Predicting financial market stress with machine learning

Iñaki Aldasoro, Peter Hördahl,
Andreas Schrimpf, Sonya Zhu[1]

Bank for International Settlements

Bank of Thailand, 31 March 2026

[1] The views expressed here are those of the authors only, and not necessarily those of the BIS.

Motivation

- The nature of liquidity provision has changed over time
  - Shift from banks to NBFIs after the GFC
  - High-speed trading on automated platforms facilitated by technology
- Episodes of market stress have become more frequent and widespread
  - Market dysfunction or conditions characterized by low liquidity
  - Impairment to price formation or breakdown of no-arbitrage relationships
  - Pressure on authorities to intervene
- Are financial market (stress) conditions predictable?
  - Out-of-sample predictability?
  - Can machine learning models help?
  - What type of indicators are important for predictability?

Measuring financial market conditions

- We expand the market conditions indicators (MCIs) for the US Treasury, Money and FX markets (Aldasoro et al. (2022))
  - Each MCI is estimated with PCA using a broad range of variables that capture:
    - *market illiquidity*, *Volatility*, and *price dislocations*
  - Real-time monitoring: **daily** frequency
  - Automated process of data collection and aggregation
  - MCIs are updated and published in the BIS data dashboard

- The MCIs are distinct from existing financial stress (FSIs) and conditions indices (FCIs)

Our paper in a nutshell

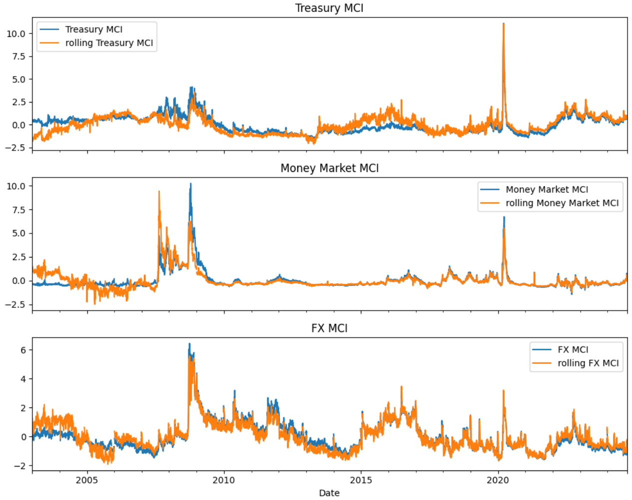
- Construct out-of-sample MCIs with expanding-window PCA
  - Only consider past information and avoid look-ahead bias
- Predict the future distribution of MCIs using machine learning methods
  - ML models (random forest) outperform classic regression models, ...
    ... both in-sample and out-of-sample
- Uncover leading indicators
  - Market-implied macro and monetary policy indicators (interest rate environment and uncertainty) are important across MCIs
  - investor sentiment, GFC, and liquidity conditions matter too for individual markets

Constructing MCIs: principal component analysis

Same input variables as Aldasoro et al. (2022)

- **Treasury market**
  - *Volatility* (MOVE index), *Market liquidity* (quoted spread, time-to-quote, Bloomberg Tsy Liquidity Index), *Price dislocation* (on-the-run liquidity premium, OIS-Treasury spread, Treasury futures basis)
- **Money market**
  - Commercial paper (CP)-OIS, Repo-OIS, TED spread, LIBOR-OIS
- **FX market**
  - *Volatility* (JP Morgan FX volatility index), *Market liquidity* (quoted spread of major currency pairs), *Price dislocation* (cross-currency basis, triangular VLOOP (ie violation of the law of one price, Huang et al. (2025))

Rolling MCI estimation



Are market conditions predictable?

- Policy relevant: central banks want to anticipate stress episodes

$$MCI_t > threshold_t,$$

where $threshold_t$ is in the right tail of the distribution

- Predict stress episodes $\Leftrightarrow$
Predict right tails of the distribution of market conditions (can be time-varying)
  - Forecast 90th or 95th MCI percentiles

Fit the tails of MCIs using classic time-series models

- An auto-regressive model to fit quantile τ of each MCI (“AR model"):

$$\tau(y_{j,t}) = \alpha_j + \beta_{1,j} y_{fx,t-h} + \beta_{2,j} y_{mm,t-h} + \beta_{3,j} y_{tr,t-h},$$

  where $y_{j,t}$ is the MCI for market $j \in \{fx, mm, tr\}$, in month t, with forecast horizon $h \in \{1, 2, \cdots, 12\}$
- Estimation-test windows:
  - rolling estimation with 84 observations from y_{t-h-84} to y_{t-h-1}
  - sample period: January 2003 to May 2024
  - number of estimation-test windows:
    - 153 windows for $h = 1$ ...... 142 windows for $h = 12$

Add potential explanatory variables

A linear model with a broad list of explanatory variables $\boldsymbol{X}_t$ (“full model")

$$\tau(y_{j,t}) = \alpha_j + \beta_{1,j} y_{fx,t-h} + \beta_{2,j} y_{mm,t-h} + \beta_{3,j} y_{tr,t-h} + \boldsymbol{\gamma} \boldsymbol{X}_{t-h},$$

We consider 44 candidate explanatory variables, under 6 broad categories

- Funding liquidity (7 variables)
  - net monthly purchase of US government bonds by the Fed
  - US Treasury's General Account
  - primary dealer fails in US Treasury settlements
  - primary dealer net coupon treasury holdings, as in Du et al. (2023)
  - dollar volume of total commercial paper issuance

- Fund flows (18 variables)
  - EPFR monthly net fund flows to
    - developed/emerging bond and equity markets
    - high-yield bonds
    - investment grade bonds
  - Intra-family flows into
    - money market funds
    - risk appetite: shifts towards high-yield bond funds (Ben-Rephael et al. (2012))
    - sentiment: shifts towards equity funds (Ben-Rephael et al. (2021))

- Risk preference (4 variables)
  - excess bond premium from Gilchrist and Zakrajšek (2012)
  - DXY and global financial cycle (GFC) index (Rey (2013))
  - maintenance margin of 10-year treasury futures
- Macroeconomic conditions (5 variables)
  - Citi Economic Surprise indices for the US and the global economy
  - 3-month US Treasury bill
  - differences between 10-year and 2-year Treasury yields
  - differences between 2-year and 3-month Treasury yields
- Perceived market volatility and monetary policy uncertainty (7 variables)
  - VIX, MOVE, SKEW indices
  - three-month EUR/GBP/JPY implied volatility (IV EUR/GBP/JPY 3m)
  - interest rate swaption implied vol (USD 2y)
- Lagged values of market conditions indicators (3 variables)

Predict MCI tails with machine learning approaches

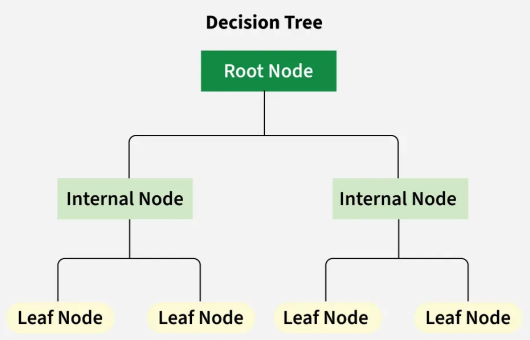
- Grinsztajn et al. (2022): tree-based models outperform deep learning models (neural networks, etc) for regression data

- Tree-based models are also more interpretable than deep learning models

- We evaluate the performance of tree-based ML models
  - Random Forest
  - and a robustness check with Extreme Gradient Boosting (XGBoost)

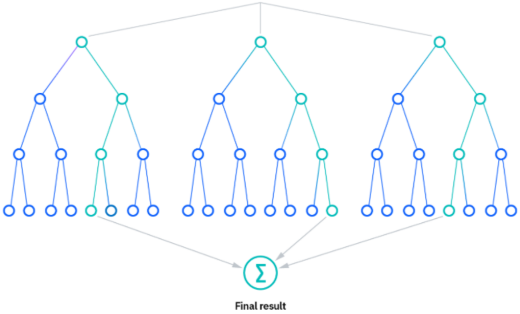
The random forest model

- A random forest is made up of multiple decision trees
- A decision tree is a collection of sequential choices and their results
  - Classification trees: predict categorical outcomes (e.g. loan approval)
  - Regression trees: predict continuous outcomes (e.g. value of a house or MCI value)



The random forest model

- RF is an 'ensemble learning method' where the predictions of individual decision trees are aggregated to identify an optimal result
- Randomness comes from subsampling of data in the training phase
- ...and selecting a random subset of features
- The final result is obtained by averaging across decision trees



Evaluate prediction errors using pinball loss

The pinball loss (for quantile τ) is:

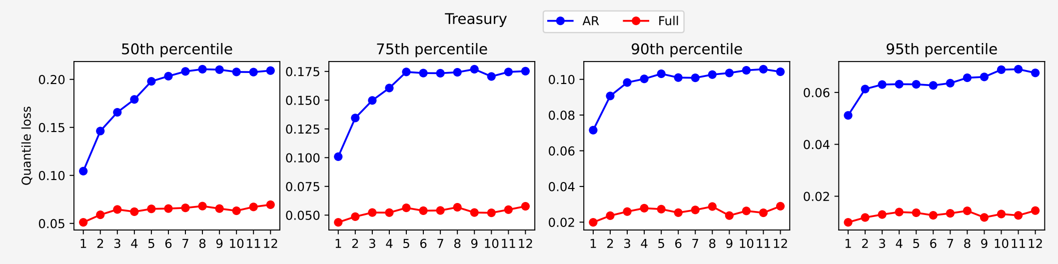
$$L_\tau(y, \hat{y}) = \begin{cases} \tau(y - \hat{y}) & \text{if } y \geq \hat{y} \\ (1 - \tau)(\hat{y} - y) & \text{if } y < \hat{y} \end{cases}$$

for observed and predicted values y and $\hat{y}$

- Asymmetric loss function (for $\tau \neq 0.5$)
- For high τ (e.g. 0.9), underpredictions are penalized more than overpredictions
- The total quantile loss averages the pinball losses across all observations

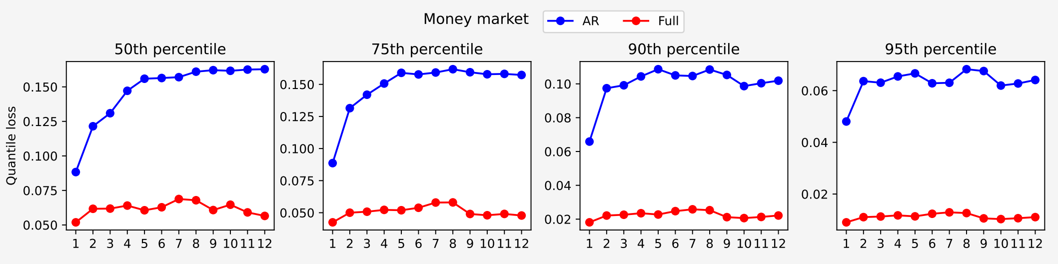
Benchmark models: in-sample quantile loss

Treasury market



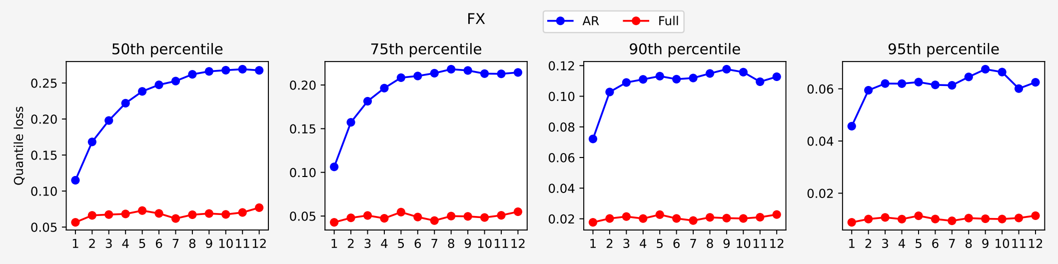
Benchmark models: in-sample quantile loss

Money market



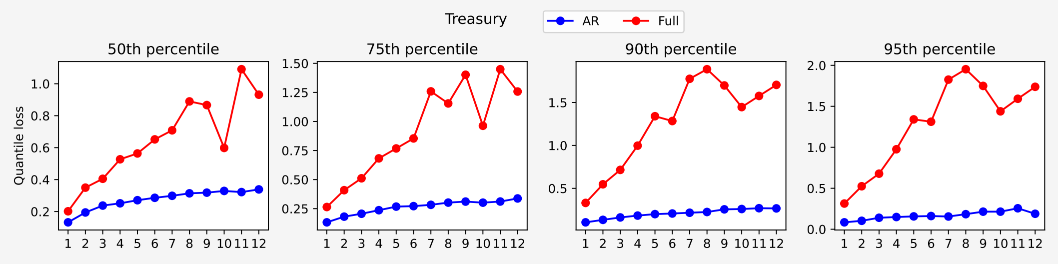
Benchmark models: in-sample quantile loss

FX market



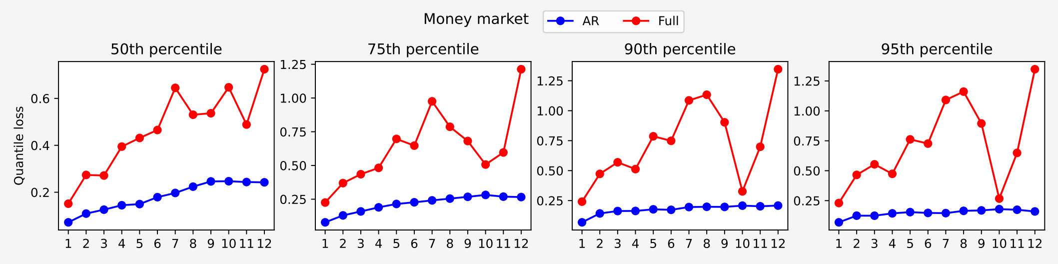
Benchmark models: out-of-sample quantile loss

Treasury market



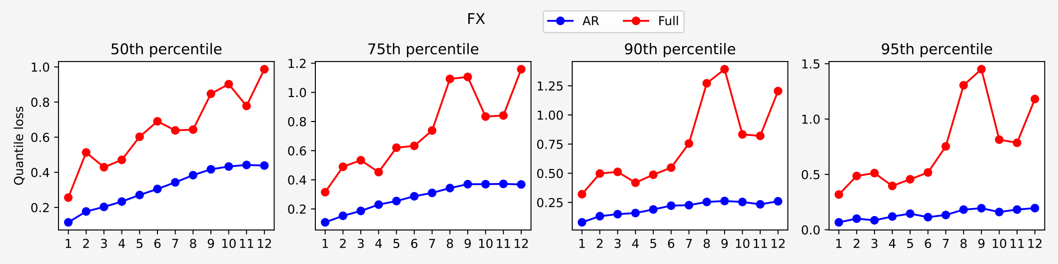
Benchmark models: out-of-sample quantile loss

Money market

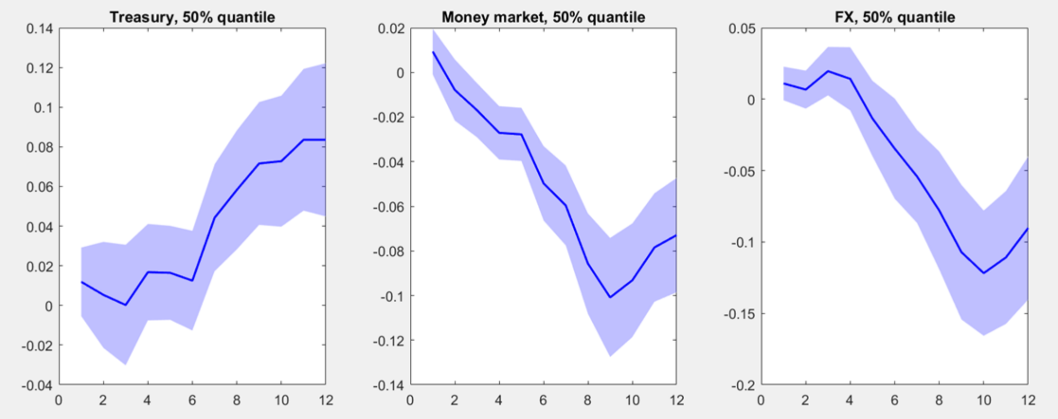


Benchmark models: out-of-sample quantile loss

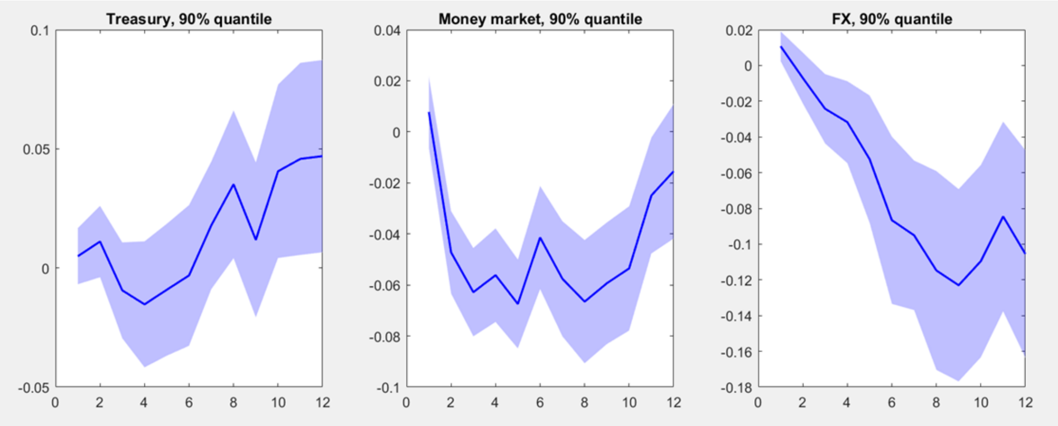
FX market



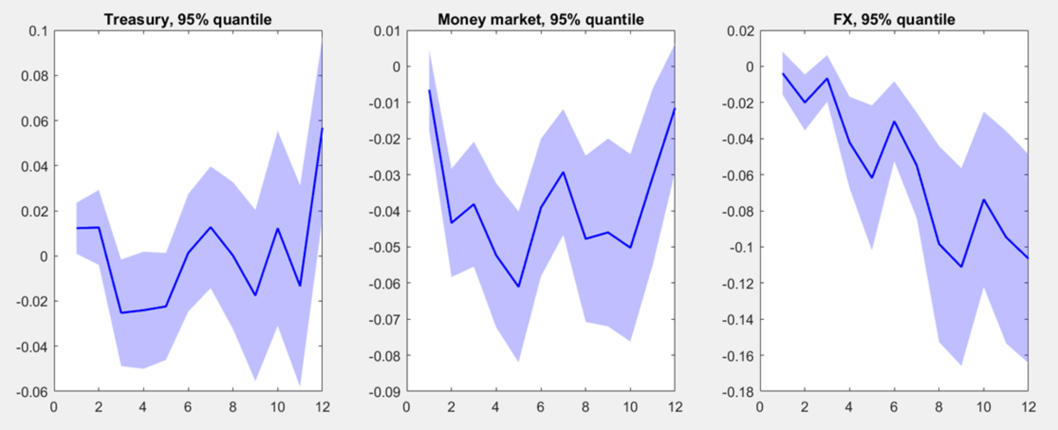
Random Forest: out-of-sample loss relative to AR model (50pctl)



Random Forest: out-of-sample loss relative to AR model (90pctl)



Random Forest: out-of-sample loss relative to AR model (95pctl)



Which predictors are most important? – Shapley value

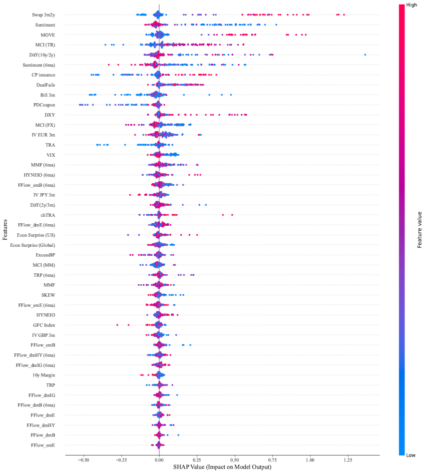
Given a feature set $\mathbf{x}$ and a model f, the Shapley value for a feature i represents the average contribution of that feature to the prediction across all possible subsets of features. Specifically, the Shapley value ϕ_i for feature i is given by:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)]$$

where:

- N is the set of all features.
- S is a subset of features not containing i.
- $f(S)$ is the model prediction using the features in subset S.
- $|S|$ is the number of features in subset S.
- $|N|$ is the total number of features.

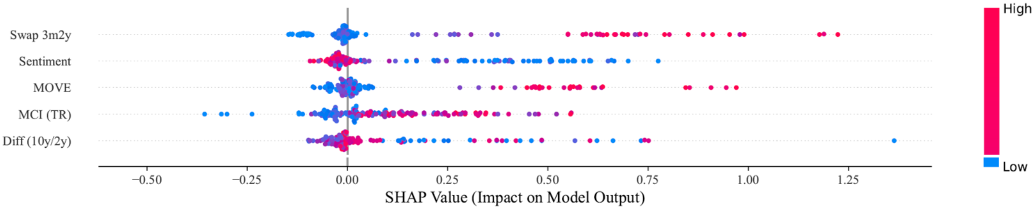
Beeswarm summary plots of Shapley values



- Plots Shapley values by feature and quantile for each forecast date
  - features are ordered by mean absolute value
- value on x-axis: impact on model forecast
- thickness of swarm shows density
- color shows value of feature (red = high; blue = low)

Next: focus on top 5 features

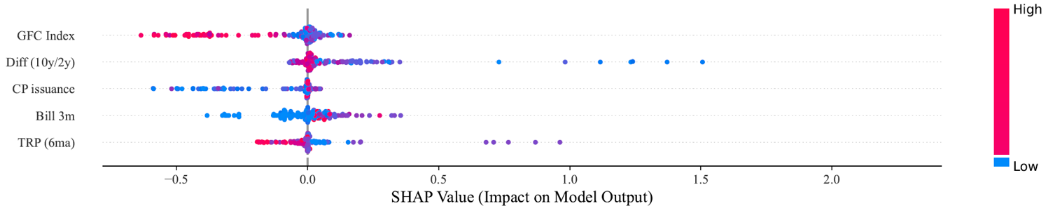
Top predictors for Treasury MCI (90th percentile, 6m ahead)



Rate uncertainty & slope, and sentiment are important predictors

- interest rate volatility
- yield slope
- investor sentiment (Ben-Rephael at al, 2012)

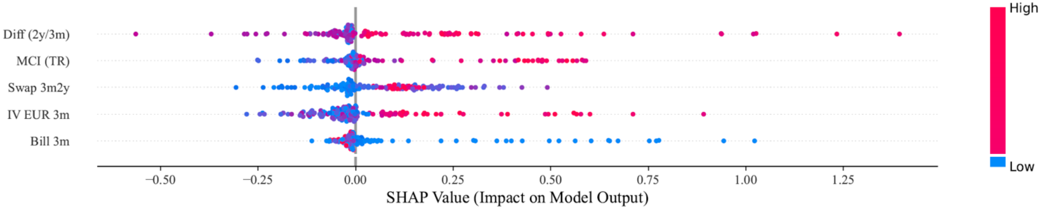
Top predictors for MM MCI (90th percentile, 6m ahead)



Global financial cycle, interest rate environment, and liquidity conditions are important predictors

- GFC index
- interest rate slope and level
- CP issuance and Fed Treasury purchases

Top predictors for FX MCI (90th percentile, 6m ahead)



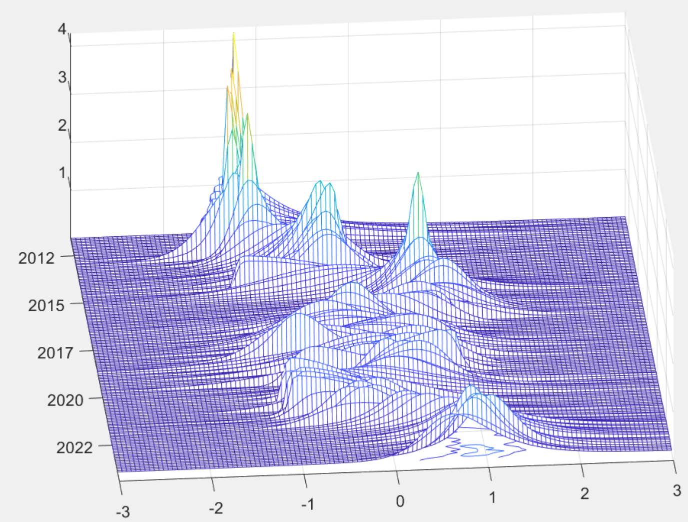
Interest rate environment, and rate & FX uncertainty are important predictors

- interest rate slope & level
- interest rate and FX implied vol
- cross-market influence (Treasury MCI)

Full distribution of forecasted MCIs

- Fit full MCI distribution based on the Random Forest results
- Follow Adrian et al. (2019): using estimated quantiles (0.05, ..., 0.95), at each forecast date we fit a skewed *t*-distribution to capture the full forecast distribution for each MCI and each horizon
- Allows us to study how MCI distributions evolve over time and their sensitivity to predictors

MCI distributions are time-varying: e.g. Treasury MCI 6m ahead

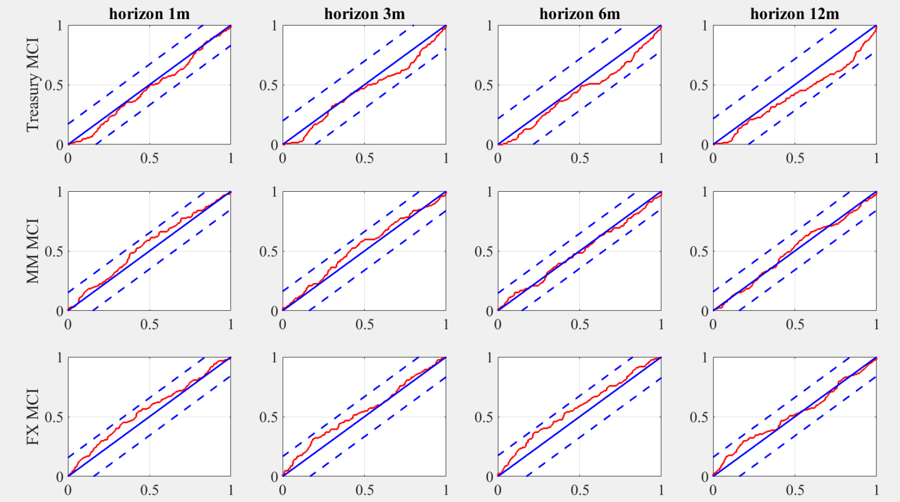


Evaluating the forecasted MCIs distributions

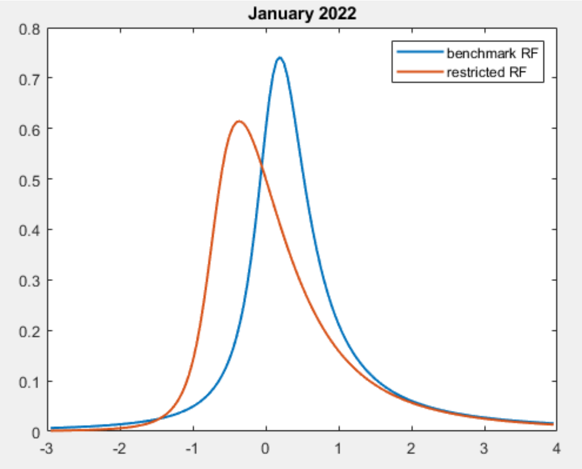
Fitting the full distribution gives us another way of evaluating the out-of-sample accuracy of the random forest method

- Use probability integral transform (PIT): evaluate the predicted cumulative distribution at the realized value
- I.e.: measure the percentage of observations that are below a given model-implied quantile
- If the model is ‘right’, the cumulative distribution of the PITs will lie on a 45-degree line

Evaluating fitted MCI distributions (out-of-sample)

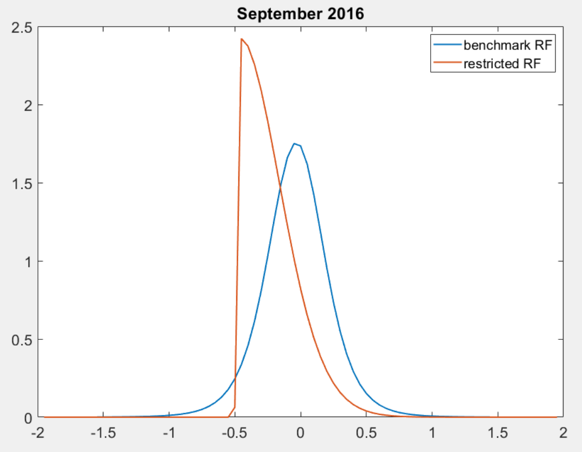


Example: forecasted Treasury MCI PDFs 6m ahead



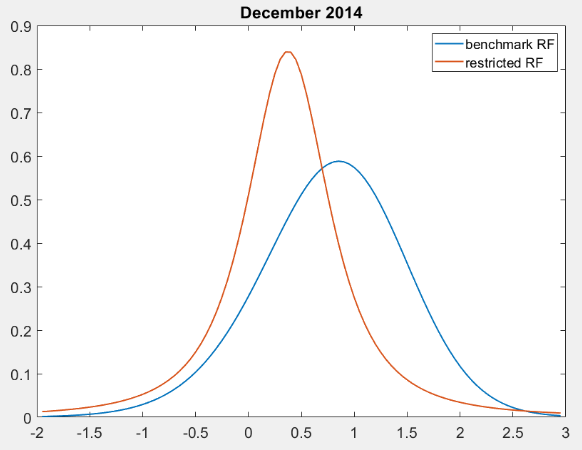
- Compare unrestricted benchmark PDF with restricted forecast
  - restricted: fixes top feature (swaption IV)
  - also restricts highly correlated features (MOVE, lagged MCI, CP issuance)
- Example: Jan 2022 - before start of Fed hike cycle
- benchmark model: higher mean

Example: forecasted Money market MCI PDFs 6m ahead



- Compare unrestricted benchmark PDF with restricted forecast
  - restricted: fixes top feature (GFC index)
  - also restricts TGA, PD Treasury holdings, futures margin requirements
- Example: Sep 2016 - before SEC MM reform
- benchmark model: higher mean and variance

Example: forecasted FX MCI PDFs 6m ahead



- Compare unrestricted benchmark PDF with restricted forecast
  - restricted: fixes top feature (2y/3m yield slope)
  - also restricts interest rate level, 10y/2y yield slope
- Example: Dec 2014 - just before CHF de-peg
- benchmark model: higher mean and variance

Conclusions

- Tree-based machine learning models outperform classic time series models in predicting the distribution of market conditions out of sample
- Based on RF quantile forecasts, we are able to accurately fit the full forecast distribution across MCIs and horizons
- We quantify the importance of individual predictors for each market using Shapley values
  - Market-implied macro and monetary policy indicators are important across MCIs
  - investor sentiment, GFC, and liquidity conditions also matter for individual market segments

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone, "Vulnerable Growth," *American Economic Review*, 2019, *109* (4), 1263–1289.

Aldasoro, Iñaki, Peter Hördahl, and Sonya Zhu, "Under pressure: market conditions and stress," *BIS Quarterly Review*, 2022, (19), 31–45.

Ben-Rephael, Azi, Jaewon Choi, and Itay Goldstein, "Mutual fund flows and fluctuations in credit and business cycles," *Journal of Financial Economics*, 2021, *139*, 84–108.

__ **, Shmuel Kandel, and Avi Wohl**, "Measuring investor sentiment with mutual fund flows," *Journal of Financial Economics*, 2012, *104*, 363–382.

Du, Wenxin, Benjamin Hébert, and Wenhao Li, "Intermediary balance sheets and the treasury yield curve," *Journal of Financial Economics*, 2023, *150* (3), 103722.

Gilchrist, Simon and Egon Zakrajšek, "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, June 2012, *102* (4), 1692–1720.

Grinsztajn, Leo, Edouard Oyallon, and Gael Varoquaux, "Why do tree-based models still outperform deep learning on typical tabular data?," in "Advances in Neural Information Processing Systems," Vol. 35 Curran Associates, Inc. 2022, pp. 507–520.

Huang, Wenqian, Angelo Ranaldo, Andreas Schrimpf, and Fabricius Somogyi, "Constrained liquidity provision in currency market," *Journal of Financial Economics*, 2025. forthcoming.

Rey, Hélène, "Dilemma not trilemma: the global cycle and monetary policy independence," *Proceedings - Economic Policy Symposium - Jackson Hole*, 2013.