Discussion Paper 3/2022
Guatemalan Econometric Study Group
Universidad Francisco Marroquín
March 12, 2022

A short note on the scaling parameter in score-driven filters

Astrid Ayala, Szabolcs Blazsek*, and Adrian Licht

School of Business, Universidad Francisco Marroquín, Guatemala City, Guatemala

Abstract:

In the literature on score-driven models, the choice of the scaling parameter of the updating term of the score-driven filter is an open question. It is presented in the literature that different choices of the scaling parameter provide different score-driven models. There are no results on the optimal choice of the scaling parameter. For potential users of score-driven models this ambiguity is problematic because one must always choose a certain scaling parameter. The present short note is motivated by this issue. We show that the commonly used alternative choices of the scaling parameter in the literature result exactly in the same filtered estimates of signals. The empirical results reported in this letter are for two very popular score-driven models, and they are easily generalizable to many other score-driven models. This sends a message to potential users of score-driven models that the ambiguous choice of the scaling parameter for score-driven models is not as problematic as it may initially seem, and it may motivate more researchers in economics and finance to apply score-driven models in the future.

Keywords: dynamic conditional score (DCS); generalized autoregressive score (GAS); scaling parameter of conditional score function; quasi-autoregressive (QAR) model; Beta-t-EGARCH (exponential generalized autoregressive conditional heteroskedasticity)

JEL classification: C22; C51; C52

*Corresponding author. Address: School of Business, Universidad Francisco Marroquín, Calle Manuel F. Ayau (6 Calle final), zona 10, Guatemala City 01010, Guatemala, E-mail: sblazsek@ufm.edu

I. Introduction

In the literature on score-driven models, started by the works of Creal, Koopman, and Lucas (2008) and Harvey and Chakravarty (2008), the choice of the scaling parameter of the updating term of the score-driven filter is an open question. Different choices of the scaling parameter provide different score-driven models (Creal, Koopman, and Lucas 2013, p. 779), and there are no results on the optimal choice of the scaling parameter. In the literature, either unit scaling or inverse information matrix-based scaling is used. For potential users of score-driven models this ambiguity is problematic, because one always has to choose a certain scaling parameter. This letter is motivated by this issue.

We use score-driven models for the Student's t-distribution, because this is the most common choice in economic and financial applications of the score-driven models in the literature. We focus on: (i) dynamic Student's t location model (Harvey 2013), i.e. quasi-autoregressive (QAR) model; (ii) Beta-t-EGARCH (exponential generalized autoregressive conditional heteroskedasticity) model of conditional volatility (Harvey and Chakravarty 2008). Model (i) provides a robust estimate of the filtered signal of the dependent variable, and model (ii) provides a robust estimate of the filtered signal of the conditional volatility of the dependent variable. For robustness results, we refer to the works of Creal, Koopman, and Lucas (2011, 2013), Harvey (2013), and Blasques, Koopman, and Lucas (2015).

We find that the commonly used choices for the scaling parameter result exactly in the same filtered estimates of the signal. This result generalizes to a large number of score-driven models, indicating that the ambiguous choice of the scaling parameter is not as problematic as it may initially seem.

The remainder of this letter is organized as follows: Section II describes the data, Section III presents the score-driven models, and Section IV summarizes the results.

II. Data

We use data on (i) seasonally adjusted quarterly United States (US) real gross domestic product (GDP) growth $y_{\text{GDP},t}$ for the period of 1947 Q2 to 2019 Q4 (data source: Federal Reserve Economic Data; ticker: GDPC1), and (ii) daily Vanguard Standard & Poor's (S&P 500) exchange traded fund (ETF) return $y_{\text{VOO},t}$ for the period of September 10, 2010 to February 9, 2022 (data source: Yahoo Finance; ticker: VOO). Variable $y_{\text{GDP},t}$ is the log percentage change in the US real GDP level. Variable $y_{\text{VOO},t}$ is the log percentage change in the opening price of VOO. Descriptive statistics are presented in Table 1(a). We use the QAR(1) model for $y_{\mathrm{GDP},t}$, which is motivated by likelihood-based model selection metrics for AR(p) specifications with $p = 0, 1, 2$ (Table 1(b)). We use the constant mean plus Beta-t-EGARCH(1,1) model for $y_{\mathrm{VOO},t}$, which is motivated by likelihood-based model selection metrics for AR(p)-normal-GARCH(1,1) specifications with $p = 0, 1, 2$ (Table 1(c)).

[APPROXIMATE LOCATION OF TABLE 1]

III. Score-driven models

First, the dynamic Student's t location model, i.e. QAR(1), is given by

$$y_{\mathrm{GDP},t} = \mu_t + \exp(\lambda)\epsilon_t \tag{1}$$

$$\mu_t = c + \phi\mu_{t-1} + \kappa S_{\mu,t-1}\nabla_{\mu,t-1} \tag{2}$$

$$\epsilon_t \sim t(\nu) \text{ i.i.d.} \tag{3}$$

where μ_t is the filtered signal of $y_{\mathrm{GDP},t}$, $\exp(\lambda)$ is the constant scale parameter of the t-distribution, $\nabla_{\mu,t}$ is the conditional score with respect to μ_t, and $S_{\mu,t}$ is the scaling parameter of the conditional score. The parameters c, ϕ, κ, and λ are $\in \mathbb{R}$, and $\nu > 0$. We assume that $|\phi| < 1$ (i.e. μ_t is asymptotically covariance stationary), and we initialize μ_t by using its unconditional mean: $E(\mu_t) = c/(1 - \phi)$.

The log conditional density of $y_{\mathrm{GDP},t}|(y_{\mathrm{GDP},1}, \ldots, y_{\mathrm{GDP},t-1}, \mu_1) = y_{\mathrm{GDP},t}|\mathcal{F}_{t-1}$ is

$$\ln f(y_{\mathrm{GDP},t}|\mathcal{F}_{t-1}) = \ln\Gamma\left(\frac{\nu+1}{2}\right) - \ln\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\ln(\pi\nu) - \lambda - \frac{\nu+1}{2}\ln\left[1 + \frac{1}{\nu} \times \frac{(y_{\mathrm{GDP},t} - \mu_t)^2}{\exp(2\lambda)}\right] \tag{4}$$

As a function of $\epsilon_t = (y_{\mathrm{GDP},t} - \mu_t)\exp(-\lambda)$, the conditional score with respect to μ_t is:

$$\nabla_{\mu,t} = \frac{\partial \ln f(y_{\mathrm{GDP},t}|\mathcal{F}_{t-1})}{\partial \mu_t} = \frac{(\nu+1)\epsilon_t}{\exp(\lambda)(\nu + \epsilon_t^2)} \tag{5}$$

We consider the following five alternatives for the scaling parameter $S_{\mu,t}$ of the conditional score:

$$S_{\mu,t} = (I^{-1})^0 = 1 \tag{6}$$

$$S_{\mu,t} = (I^{-1})^1(\mathrm{DCS}) = \frac{\nu}{\nu+1}\exp(2\lambda) \qquad \text{(Harvey 2013, p. 60)} \tag{7}$$

$$S_{\mu,t} = (I^{-1})^1(\text{GAS}) = \frac{\nu+3}{\nu+1}\exp(2\lambda) \qquad \text{(Creal, Koopman, and Lucas 2013, p. 780)} \tag{8}$$

$$S_{\mu,t} = (I^{-1})^{1/2}(\text{GAS}) = \left[\frac{\nu+3}{\nu+1}\exp(2\lambda)\right]^{1/2} \qquad \text{(Creal, Koopman, and Lucas 2013, p. 780)} \tag{9}$$

$$S_{\mu,t} = (I^{-1})^d = \left[\frac{\nu+3}{\nu+1}\exp(2\lambda)\right]^d \tag{10}$$

for $d > 0$, where for all scaling parameters I^{-1} refers to the inverse of the information quantity with respect to μ_t (Harvey 2013). The last alternative scaling parameter for $S_{\mu,t}$ is new in the literature. Parameter d is jointly estimated with the rest of the parameters and it is identified.

Second, the Beta-t-EGARCH(1,1) model is given by

$$y_{\text{VOO},t} = c + \exp(\lambda_t)\epsilon_t \tag{11}$$

$$\lambda_t = \omega + \beta\lambda_{t-1} + \alpha S_{\lambda,t-1}\nabla_{\lambda,t-1} \tag{12}$$

$$\epsilon_t \sim t(\nu) \text{ i.i.d.} \tag{13}$$

where $\exp(\lambda_t)$ is the score-driven scale parameter of the t-distribution, $\nabla_{\lambda,t}$ is the conditional score with respect to λ_t, and $S_{\lambda,t}$ is the scaling parameter of the conditional score. We initialize λ_t by using parameter λ_0. The parameters c, ω, β, α, and λ_0 are $\in \mathbb{R}$, where $|\beta| < 1$ (i.e. λ_t is asymptotically covariance stationary). Moreover, $\nu > 2$, hence the conditional volatility is $\sigma_t = \exp(\lambda_t)[\nu/(\nu-2)]^{1/2}$.

The log conditional density of $y_{\text{VOO},t}|(y_{\text{VOO},1}, \ldots, y_{\text{VOO},t-1}, \lambda_1) = y_{\text{VOO},t}|\mathcal{F}_{t-1}$ is

$$\ln f(y_{\text{VOO},t}|\mathcal{F}_{t-1}) = \ln\Gamma\left(\frac{\nu+1}{2}\right) - \ln\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\ln(\pi\nu) - \lambda_t - \frac{\nu+1}{2}\ln\left[1 + \frac{1}{\nu} \times \frac{(y_{\text{VOO},t}-c)^2}{\exp(2\lambda_t)}\right] \tag{14}$$

As a function of $\epsilon_t = (y_{\text{VOO},t} - c)\exp(-\lambda_t)$, the conditional score with respect to λ_t is:

$$\nabla_{\lambda,t} = \frac{\partial \ln f(y_{\text{VOO},t}|\mathcal{F}_{t-1})}{\partial\lambda_t} = \frac{(\nu+1)\epsilon_t^2}{\nu+\epsilon_t^2} - 1 \tag{15}$$

We consider the following four alternatives for the scaling parameter $S_{\lambda,t}$:

$$S_{\lambda,t} = (I^{-1})^0(\text{DCS}) = 1 \qquad \text{(Harvey 2013, p. 99)} \tag{16}$$

$$S_{\lambda,t} = (I^{-1})^{1}(\text{GAS}) = \frac{\nu+3}{2\nu} \qquad \text{(Creal, Koopman, and Lucas 2013, p. 780)} \tag{17}$$

$$S_{\lambda,t} = (I^{-1})^{1/2}(\text{GAS}) = \left[\frac{\nu+3}{2\nu}\right]^{1/2} \qquad \text{(Creal, Koopman, and Lucas 2013, p. 780)} \tag{18}$$

$$S_{\lambda,t} = (I^{-1})^{d} = \left[\frac{\nu+3}{2\nu}\right]^{d} \tag{19}$$

for $d > 0$, where for all scaling parameters I^{-1} refers to the inverse of the information quantity with respect to λ_t (Harvey 2013). The last alternative scaling parameter for $S_{\lambda,t}$ is new in the literature. Parameter d is jointly estimated with the rest of the parameters and it is identified.

Both models are estimated by using the maximum likelihood (ML) method (Blasques et al. 2022).

IV. Results

The ML estimates for all models are presented in Table 2. The results show that the LL estimates are identical for all alternative choices of $S_{\mu,t}$ (Table 2(a)) or $S_{\lambda,t}$ (Table 2(b)); see the κ and α parameters, respectively, which ensure this. We also find that the filtered signal estimates, i.e. μ_t for QAR(1) and λ_t for AR(1) plus Beta-t-EGARCH(1,1) (Figure 1), coincide for all choices of $S_{\mu,t}$ or $S_{\lambda,t}$. The robust standard error estimates-based z-ratios of parameters which correspond to alternative choices of $S_{\mu,t}$ or $S_{\lambda,t}$ are very similar; different choices of the scaling parameter provide similarly precise filtered signal estimates. The conclusions for QAR and Beta-t-EGARCH can be extended to the score-driven models for which the information matrix is time-invariant (see Theorem 1 in Harvey 2013 and its generalizations in the same work). The results of this letter answer doubts on score-driven models that may arise from the ambiguous choice of the scaling parameters, and our results may also motivate researchers and practitioners in economics and finance to apply score-driven models in empirical works.

[APPROXIMATE LOCATION OF TABLE 2 AND FIGURE 1]

Acknowledgements

The authors gratefully appreciate the helpful comments of Matthew Copley. All remaining errors are our own. Funding from the School of Business of Universidad Francisco Marroquín is gratefully acknowledged. Data source is reported in the paper, and data are available from the authors upon request. Computer codes are available from the authors upon request.

References

Blasques, F., J. van Brummelen, S. J. Koopman, and A. Lucas. 2022. "Maximum Likelihood Estimation for Score-Driven Models." *Journal of Econometrics* 227 (2): 325–346 doi:10.1016/j.jeconom.2021.06.003.

Blasques, F., S. J. Koopman, and A. Lucas. 2015. "Information-Theoretic Optimality of Observation-Driven Time Series Models for Continuous Responses." *Biometrika* 102 (2): 325–343. doi:10.1093/biomet/asu076.

Bollerslev, T. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31 (3):307–327. doi:10.1016/0304-4076(86)90063-1.

Box, G. E. P., and G. M. Jenkins. 1970. *Time Series Analysis, Forecasting and Control.* San Francisco: Holden-Day.

Creal, D., S. J. Koopman, and A. Lucas. 2008. "A General Framework for Observation Driven Time-Varying Parameter Models." Tinbergen Institute Discussion Paper 08-108/4. https://papers.tinbergen.nl/08108.pdf.

Creal, D., S. J. Koopman, and A. Lucas. 2011. "A Dynamic Multivariate Heavy-Tailed Model for Time-Varying Volatilities and Correlations." *Journal of Business & Economic Statistics* 29 (4): 552–563. doi:10.1198/jbes.2011.10070

Creal, D., S. J. Koopman, and A. Lucas. 2013. "Generalized Autoregressive Score Models with Applications." *Journal of Applied Econometrics* 28 (5): 777–795. doi:10.1002/jae.1279.

Engle, R. F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50 (4): 987–1007. doi:10.2307/1912773.

Harvey, A. C. 2013. *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series.* Econometric Society Monographs. Cambridge: Cambridge University Press.

Harvey, A. C., and T. Chakravarty. 2008. Beta-t-(E)GARCH. Cambridge Working Papers in Economics 0840, Faculty of Economics, University of Cambridge. https://econpapers.repec.org/paper/camcamdae/0840.htm.

Newey, W.K., and K. D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3): 703–708. doi:10.2307/1913610.

Table 1. Descriptive statistics.

(a) Statistics	US real GDP growth	VOO return	
Start date	1947 Q2	September 10, 2010	
End date	2019 Q4	February 9, 2022	
Sample size	291	2875	
Minimum	−0.0263	−0.0884	
Maximum	0.0385	0.0612	
Average	0.0077	0.0005	
Standard deviation	0.0093	0.0100	
Skewness	−0.0671	−0.9409	
Excess kurtosis	1.6490	8.6957	
(b) AR for US real GDP	constant mean	AR(1)	AR(2)
c	0.0077***(0.0007)	0.0050***(0.0007)	0.0045***(0.0007)
ϕ_1	NA	0.3614***(0.0607)	0.3184***(0.0643)
ϕ_2	NA	NA	0.1113*(0.0667)
AIC	−1896.7091	**−1929.2001**	−1923.6559
BIC	−1893.0358	**−1921.8604**	−1912.6566
HQC	−1895.2375	**−1926.2595**	−1919.2486
(c) GARCH for VOO	constant plus normal-GARCH(1,1)	AR(1)-normal-GARCH(1,1)	AR(2)-normal-GARCH(1,1)
c	0.0007***(0.0007)	0.0007***(0.0001)	0.0007***(0.0001)
ϕ_1	NA	−0.0522**(0.0209)	−0.0523**(0.0210)
ϕ_2	NA	NA	−0.0176(0.0216)
ω	0.0000***(0.0000)	0.0000***(0.0000)	0.0000***(0.0000)
α	0.1717***(0.0327)	0.1722***(0.0330)	0.1724***(0.0331)
β	0.8040***(0.0328)	0.8029***(0.0332)	0.8030***(0.0333)
AIC	**−19462.2993**	−19459.8300	−19452.6475
BIC	**−19432.4803**	−19424.0492	−19410.9058
HQC	**−19451.5503**	−19446.9316	−19437.6001

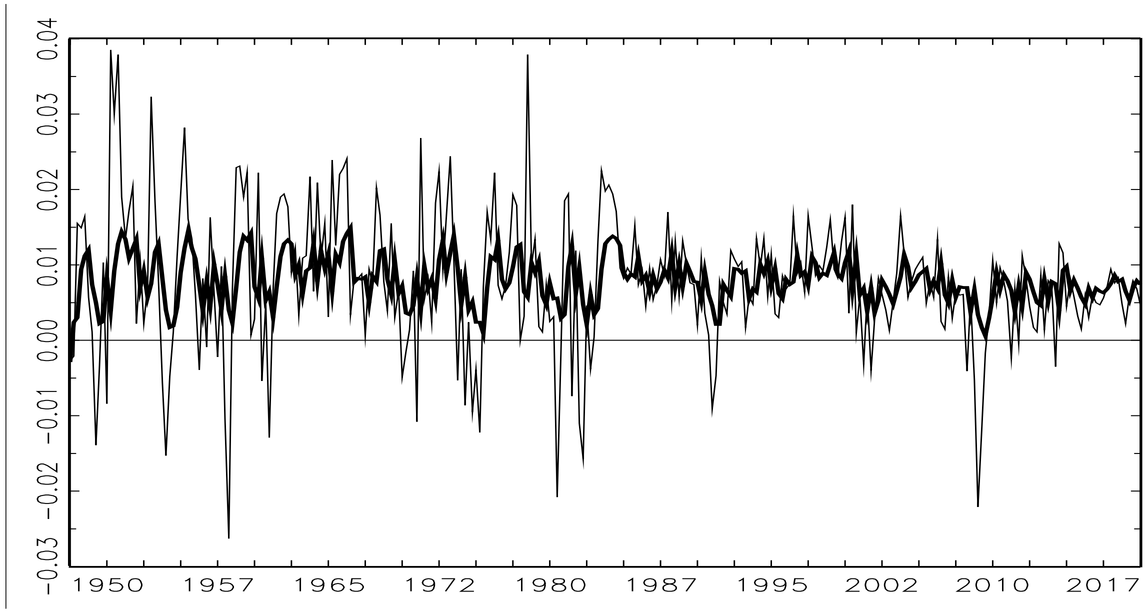
Notes: United States (US); gross domestic product (GDP); autoregressive (AR); generalized AR conditional heteroskedasticity (GARCH); Akaike information criterion (AIC); Bayesian information criterion (BIC); Hannan–Quinn criterion (HQC); not available (NA). The AR(2) model in Panel (b) is: $y_{\text{GDP}} = \mu_t + v_t = \mu_t + \sigma\epsilon_t$; $\epsilon_t \sim N(0,1)$ i.i.d.; $\mu_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2}$ (Box and Jenkins 1970). The AR(2)-normal-GARCH(1,1) model in Panel (c) is: $y_{\text{VOO}} = \mu_t + v_t = \mu_t + \sigma_t\epsilon_t$; $\epsilon_t \sim N(0,1)$ i.i.d.; $\mu_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2}$; $\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha v_{t-1}^2$ (Engle 1982; Bollerlev 1986). Robust standard errors, by using the heteroskedasticity and autocorrelation consistent (HAC) (Newey and West 1987) and Huber sandwich estimators for Panels (b) and (c), respectively, are reported in parentheses. *, **, and *** show significance at the 10%, 5%, and 1% levels, respectively.

Table 2. Parameter estimates and model performance metrics.

(a) $y_{\mathrm{GDP},t}$	$S_t = (I^{-1})^0$	$S_t = (I^{-1})^1$ (DCS)	$S_t = (I^{-1})^1$ (GAS)	$S_t = (I^{-1})^{1/2}$ (GAS)	$S_t = (I^{-1})^d$
c	0.0038***(0.0007)	0.0038***(0.0007)	0.0038***(0.0007)	0.0038***(0.0007)	0.0038***(0.0007)
ϕ	0.5205***(0.0732)	0.5205***(0.0731)	0.5205***(0.0732)	0.5205***(0.0731)	0.5205***(0.0732)
κ	**0.0000*****(**0.0000**)	**0.4878*****(**0.0817**)	**0.3253*****(**0.0623**)	**0.0026*****(**0.0005**)	**0.0031*****(**0.0005**)
λ	−4.9569***(0.0757)	−4.9570***(0.0754)	−4.9570***(0.0753)	−4.9570***(0.0752)	−4.957***(0.0752)
ν	6.0085***(2.0944)	6.0081***(2.0874)	6.0079***(2.0809)	6.0080***(2.0771)	6.0079***(2.0769)
d	NA	NA	NA	NA	0.5189***(0.05234)
LL	**3.3655**	**3.3655**	**3.3655**	**3.3655**	**3.3655**
AIC	−6.6966	−6.6966	−6.6966	−6.6966	−6.6897
BIC	−6.6335	−6.6335	−6.6335	−6.6335	−6.6140
HQC	−6.6713	−6.6713	−6.6713	−6.6713	−6.6594
(b) $y_{\mathrm{VOO},t}$	$S_t = (I^{-1})^0$ (DCS)	$S_t = (I^{-1})^1$ (GAS)	$S_t = (I^{-1})^{1/2}$ (GAS)	$S_t = (I^{-1})^d$	
c	0.0009***(0.0001)	0.0009***(0.0001)	0.0009***(0.0001)	0.0009***(0.0001)	
ω	−0.1660***(0.0417)	−0.1660***(0.0417)	−0.1660***(0.0417)	−0.1660***(0.0417)	
α	**0.0897*****(**0.0105**)	**0.1199*****(**0.0134**)	**0.1037*****(**0.0117**)	**0.1071*****(**0.0120**)	
β	0.9668***(0.0083)	0.9668***(0.0083)	0.9668***(0.0083)	0.9668***(0.0083)	
ν	6.0506***(0.6700)	6.0507***(0.6697)	6.0507***(0.6707)	6.0507***(0.6731)	
λ_0	−4.8936***(0.2390)	−4.8937***(0.2390)	−4.8933***(0.2389)	−4.8936***(0.2390)	
d	NA	NA	NA	0.6118***(0.0621)	
LL	**3.4189**	**3.4189**	**3.4189**	**3.4189**	
AIC	−6.8337	−6.8337	−6.8337	−6.8330	
BIC	−6.8212	−6.8212	−6.8212	−6.8184	
HQC	−6.8292	−6.8292	−6.8292	−6.8277	

Notes: Gross domestic product (GDP); dynamic conditional score (DCS); generalized autoregressive score (GAS); log-likelihood (LL); Akaike information criterion (AIC); Bayesian information criterion (BIC); Hannan–Quinn criterion (HQC). The most relevant estimates are highlighted by bold numbers. Robust standard errors, by using the Huber sandwich estimator, are in parentheses. *** indicates significance at the 1% level.

(a) Quarterly US real GDP growth $y_{\mathrm{GDP},t}$ (thin line), and QAR(1) filtered signal μ_t for all $S_{\mu,t}$ (thick line)



(b) Daily VOO return $y_{\mathrm{VOO},t}$ (thick line), and $c \pm 3\sigma_t$ interval (thin lines) for Beta-t-EGARCH(1,1) filtered volatility for all $S_{\lambda,t}$

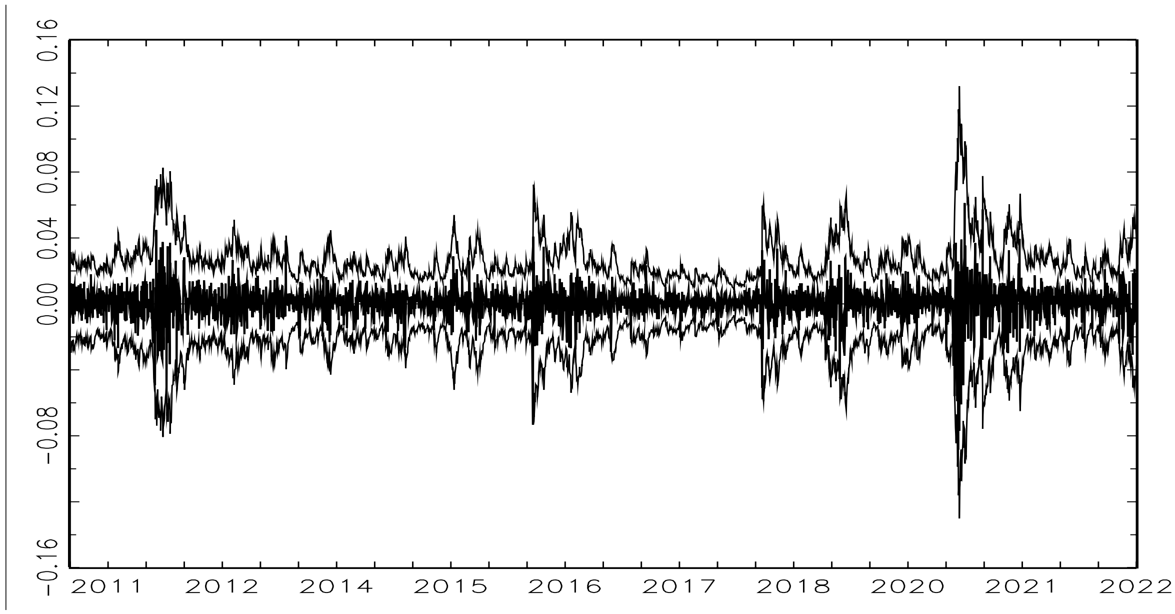


Figure 1. US real GDP growth (from 1947 Q2 to 2019 Q4), and VOO return (from September 10, 2010 to February 9, 2022). *Notes:* United States (US); gross domestic product (GDP); for VOO, c is mean return and $\sigma_t = \exp(\lambda_t)[\nu/(\nu-2)]^{1/2}$ is volatility.