Forecast Adjustment Under Shocks and other Similarity-based Solutions to Unprecedented Events

Doctoral Defense

David Lundquist[1], Daniel Eck[2] (advisor)

December 19th, 2024

[1] davidl11@ilinois.edu
[2] dje13@illinois.edu

A High-level Summary Uniting All My Projects

Forecasting under non-ideal conditions like a change in the DGP or a lack of past information for units we wish to forecast: what do we do?

The unifying technique we explored was *correction* of model parameters using an aggregation strategy, where aggregation means that we make use of external data.

We began with particular models and then built upwards, establishing an abstract, general framework for correcting forecasts using external data that is widely applicable for forecasting and beyond.

Why is this valuable?

Canonical problem: forecasting when there's a location shift - here we quote clements and Hendry about the devastating impact of location shifts on forecasts

Common problem: predicting based on a limited feature space and/or small n

The framework we've offered is something others can 'plug-in' to easily and add to – in so many directions.

Bringing the causal inference literature into forecasting

Software developed

Questions guiding these slides (remove later)

1. Most common question asked is what you learned from the study you have done. You have to sum up your entire study in a few sentences and remember the technical terms you have mentioned in your research because that is what your examiner wants to hear from you. Forecasting is difficult; Heterogeneity of DGP
2. Why did you choose this topic? Although it seems narrow, the question of "what if you had a new unit and had very little past information to go on" is a common phenomenon in statistics. Examples: Shock to a time series Unscheduled scheduled A new seller joins amazon's platform, i.e. cold starts (Fatemi et al. 2023)
3. What does this contribute to the literature and econometrics more broadly?
4. How would you improve your work? Many ideas here. There were more directions than I could pursue.
5. Why did you choose this particular topic or what your inspiration behind this study was. This is one of the trickiest questions as you have to prove your convincing power to the panel of the teachers that what you did is valuable for the society and was worth their time.
6. Tell about how zealous you were about this particular problem.

Questions guiding these slides (remove later)

1. What is the importance of your study or how will it contribute or add up to the existing body of knowledge? Two entirely separate perspectives: (1) Post-shock forecasting is a novel research framework. (2) Post-shock forecasting builds on intercept corrections and other canonical questions.
2. What type of background research have you done for the study? In no particular order: synthetic control; convex geometry; convex optimization; p-value combination; FWER/FDR; fixed/mixed/random effects; panel data; linear time series models; RNN; econometrics
3. What are the limitations you have faced while writing? A ton of hyperparameters
4. What will you include if you are told to add something extra to the study? LLM-generated donors and covariates
5. What are the recommendations of your study? Relatively easy to answer: for an unprecedented event, locate it in the space of previous events.
6. 10. What was your hypothesis and how did you framed it? Signal to noise

Questions guiding these slides (remove later)

1. 12. If given a chance, would like to do something different with your work?
2. 13. What are the limitations you faced while dealing with your samples? Realized volatility is something that can be estimated with HF data
3. 14. How did you relate your study to the existing theories?
4. 15. What is the future scope of this study?
5. 19. How did you evaluate your work? Simulations and real data examples
6. How would you improve your work? By design, synthetic control cannot extrapolate. In the causal inference context, that may very well be a virtue. However, in the prediction context, it may hinder us. Note that random forest has the same problem.

A seemingly unprecedented event might make one ask

1. What does it resemble from the past?

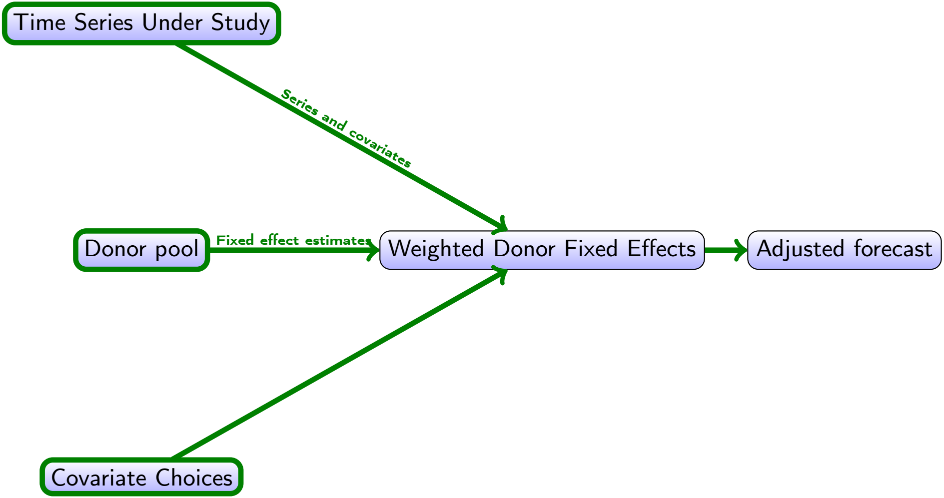
A seemingly unprecedented event might make one ask

1. What does it resemble from the past?
2. What past events are most relevant for our objectives?

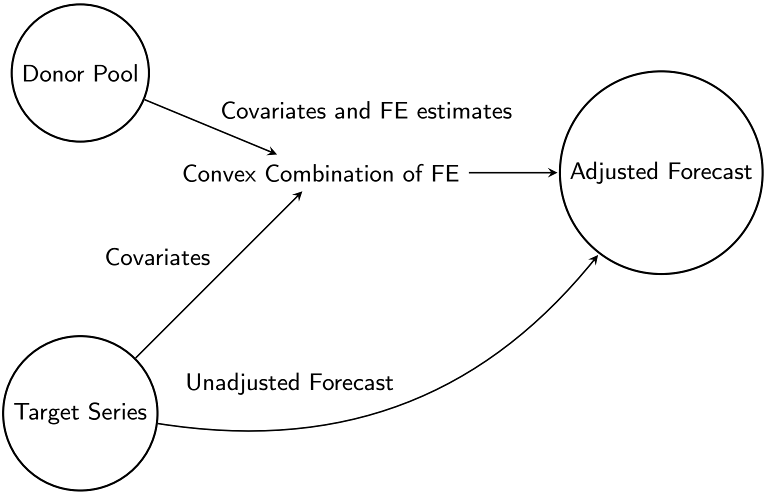
A seemingly unprecedented event might make one ask

1. What does it resemble from the past?
2. What past events are most relevant for our objectives?
3. Can we incorporate past events in a systematic, principled manner?

My Prelim: An incredibly Brief Recap



My Prelim: An Incredibly Brief Recap



My Prelim: An Incredibly Brief Recap

- Paper is under review at the *International Journal of Forecasting*

My Prelim: An Incredibly Brief Recap

- Paper is under review at the *International Journal of Forecasting*
- Event-driven investing strategies (unscheduled news shock)

My Prelim: An Incredibly Brief Recap

- Paper is under review at the *International Journal of Forecasting*
- Event-driven investing strategies (unscheduled news shock)
- Scheduled macroeconomic news possibly pre-empted by a news leak

ECONOMY

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)

Coronavirus   Health   Executive Order

Public Safety

MARCH 7, 2020 | Albany, NY

At Novel Coronavirus Briefing, Governor Cuomo Declares State of Emergency to Contain Spread of Virus

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

By Matt Egan, CNN Business

3 minute read · Updated 3:21 PM EDT, Mon March 9, 2020

The Post-prelim Direction of my research

- Extending similarity-based parameter correction to the state-of-the-art HAR model
- Extending similarity-based parameter correction to non-linear shock models
- Building out a general framework for parameter correction

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.

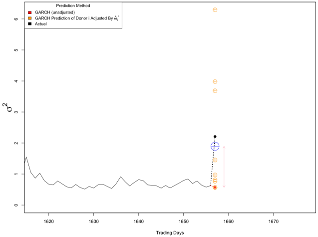


Figure: Adjusting our One-Step-Ahead Forecast Using Only Arithmetic Mean of Donors

Literature Review: if we want it, it goes here

uncomment the blocks below

What we will discuss in this section

1. Role of outside information
2. The Meaning and Use of Similarity

Outline

Premise: There is a breaking news at some fractional lag T-ϵ

- 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

$$\sigma_t^2 = \omega + \sum_{k=1}^{m} \alpha_k a_{t-k}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

$$a_t = \sigma_t \epsilon_t$$

$$\epsilon_t \overset{iid}{\sim} E[\epsilon_t] = 0, Var[\epsilon_t] = 1$$

$$\forall k, j, \alpha_k, \beta_j \geq 0$$

$$\forall t, \omega, \sigma_t > 0$$

Volatility Equation with an exogenous term: GARCH-X

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

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 $i = 1, \ldots, n+1$.

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

$$\mathcal{M}_1: \begin{aligned} \sigma^2_{i,t} &= \omega_i + \omega^*_{i,t} + \sum_{k=1}^{m_i} \alpha_{i,k} a^2_{i,t-k} + \sum_{j=1}^{s_i} \beta_{i,j} \sigma^2_{i,t-j} + \gamma_i^T \mathbf{x}_{i,t} \\ a_{i,t} &= \sigma_{i,t}((1 - D^{return}_{i,t})\epsilon_{i,t} + D^{return}_{i,t} \epsilon^*_i) \\ \omega^*_{i,t} &= D^{vol}_{i,t}[\mu_{\omega^*} + \delta' \mathbf{v}_{i,t} + u_{i,t}], \end{aligned}$$

with error structure

$$\begin{aligned} &\epsilon_{i,t} \overset{iid}{\sim} \mathcal{F}_\epsilon \text{ with } \mathrm{E}_{\mathcal{F}_\epsilon}(\epsilon) = 0, \mathrm{Var}_{\mathcal{F}_\epsilon}(\epsilon) = 1 \\ &\epsilon^*_{i,t} \overset{iid}{\sim} \mathcal{F}_{\epsilon^*} \text{ with } \mathrm{E}_{\mathcal{F}_{\epsilon^*}}(\epsilon) = \mu_{\epsilon^*}, \mathrm{Var}_{\mathcal{F}_{\epsilon^*}}(\epsilon^*) = \sigma^2_{\epsilon^*} \\ &u_{i,t} \overset{iid}{\sim} \mathcal{F}_u \text{ with } \mathrm{E}_{\mathcal{F}_u}(u) = 0, \mathrm{Var}_{\mathcal{F}_u}(u) = \sigma^2_u \\ &\epsilon_{i,t} \perp\!\!\!\perp \epsilon^*_{i,t} \perp\!\!\!\perp u_{i,t} \end{aligned}$$

where $D^{return}_{i,t} = I(t \in \{T^*_i + 1, \ldots, T^*_i + L_{i,return}\})$ and $D^{vol}_{i,t} = I(t \in \{T^*_i + 1, \ldots, T^*_i + L_{i,vol}\})$ 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 $\mathcal{M}_1$ in the context of the factor model from Abadie, Diamond, and Hainmueller 2010, where an untreated unit is governed by:

$$Y_{i,t}^{N} = \delta_t + \boldsymbol{\theta}_t^{'}\mathbf{Z}_i + \boldsymbol{\lambda}_t^{'}\boldsymbol{\mu}_i + \varepsilon_{i,t}$$

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

$\boldsymbol{\theta}_t \sim \boldsymbol{\alpha}_k$, a vector of ARCH parameters and other coefficients shared across donors

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

$\boldsymbol{\lambda}_t \sim \boldsymbol{\beta}_j$, a vector of GARCH parameters shared across donors

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

Covariate Profile of a Time Series

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

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

where RV denotes realized variance and IV the implied volatility

Significance of the Covariates

Covariates chosen for inclusion 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?

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 Benzoni 2008).

Estimate variance of ith 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. $\tilde{r}(t,1) \overset{iid}{\sim} N(\mu, \delta^2)$, 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.

Forecasting

We present two forecasts:

Forecast 1: $\hat{\sigma}^2_{unadjusted} = \hat{\mathbb{E}}[\sigma^2_{1,T^*_1+1}|\mathcal{F}_{T^*}] = \hat{\omega}_i + \sum_{k=1}^{m_i} \hat{\alpha}_{i,k} a^2_{i,t-k} + \sum_{j=1}^{s_i} \hat{\beta}_{i,j}\sigma^2_{i,t-j} + \hat{\gamma}_i^T \mathbf{x}_{i,t}$

Forecast 2: $\hat{\sigma}^2_{adjusted} = \hat{\mathbb{E}}[\sigma^2_{1,T^*_1+1}|\mathcal{F}_{T^*}] + \hat{\omega}^* = \hat{\omega}_i + \sum_{k=1}^{m_i} \hat{\alpha}_{i,k} a^2_{i,t-k} + \sum_{j=1}^{s_i} \hat{\beta}_{i,j}\sigma^2_{i,t-j} + \hat{\gamma}_i^T \mathbf{x}_{i,t} + \hat{\omega}^*$.

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 $\hat{\omega}^* := \sum_{i=2}^{n+1} \pi_i \hat{\omega}_i^*$, 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 and Gardeazabal 2003,Abadie, Diamond, and Hainmueller 2010, let $\|\cdot\|_{\mathbf{S}}$ denote any semi-norm on $\mathbb{R}^p$, and define

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

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.

IYG **iShares U.S. Financial Services ETF**

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

1. **Model choice** GARCH(1,1) on the daily log return series of IYG in each donor
2. **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
3. **Donor pool construction** US Elections from 2004, 2008, 2012
4. **Choice of estimator for volatility** Sum of 77 squared five-minute returns generated between 9:35am and 4pm on November 9th, 2016.

2016 Election

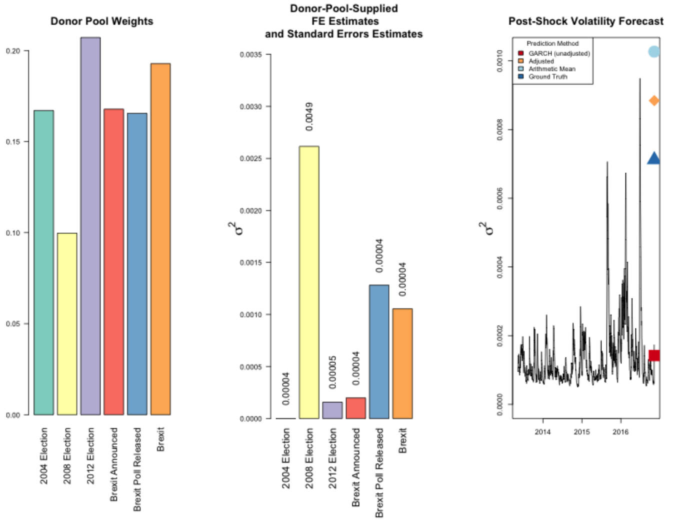


Figure: The volatility induced by the 2016 US election

The forecaster's decision tree

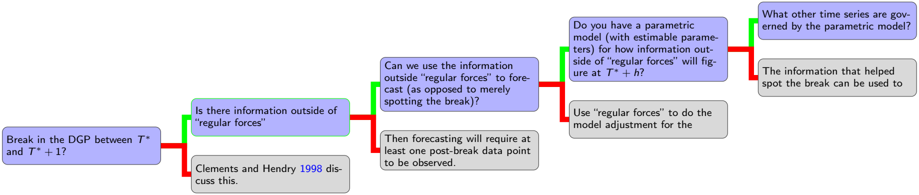


Figure: Forecast Model Adjustment: A Decision Tree

Formal Results

Proposition

Assume

1. *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.*
2. *For each $i, \{\omega^*_{i,t}\}_{t=0,...,T_i}$ is potentially non-zero at $\{T^*_i + 1, ..., T^*_i + k\}$, $\omega^*_{i,T^*+1} \equiv ... \equiv \omega^*_{i,T^*+k}$, and zero otherwise, where the arrival of T^*_i is governed by a time-invariant distribution on $\{a_{i,t}\}_{t=0,...,T_i-1}$.*
3. *The conditions in Assumption 0 of Han and Kristensen 2014 prevail.*

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

Proposition

Assume

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

Simplest Simulation Setup

Most elementary simulation regime tests Hypothesis 1 and 2 by varying δ 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:

$$\mathcal{M}_1 : \begin{aligned} \sigma^2_{i,t} &= \omega_i + \omega^*_{i,t} + \textstyle\sum_{k=1}^{m_i} \alpha_{i,k} a^2_{i,t-k} + \sum_{j=1}^{s_i} \beta_{i,j} \sigma^2_{i,t-j} + \gamma_i^T \mathbf{x}_{i,t} \\ a_{i,t} &= \sigma_{i,t}((1 - D^{return}_{i,t})\epsilon_{i,t} + D^{return}_{i,t} \epsilon^*_i) \\ \omega^*_{i,t} &= D^{vol}_{i,t}[\mu_{\omega^*} + \delta' \mathbf{v}_{i,t} + u_{i,t}] \\ D^{return}_{i,t} &\equiv 0 \end{aligned}$$

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?
- Can we extend the results of (Bodilsen and Lunde 2023)

How much can we automate?

Use NLP to identify donors.

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:

1. **Huawei Ban in the United States**: The United States government banned Huawei, 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 Huawei's business globally.
2. **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 China.
4. **Turkey's Wikipedia Ban**: In 2017, the Turkish government banned access to Wikipedia, alleging that the website was involved in a smear campaign against Turkey. The ban stemmed from Wikipedia'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 lifted 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.

Message ChatGPT...

How much can we automate?

What if the covariates 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 used as covariates
- 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

- Integrate lessons from literature on under/over reactions to information shocks (Jiang and Zhu 2017)
- Distance-based Weighting of Impulse Response Functions

Distance-based Weighting of Impulse Response Functions

Suppose

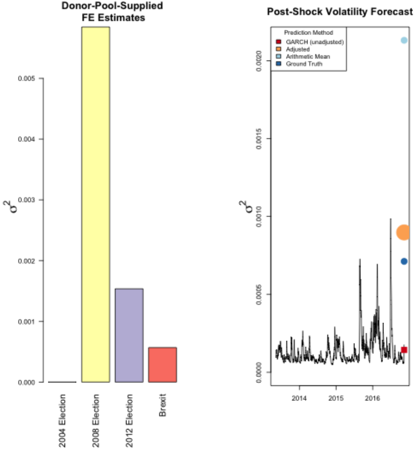
- We have a collection of p-variate time series of lengths $T_i, i = 1, 2, ...n + 1$.
- We are interested in the response of variable r to shocks in variable j, $1 \leq r \leq j \leq p$.

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

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program". In: *Journal of the American Statistical Association* 105.490, pp. 493–505.

Abadie, Alberto and Javier Gardeazabal (2003). "The Economic Costs of Conflict: A Case Study of the Basque Country". In: *American Economic Review* 93.1, pp. 113–132.

Andersen, Torben Gustav and Luca Benzoni (2008). "Realized Volatility, Working Paper 2008-14". In.

Bodilsen, Simon Tranberg and Asger Lunde (2023). "Exploiting news analytics for volatility forecasting". In: *Available at SSRN 4401032*.

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.

Fatemi, Zahra et al. (2023). "Mitigating cold-start forecasting using cold causal demand forecasting model". In: *arXiv preprint arXiv:2306.09261*.

References II

- Ferwana, Ibtihal and Lav R Varshney (2022). "Optimal Recovery for Causal Inference". In: *arXiv preprint arXiv:2208.06729*.
- Francq, Christian and Jean-Michel Zakoian (2019). *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons.
- 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.
- 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.