Market Response to a VIX Impulse

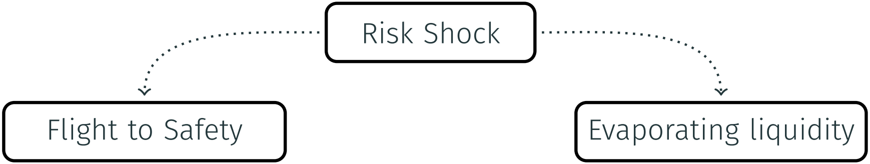
Nikolaus Hautsch, Albert J. Menkveld & Stefan Voigt

May 2022

Future of Financial Information 2022

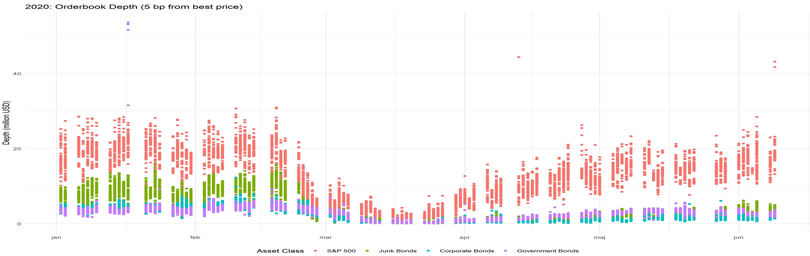
How do markets respond to a sudden VIX impulse?

- Risk shock triggers re-allocations
- Flight to safety episodes are an important source of financial instability (Caballero and Krishnamurthy, 2008)



- Costly liquidity demand if re-allocation triggers liquidity spiral
- Liquidity evaporates entirely during periods of major market turmoil (Brunnermeier, 2009; Pedersen, 2009)
- Vivid examples: GFC and market turmoils in 2020

Worrisome liquidity dynamics?



Market liquidity trends concern regulators

- Central banks focus on liquidity risk management (BIS and IOSCO, 2012)
- SEC (2016) demands open-end funds report their liquidity risk
- ESMA (2019) contains guidelines on liquidity stress testing

This paper: Nature of VIX impulse matters

- We go high-frequency: 2007-2021 sample of *all* Nasdaq trading messages for four exchange-traded funds (ETFs)
- Impulse response analysis for surprise shock in VIX

⇒ How does an implied volatility (*VIX*) shock ripple through financial markets?

- active selling of equities and active buying of government bonds
- market liquidity responds but remains solid

⇒ Which channel drives the flight to safety dynamics?

- Decomposition of VIX changes into two canonical components: variance premium (VRP) and expected realized volatility (ERV)

Main finding: ERV and VRP trigger flight-to-safety *or* market fragility

- *VRP* leads to a market response that mirrors the response to VIX shocks
- *ERV* yields active *buying*(!) of equities and fragile market liquidity

A model to reconcile the responses

- Two extensions of Vayanos and Wang (2012):
  1) letting some agents experience an increase in their risk aversion ("VRP channel")
  2) model a market with elevated cash-flow risks ("ERV channel")
- Formal analysis of the two types of shocks in the model yields results that are consistent with the empirical findings
- VRP channel: Unsurprisingly, investors who suddenly become more risk averse re-allocate risk to those who do not
- ERV Channel: *Active buying of the risky asset is an equilibrium result*
- As in Grossman and Stiglitz (1980), liquidity deteriorates because of the adverse-selection cost. Liquidity demanders effectively bear less *posterior* risk and are the natural *holders* of the risky asset

Responses to a VIX Impulse

Our data sample

asset classes / exchange-traded funds (ETFs)

1. **SPY for equities (S&P 500)**
2. **TLT for government bonds**
3. HYG for junk bonds
4. LQD for corporate bonds

Sample period: July 1st, 2007 - April 7th, 2021

- Event-time Nasdaq order book snapshots from data provider LOBSTER
- *All* Trades, Submissions, Cancellations, and Adjustments
- first 50 order book levels ($\approx$ 36 billion messages)

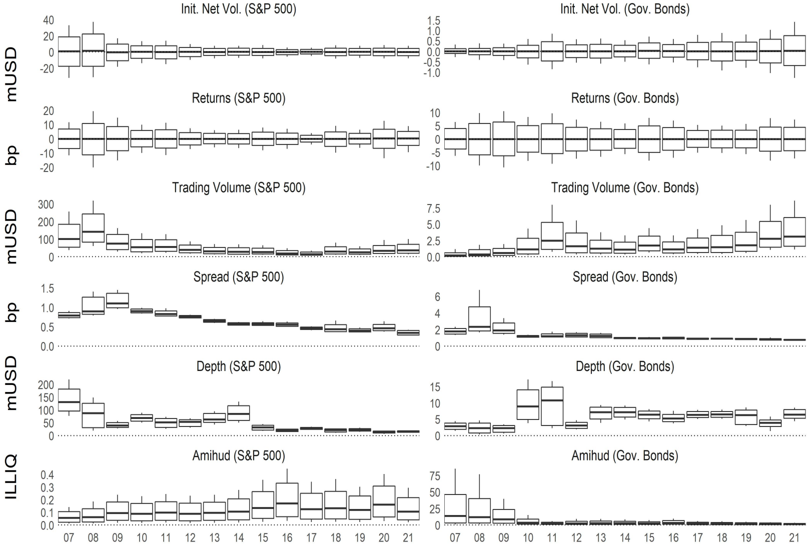
- VIX minute level data from data provider pitrading

5-minute observations: Variable definitions

- Aggregation to 5-minute sample frequency

1. *Initiator Net Volume* (in million USD) Signing based on initiating side
2. *Return* (basis points) Log difference of prevailing midquotes
3. *Trading volume* (in million USD) Cumulative trading volume
4. *Bid ask spread* (basis points) time-weighted relative bid-ask spread
5. *Depth* (million USD) time-weighted number of shares within 5 bps of best quote
6. *Amihud illiquidity measure*, computed as $ILLIQ_{t,\tau} := \frac{\left|\log(p_{t,\tau}) - \log(p_{t,\tau-1})\right|}{V_{t,\tau}}$

Summary statistics: Box plots on an annual basis



Impulse responses with local projection (Jorda, 2005)

- Impulse response function $\text{ir}_i(t, h, \boldsymbol{d})$: response of $y_{t+h,i}$ at time $t+h$ on a (reduced form) shock $\boldsymbol{d}$ at time t:

  $$\text{ir}_i(t, h, \boldsymbol{d}) := E\left(y_{t+h,i} | \boldsymbol{v}_t = \boldsymbol{d}, \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_1\right) - E\left(y_{t+h,i} | \boldsymbol{v}_t = \boldsymbol{0}_N, \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_1\right)$$

  where $\boldsymbol{v}_t$ are the reduced form disturbances
- $\boldsymbol{y}_t = (y_{t,1}, \ldots, y_{t,N})'$: $(N \times 1)$ vector of observations at time t
- We estimate $\text{ir}_i(t, h, \boldsymbol{d})$ from lag-augmented regressions of $y_{t+h,i}$ on $\boldsymbol{y}_t$ *and* $\boldsymbol{y}_{t-1}$ as additional control (Montiel Olea and Plagborg-Møller, 2021)

  $$y_{t+h,i} = \boldsymbol{\beta}(h)'\boldsymbol{y}_t + \boldsymbol{\gamma}'\boldsymbol{y}_{t-1} + \varepsilon_t$$

- Then,

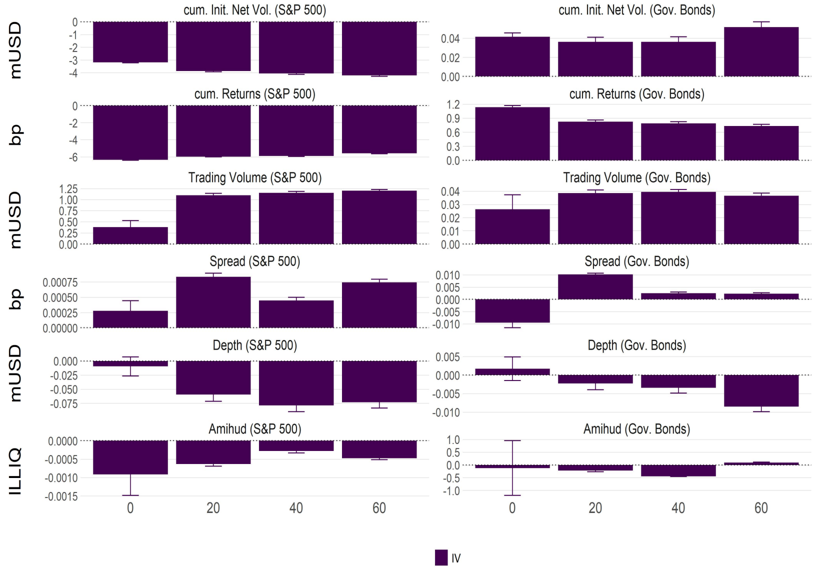
  $$\text{ir}_i(t, h, \boldsymbol{d}) = \hat{\boldsymbol{\beta}}(h)'\mathbf{d}$$

- We integrate out contemporaneous effects of ε_t other shocks using the historically observed distribution of the errors (Pesaran and Shin, 1998)

Shocking the fear gauge

- Our sample consists of the time series order book related variables and the *implied variance (IV) changes*
- We compute the responses to a (SD) shock in *IV* changes
- We report estimated cumulative impulse response functions for initiator net volume (in million USD) and return (in basis points) and impulse response functions for trading volume and depth (all in million USD), bid-ask spreads (in basis points) and the Amihud measure (ILLIQ)
- We report 95% confidence intervals with heteroscedasticity-robust standard errors

Shocking the fear gauge



Delineate the responses to a VIX impulse

The variance premium

- *VIX*: implied volatility of the S&P 500 index, constructed from market prices of out-of-the-money puts and calls with maturity of one month
- A shock of the implied volatility relates to the risk-neutral measure: *Heightened future cash-flow risk? Increase in the variance premium?*

- Variance premium: difference between *implied* variance $IV_{t,\tau}$ as the risk-neutral expectation of the future return variation and the expected *realized* variance as its "physical" counterpart
- Interpretation of a VRP increase: Higher crash risk probabilities or risk aversion / habit formation (Bekaert et al., 2013)
- Decomposition:

$$IV_{t,\tau} := VRP_{t,\tau} + ERV_{t,\tau}$$

  where $ERV_{t,\tau}$ is an estimate of the conditional expected realized variance of the S&P 500 index over the next month.

Intraday ERV estimation

- Notation: 5-minute returns $r_{t,\tau} = \log(p_{t,\tau}) - \log(p_{t,\tau-1})$ for $\tau > 0$, where $p_{t,\tau}$ is the S&P 500 index value on the day t at timestamp τ. $r_{t,0} = \log(p_{t,0}) - \log(p_{t-1,78})$ is the overnight return
- Realized variance of the S&P 500 index, computed from closure on day $t-1$ to the end of the τ-th 5 minute interval at day t:

$$\widetilde{RV}_{t,\tau} := \sum_{k=0}^{\tau} r_{t,k}^2$$

- Model (Corsi, 2009) for the log one-month-ahead $RV_{t,\tau}^{(22)}$

$$\log\left(RV_{t,\tau}^{(22)}\right) = c_\tau + \beta_\tau \log\left(RV_{t-22,0}^{(21)} + \widetilde{RV}_{t,\tau}\right) + \gamma_\tau \log\left(RV_{t-5,0}^{(4)} + \widetilde{RV}_{t,\tau}\right) + \delta_\tau \log\left(\widetilde{RV}_{t,\tau}\right) + \varepsilon_{t,\tau}$$

with $\varepsilon_{t,\tau}$ is zero mean white noise

Intraday variance risk premium estimation

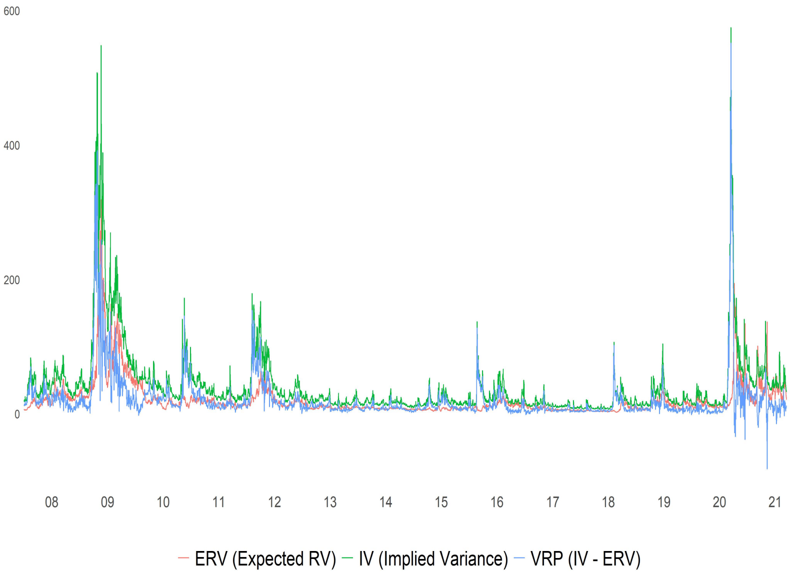
- Least squares estimation for $\tau \in \{1, \ldots, 78\}$ yields estimates $\{\hat{c}_\tau, \hat{\beta}_\tau, \hat{\gamma}_\tau, \hat{\delta}_\tau\}$ and thus

$$\widehat{E}_{t,\tau}\left(RV_{t,\tau}^{(22)}\right) = \exp\left(\hat{c}_\tau + \hat{\beta}_\tau \log\left(RV_{t-22,0}^{(21)} + \widetilde{RV}_{t,\tau}\right) + \hat{\gamma}_\tau \log\left(RV_{t-5,0}^{(4)} + \widetilde{RV}_{t,\tau}\right) + \hat{\delta}_\tau \log\left(\widetilde{RV}_{t,\tau}\right)\right)$$

  Accordingly, for each day t and time stamp τ we obtain the estimated variance risk premium $VRP_{t,\tau}$ as

$$\widehat{VRP}_{t,\tau} = IV_{t,\tau} - \widehat{E}_{t,\tau}\left(RV_{t,\tau}^{(22)}\right)$$

Intraday variance risk premium estimation



What drives the flight-to-safety dynamics?

Further empirical results

- *Institutional client responses* to a VIX shock largely mirror initiator net volume: Institutional trading responds to an IV shock with selling of equity at large magnitudes and turns positive after an ERV shock
- Net buying of (US) equity is *not* a response of net selling of global equity. Results are qualitatively similar when including MSCI World index tracking ETF
- The results are *not* driven by estimation error in ERV (and thus in VRP): Imposing *perfect foresight* ($RV_t^{(22)}$ instead of $\hat{E}(RV_t^{(22)})$ delivers similar results
- *Sub-period analysis* for crisis periods shows that the responses magnify during crisis periods.

A model to disentangle the responses

The theoretical framework in a nutshell

- Three periods, risk averse initially identical agents who derive (exponential) utility from consumption
- Market: a riskless and a risky asset in supply of $\bar{\theta}$ shares which pays $D \sim N(\bar{D}, \sigma^2)$ units of consumption in $t = 2$
- Just before Period 1, a fraction π of the agents learns that they will receive an endowment $z(D - \bar{D})$ of the consumption good in Period 2, with $z \sim N(0, \sigma_z^2)$ independent of D
- **Baseline insight:** Price impact from initiated volume to incentivize liquidity providers + liquidity premium already in $t = 0$

Two extensions: Risk shocks in addition to endowment

1. *risk aversion shock (VRP channel)*: the fraction π of agents with endowment shock simultaneously experiences increase of their risk aversion
2. *cash-flow risk (ERV channel)* for the (identical) agents in $t = 0$: common knowledge that liquidity demanders learn additional information in $t = 1$ from a private signal correlated with D

Equilibrium outcomes from the model

1. Risk aversion shocks trigger *net selling* of the risky asset and simultaneously net buying of government bonds
2. Cash-flow shocks generate *net buying* for the risky asset (resemble relative increase in risk for liquidity providers)
3. Larger price impact in response to cash-flow risk shocks than for risk aversion shocks

Conclusions

Conclusions

- In this paper we study the connection between flight-to-safety and market fragility
- Responses to risk shocks depend on the *nature* of the shock
  1. Variance premium shocks predominantly trigger net selling of equities and net buying of government bonds
  2. Expected realized variance shocks create *net buying of risky assets* and *illiquidity*, pointing towards asymmetric information

Ait-Sahalia, Y., F. Matthys, E. Osambela, and R. Sircar (2021). When Uncertainty and Volatility Are Disconnected: Implications for Asset Pricing and Portfolio Performance. Working paper.

Angel, J. J., L. E. Harris, and C. S. Spatt (2015). Equity Trading in the 21st Century: An Update. *Quarterly Journal of Finance 5*.

Beber, A., M. W. Brandt, and K. A. Kavajecz (2009). Flight-to-Quality or Flight-to-Liquidity? Evidence from the Euro-Area Bond Market. *Review of Financial Studies 22*, 925–957.

Bekaert, G. and M. Hoerova (2014). The VIX, the Variance Premium and Stock Market Volatility. *Journal of Econometrics 183*, 181–192.

Bekaert, G., M. Hoerova, and M. L. Duca (2013). Risk, Uncertainty and Monetary Policy. *Journal of Monetary Policy 60*, 771–788.

Ben-David, I., F. Franzoni, and R. Moussaw (2017). Exchange-Traded Funds. *Annual Review of Financial Economics 9*, 169–189.

BIS and IOSCO (2012). Principles for Financial Market Infrastructures. Manuscript, Committee on Payment and Settlement Systems and Technical Committee of the International Organization of Securities Commissions.

Bollerslev, T., J. Marrone, L. Xu, and H. Zhou (2014). Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence. *Journal of Financial and Quantitative Analysis 49*, 633–661.

Bollerslev, T., N. Sizova, and G. Tauchen (2012). Volatility in Equilibrium: Asymmetries and Dynamic Dependencies. *Review of Finance 16*, 31–80.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected Stock Returns and Variance Risk Premia. *Review of Financial Studies 22*, 4463–4492.

Brenner, M. and Y. Izhakian (2018). Asset Pricing and Ambiguity: Empirical Evidence. *Journal of Financial Economics 130*(3), 503–531.

Brunnermeier, M. K. (2009). Deciphering the Liquidity and Credit Crunch 2007-2008. *Journal of Economic Perspectives 23*(1), 77–100.

Caballero, R. J. and A. Krishnamurthy (2008). Collective risk management in a flight to quality episode. *The Journal of Finance 63*(5), 2195–2230.

Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics 7*(2), 174–196.

Drechsler, I. (2013). Uncertainty, Time-Varying Fear, and Asset Prices. *The Journal of Finance 68*(5), 1843–1889.

ESMA (2019). Guidelines on liquidity stress testing in UCITS and AIFs. Manuscript, European Securities and Markets Authority.

Falato, A., I. Goldstein, and A. Hortacsu (2020). Financial Fragility in the COVID-19 Crisis: The Case of Investment Funds in Corporate Bond Markets. Manuscript.

Grossman, S. J. and J. E. Stiglitz (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review 70*, 393–408.

Jorda, O. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review 95*(1), 161–182.

Koop, G., M. Pesaran, and S. M. Potter (1996). Impulse Response Analysis in Nonlinear Multivariate Models. *Journal of Econometrics 74*(1), 119 – 147.

Lettau, M. and A. Madhavan (2017). Exchange-Traded Funds 101 for Economists. *Journal of Economic Perspectives 32*, 135–154.

Liu, J., J. Pan, and T. Wang (2004). An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks. *The Review of Financial Studies 18*(1), 131–164.

Longstaff, F. A. (2004). The Flight-to-Liquidity Premium in U.S. Treasury Bond Prices. *Journal of Business 77*, 511–526.

Montiel Olea, J. L. and M. Plagborg-Møller (2021). Local Projection Inference Is Simpler and More Robust Than You Think. *Econometrica 89*(4), 1789–1823.

Nagel, S. (2012). Evaporating liquidity. *Review of Financial Studies 25*, 2005–2039.

Pedersen, L. H. (2009). When Everyone Runs for the Exit. *International Journal of Central Banking 5*, 177–199.

Pesaran, H. and Y. Shin (1998). Generalized Impulse Response Analysis in Linear Multivariate Models. *Economics Letters 58*(1), 17 – 29.

SEC (2016). Investment Company Liquidity Risk Management Programs. Manuscript, SEC Release No. 33–10233.

Vayanos, D. (2004). Flight-to-Quality or Flight-to-Liquidity? Evidence from the Euro-Area Bond Market. Manuscript, NBER Working Paper No. 10327.

Vayanos, D. and J. Wang (2012). Liquidity and Asset Returns Under Asymmetric Information and Imperfect Competition. *Review of Financial Studies 25*, 1339–1365.

Literature (incomplete)

- *Flight to quality/liquidity* Longstaff (2004); Vayanos (2004); Beber et al. (2009)
- Ait-Sahalia et al. (2021) rationalize *asset-price dynamics* by introducing two "disconnected" stochastic processes: one that drives (realized) volatility and another that drives uncertainty (aversion) (same spirit: Liu et al. (2004), Drechsler (2013), Brenner and Izhakian (2018))
- *Variance risk premium* (Bollerslev et al., 2009, 2012, 2014; Bekaert and Hoerova, 2014; Nagel, 2012)
- *Crisis periods*: GFC (Brunnermeier, 2009) and COVID-19 (Falato et al., 2020)
- *Exchange-traded funds* (Ben-David et al., 2017; Lettau and Madhavan, 2017)
- *Market liquidity trends* (Angel et al., 2015; SEC, 2016; ESMA, 2019)

Impulse responses with local projection

Impulse response functions (IRF)

- Denote $(N \times 1)$ vector of observations at time t as $\boldsymbol{y}_t = (y_{t,1}, \ldots, y_{t,N})'$
- Impulse response function $\mathrm{ir}_i\,(t, h, \boldsymbol{d})$ measures the response of $y_{t+h,i}$ at time $t+h$ on a (reduced form) shock $\boldsymbol{d}$ at time t:

  $$\mathrm{ir}_i\,(t, h, \boldsymbol{d}) := E\,(y_{t+h,i}|\boldsymbol{v}_t = \boldsymbol{d}, \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_1) - E\,(y_{t+h,i}|\boldsymbol{v}_t = \boldsymbol{0}_N, \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_1)$$

  where $\boldsymbol{v}_t$ are the reduced form disturbances, $\boldsymbol{0}_N$ denotes a $(N \times 1)$ vector of zeros

Common approach: vector autoregressive model (VAR)

- specify VAR: $\boldsymbol{y}_t = \boldsymbol{A}\boldsymbol{y}_{t-1} + \boldsymbol{u}_t$ and obtain $\mathrm{ir}_i\,(t, h, \boldsymbol{d})$ as non-linear transformation of the estimated coefficients
- $\mathrm{ir}_i\,(t, h, \boldsymbol{d}) = \boldsymbol{e}_i'\boldsymbol{\Phi}_h\boldsymbol{d}$ where e_i is a $(N \times 1)$ unit vector and $\boldsymbol{\Phi}_h$ is computed recursively from the canonical $MA(\infty)$ representation
- **Problem:** High-dimensional nonlinear functions of parameter estimates $\hat{\boldsymbol{A}}$ are prone to estimation and misspecification error

Local Projections (Jorda, 2005)

- Simple equation-by-equation least squares estimation
- We obtain an estimate of $\mathrm{ir}_i(t, h, \boldsymbol{d})$ from baseline lag-augmented regression of $y_{t+h,i}$ on $\boldsymbol{y}_t$ *and* $\boldsymbol{y}_{t-1}$ as additional control

$$y_{t+h,i} = \boldsymbol{\beta}(h)'\boldsymbol{y}_t + \boldsymbol{\gamma}'\boldsymbol{y}_{t-1} + \varepsilon_t$$

- For a given horizon h, the lag-augmented local projection estimates are

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}(h) \\ \hat{\gamma} \end{pmatrix} = \left(\sum_{t=1}^{T-h} \boldsymbol{x}_t \boldsymbol{x}_t' \right)^{-1} \left(\sum_{t=1}^{T-h} \boldsymbol{x}_t y_{t+h,i} \right)$$

where $\boldsymbol{x}_t = \begin{pmatrix} \boldsymbol{y}_t, & \boldsymbol{y}_{t-1} \end{pmatrix}'$ is a $(2N \times 1)$ vector, $\hat{\boldsymbol{\beta}}(h)$ is *a consistent estimator of* $\mathrm{ir}_i(t, h, \boldsymbol{d})$ and $\hat{\boldsymbol{\gamma}}$ is a nuisance parameter

What makes the lag-augmented estimator $\hat{\boldsymbol{\beta}}(h)$ appealing?

- Covariance of the estimator $\hat{\boldsymbol{\beta}}(h)$ does not require Newey-West inference. (Eicker-Huber-White) heteroscedasticity-robust standard errors suffice for inference (Montiel Olea and Plagborg-Møller, 2021)
- Lag augmentation handles nearly persistent series particularly well

Choice of the shock

- Common problem: contemporaneous correlation of the errors ε_t
- Our (current) approach: *pseudo* generalized IRFs in spirit of Pesaran and Shin (1998) and Koop et al. (1996)
- Two steps to integrate out the effects of other shocks using the historically observed distribution of the errors
  1. estimate the parameters of a VAR(p) regression for $\boldsymbol{y}_t$ of the standard form $\boldsymbol{y}_t = \boldsymbol{\mu} + \sum_{l=1}^{p} \tilde{\boldsymbol{A}}_l \boldsymbol{y}_{t-l} + \boldsymbol{\varepsilon}_t$ where $\tilde{\boldsymbol{A}}_l$ is a $(N \times N)$ matrix
  2. Set $\boldsymbol{d} = E(\boldsymbol{\varepsilon}_t | \varepsilon_{t,j} = \delta) = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\varepsilon}} \boldsymbol{e}_j \hat{\sigma}_{jj}^{-1} \delta$ where $\hat{\sigma}_{jj}$ is the j-th diagonal element of $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\varepsilon}}$ and $\delta \neq 0$ denotes the shock in element j

Intraday variance risk premium estimation