Market. Working paper, City University of Hong Kong.