# Synthetic Volatility Forecasting and Other Aggregation Techniques for Time Series Forecasting Preliminary Exam

David Lundquist<sup>1</sup>, Daniel Eck<sup>2</sup> (advisor)

April 10th, 2024

<sup>&</sup>lt;sup>1</sup>davidl11@ilinois.edu

<sup>&</sup>lt;sup>2</sup>dje13@illinois.edu

# A seemingly unprecedented event might make one ask

What does it resemble from the past?



# A seemingly unprecedented event might make one ask

- What does it resemble from the past?
- What past events are most relevant for our objectives?



# A seemingly unprecedented event might make one ask

- What does it resemble from the past?
- What past events are most relevant for our objectives?
- Oan we incorporate past events in a systematic, principled manner?

#### When would we ever have to do this?

• Event-driven investing strategies (unscheduled news shock)



#### When would we ever have to do this?

- Event-driven investing strategies (unscheduled news shock)
- Pairs trading strategies



#### When would we ever have to do this?

- Event-driven investing strategies (unscheduled news shock)
- Pairs trading strategies
- Scheduled macroeconomic news possibly pre-empted by a news leak

#### Fed Likely to Consider 0.75-Percentage-Point Rate Rise This Week

Officials had signaled plans to raise interest rates in half-point increments before recent deterioration in data

By Nick Timiraos Follow

Updated June 13, 2022 7:47 pm ET



Example (Weekend of March 6th - 8th, 2020)



# Oil nose-dives as Saudi Arabia and Russia set off 'scorched earth' price war

PUBLISHED SUN, MAR 8 2020+9:01 AM EDT | UPDATED MON, MAR 9 2020+5:33 PM EDT

# Oil crashes by most since 1991 as Saudi Arabia launches price war



# Punchline of the paper

Credible forecasting is possible under news shocks, so long as we incorporate external information to account for the nonzero errors.

# Punchline of the paper

Credible forecasting is possible under news shocks, so long as we incorporate external information to account for the nonzero errors.

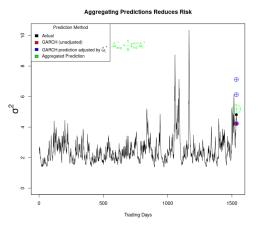


Figure: Adjusting our One-Step-Ahead Forecast Using Only Arithmetic Mean of Donors

April 10th, 2024

# Background and related methods

#### Volatility Modeling

- GARCH is slow to react to shocks (Andersen et al. 2003)
- Asymmetric GARCH models catch up faster but need post-shock data
- Realized GARCH (Hansen, Huang, and Shek 2012), in our setting, would require post-shock information and/or high-frequency data in order to outperform, and Realized GARCH is highly parameterized

# Background and related methods

#### Forecast Adjustment

- Clements and Hendry 1998; Clements and Hendry 1996 laid the groundwork for modeling nonzero errors in time series forecasting
- Guerrón-Quintana and Zhong 2017 use a series' own errors to correct the forecast for that series
- ullet Dendramis, Kapetanios, and Marcellino 2020 use a similarity-based procedure to correct  $\hat{eta}$  in time series forecasts
- Foroni, Marcellino, and Stevanovic 2022 adjust pandemic-era forecasts using intercept correction techniques and data from Great Financial Crisis
- Lin and Eck 2021 use distanced-based weighting (a similarity approach) to aggregate and weight fixed effects from a donor pool

# Outline

- Introduction
- Setting
- Post-shock Synthetic Volatility Forecasting Methodology
- 4 Properties of Volatility Shock and Shock Estimators
- Real Data Example
- Mumerical Examples
- Future directions for Synthetic Volatility Forecasting
- Supplement



#### Setting

### Premise: News has broken but markets are closed

- After-hours trading provides a poor forum in which to digest news
- News constitutes public, material information for one or more traded assets
- The qualitative aspects of the news provide a basis upon which to
  - match to past news shocks
  - match in a p-dimensional covariate space

#### A Primer on GARCH

#### Definition

Let  $\{a_t\}$  denote an observable, real-valued discrete-time stochastic process.

We call  $\{a_t\}$  a strong GARCH process (Francq and Zakoian 2019) with respect to  $\{\epsilon_t\}$  iff

$$\begin{split} \sigma_t^2 &= \omega + \sum_{k=1}^m \alpha_k a_{t-k}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \\ a_t &= \sigma_t \epsilon_t \\ \epsilon_t \stackrel{\textit{iid}}{\sim} E[\epsilon_t] &= 0, \textit{Var}[\epsilon_t] = 1 \\ \forall k, j, \alpha_k, \beta_j &\geq 0 \\ \forall t, \omega, \sigma_t &> 0 \end{split}$$

# Volatility Equation with an exogenous term: GARCH-X

$$\sigma_t^2 = \omega + \sum_{k=1}^m \alpha_k \mathbf{a}_{t-k}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 + \gamma^T \mathbf{x}_t \ .$$

We will be looking at only one exogenous term.



#### Model Preliminaries

Let  $I(\cdot)$  be an indicator function.

Let  $T_i$  denote the time length of the time series i for i = 1, ..., n + 1.

Let  $T_i^*$  denote the largest time index prior to news shock, with  $T_i^* < T_i$  (i.e. we assume at least one post-shock observation).

Let  $\delta, \mathbf{v}_{i,t} \in \mathbb{R}^p, \mathbf{x}_{i,t} \in \mathbb{R}^d$ .

# Model Setup

For  $t = 1, ..., T_i$  and i = 1, ..., n + 1, the model  $\mathcal{M}_1$  is defined as

$$\begin{split} \sigma_{i,t}^2 &= \omega_i + \omega_{i,t}^* + \sum_{k=1}^{m_i} \alpha_{i,k} \mathbf{a}_{i,t-k}^2 + \sum_{j=1}^{s_i} \beta_{i,j} \sigma_{i,t-j}^2 + \gamma_i^\mathsf{T} \mathbf{x}_{i,t} \\ \mathcal{M}_1 \colon &\quad \mathbf{a}_{i,t} = \sigma_{i,t} ((1 - D_{i,t}^{\mathsf{return}}) \epsilon_{i,t} + D_{i,t}^{\mathsf{return}} \epsilon_i^*) \\ &\quad \omega_{i,t}^* = D_{i,t}^{\mathsf{vol}} [\mu_{\omega^*} + \delta' \mathbf{v}_{i,t} + u_{i,t}], \end{split}$$

with error structure

$$\begin{split} & \epsilon_{i,t} \stackrel{iii}{\sim} \mathcal{F}_{\epsilon} \text{ with } \ \mathbb{E}_{\mathcal{F}_{\epsilon}}(\epsilon) = 0, \mathrm{Var}_{\mathcal{F}_{\epsilon}}(\epsilon) = 1 \\ & \epsilon_{i,t}^* \stackrel{iiid}{\sim} \mathcal{F}_{\epsilon^*} \text{ with } \ \mathbb{E}_{\mathcal{F}_{\epsilon^*}}(\epsilon) = \mu_{\epsilon^*}, \mathrm{Var}_{\mathcal{F}_{\epsilon^*}}(\epsilon^*) = \sigma_{\epsilon^*}^2 \\ & u_{i,t} \stackrel{iiid}{\sim} \mathcal{F}_{u} \text{ with } \ \mathbb{E}_{\mathcal{F}_{u}}(u) = 0, \mathrm{Var}_{\mathcal{F}_{u}}(u) = \sigma_{u}^2 \\ & \epsilon_{i,t} \perp \!\!\! \perp \!\!\! \perp \epsilon_{i,t}^* \perp \!\!\! \perp u_{i,t} \end{split}$$

where  $D_{i,t}^{return} = I(t \in \{T_i^* + 1, ..., T_i^* + L_{i,return}\})$  and  $D_{i,t}^{vol} = I(t \in \{T_i^* + 1, ..., T_i^* + L_{i,vol}\})$  and  $L_{i,return}, L_{i,vol}$  denote lengths of log return and volatility shocks, respectively.

Note: we will be looking GARCH(1,1) only in this presentation.



#### Our Model is Nested inside a Factor Model

Consider  $M_1$  in the context of the factor model from Abadie, Diamond, and Hainmueller 2010, where an untreated unit is governed by:

$$Y_{i,t}^{N} = \delta_{t} + \boldsymbol{\theta}_{t}^{\prime} \mathbf{Z}_{i} + \boldsymbol{\lambda}_{t}^{\prime} \boldsymbol{\mu}_{i} + \varepsilon_{i,t}$$

which nests the GARCH model's volatilty equation as well as the ARMA representation of a GARCH model, where

 $\delta_t \sim \omega$  , a location parameter shared across donors

 $m{ heta}_t \sim m{lpha}_k$  , a vector of ARCH parameters and other coefficients shared across donors

 $\mathbf{Z}_i \sim \mathbf{\emph{a}}_{i,t-k},$  a vector of observable quantities specific to each donor

 $oldsymbol{\lambda}_t \sim oldsymbol{eta}_i,$  a vector of GARCH parameters shared across donors

 $oldsymbol{\mu}_i \sim \sigma_{i,t-j}^2$  , a vector of latent quantities specific to each donor



### Volatility Profile of a Time Series

Consider the  $p \times n$  matrix that stores donor and covariate information at time t

$$\mathbf{V}_{t} = \begin{pmatrix} \hat{\alpha}_{1,t} & \hat{\alpha}_{t,2} & \cdots & \hat{\alpha}_{t,n} \\ \hat{\beta}_{1,t} & \hat{\beta}_{t,2} & \cdots & \hat{\beta}_{t,n} \\ \vdots & \vdots & \ddots & \vdots \\ RV_{1,t} & RV_{2,t} & \cdots & RV_{n,t} \\ RV_{1,t-1} & RV_{2,t-1} & \cdots & RV_{n,t-1} \\ \vdots & \vdots & \ddots & \vdots \\ IV_{1,t-1} & IV_{2,t-1} & \cdots & IV_{n,t} \\ IV_{1,t-1} & IV_{2,t-1} & \cdots & IV_{n,t-1} \\ \vdots & \vdots & \ddots & \vdots \\ Ir_{1,t} & |r_{2,t}| & \cdots & |r_{n,t}| \\ |r_{1,t-1}| & |r_{2,t-1}| & \cdots & |r_{n,t-1}| \end{pmatrix}$$

where RV denotes realized variance and IV the implied volatility



Covariates chosen for inclusion in a given volatility profile may be any  $\mathcal{F}_t$ -measurable function, for example

- levels
- differences in levels
- log returns
- percentage returns
- measurable transformations of the above

Key criterion for inclusion: how plausible is the covariate as a proxy for risk conditions for the volatility series to be forecasted?

# Forecasting

#### We present two forecasts:

$$\text{Forecast 1: } \hat{\sigma}_{\textit{unadjusted}}^2 = \hat{\mathbb{E}}[\sigma_{1,T_1^*+1}^2|\mathcal{F}_{\mathcal{T}^*}] = \hat{\omega}_i + \sum_{k=1}^{m_i} \hat{\alpha}_{i,k} \mathbf{a}_{i,t-k}^2 + \sum_{j=1}^{s_i} \hat{\beta}_{i,j} \sigma_{i,t-j}^2 + \hat{\gamma}_i^\mathsf{T} \mathbf{x}_{i,t}]$$

Forecast 2: 
$$\hat{\sigma}^2_{adjusted} = \hat{\mathbb{E}}[\sigma^2_{1,T_1^*+1}|\mathcal{F}_{T^*}] + \hat{\boldsymbol{\omega}}^* = \hat{\omega}_i + \sum_{k=1}^{m_i} \hat{\alpha}_{i,k} a^2_{i,t-k} + \sum_{i=1}^{s_i} \hat{\beta}_{i,j} \sigma^2_{i,t-j} + \hat{\gamma}^T_i \mathbf{x}_{i,t} + \hat{\boldsymbol{\omega}}^*$$
.

# Excess Volatility Estimators

• Observe the pair  $(\{\hat{\omega}_i^*\}_{i=2}^{n+1}, \{\mathbf{v}_i\}_{i=2}^{n+1}).$ 

# Excess Volatility Estimators

- Observe the pair  $(\{\hat{\omega}_i^*\}_{i=2}^{n+1}, \{\mathbf{v}_i\}_{i=2}^{n+1}).$
- Goal: recover weights  $\{\pi_i\}_{i=2}^{n+1} \in \Delta^n$  and compute  $\hat{\omega}^* \coloneqq \sum_{i=2}^{n+1} \pi_i \hat{\omega}_i^*$ , our forecast adjustment term.

# Excess Volatility Estimators

- Observe the pair  $(\{\hat{\omega}_i^*\}_{i=2}^{n+1}, \{\mathbf{v}_i\}_{i=2}^{n+1}).$
- Goal: recover weights  $\{\pi_i\}_{i=2}^{n+1} \in \Delta^n$  and compute  $\hat{\omega}^* := \sum_{i=2}^{n+1} \pi_i \hat{\omega}_i^*$ , our forecast adjustment term.
- Following Abadie, Diamond, and Hainmueller 2010; Abadie and Gardeazabal 2003, let  $\|\cdot\|_S$  denote any semi-norm on  $\mathbb{R}^p$ , and define

$$\{\pi\}_{i=2}^{\mathit{n}+1} = \operatorname*{arg\,min}_{\pi} \|\mathbf{v}_{1,\mathit{T}^*} - \mathbf{V}_{\mathit{T}^*}\pi\|_{\mathbf{S}} \ .$$

#### **Ground Truth Estimators**

We use realized volatility (RV)

- "model-free" in the sense that it requires no modeling assumptions (Andersen and Benzoni 2010).
- RV can be decomposed into the sum of a continuous component and a jump component, with the latter being less predictable and less persistent (Andersen, Bollerslev, and Diebold 2007), cited in De Luca et al. 2006, two facts that further motivate our method

# Realized Volatility Estimation

Examine K units of of time; each unit is divided into m intervals of length  $\frac{1}{m}$ . Let  $p_t = \log P_t$ , and let  $\tilde{r}(t, \frac{1}{m}) = p_t - p_{t-\frac{1}{m}}$  (Andersen and Teräsvirta 2009).

Estimate variance of *i*th log return series using Realized Volatility of the K consecutive trading days that conclude with day t, denoted  $RV_{i,t}^{K,m}$ , using

$$RV_{i,t}^{K,m} = \frac{1}{K} \sum_{v=1}^{Km} \tilde{r}^2(v/m, 1/m),$$

where the K trading days have been chopped into Km equally-sized blocks.

Assuming the K units  $\tilde{r}(t,1)=p_t-p_{t-1}$  are s.t.  $\tilde{r}(t,1)\stackrel{iid}{\sim} N(\mu,\delta^2)$ , it is easily verified that

$$\mathbb{E}[RV^{K,m}] = \frac{\mu^2}{m} + \delta^2,$$

which is a biased but consistent estimator of the variance. We pick m=77, corresponding to the 6.5-hour trading day chopped into 5-minute blocks, omitting first five-minutes of the day.

#### Loss Functions

Aim: point forecasts for  $\sigma_{1,T^*+h}^2|\mathcal{F}_{T^*}$ , h=1,2,..., the h-step ahead conditional variance for the time series under study

Let  $L^h$  with the subscripted pair {prediction method, ground truth estimator}, denote the loss function for an h-step-ahead forecast using a given prediction function and ground truth estimator.

# Loss Function Examples

For example, one loss possible function of interest in this study is the 1-step-ahead MSE using Synthetic Volatility Forecasting and Realized Volatility:

$$\mathsf{MSE}^1_{\mathsf{SVF},\;\mathsf{RV}} = (\hat{\sigma}^2_{\mathit{SVF}} - \hat{\sigma}^2_{\mathit{RV}})^2$$

Also of interest in mean absolute percentage error for an h-step-ahead forecast, defined as

$$\mathsf{MAPE}^h_{method,groundtruth} = \frac{|\hat{\sigma}^2_{h,method} - \hat{\sigma}^2_{h,groundtruth}|}{\hat{\sigma}^2_{h,groundtruth}}$$

#### Our choice of Loss Function

Finally, we introduce the QL (quasi-likelihood) Loss (Brownlees, Engle, and Kelly 2011):

$$\mathsf{QL}^h_{\mathsf{method},\mathsf{groundtruth}} = \frac{\hat{\sigma}_{h,\mathsf{method}}^2}{\hat{\sigma}_{h,\mathsf{groundtruth}}^2} - \log \frac{\hat{\sigma}_{h,\mathsf{method}}^2}{\hat{\sigma}_{h,\mathsf{groundtruth}}^2} - 1 \; .$$

What distinguishes QL Loss?

- Multiplicative rather than additive
- As Brownlees, Engle, and Kelly 2011 explain, "[a]mid volatility turmoil, large MSE losses will be a consequence of high volatility without necessarily corresponding to deterioration of forecasting ability. The QL avoids this ambiguity, making it easier to compare losses across volatility regimes."

# Two Consistency Results

#### **Proposition**

#### Assume

- For each i,  $\{a_{i,t}\}_{t=0,...,T_i}$  obeys a GARCH-X(m, s) with volatility shocks found in  $\mathcal{M}_1$ , where  $T_i$  is the length of the ith series.
- $\bullet$  For each  $i, \{\omega_{i,t}^*\}_{t=0,...,T_i}$  is potentially non-zero at  $\{T_i^*+1,...,T_i^*+k\}$ ,  $\omega_{iT^*+1}^* \equiv ... \equiv \omega_{iT^*+k}^*$ , and zero otherwise, where the arrival of  $T_i^*$  is governed by a time-invariant distribution on  $\{a_{i,t}\}_{t=0,...,T_i-1}$ .
- The conditions in Assumption 0 of Han and Kristensen 2014 prevail.

Then for any i,  $\hat{\omega}_{i}^{*}$ ,  $\xrightarrow{p} \omega_{i}^{*}$ .

#### **Proposition**

#### Assume

- All conditions from the previous proposition.
- 2 There exist weights  $\{\pi_i\}_{i=2}^{n=1}$  such that  $\mathbf{v}_{1,T^*} = \sum_{i=2}^{n+1} \pi_i \mathbf{v}_{i,T^*}$ .

Then  $\hat{\sigma}^2_{adjusted} \xrightarrow{p} \sigma^2_{1,T^*+1}$ .

Synthetic Volatility Forecasting and Other Aggregation

# Why apply our method to the 2016 US Election?

- You can win the US Presidency without a majority.
- No incumbent candidate
- Donald J. Trump espoused unorthodox, populist positions on healthcare, trade, foreign policy
- Donald J. Trump had no record to assess or criticize
- It was not predicted hence it delivered news.



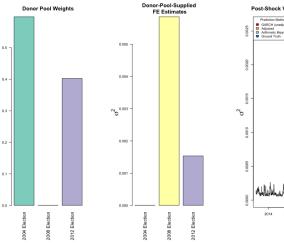
# iShares U.S. Financial Services ETF

Figure: IYG includes JPM, BAC, WF, CITI, among other financial majors

- **Model choice** GARCH(1,1) on the daily log return series of IYG in each donor
- Covariate Choice
  - previous 30 log returns of IYG (large pre-treatment period, in the language of SC)
  - log return Crude Oil (CL.F)
  - VIX
  - log return of the VIX
  - log returns of the 3-month, 5-year, 10-year, and 30-year US Treasuries
  - return of the most recently available monthly spread between AAA and BAA corporate debt
  - log return in the trading volume of the ETF IYG itself
- Onor pool construction US Elections from 2004, 2008, 2012
- Choice of estimator for volatility Sum of 77 squared five-minute returns generated between 9:35am and 4pm on November 9th, 2016.



### 2016 Election



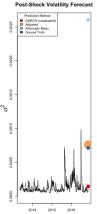


Figure: The volatility induced by the 2016 US election

# Hypotheses to Test Via Simulations

#### Ceteris paribus...

lacktriangle Distance-based weighting should outperform the unadjusted forecast as  $\delta$  increases (signal-to-noise)



# Hypotheses to Test Via Simulations

#### Ceteris paribus...

- ullet Distance-based weighting should outperform the unadjusted forecast as  $\delta$  increases (signal-to-noise)
- **②** Distance-based weighting should underperform as the variability in  $u_{i,t}$  increases (signal-to-noise)

# Hypotheses to Test Via Simulations

#### Ceteris paribus...

- ullet Distance-based weighting should outperform the unadjusted forecast as  $\delta$  increases (signal-to-noise)
- **②** Distance-based weighting should underperform as the variability in  $u_{i,t}$  increases (signal-to-noise)
- **3** Distance-based weighting should outperform the arithmetic mean as the variability of  $\mathbf{v}_{i,T^*}$  increases between donors (importance of linear signal)

# Hypotheses to Test Via Simulations

#### Ceteris paribus...

- lacktriangle Distance-based weighting should outperform the unadjusted forecast as  $\delta$  increases (signal-to-noise)
- ② Distance-based weighting should underperform as the variability in  $u_{i,t}$  increases (signal-to-noise)
- **3** Distance-based weighting should outperform the arithmetic mean as the variability of  $\mathbf{v}_{i,T^*}$  increases between donors (importance of linear signal)
- Distance-based weighting should underperform the arithmetic mean as  $\mu_{\omega^*}$  increases (importance of linear signal)

# Hypotheses to Test Via Simulations

#### Ceteris paribus...

- ullet Distance-based weighting should outperform the unadjusted forecast as  $\delta$  increases (signal-to-noise)
- ② Distance-based weighting should underperform as the variability in  $u_{i,t}$  increases (signal-to-noise)
- **3** Distance-based weighting should outperform the arithmetic mean as the variability of  $\mathbf{v}_{i,T^*}$  increases between donors (importance of linear signal)
- Distance-based weighting should underperform the arithmetic mean as  $\mu_{\omega^*}$  increases (importance of linear signal)
- When  $D_{i,T^*+1}^{return}=1=D_{i,T^*+1}^{vol}$ , i.e. when there is both a return shock and volatility shock, our adjustment methods should underperform due to failed identification in  $a_{T^*+1}=\sigma_{T^*+1}\epsilon_{T^*+1}$

# Simplest Simulation Setup

Most elementary simulation regime tests Hypothesis 1 and 2 by varying  $\delta$  and  $u_{i,t}$ .

Recall an  $\mathcal{M}_1$  model on the volatility, which is characterized by an exogenous shock to the volatility equation generated by an affine function of the covariates:

$$\sigma_{i,t}^{2} = \omega_{i} + \omega_{i,t}^{*} + \sum_{k=1}^{m_{i}} \alpha_{i,k} a_{i,t-k}^{2} + \sum_{j=1}^{s_{i}} \beta_{i,j} \sigma_{i,t-j}^{2} + \gamma_{i}^{T} \mathbf{x}_{i,t}$$

$$a_{i,t} = \sigma_{i,t} ((1 - D_{i,t}^{return}) \epsilon_{i,t} + D_{i,t}^{return} \epsilon_{i}^{*})$$

$$\omega_{i,t}^{*} = D_{i,t}^{vol} [\mu_{\omega^{*}} + \delta' \mathbf{v}_{i,t} + u_{i,t}]$$

$$D_{i,t}^{return} \equiv 0$$

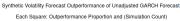




Figure: Fixed parameter values:  $\alpha = .1, \beta = .82, \mu_x = 1, \sigma_x = .1$ 

#### If we switch the values of $\alpha$ and $\beta$ , we see similar behavior.

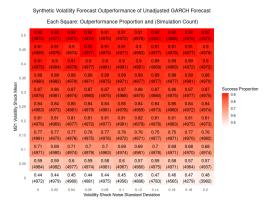


Figure: Fixed parameter values:  $\alpha = .82, \beta = .1, \mu_x = 1, \sigma_x = .1$ 

#### We shall group the extensions into five buckets:

- How much can we automate?
- Alternatives for fixed effect estimation
- Alternative estimators and estimands
- What can you do with a volatility forecast?
- Where else is distanced-based weighting useful?

## How much can we automate?

#### Use NLP to identify donors.

#### O You

What are some events that are similar to Chinese government's ban on iPhones for government officials in September 2023?

#### ChatGPT

The Chinese government's ban on iPhones for government officials in September 2023 is a notable example of government-led restrictions or bans on certain technology products or services due to security concerns or political reasons. Here are some events or actions that are somewhat similar:

- Huawel Ban in the United States: The United States government banned Huawel, a Chinese telecommunications company, from selling its equipment and infrastructure in the U.S. due to concerns about national security and espionage.
   This ban was initiated in 2019 and has had significant repercussions for Huawel's business oldobar.
- Russian Ban on Telegram: In 2018, the Russian government banned Telegram, a
  popular messaging app, due to the company's refusal to provide encryption keys to
  Russian security agencies. The ban resulted in the disruption of Telegram services
  within Russia and sparked debates about privacy and government surveillance.
- 3. India's Ban on Chinese Apps: In 2020, the Indian government banned several Chinese mobile apps, including TikTok and WeChat, citing concerns about data privacy and national security after a border clash between Indian and Chinese troops. The ban affected millions of users and led to diplomatic tensions between India and Chin.
- 4. Turkey's Wildpedia Ban. 1 2017, the Turkish government hanned access to Wildpedia, alleging that the website was involved in a smear campaign against Turkey. The ban stemmed from Wildpedia's refusal to remove content that the Turkish government deemed offensive. Despite international criticism, the ban remained in place for over two years before being lifet in 2020.
- 5. Iran's Ban on Messaging Apps: The Iranian government has intermittently banned messaging apps such as Telegram and WhatsApp over concerns about the spread of misinformation, political dissent, and national security threats. These bans have been part of broader efforts to control online communication and social media platforms within Iran.

These examples demonstrate how governments may impose bans or restrictions on technology products, services, or platforms for various reasons, including national security, political control, and regulatory compliance.

Synthetic Volatility Forecasting and Other Aggregation

000



## How much can we automate?

What if the variables in the volatility profile are difficult to specify?

Proposed solution:

Use shrinkage estimation to detect fleeting signals in the cross section of  $a_t^2$  (Chinco, Clark-Joseph, and Ye 2019).

# Alternative Ways of Estimating Fixed Effects

### High-frequency data?

- Realized GARCH with High-Frequency Data
- Stochastic Volatility

# Alternative Estimators and Estimands in Volatility Modeling

- Factors in volatility profile
- Overnight returns instead of open-to-close
- Signal Recovery Perspective (Ferwana and Varshney 2022)
- Stochastic Volatility: Correlation between errors
- Multivariate GARCH

# What can you do with a volatility forecast?

• Value-at-Risk using SVF-based  $\hat{\sigma}_t^2$ 

# New Frontiers in Distance-based Weighting

- ullet Integrate lessons from literature on under/over reactions to information shocks (Jiang and Zhu 2017)
- Synthetic Impulse Response Functions

# Synthetic Impulse Response Functions: A Proposal

### Suppose

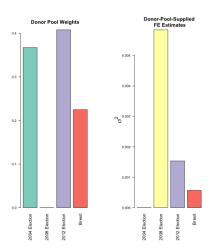
- We have a collection of *p*-variate time series of lengths  $T_i$ , i = 1, 2, ... n + 1).
- We are interested in the response of variable r to shocks in variable j,  $1 \le r \le j \le p$ .

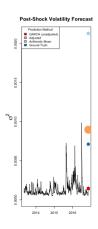
There many ways to estimate  $IRF_1(r, j)$ .

Can we somehow aggregate the estimates  $\widehat{IRF}_i(r,j)$ , i=2,3,...,n+1? Additional research questions:

- What DGP would best motivate/justify such a method?
- Which method of IRF estimation would perform best?

We analyze the real data example with Brexit included.





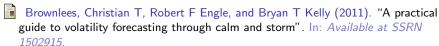
#### References I





- Andersen, Torben G and Luca Benzoni (2010). "Stochastic volatility". In: *CREATES Research Paper* 2010-10.
- Andersen, Torben G, Tim Bollerslev, and Francis X Diebold (2007). "Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility". In: *The review of economics and statistics* 89.4, pp. 701–720.
- Andersen, Torben G and Timo Teräsvirta (2009). "Realized volatility". In: Handbook of financial time series. Springer, pp. 555–575.
- Andersen, Torben G et al. (2003). "Modeling and forecasting realized volatility". In: *Econometrica* 71.2, pp. 579–625.

#### References II



- Chinco, Alex, Adam D Clark-Joseph, and Mao Ye (2019). "Sparse signals in the cross-section of returns". In: *The Journal of Finance* 74.1, pp. 449–492.
- Clements, Michael and David F Hendry (1998). Forecasting economic time series. Cambridge University Press.
- Clements, Michael P and David F Hendry (1996). "Intercept corrections and structural change". In: *Journal of Applied Econometrics* 11.5, pp. 475–494.
- De Luca, Giovanni et al. (2006). "Forecasting Volatility using High-Frequency Data". In: Statistica Applicata 18.
- Dendramis, Yiannis, George Kapetanios, and Massimiliano Marcellino (2020). "A similarity-based approach for macroeconomic forecasting". In: *Journal of the Royal Statistical Society Series A: Statistics in Society* 183.3, pp. 801–827.
- Ferwana, Ibtihal and Lav R Varshney (2022). "Optimal Recovery for Causal Inference". In: arXiv preprint arXiv:2208.06729.

#### References III



Foroni, Claudia, Massimiliano Marcellino, and Dalibor Stevanovic (2022). "Forecasting the Covid-19 recession and recovery: Lessons from the financial crisis". In: International Journal of Forecasting 38.2, pp. 596–612.



Francq, Christian and Jean-Michel Zakoian (2019). GARCH models: structure, statistical inference and financial applications. John Wiley & Sons.



Guerrón-Quintana, Pablo and Molin Zhong (2017). "Macroeconomic forecasting in times of crises". In.



Han, Heejoon and Dennis Kristensen (2014). "Asymptotic theory for the QMLE in GARCH-X models with stationary and nonstationary covariates". In: Journal of business & economic statistics 32.3, pp. 416–429.



Hansen, Peter Reinhard, Zhuo Huang, and Howard Howan Shek (2012). "Realized GARCH: a joint model for returns and realized measures of volatility". In: Journal of Applied Econometrics 27.6, pp. 877-906.



Jiang, George J and Kevin X Zhu (2017). "Information shocks and short-term market underreaction". In: Journal of Financial Economics 124.1, pp. 43-64.



Lin, Jilei and Daniel J Eck (2021). "Minimizing post-shock forecasting error through aggregation of outside information". In: International Journal of Forecasting.

April 10th, 2024