UNIVERSITY OF CAMBRIDGE

CCIMI
CANTAB CAPITAL INSTITUTE FOR THE MATHEMATICS OF INFORMATION

Adaptive Learning on Time Series: Theory and Practice

Parley Ruogu Yang[#†*], Raena McElwee[‡*], Ryan Lucas[‡*], Hugo Dolan[‡*]

[#] Correspondent [†] CCIMI, Faculty of Mathematics, University of Cambridge
[‡] University College Dublin [*] Optimal Portfolio Research Group optimalportfolio.github.io

Introduction

- In a discrete time series, we consider ordered and irreversible processes: for time $t \in \mathbb{N}$, write the vector of potential explanatory variables $x_t \in \mathbb{X}$ and the vector of dependent variables $y_t \in \mathbb{Y}$, where $\dim(\mathbb{X}) \gg \dim(\mathbb{Y})$. At time t, we have access to an information set representing all data available up to t, written as: $\Phi_t := \{(x_\tau, y_\tau) : \tau \in [t]\}$. We are interested in producing a conditional forecast of future values of the dependent variables k periods ahead, written as:
$$y_{t+k|t} := \mathbb{E}[y_{t+k}|\Phi_t] = f_k(\Phi_t; \theta_t; h_t)$$
The right-hand-side acts as general notation for such a forecast — θ_t represents the contemporary parameters and h_t is the contemporary functional form.
- **20^{th} century problems:** We design a finite model space based on intuition. How to estimate $\theta_t(h_t)$ and make forecasts? How to capture time-varying parameters? Solution: MLE estimation; SARIMA models; windowed time series; model selection via information criteria (AIC, BIC, etc.). [1,3,5]
- **21^{st} century problems:** Is there a better way to determine $h_t \in H$? What if $|H| = \infty$? What if $h_t \neq h_{t-1}$? How to efficiently capture the right model at the right time? How to effectively penalise the complexity under small data size? Do we even need to penalise the complexity if far-order autocorrelation were to exist? [4,5,6,7,9,10]

Traditional algorithm (SARIMA + AIC) [3: pp.48-50]

1. Determine seasonality and stationarity from the content of the data (e.g. simple intuition and ADF tests).
2. Find the maximum lags via ACF and PACF. Usually, total number of explanatory variables are upper-bounded by $\sqrt{T}$ to ensure efficiency and to discourage over-fitting.
3. Pin down ARMA lags by finitely searching the upper-bounded model.

Simulated example (similar to [7])

- Consider an SAR process $(1 - \gamma L^7)(1 - \rho L) = \varepsilon_t$, $\varepsilon_t \sim N(0, 1)$ With significant γ and significant ρ where $|\rho| < 1$.
- Without specific knowledge, a traditional algorithm is likely to output $AR(p)$, where $p \approx 8$. Inefficient estimation and poorer forecasts.
- With appropriate training, we are able to find better forecasts by finding a functional form closer to the underlying process.

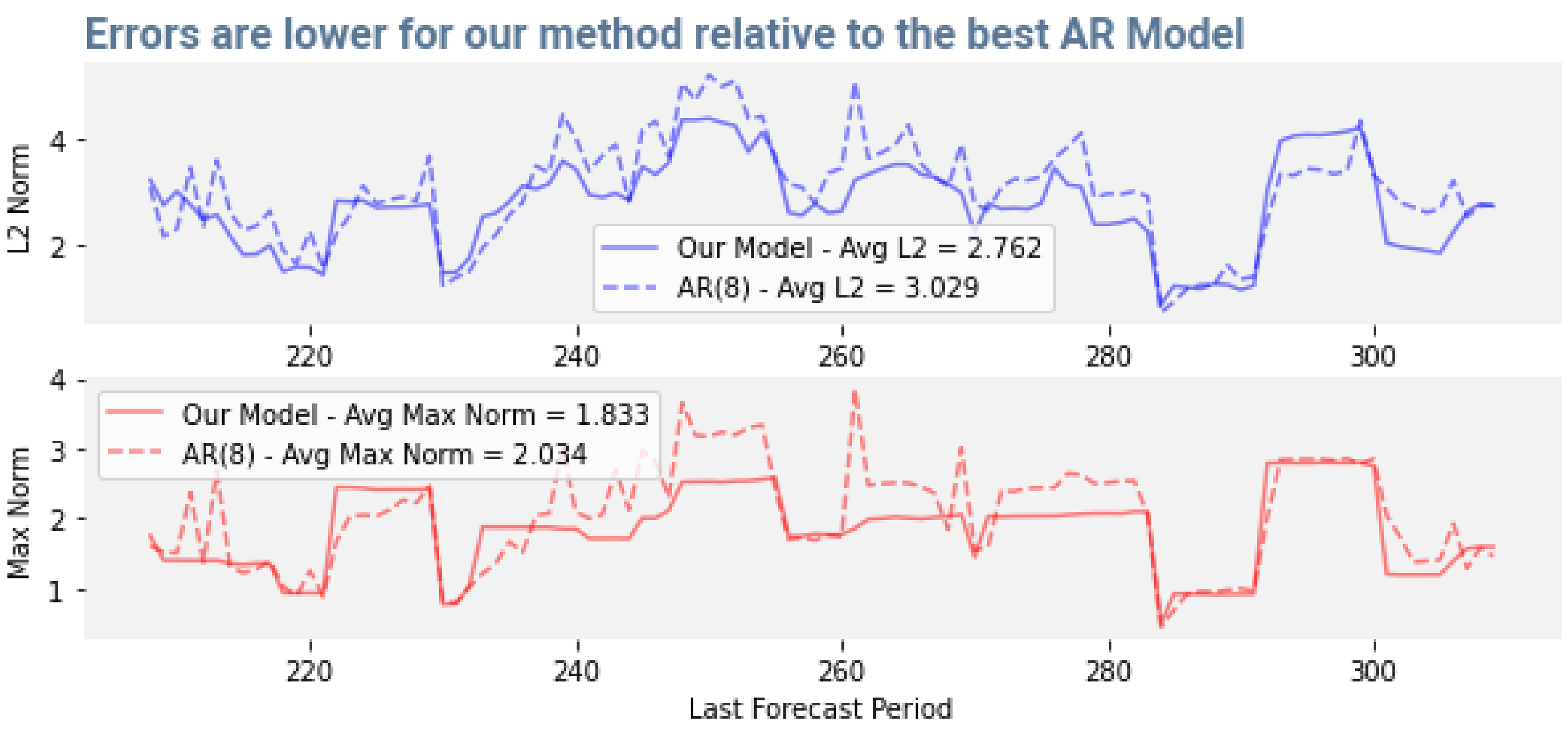


Figure: Our proposed model performs better than an AR(8) model chosen by the AIC

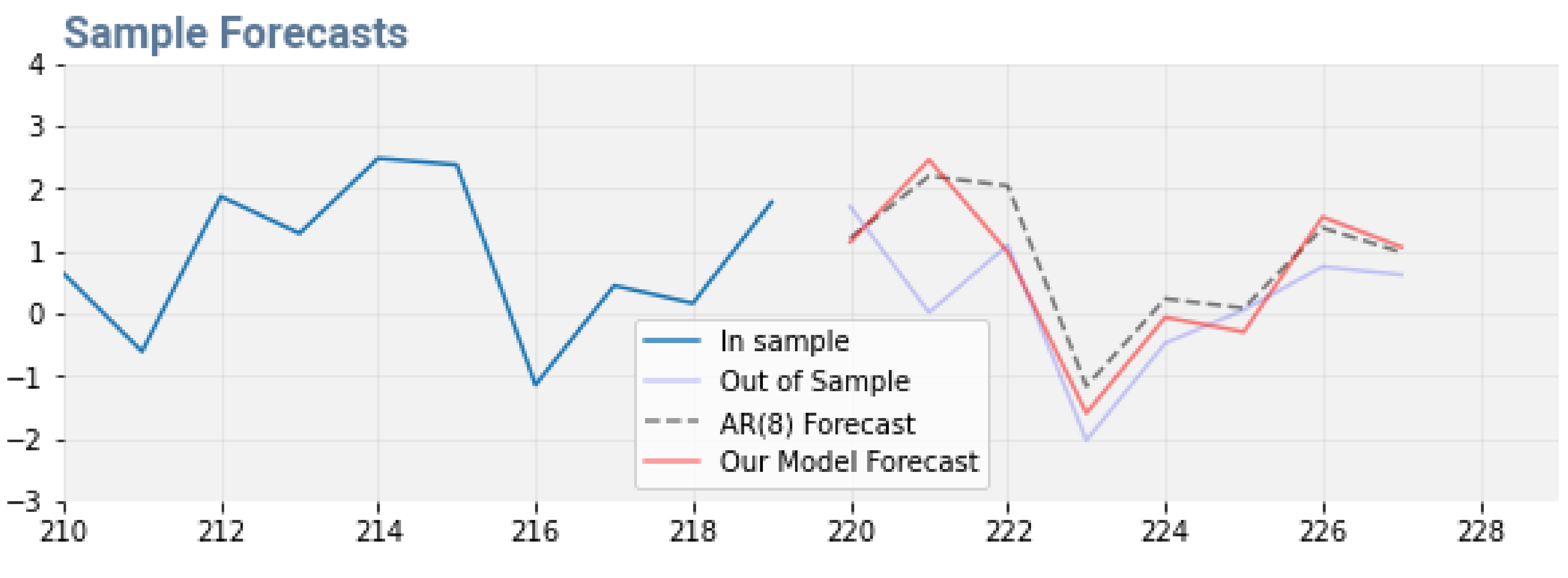
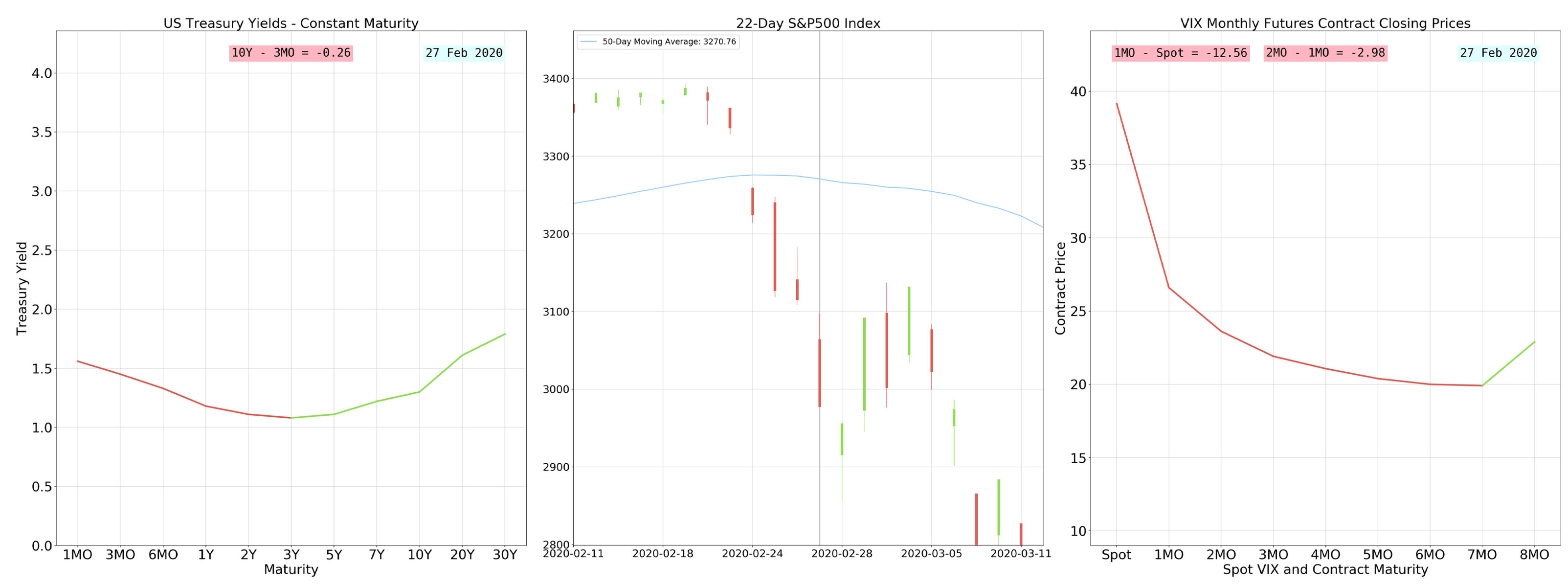


Figure: Forecasts from AR(8) and our proposed Model

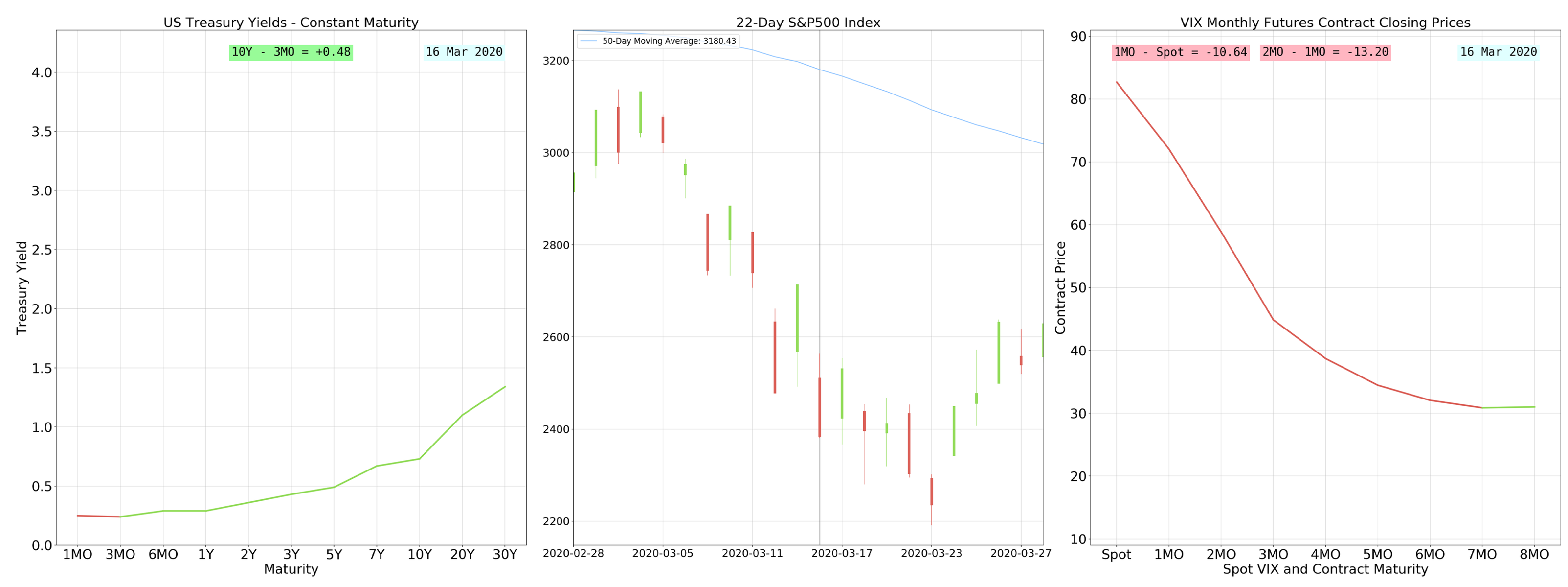
Motivating example: Yield curve, SP500, and VIX curve [9, 11]



OPTIMAL PORTFOLIO Research Group
optimalportfolio.github.io

Yield Curve &
VIX Term Structure Inversion
27 Feb 2020

Spot VIX: 39.16
VIX Futures Front Month: 26.60
Treasury, 1 Month: 1.56
Treasury, 10 Year: 1.30



OPTIMAL PORTFOLIO Research Group
optimalportfolio.github.io

Market Wide Circuit-Breaker
16 Mar 2020 - Level 1

Spot VIX: 82.69 (All-Time Closing High)
VIX Futures: 72.05 (March 18th Expiry)
Treasury, 1 Month: 0.25

Figure: The February 2020 US Treasury yield curve inversion, coupled with a negatively sloping VIX futures term structure, preceded the pandemic-induced market crash of March 2020 (above). During the crisis, the VIX term structure exhibited extreme inversion, then accompanied by a drop in the overall level of Treasury yields (below).

There is widespread existing literature on the predictive power of the VIX and interest rate term structure [2,7,8,9], and we can observe from preliminary data analysis that a relationship likely exists between the shape of the VIX futures term structure and the US Treasury yield curve, which offers the potential to forecast significant macroeconomic regime changes. However, this relationship may not remain static over time: parameters may change, and the functional form of this relationship may vary over different windowed periods. This motivates the adoption of a novel learning technique to best understand h_t and associate high-dimensional modelling to time series.

Adaptive learning: theory [11]

An adaptive learning on a time series dataset is a learning method in which there is a learning function l, a set of models H, and estimation techniques $\Lambda_h : \Theta(h) \mapsto \theta(h) \quad \forall h \in H$ such that for all time $t \in T^{\text{validation}} \cup T^{\text{test}}$, we are able to obtain the following:

- $\forall h \in H$, an estimated statistic $\theta_t(h)$
- $\forall h \in H$, an estimated forecast $y_{t+k|t}(h)$
- an optimal model $h^* \in H$ with an induced optimal forecast $y_{t+k|t}(h^*)$

Adaptive learning: algorithm [10: p.18]

1. Produce $\tilde{H}_t \subset H$ subject to computational capacity and statistical restrictions (e.g. flat or boundary MLE due to stationarity).
2. Obtain $\theta_t(h_t)$ for all $h_t \in \tilde{H}_t$, and subsequently: $h^*_t = \arg\min_{h \in \tilde{H}_t} \ell(\Phi_t, h, H \setminus \{h\})$

References

1. Akaike, H. (1974). "A new look at the statistical model identification", *IEEE Transactions on Automatic Control*, 19 (6): 716–723.
2. Estrella, A. (2005). "Why Does the Yield Curve Predict Output and Inflation?", *The Economic Journal*, 2005; 115: 722-744.
3. Prado, R. and West, M. (2010). Time Series: Modeling, Computation, and Inference. *Texts in Statistical Science*
4. Barber, D., Cemgil, A. and Chiappa, S. (2011). Bayesian Time Series Models. *Cambridge University Press*.
5. Efron, Bradley and Hastie, Trevor (2016). Computer age statistical inference. *Cambridge University Press*.
6. Tsay, R.S. and Chen, R. (2018). Nonlinear Time Series Analysis. *Wiley*.
7. Yang, Parley Ruogu (2019). *Bank of England: Modelling with Big Data and Machine Learning Conference* https://parleyyang.github.io/prev/BoESlides.pdf
8. Fassas, Athanasios P. and Nikolas Hourvouliades (2019). "VIX Futures as a Market Timing Indicator". *Journal of Risk and Financial Management*
9. Yang, Parley Ruogu (2020). Using The Yield Curve To Forecast Economic Growth. *Journal of Forecasting*. 2020; 39: 1057– 1080. https://doi.org/10.1002/for.2676
10. Yang, Parley Ruogu (2021). Forecasting High-Frequency Financial Time Series: An Adaptive Learning Approach With the Order Book Data. https://arxiv.org/abs/2103.00264
11. Yang, Parley Ruogu. McElwee, R. Lucas, R. Dolan, H. Shah, N. Schelpe, C (2021). Adaptive Learning On Time Series: Methodology and Application to Finance (Work In Progress)