Discussion of “Harnessing Artificial Intelligence For Monitoring Financial Markets”

CEMLA Conference

November 13, 2026

Greg Hopper

Bank Policy Institute

Overview

- Authors have set themselves an incredibly difficult problem
- Can machine learning and AI be used to help regulatory supervisors monitor financial conditions?
- Very interesting methodology
- I'll discuss some of the difficulties and empirical questions

Very Difficult Problem

- My custom AI harness read about 10 million words of all ECB, FED, and BOE Financial Stability reports, concluding that they were not able to identify financial vulnerabilities in advance
- My previous attempts to predict one-off financial events for stress testing not successful

Very Interesting methodology of the paper

- Use 20-day moving average of standard deviation of discrepancies in triangular arbitrage of dollar/yen/euro pairs measured at very high frequencies to measure market disfunction
- Use (recurrent neural network) RNN to learn which explanatory variables are good a predicting the 20-day moving average 60 trading days ahead (in conjunction with Long Short –Term Memory (LTSM) network)
- Feed the most important explanatory variables (which are time-varying) to an LLM to look for current market analysis that could explain predictions of future stress

Unanswered Empirical Questions

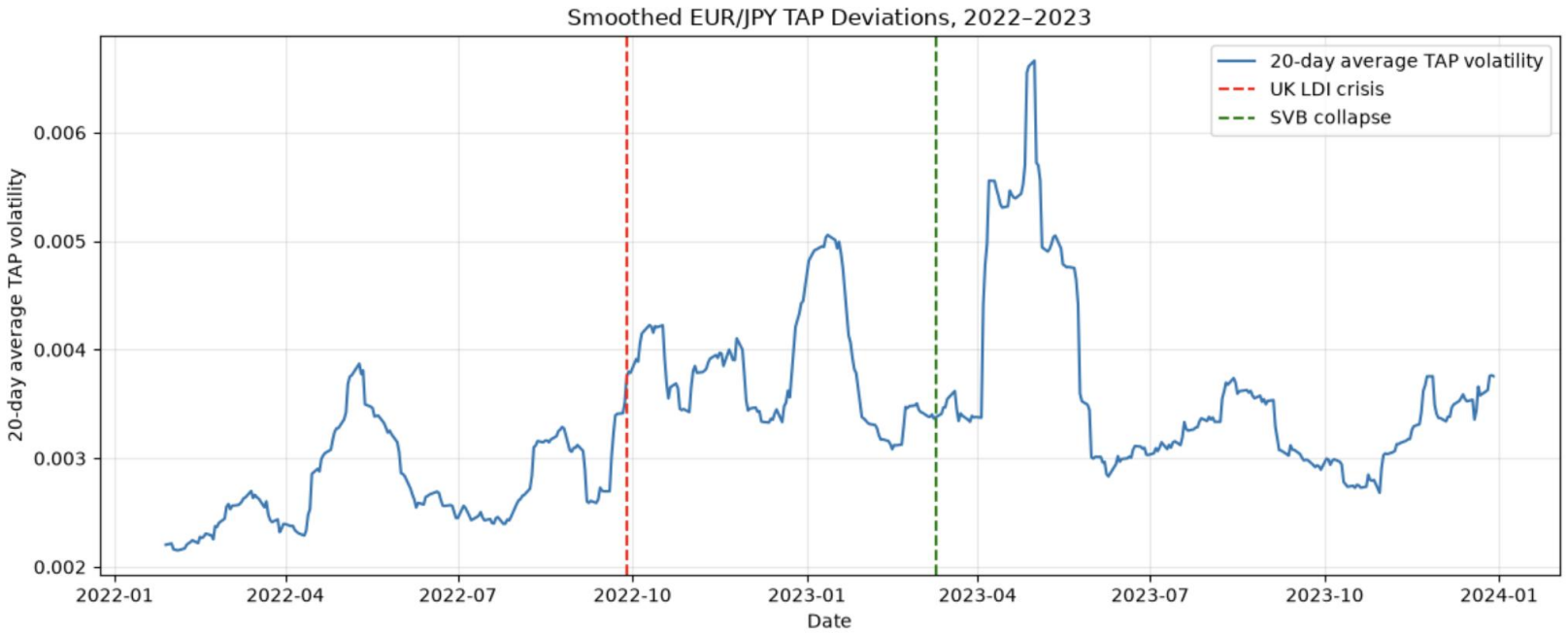
- Is the 20-day moving average of standard deviations of discrepancies in triangular arbitrage really a predictor of financial distress?
  - Many market micro-structure questions
- Is the RNN really optimal or necessary?
- Do the neural network weights provide meaningful economic explanations?
- Do the weights really help the LLMs to pick out relevant news articles?
- What about false positives and false negatives?

20-Day Moving Average Analysis

- Download Dukascopy foreign exchange data 2022-2023
  - Approximately 739,000 matched minute observations of EUR/USD, USD/JPY, and EUR/JPY
- Construct 20-day moving average of standard deviation of triangular arbitrage discrepancies as in paper
- Looked at different data set to understand how the measure behaves as a predictor of financial distress

Does the Measure Predict Financial Distress?

- TAP volatility did not spike at the right places



- TAP vol is very sensitive to trading hours and holidays

Arbitrage Opportunities Suspiciously Frequent

- Using bid-ask quotations, apparent arbitrage opportunities were surprisingly frequent
- Some days showed arbitrage opportunities every minute
- Indicates data discrepancies or other micro-structure effects

Statistic	Result
Matched minute observations	202,501
Apparent arbitrage opportunities	11,188
Frequency	5.52%
Median apparent profit	0.75 basis points
95th percentile profit	5.01 basis points

Weak Correlation Between TAP vol and Apparent Arbitrage Opportunities

Sample	Observations (days)	Correlation
All eligible trading days	519	0.083
Excluding holidays	497	0.169

- Should a different measure be used, such as failure of covered interest rate parity?
- Violations of CIP may not always indicate financial distress, e.g., 2016 BIS paper “Covered Interest Parity Lost: Understanding the Cross Currency Basis” documented violations after 2014

Other Crucial Empirical Questions

- How does the RNN compare to
  - Time-varying coefficient regression?
  - Markov-switching regression
  - XGBoost with SHAP
- RNNs are non-linear.
  - Why would the weights that matter be necessarily the largest?
  - What about DeepSHAP, GradientSHAP, KernelSHAP, Grouped Shapley values?
- Do the weights really help the LLM pick out the most relevant news articles?
  - Need independent verification that the news picked out is really useful for understanding financial distress 60 days ahead
- What about false positives and negatives?
  - Measured in my LLM analysis of central bank Financial Stability reports