Synthetic Volatility Forecasting and Other Aggregation Techniques for Time Series Forecasting

Preliminary Exam

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

March 27, 2024

[1]davidl11@ilinois.edu

[2]dje13@illinois.edu

A seemingly unprecedented event might make one ask

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

When would we ever have to do this?

- Event-driven investing strategies (unscheduled news shock)
- Pairs trading strategies
- Structural shock to macroeconomic conditions (scheduled news possibly pre-empted by news shock)
- Biomedical panel data subject to exogenous shock or interference

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

Punchline of the paper

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

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 Augmentation

- 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
- Dendramis, Kapetanios, and Marcellino 2020 use a similarity-based procedure to correct linear parameters 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

Premise: News has broken but markets are closed

- After-hours trading provides a poor forum in which to digest news
- The news constitutes public, material information relevant to one or more traded assets
- The qualitative aspects of the news provide basis upon which to match to past events

A Primer on GARCH

Let $\{a_t\}$ denote an observable, real-valued discrete-time stochastic process. We say $\{a_t\}$ is a strong GARCH process 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

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

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 $T_i^* < T_i$.

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

Model Setup

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

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

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.

Model Details

Let $\mathcal{M}_0$ denote the subclass of $\mathcal{M}_1$ models such that $\delta \equiv 0$.

Note that $\mathcal{M}_0$ assumes that ω_i^* have no dependence on the covariates and are i.i.d. with $\mathbb{E}[\omega_i^*] = \mu_{\omega^*}$.

Our Model is Nested inside a Factor Model

Consider the preceding with the factor model from Abadie, Diamond, and Hainmueller 2010, where an untreated unit is governed by:

$$Y^N_{i,t} = \delta_t + \boldsymbol{\theta}_t \mathbf{Z}_i + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \varepsilon_{i,t}$$

which happens to nest the GARCH model's volatilty equation (putting aside that σ_t is latent in the GARCH model) as well as 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

Volatility Profile of a Time Series

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

where RV denotes realized variance, IV the implied volatility

Significance of the Volatility Profile

Covariates chosen for inclusion in a given volatility profile may be levels, log differences in levels, percentage changes in levels, or absolute values thereof, among many choices.

Forecasting

We present two forecasts:

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

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

Excess Volatility Estimators

Observe the pair $(\{\hat{\omega}_i^*\}_{i=2}^{n+1}, \{\mathbf{v}_i\}_{i=2}^{n+1})$.
Goal: recover weights $\{\pi_i\}_{i=2}^{n+1} \in \Delta^n$

Compute: $\hat{\omega}^* := \sum_{i=2}^{n+1} \pi_i \hat{\omega}_i^*$, our forecast adjustment term.

the set $\{\pi_i\}_{i=2}^{n+1}$ is deterministic

Following Abadie, Diamond, and Hainmueller 2010; Abadie and Gardeazabal 2003, let $\|\cdot\|_{\mathbf{S}}$ denote any semi-norm on $\mathbb{R}^p$, and define

$$\{\pi\}_{i=2}^{n+1} = \arg\min_{\pi} \|\mathbf{v}_1 - \mathbf{V}_t \pi\|_{\mathbf{S}} \ .$$

Ground Truth Estimators

We use realized volatility (RV); has virtue of being "model-free" in the sense that it requires no modeling assumptions (Andersen and Benzoni 2010). RV can be decomposed into the sum of a continous component and a jump component, with the former being less predictable and less persistent (Andersen, Bollerslev, and Diebold 2007), cited in De Luca et al. 2006, two factors that further motivate the method employed herein.

Realized Volatility Estimation

Suppose we examine K units of of time, where 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 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

Loss Functions

Goal: point forecasts for $\sigma^2_{1,T^*+h}|\mathcal{F}_{T^*}$, $h = 1, 2, ..., H$, the h-step ahead conditional variance for the time series under study, up to a forecast length of H.

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 function of interest in this study is the 1-step-ahead MSE using Synthetic Volatility Forecasting and Realized Volatility:

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

In more generality, for a volatility forecast with forecast length H, the MSE is

$$\text{MSE}^{H}_{method,groundtruth} = \frac{1}{H}\sum_{h=1}^{H}(\hat{\sigma}^2_{h,method} - \hat{\sigma}^2_{h,groundtruth})^2$$

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

$$\text{MAPE}^{H}_{method,groundtruth} = \frac{1}{H}\sum_{h=1}^{H}\frac{|\hat{\sigma}^2_{h,method} - \hat{\sigma}^2_{h,groundtruth}|}{\hat{\sigma}^2_{h,groundtruth}}$$

Our choice of Loss Function

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

$$\mathrm{QL}^{H}_{method,groundtruth} = \frac{1}{H}\sum_{h=1}^{H}\left(\frac{\hat{\sigma}^2_{h,method}}{\hat{\sigma}^2_{h,groundtruth}} - \log\frac{\hat{\sigma}^2_{h,method}}{\hat{\sigma}^2_{h,groundtruth}} - 1\right) .$$

What distinguishes QL Loss is that it is multiplicative rather than additive. This has benefits, both practical and theoretical. As Brownlees, Engle, and Kelly 2011 explains, "[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." For this reason, we proceed to evaluate the method, both in simuations and real data examples, using the QL loss.

1. **Model choice** GARCH(1,1) on the daily log return series of IYG in each donor
2. **Covariate Choice** log return Crude Oil (CL.F), the VIX (VIX) and the log return of the VIX, the 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, widely considered a proxy for lending risk (Goodell and Vähämaa 2013; Kane, Marcus, and Noh 1996). We also include the log return in the trading volume of the ETF IYG itself, which serves as a proxy for panic.
3. **Donor pool construction** the three most recent US presidential elections prior to the 2016 election. The three US presidential elections are the only presidential elections since the advent of the ETF IYG. We exclude the midterm congressional elections in the US, which generate far lower voter turnout and feature no national races.
4. **Choice of estimator for volatility** Sum of squared 5-minute log returns of IYG on November 9th, 2016, otherwise known as the Realized Volatility estimator of volatility (Andersen and Teräsvirta

2016 Election

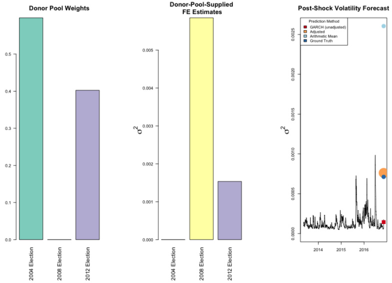


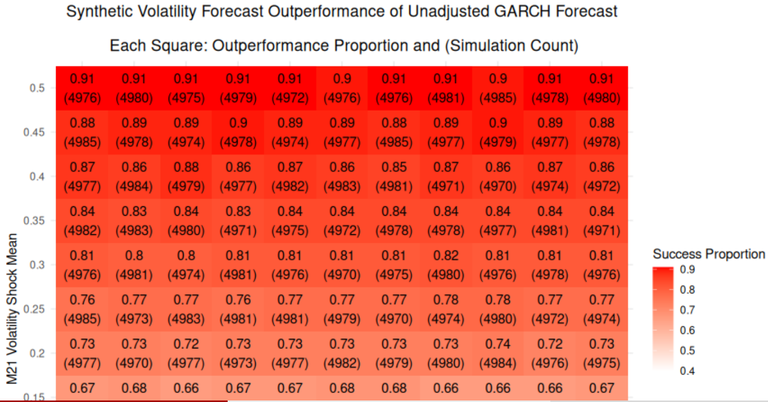
Figure: The volatility induced by the 2016 US election

Simplest Simulation Setup

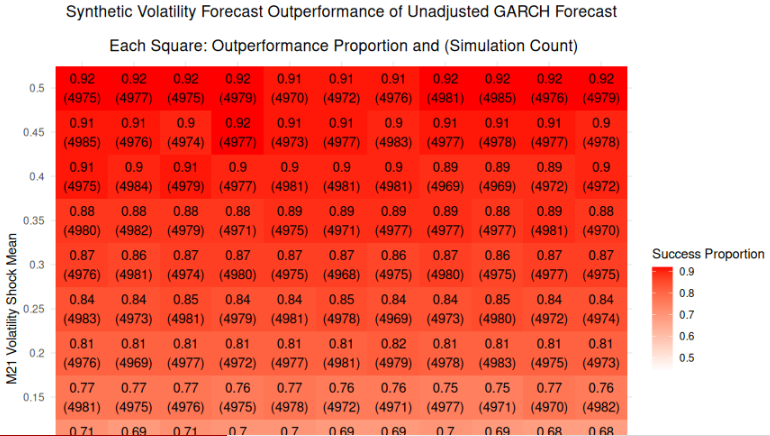
In order to investigate the Synthetic Volatility Forecasting method, our most elementary simulation uses $\mathcal{M}_1$. We vary only two parameters. Recall an $\mathcal{M}_1$ model on the volatility, which is characterized by an exogenous shock to the volatility equation generated by an affine function of the covariates:

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

In Figure 2, when only two parameters are varied, the volatility shock signal and the volatility shock noise, we observe several encouraging phenomena. First, for any column selected, an increasing trend exists as the shock signal increases. Second, for almost all small values of the shock signal, the outperformance rate hovers around .5, supporting the hypothesis that in the absence of a signal, any level of noise renders the method no better at GARCH than a flip of a coin.



If we switch the values of α and β, we see similar behavior, as in 3. Here we cite Francq and Zakoian 2019, p. 22 in motivating simulations for large α values and then large β values. However, for small values of the shock mean, increasing noise does lead to fewer converged simulations, likely due to large negative realizations of the noise term, which in turn lead to near-zero and even negative estimates of the terms ω_i^*.



Alternative Data-Generating Processes

- Could we do all of the above with high-frequency data?
- Realized GARCH with High-Frequency Data
- Stochastic Volatility

Alternative Estimators and Estimands in Volatility Modeling

- Realized GARCH with High-Frequency Data
- Overnight returns instead of open-to-close
- Value-at-Risk using SVF-based $\hat{\sigma}^2_t$
- Signal Recovery Perspective (Ferwana and Varshney 2022)
- Stochastic Volatility: Correlation between errors

New Frontiers in Aggregation Methods

- 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 multivariate time series of dimension $ptimesT$ subject to shocks from a common shock distribution
- Using an IRF estimate aggregated from the first n shocks of interest, we predict the response of variable i from variable j, $1 \leq i \leq j \leq p$.

We analyze the real-world example with Brexit included.

Bibliography

- 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 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