Learning from data interacts with experience

- **Main empirical finding:** availability of additional predictive variables is associated with a steepening of the learning dynamics...
- with the beneficial effects of additional data availability found only for the more experienced investors.
- Surprising because all investors should rationally be making use of all available signals to attain their (common) objectives.

![Learning dynamics in our main sample, split into the Treatment & Control groups. Bars represent standard errors.](image)

Model uncertainty as explanation for results

- Experienced investors appear to benefit from wider data
- Why don’t inexperienced investors also take advantage?
- Explanation rooted in model uncertainty:
  - Inexperienced investors face model uncertainty more, leading them to discard some prediction signals that are available to them.
  - As they gain in experience, investors shed some model uncertainty.
  - This mechanism is captured by the following model of investor learning.

Investor learning under model uncertainty

- Follow Martin and Nagel (2021) in modeling each investor as behaving like an econometrician when using historical data.
- Recall Quanticians investors are incentivized to maximize out-of-sample (i.e. future) Sharpe Ratios over a fixed horizon, 

\[
\max_{\hat{\beta}} \left( \left( S_{t+1} - U_{t+1} \right) / \left( \sigma_{t+1} \right) \right)
\]

- Assume the variance is known (Merton 1980) and that the expected return is a linear combination of the given predictive signals, 

\[
\mu_i = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p
\]

- Then the investor must learn \( \hat{\beta} \) based on historical expected returns from (similar but not identical) futures contracts that expired in the past and corresponding historical signals \( S \).

- Fearing worst-case model uncertainty, her learning problem is thus to

\[
\max_{\hat{\beta}} \min_p \left( \mu_i - \left( S_{i,t} - U_{i,t} \right) / \left( \sigma_{i,t} \right) \right)
\]

where the model uncertainty can be represented as a matrix of signal-wise perturbations \( U \) that maximizes the \( \delta \) norm, for any choice of \( \beta \) and is constrained by an uncertainty set.

- Assuming orthornormal, \( S \), it follows from Xu, Caramanis, and Mannor (2010) and Tibshirani (2019) that the investor should use

\[
\hat{\beta} = \frac{-\lambda}{\sqrt{\|S\|^2}}
\]

in her portfolio choice problem, with elements of \( \hat{\beta} \) being

\[
\lambda_k = \frac{\hat{\beta}_k / \|S\|}{\sqrt{\|S\|^2 - \lambda_k^2}}
\]

where \( \lambda \geq 0 \) is a scaling of \( \|S\|^2 - \lambda \|S\| \).

Imlications of Egn. (5)

- The investor should ignore signals whose historical predictive contribution is less than her subjective model uncertainty threshold \( \lambda \).
- The higher her fear of model uncertainty, the fewer predictive signals she should use (informal statement).
- Conjecture: investor’s fear of model uncertainty \( \lambda \) falls with experience.
- Therefore, the number of predictive variables she uses should increase with her experience.

Estimating investors’ usage of predictive variables

- Investors use more predictive variables as they gain in experience
- Once again, highlights the interaction between the complementary channels of learning with experience & learning from data.

![The dynamics of the estimated number of predictive variables used by investors to solve their portfolio choice problem. Bars represent standard errors.](image)

More results in the paper

- Identification by exploiting the fact that all the new predictive variables happen to be lower-frequency macroeconomic variables
- Secondary results on: realized ex-post moments of returns, dispersions (within-investor & across-investor), overconfidence.

References