Opschoor, Anne; Lucas, André

Working Paper

Time-varying tail behavior for realized kernels

Tinbergen Institute Discussion Paper, No. TI 2019-051/IV

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Opschoor, Anne; Lucas, André (2019) : Time-varying tail behavior for realized kernels, Tinbergen Institute Discussion Paper, No. TI 2019-051/IV, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/205341

TI 2019-051/IV
Tinbergen Institute Discussion Paper

Time-varying tail behavior for realized kernels

Anne Opschoor[1]
André Lucas[1]

[1] Vrije Universiteit Amsterdam

Time-varying tail behavior for realized kernels*

Anne Opschoor André Lucas

Vrije Universiteit Amsterdam and Tinbergen Institute

This version: July 23, 2019

Abstract

We propose a new score-driven model to capture the time-varying volatility and tail behavior of realized kernels. We assume realized kernels follow an F distribution with two time-varying degrees-of-freedom parameters, accounting for the Vol-of-Vol and the tail shape of the realized kernel distribution. The resulting score-driven dynamics imply that the influence of large (outlying) realized kernels on future volatilities and tail-shapes is mitigated. We apply our model to 30 stocks from the S&P 500 index over the period 2001-2014. The results show that tail shapes vary over time, even after correcting for the time-varying mean and Vol-of-Vol of the realized kernels. The model outperforms a number of recent competitors, both in-sample and out-of-sample. In particular, accounting for time-varying tail shapes matters for both density forecasts and forecasts of volatility risk quantiles.

Keywords: realized kernel; heavy tails; F distribution; time-varying shape-parameter; Vol-of-Vol, score-driven dynamics.

Classification codes: C32, C58.

*We appreciate the comments of participants at the 11th Annual SoFiE Conference (Lugano, June 2018), the 12th meeting of the NESG (Amsterdam, June 2018) and seminar participants at the University of Parma (June 2018) at the Econometrics Brown Bag Seminar Series at Vrije Universiteit Amsterdam (May 2018).*Corresponding author*: Anne Opschoor, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, The Netherlands. *E-mail*: a.opschoor@vu.nl. *Phone*: +31(0)20-5982663.

1 Introduction

Volatility is a key ingredient for option pricing and volatility trading; see for instance Sinclair (2013). Therefore, volatility related risk measures such as the 'Volatility at Risk' (VolaR) (Caporin *et al.*, 2017) have been put forward to estimate the upper quantiles of the realized volatility distribution. Good estimates of the volatility distribution and the VolaR are crucial for managing the risk of volatility-based trading strategies, such as portfolios containing volatility indices or options.

Due to the availability of High-Frequency (HF) data, volatility measurement, modeling and forecasting has improved significantly over the past two decades (Andersen *et al.*, 2003). In many modern volatility models, ex-post (direct) measurements of the daily volatility are used, ranging from the Realized Variance (Andersen and Bollerslev, 1998) to the Realized Kernel of Barndorff-Nielsen *et al.* (2008). Older volatility models focus more on modeling the latent volatility via the distribution of observed returns. The current literature, however, has shifted towards modeling the ex-post realized volatility measures directly, such as the HAR model of Corsi (2009), the MEM model of Engle and Gallo (2006), or the multivariate CAW model of Golosnoy *et al.* (2012), among others. Also hybrid models based on both returns and realized measures have been put forward, such as the univariate and multivariate HEAVY models of Shephard and Sheppard (2010) and Noureldin *et al.* (2012).

As a stylized fact, realized measures are typically fat-tailed and right-skewed. This can be attributed to turbulent market periods and price jumps, even within the day (e.g. the Flash Crash in May 2011). Only very few studies take into account the fat-tailedness of realized measures. Examples include Caporin *et al.* (2017) and Opschoor *et al.* (2018). Caporin *et al.* (2017) extend the MEM model of Engle and Gallo (2006) by including jumps under the assumption of a mixture of Gamma distributions for the realized variance to capture the skewed right tail of the volatility density. Opschoor *et al.* (2018) on the other hand propose an matrix F distribution for realized covariance matrices, which is a continuous mixture of Wishart distributions. Neither of these models, however, allows the Vol-of-Vol and tail

shape to vary over time.

This paper develops a new model for the dynamics of realized kernels, allowing for time-variation in the mean volatility level, in the Vol-of-Vol, and in the tail-behavior of realized kernel volatility simultaneously. Intuitively, the tail of realized kernels might be relatively larger during crisis periods (e.g. the Internet bubble, the Global Financial Crisis or the Sovereign debt crisis) than during calm periods. Time-varying tail shapes have been studied before by for instance Gerlach *et al.* (2013) and Lucas and Zhang (2016), but both studies only focus on the return distribution. The dynamics in our new model are score-driven (see Creal *et al.*, 2013). Earlier successful models for capturing the dynamics of parameters under possibly non-standard distributions include volatility and location modeling (Harvey, 2013; Harvey and Luati, 2014), credit risk modeling (Creal *et al.*, 2014), and systemic risk modeling (Lucas *et al.*, 2014; Oh and Patton, 2018). The availability of a closed-form expression for the likelihood function allows for straightforward estimation by maximum likelihood. Combined with the optimality results for score-driven steps of Blasques *et al.* (2015), the score-driven framework is attractive for modeling time-varying parameters of the conditional distribution of realized kernels.

We account for fat-tails of the realized kernels by assuming an F distribution, where we impose score-driven dynamics for the time-varying mean and for the two degrees of freedom (DoF) parameters of the F distribution. The (matrix) F distribution was recently introduced in financial econometrics by Opschoor *et al.* (2018) and appears to fit well to realized (covariance)kernels. Although higher moments such as the variance and skewness of the F distribution depend on *both* DoF parameters of the F, we show that the second DoF parameter has the most impact on the tail shape (skewness) of the distribution, while the first DoF mostly affects the dispersion (or variance) of the distribution. This dispersion or Vol-of-Vol might be time-varying as well, see for instance Corsi *et al.* (2008).

In our empirical application, we use the new model to describe daily realized kernels of 30 constituents of the S&P 500 index over the period January 2001 to December 2014. In-sample, the statistical fit increases significantly when allowing for time-varying DoF pa-

rameters compared to benchmarks such as the MEM model of Engle and Gallo (2006) and the univariate GAS F model with HAR dynamics of Opschoor *et al.* (2018).

Out-of-sample, we assess the economic significance of our results by considering 1-step ahead Volatility-at-Risk (VolaR) predictions and density forecasts. The results strongly indicate that it is crucial to allow for a time-varying shape parameter: it produces significantly better VolaR predictions than all competing models. Moreover, we also show that a score-driven model (with a time varying 'Vol-of-Vol' parameter) for the logarithm of the realized kernel does not produce accurate VolaR predictions. Our model also performs well in terms of density forecasts. In particular it outperforms benchmarks such as the MEM model with Gamma distributions. Moreover, in 24 out of 30 stocks the average log score of our new model with time-varying shape and Vol-of-Vol parameters exceeds the log score of our benchmark score-driven model. In 50% of the cases this difference is statistically significant at a 10 % level.

The rest of this paper is set up as follows. In Section 2, we introduce the new score-driven models for the dynamic shape parameters of the realized kernel distribution. In Section 3, we give a brief overview of the data used. Section 4 covers the empirical application. We conclude in Section 5.

2 Score driven models for the realized kernel

2.1 The HEAVY GAS HAR F model

Let $RK_t \in \mathbb{R}$ denote the realized kernel on day t, $t = 1, \ldots, T$, where RK_t is computed following Barndorff-Nielsen *et al.* (2009). Guided by the literature (see for instance Opschoor *et al.*, 2018)), we assume that RK_t is fat-tailed and follows an F distribution. The

conditional density for RK_t then reads

$$p(RK_t|\mu_t, \mathcal{F}_{t-1}; \nu_1, \nu_2) = \frac{\Gamma((\nu_1+\nu_2)/2)}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \frac{RK_t^{(\nu_1-2)/2} \left(\frac{\nu_1}{\mu_t(\nu_2-2)}\right)^{\nu_1/2}}{\left(1+\frac{\nu_1}{\nu_2-2}\frac{RK_t}{\mu_t}\right)^{(\nu_1+\nu_2)/2}}, \tag{1}$$

where $\mathcal{F}_{t-1}$ denotes the information set containing all realized kernels up to and including time $t-1$, μ_t contains the time-varying mean of RK_t, and ν_1 and ν_2 are the degrees of freedom (DoF) parameters. We assume that $\nu_2 > 2$, such that the mean of RK_t exists, where $\mathbb{E}[RK_t|\mathcal{F}_{t-1}] = \mu_t$. When $\nu_2 \to \infty$, the F distribution collapses to the χ^2 or Wishart distribution. By an appropriate choice of μ_t, this also covers the Gamma distribution used by Engle and Gallo (2006) and Caporin *et al.* (2017), among others.

The corresponding (conditional) variance and skewness of the F distribution are given by

$$\text{Var}[RK_t|\mathcal{F}_{t-1}] = 2\,\frac{(\nu_1+\nu_2-2)}{\nu_1(\nu_2-4)}\,\mu_t^2, \tag{2}$$

$$\text{Skew}[RK_t|\mathcal{F}_{t-1}] = \frac{(\nu_2-2)^3(2\nu_1+\nu_2-2)\sqrt{8(\nu_2-4)}}{\nu_2^3(\nu_2-6)\sqrt{\nu_1(\nu_1+\nu_2-2)}}\,\mu_t^3, \tag{3}$$

where the conditional variance and skewness exist if $\nu_2 > 4$ and $\nu_2 > 6$ respectively. Both the variance and skewness of the F distribution depend on ν_1 *and* ν_2. It is therefore not immediately clear how these parameters affect each of these quantities separately. To disentangle these effects, Figure 1 shows a surface plot for the variance and skewness for different combinations of ν_1 and ν_2. We obtain two important insights. First, the variance is decreasing in ν_2 and ν_1. However, if both values become small, the variance increases relatively more due to ν_1 compared to ν_2. Second, the skewness is decreasing in both ν_2 and ν_1, but for small values of ν_1 and ν_2 the impact of ν_2 on the skewness is larger than the impact of ν_1. We therefore label ν_2 as the 'tail shape' or 'skewness' parameter and ν_1 as the

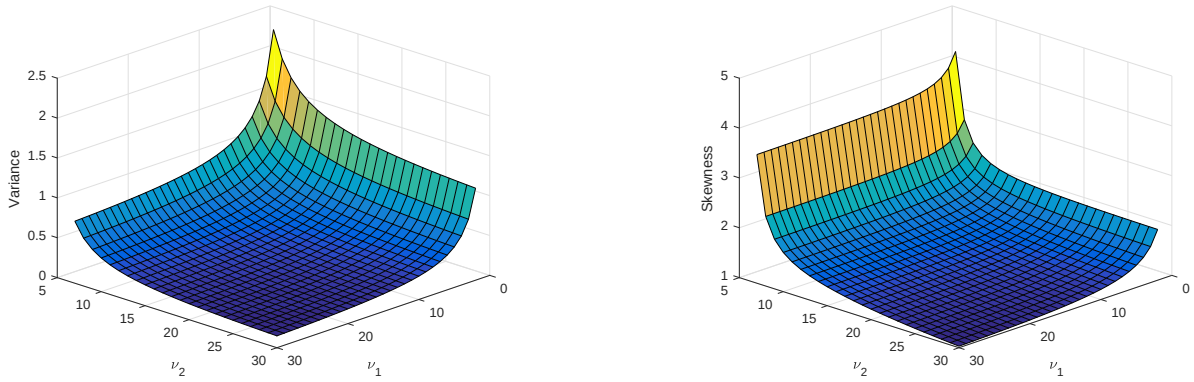

Figure 1: Variance and Skewness of the F distribution.
This figure shows the variance and skewness of an F distributed random variable with unit mean as a function of ν_1 and ν_2.

'dispersion' or 'Vol-of-Vol' parameter.

Our modeling framework starts with the univariate version of the multivariate HEAVY GAS model of Opschoor *et al.* (2018) with HAR dynamics. The dynamics of the volatility process μ_t follow the generalized autoregressive score (GAS) framework of Creal *et al.* (2013), augmented with HAR dynamics to accommodate for long-memory type persistence of the variance. Allowing for long-memory type persistence is empirically important, see for instance Corsi (2009). Using the score-driven dynamics, it is easy to derive a closed-form expression for the likelihood function, which facilitates easy parameter estimation and inference. Using Opschoor *et al.*, the GAS HAR dynamics for μ_t are given by

$$
\begin{aligned}
RK_t &= \mu_t \epsilon_t \qquad \epsilon_t \sim F(1, \nu_1, \nu_2), && (4)\\
\mu_{t+1} &= \omega_0 + \alpha_0\, s_{\mu,t} + \beta_l\, \mu_{l_1,t} + \beta_2\, \mu_{l_2,t} + \beta_3\, \mu_{l_3,t}, && (5)\\
s_{\mu,t} &= S_t \nabla_t = S_t \frac{\partial \log p(RK_t|\mu_t, \mathcal{F}_{t-1}; \nu_1, \nu_2)}{\partial \mu_t} && (6)\\
&= \frac{\nu_1}{\nu_1 + 1} \left(\frac{\frac{\nu_1+\nu_2}{\nu_2-2} RK_t}{(1 + \frac{\nu_1}{\nu_2-2} \frac{RK_t}{\mu_t})} - \mu_t \right),
\end{aligned}
$$

where $\partial \log p_{RK}(RK_t|\mu_t, \mathcal{F}_{t-1}; \nu_1, \nu_2)/\partial \mu_t$ denotes the score with respect to μ_t, S_t is a scaling factor, $\mu_{l,t} = l^{-1} \sum_{i=1}^{l} \mu_{t-i}$, and $l_1 = 1, l_2 = 12$, and $l_3 = 60$. We follow Opschoor *et al.* (2018) and scale the score by $2\mu_t^2/(\nu_1+1)$ to account for the curvature of $\log p_{RK}(RK_t|\mu_t, \mathcal{F}_{t-1}; \nu_1, \nu_2)$ with respect to μ_t. It is proportional to the inverse conditional Fisher information with re-

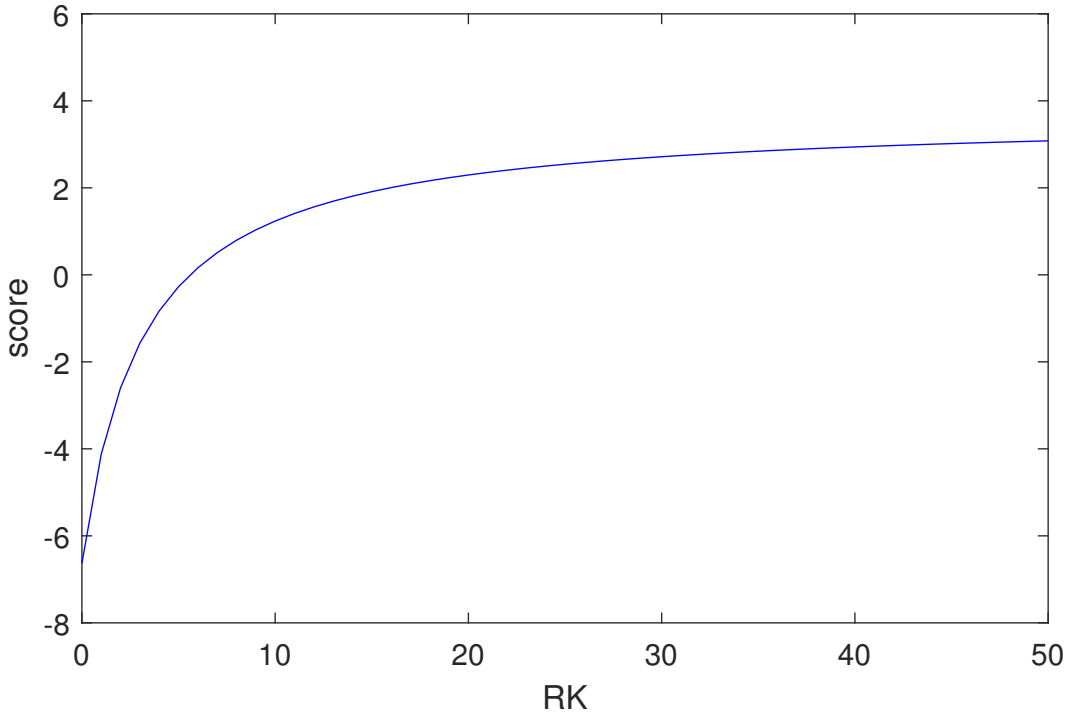

Figure 2: Score of the F distribution with respect to μ_t
This figure shows the score of the F distribution (given in (6)) with respect to μ_t for various values of RK. The score is computed by setting $\nu_1 = 18, \nu_2 = 10$ and μ_t to 7.

spect to μ_t. In the remainder of this paper, we label this model the GAS HAR model.

The scaled score (6) has an appealing interpretation for the dynamics of μ_t. It is the difference between a weighted value of RK_t and μ_t, where the weight accounts for the fat-tailedness of RK_t. That is, large values of RK_t imply a low weight such that the impact of a large RK_t on μ_{t+1} will be limited. Figure 2 visualizes this property of the score of a fat-tailed distribution. If RK_t increases, the score increases as well but the increase levels off and becomes flat for large values of RK_t, refraining μ_{t+1} from exploding if ν_2 is finite.

2.2 GAS model for the DoF parameters

The flexibility of the GAS framework is that we can easily handle time-variation in other parameters than μ_t. This includes parameters describing the higher moments of ϵ_t. Time-variation in parameters like ν_1 and/or ν_2 might be expected. For instance, during crisis periods the tail shape might become relatively fat, while in calm periods ϵ_t might have lighter tails, implying time-variation in ν_2.

We consider the following two time-varying parameter models for the DoF parameters. In the first model, we allow for a time-varying dispersion parameter $\nu_{1,t}$,

$$
\begin{aligned}
f_{1,t+1} &= \omega_1 + \alpha_1 s_{\nu_1,t} + \beta_1 f_{1,t}, \qquad \nu_{1,t} = 2 + \exp(f_{1,t}), && \text{(7a)}\\
s_{\nu_1,t} &= \frac{\partial \log p(RK_t|\mu_t, \mathcal{F}_{t-1}; \nu_{1,t}, \nu_2)}{\partial \nu_{1,t}} \frac{\partial \nu_{1,t}}{\partial f_{1,t}} \\
&= \left[\frac{1}{2}\psi\left(\frac{\nu_{1,t}+\nu_2}{2}\right) - \frac{1}{2}\psi\left(\frac{\nu_{1,t}}{2}\right) + \frac{1}{2}\left(\log\left(\frac{\nu_{1,t}}{\nu_2-1}\right)+1\right) + \frac{1}{2}\log\left(\frac{RK_t}{\mu_t}\right) - \right. \\
&\quad \left. \frac{1}{2}\log\left(1+\frac{\nu_{1,t}}{\nu_2-2}\frac{RK_t}{\mu_t}\right) - \frac{\nu_{1,t}+\nu_2}{2}\frac{\frac{1}{\nu_2-2}\frac{RK_t}{\mu_t}}{(1+\frac{\nu_{1,t}}{\nu_2-2}\frac{RK_t}{\mu_t})}\right](\nu_{1,t}-2), && \text{(7b)}
\end{aligned}
$$

where ψ denotes the digamma function $\psi(x) = \partial \log \Gamma(x)/\partial x$. We label this the GAS HAR ν_1 model. As we have seen before, this model focuses on modeling the Vol-of-Vol (see also Corsi *et al.*, 2008). In the second model, we consider time-variation in the tail shape via the parameter ν_2. We have

$$
\begin{aligned}
f_{2,t+1} &= \omega_2 + \alpha_2 s_{\nu_2,t} + \beta_2 f_{2,t}, \qquad \nu_{2,t} = 2 + \exp(f_{2,t}), && \text{(8a)}\\
s_{\nu_2,t} &= \frac{\partial \log p(RK_t|\mu_t, \mathcal{F}_{t-1}; \nu_1, \nu_{2,t})}{\partial \nu_{2,t}} \frac{\partial \nu_{2,t}}{\partial f_{2,t}} \\
&= \left[\frac{1}{2}\psi\left(\frac{\nu_1+\nu_{2,t}}{2}\right) - \frac{1}{2}\psi\left(\frac{\nu_{2,t}}{2}\right) - \frac{1}{2}\frac{\nu_1}{\nu_{2,t}-1} - \frac{1}{2}\log\left(1+\frac{\nu_1}{\nu_{2,t}-2}\frac{RK_t}{\mu_t}\right) + \right. \\
&\quad \left. \frac{\nu_1+\nu_{2,t}}{2}\frac{\frac{\nu_1}{(\nu_{2,t}-2)^2}\frac{RK_t}{\mu_t}}{(1+\frac{\nu_1}{\nu_{2,t}-2}\frac{RK_t}{\mu_t})}\right](\nu_{2,t}-2). && \text{(8b)}
\end{aligned}
$$

The parameterization $\nu_{2,t} = 2 + \exp(f_{2,t})$ ensures that $\nu_{2,t} > 2$ for all $f_{2,t} \in \mathbb{R}$, such that the mean of RK_t always exists. We label this model the GAS HAR ν_2. Note that we can easily combine models (7a)–(7b) and (8a)–(8b) into a model with both time-varying mean, Vol-of-Vol, and tail shape. We call this combined model the GAS HAR ν_{12} model.

Even though ν_1 and ν_2 relate to the higher order moments of the conditional distribution, the score expressions in equations (7b) and (8b) show that the dynamics of $f_{1,t}$ and $f_{2,t}$ are *not* driven by high order powers of RK_t such as RK_t^k for $k \geq 3$. A similar result is found by Lucas and Zhang (2016), who model the tails of asset returns by varying the degrees of

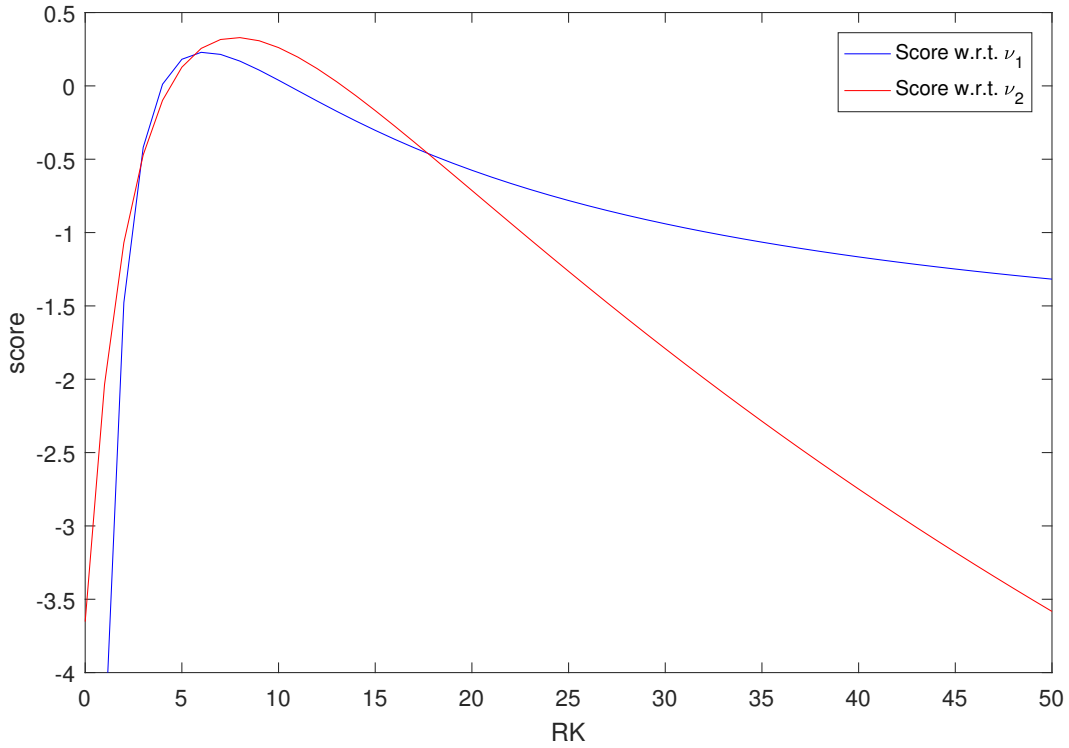

Figure 3: Score of the F distribution with respect to ν_1 and ν_2
This figure shows the score of the F distribution (given in (8b)) with respect to ν_1 (blue line) and ν_2 (red line) for various values of RK. The former (latter) score is computed by setting $\nu_1 = 17(19), \nu_{2,t} = 13(17)$ and μ_t to 7.

freedom parameter of a (skewed) t distribution. Figure 3 visualizes the impact of the score on both $\nu_{1,t}$ and $\nu_{2,t}$ by varying RK_t. The figure clearly shows that when RK becomes large, both scores become more negative, and hence the values of $\nu_{1,t}$ and $\nu_{2,t}$ decrease. Note that the impact of large values of the realized kernel on $\nu_{2,t}$ is much larger than for $\nu_{1,t}$. This is in line with Figure 1, which confirms our finding that $\nu_{2,t}$ is more related to the tail of the distribution and changes more than the Vol-of-Vol related parameter $\nu_{1,t}$ for large RK_t. However, the *rate* of change is much more moderate than compared to a power RK_t^k of RK_t for $k \geq 3$. This robustness feature is very important for the time-series dynamics of $\nu_{1,t}$ and $\nu_{2,t}$, which will react much less violently to incidental outliers. Figure 1 also shows that small values (near zero) of RK result in lower values of $\nu_{1,t}$ and $\nu_{2,t}$. The reason is that for low values of these parameters the F distribution also exhibits more leptokurtosis, i.e., a higher peak near zero, coinciding with more realizations of RK_t close to zero.

A direct extension of the current models would be to allow the dynamics of μ_t and/or

$\nu_{1,t}$ to spill over into those of $\nu_{2,t}$, or vice versa. The updating mechanism for such a model is given by

$$\begin{pmatrix} \mu_{t+1} \\ f_{1,t+1} \\ f_{2,t+1} \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \begin{pmatrix} s_{\mu,t} \\ s_{\nu_1,t} \\ s_{\nu_2,t} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{pmatrix} \begin{pmatrix} \mu_t \\ f_{1,t} \\ f_{2,t} \end{pmatrix}, \tag{9}$$

For instance, the coefficients α_{31} and α_{32} now allow for a possible spillover effect of the score of the mean and Vol-of-Vol to the tail shape parameter. Similarly, β_{ij} enables for a spillover effect of the current level of the mean, volatility, and tail shape parameter to the other parameters in the next period. In our empirical application, such spillovers did not result in substantial model improvements. In other settings, however, such spillovers might be relevant.

The static parameters in all models can be estimated by maximum likelihood using a standard prediction error decomposition. We first estimate the GAS model with time-varying μ_t only as our benchmark. Next, we make $\nu_{1,t}$ and/or $\nu_{2,t}$ time-varying. We parameterize the intercept ω for the GAS HAR models as $(1-\beta)\cdot\bar{f}$, where $\bar{f}$ denotes unconditional mean of the time-varying parameter f_t, and estimate $\bar{f}$ rather than ω.

3 Data

The data consist of daily realized kernels of 30 U.S. equities from January 2, 2001 until December 31, 2014 and contains $T = 3521$ trading days. We retrieve consolidated trades (transaction prices) from the Trade and Quote (TAQ) database from 9:30 until 16:00 with a time-stamp precision of one second. After cleaning the high-frequency data following the guidelines of Barndorff-Nielsen *et al.* (2009) and Brownlees and Gallo (2006), we construct realized kernels based on 5-minute returns following again Barndorff-Nielsen *et al.* (2009).

Table 1 provides summary statistics of the data. All realized kernels are multiplied by a factor 10,000 and are thus denominated in basis points. The 'skew' column shows

Table 1: Summary statistics of realized kernels(in basis points) of 30 assets from the S&P 500 index

This table shows summary statistics of daily realized kernels of 30 stocks from the S&P 500 index. All realized kernels are multiplied by 10,000 and are thus in terms of basis points. We show per company the ticker symbol, the mean, standard deviation, minimum, maximum and the skewness. A bold number indicates that the time series has been winsorized (upper tail) using the 99.75% quantile. The sample covers the period January 2, 2001 until December 31 2014 and contains 3521 observations.

Ticker	mean	std.	min	max	skew	Ticker	mean	std.	min	max	skew
AES	9.42	25.42	0.14	**291**	7.20	LLY	1.88	3.40	0.07	120	15.86
AIG	8.46	30.84	0.13	**391**	8.80	MCO	3.65	6.48	0.07	139	6.80
BAC	5.54	16.00	0.08	**171**	6.70	MDT	1.97	3.63	0.13	143	19.66
BXP	2.97	7.18	0.03	147	7.26	MMC	2.60	5.09	0.08	118	10.42
C	6.48	19.61	0.10	**240**	8.03	MO	1.74	3.11	0.09	55	7.70
CNX	8.06	15.69	0.14	339	9.13	MS	7.62	24.07	0.23	**360**	10.73
CSC	3.16	5.17	0.06	129	8.19	MTB	3.03	6.96	0.05	127	7.04
EQR	3.39	7.70	0.04	120	6.17	NLY	3.41	9.36	0.06	181	9.85
FLR	4.91	8.74	0.09	159	7.35	POT	4.41	9.57	0.06	207	9.99
GS	4.21	13.85	0.16	429	18.78	PXD	5.53	8.22	0.28	191	8.02
HAL	6.23	11.89	0.27	295	13.43	SU	3.97	7.86	0.08	200	11.24
HON	3.08	5.93	0.09	191	13.01	TV	3.18	5.28	0.09	167	13.19
JCP	6.70	8.35	0.31	187	6.08	USB	4.10	11.47	0.09	351	12.90
KEY	6.81	19.82	0.13	**240**	7.35	WMB	8.17	21.76	0.18	**274**	7.97
KO	1.39	2.54	0.05	89	15.02	XL	6.71	22.56	0.06	**269**	7.57

the huge positive skewness of each time series, which motivates the use of the fat-tailed F distribution. Figure 4 shows the evolution of RK_t for four random companies: Fluor Corp, KeyCorp, Boston Properties INC, and Consol Energy Inc. For all stocks, there are quite some peaks (outliers) in the data, especially during the 2007 Global Financial Crisis (GFC). This can also be seen in Table 1, as the maxima of most time series are large. For example, we have a maximum value of 159 for Fluor Corp. (FLR), implying a daily volatility of about 12.6%, which is large. Such high values may disrupt the estimated dynamic pattern of the realized kernel and therefore require careful handling. For some stocks, the maximum RK_t becomes unrealistically large, such that we first winsorize the data using an upper tail level of 99.75%. Despite this, the use of the F distribution will still turn out to be empirically highly relevant.

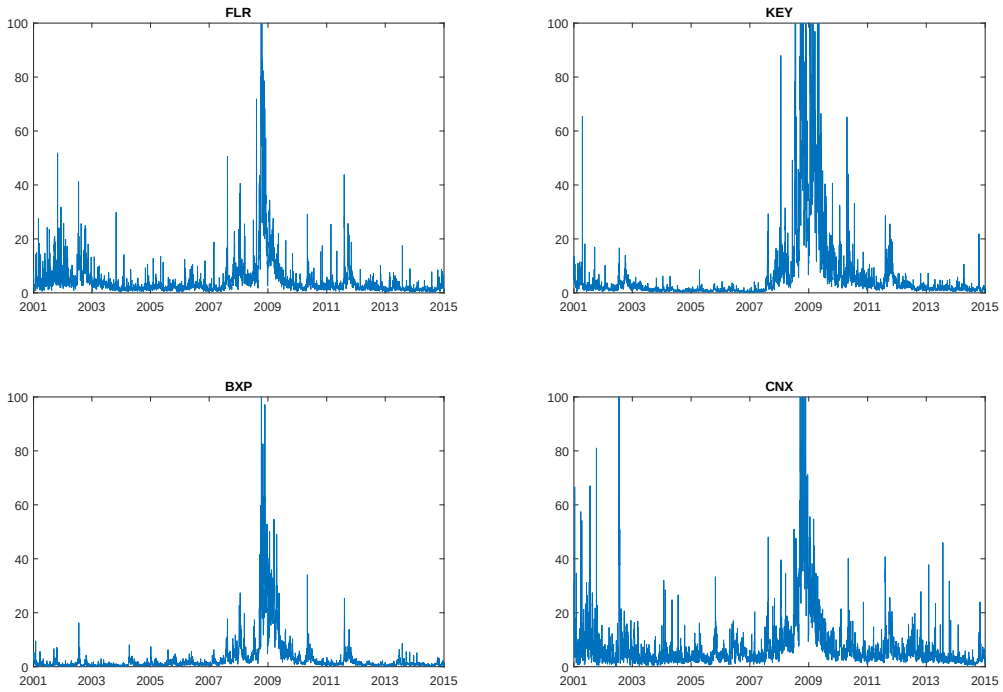

Figure 4: Realized kernel time series
This figure shows the time series of the realized kernels (RK) of Fluor Corp, KeyCorp, Boston Properties INC, and Consol Energy Inc. For visual purposes, the vertical axes has been capped. The sample covers the period January 2, 2001 until December 31 2014 and contains 3521 observations.

4 Empirical application

In this section, we first provide an in-sample analysis by estimating our models using the whole sample to assess the possibly time-varying tail-behavior of the realized kernels. We then show the forecasting power of the newly introduced score-driven model against some benchmarks by computing 1-step-ahead Volatility-at-Risk (VolaR) forecasts and density forecasts.

4.1 Benchmark models

We use two benchmark models in both the in-sample and out-of-sample analysis. First, we compare the GAS HAR model with dynamic DoF parameters against the GAS HAR F model (5) with static Vol-of-Vol and tail shape behavior. Second, we consider the MEM model of Engle and Gallo (2006) to assess the value-added of a fat-tailed F distribution

against the thinner-tailed Gamma distribution for the realized kernel. The MEM model is given by

$$RK_t = \mu_t \epsilon_t, \qquad \epsilon_t \sim \text{Gamma}(1, \nu_M), \qquad \mu_{t+1} = \omega_M + \alpha_M RK_t + \beta_M \mu_t. \tag{10}$$

Third, we benchmark our GAS model out-of-sample against a model for the logarithm of RK_t. A model for the log volatility using a more standard normal or Student's t density for the innovations could arguably be labeled as a simpler alternative to a model for RK_t directly using the more non-standard F distribution. To endow the model for log volatility with sufficient dynamic flexibility, we use a score-driven time-varying mean and Vol-of-Vol, following Corsi *et al.* (2008). The model is given by

$$\log RK_t = \mu_t + \sigma_t \epsilon_t, \qquad \epsilon_t \sim t(0, 1, \nu_0), \tag{11a}$$

$$\mu_{t+1} = \omega_1 + \alpha_1 s_{\mu,t} + \beta_1 \mu_t, \qquad \sigma^2_{t+1} = \omega_2 + \alpha_2 s_{\sigma^2,t} + \beta_2 \sigma^2_t, \tag{11b}$$

$$s_{\mu,t} = S_\mu \frac{\partial \log p(\log RK_t | \mu_t, \mathcal{F}_{t-1}; \nu_0)}{\partial \mu_t} = \frac{(\nu_0 + 1)\epsilon_t}{(\nu_0 - 2)\sigma^2 + \epsilon^2_t} - \mu_t, \tag{11c}$$

$$s_{\sigma^2,t} = S_{\sigma^2} \frac{\partial \log p(\log RK_t | \mu_t, \sigma^2, \mathcal{F}_{t-1}; \nu_0)}{\partial \sigma^2_t} = \left(\frac{\nu_0 + 1}{(\nu_0 - 2) + \frac{\epsilon^2_t}{\sigma^2_t}} \right) \epsilon^2_t - \sigma^2_t, \tag{11d}$$

where σ^2_t denotes the Vol-of-Vol. We also estimate a restricted version of this model by setting $\sigma^2_t = \sigma^2$. Fat-tails are incorporated in this model by assuming a Student's t distribution for $\log RK_t$. Note that this model implies no integer moments exist for RK_t itself. This may qualify as an undesirable feature of benchmark model (11a), which we label as the GAS t (Vol-of-Vol) model. All benchmark models are again estimated by Maximum Likelihood.

4.2 In-sample results

Table 2 shows summary statistics of the estimated parameters using all observations for all 30 individual stocks from the S&P 500 index. We report the means, standard deviations,

and the 5th and 95th quantiles across the all assets.

The table shows the well-known high persistence in volatility, which is measured by the sum of α_{vol} and β_{vol} in case of the MEM model, and by $\sum_{j=1}^{3} \beta_{j,vol}$ in case of the GAS HAR models. There is also persistence in the time-varying $\nu_{1,t}$ and $\nu_{2,t}$ processes, although the Vol-of-Vol time series $\nu_{1,t}$ appears to be overall less persistent than the time series of the 'shape' parameter $\nu_{2,t}$: the 5% quantile (across all assets) of β_{ν_1} of the GAS H ν_1 model equals 0.540, while the 5% quantile of β_{ν_2} of the GAS H ν_2 model equals 0.954. The same pattern is apparent for the GAS H ν_{12} model.

Figure 5 shows differences in the maximized log-likelihood versus a MEM (upper panel) or GAS HAR F (lower panel) benchmark for all 30 stocks. The stocks are ordered in alphabetical order as in Table 1. The high values in the upper graph clearly suggests that the conditional F distribution of the GAS HAR model fits the Realized Kernel considerably better than the Gamma distribution of the MEM model. The bottom figure shows that allowing for a time-varying Vol-of-Vol or shape further increases the statistical fit for each stock. Note that two times the log-likelihood difference always exceeds the values 5.99 and 9.48, i.e., the 95% critical values χ^2 distribution with two or four degrees of freedom, respectively. Hence the increase in the likelihood of the models with a time-varying shape or Vol-of-Vol is statistically significant w.r.t. the benchmark GAS H model. Allowing for both time-varying DoFs at the same time again increases the fit, but less strong.

Figure 6 shows the implied standardized conditional skewness and variance of ϵ_t in (4) for three specific stocks: WMB, BXP and AIG. Note that we exclude the effect of μ_t on the skewness and Vol-of-Vol by dividing by μ_t^3 and μ_t^2, respectively. The graphs for all three stocks show the same picture: allowing only for a time-varying Vol-of-Vol parameter $\nu_{1,t}$ indeed produces a time-varying Vol-of-Vol, but hardly results in any time-variation in the skewness. By contrast, allowing for a time-varying shape parameter $\nu_{2,t}$ has a relatively huge impact on the skewness, but also produces time-variation in the Vol-of-Vol graphs. The skewness is particularly high during turbulent periods, such as the 2001 U.S. recession, the 2007-2008 Global Financial Crisis and/or the 2013 Sovereign Debt Crisis. If only $\nu_{2,t}$

Table 2: In-sample parameter estimates

This table reports summary statistics of maximum likelihood parameter estimates of the MEM, GAS HAR, GAS HAR ν_1, GAS HAR ν_2, and the GAS HAR ν_{12} model, applied to daily realized kernels of 30 stocks from the S&P 500 index. We show the mean, standard deviation and the 5% and 95% quantile of the parameters across all 30 stocks. The sample covers the period January 2, 2001 until December 31 2014 and contains 3521 observations.

model		ω	α_{vol}	$\beta_{1,vol}$	$\beta_{2,vol}$	$\beta_{3,vol}$	$\nu_1(\nu_M)$	ν_2	f_1	α_{ν_1}	β_{ν_1}	f_2	α_{ν_2}	β_{ν_2}
MEM	mean	0.089	0.417			0.567	6.73							
	std	(0.044)	(0.077)			(0.067)	(1.38)							
	5%	0.043	0.309			0.466	4.72							
	95%	0.189	0.542			0.668	9.08							
GAS H	mean	0.043	0.864	0.845	0.100	0.041	16.15	16.17						
	std	(0.024)	(0.122)	(0.036)	(0.035)	(0.011)	(4.31)	(3.46)						
	5%	0.020	0.642	0.789	0.040	0.016	10.13	10.79						
	95%	0.085	1.000	0.891	0.173	0.059	24.09	22.01						
GAS H ν_1	mean	0.046	0.828	0.844	0.097	0.039		18.42	2.298	0.223	0.910			
	std	(0.030)	(0.120)	(0.028)	(0.029)	(0.011)		(4.87)	(0.695)	(0.213)	(0.164)			
	5%	0.017	0.585	0.796	0.044	0.013		11.51	0.926	0.042	0.540			
	95%	0.101	1.000	0.881	0.161	0.056		24.15	3.012	0.684	0.999			
GAS H ν_2	mean	0.042	0.874	0.856	0.085	0.042	17.20					2.287	0.085	0.985
	std	(0.020)	(0.112)	(0.032)	(0.033)	(0.012)	(3.91)					(0.766)	(0.046)	(0.017)
	5%	0.020	0.603	0.789	0.037	0.018	11.54					0.586	0.026	0.954
	95%	0.082	1.000	0.902	0.150	0.063	24.21					3.051	0.199	1.000
GAS H ν_{12}	mean	0.046	0.859	0.847	0.091	0.041			2.388	0.249	0.862	2.426	0.075	0.937
	std	(0.025)	(0.120)	(0.027)	(0.028)	(0.012)			(0.625)	(0.222)	(0.205)	(0.597)	(0.057)	(0.186)
	5%	0.020	0.627	0.810	0.041	0.013			1.299	0.026	0.338	1.143	0.026	0.739
	95%	0.094	1.000	0.888	0.153	0.059			3.100	0.684	1.000	3.027	0.161	1.000

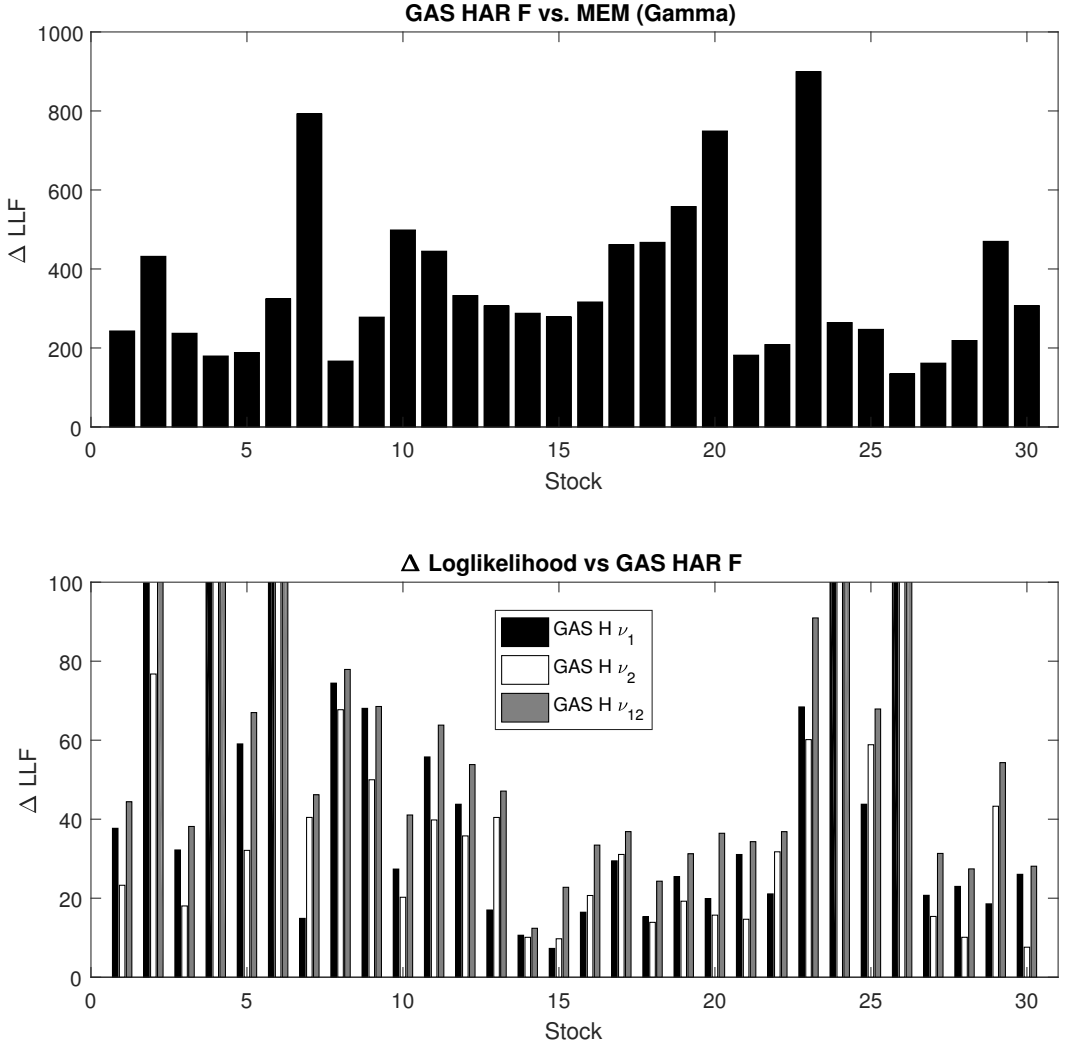

Figure 5: Differences in maximized log-likelihood values
This figure shows the differences in the maximized log-likelihood between several score-driven models and a benchmark model to model the dynamics of RK_t. The upper figure shows the difference between the log-likelihood of the GAS HAR model and the MEM model. The lower figure shows the differences between the maximized log-likelihood values of the GAS HAR ν_1, GAS HAR ν_2, and GAS HAR ν_{12} models vis-á-vis the GAS HAR model. The sample covers the period January 2, 2001 until December 31 2014 and contains 3521 observations.

varies and ν_1 is kept constant, $\nu_{2,t}$ also tries to capture the empirical variation in the Vol-of-Vol. This causes excessive variation in estimated skewness from time to time, which appears attributable to Vol-of-Vol increases (see WMB in 2001-2003 and AIG in 2009-2010) rather than to genuine variation in the skewness itself. These effects disappear if we let both the Vol-of-Vol ($\nu_{1,t}$) and tail shape ($\nu_{2,t}$) parameters vary over time. In that case, we obtain less excessive variation in skewness than if only ν_2 varies, but substantially more than if only

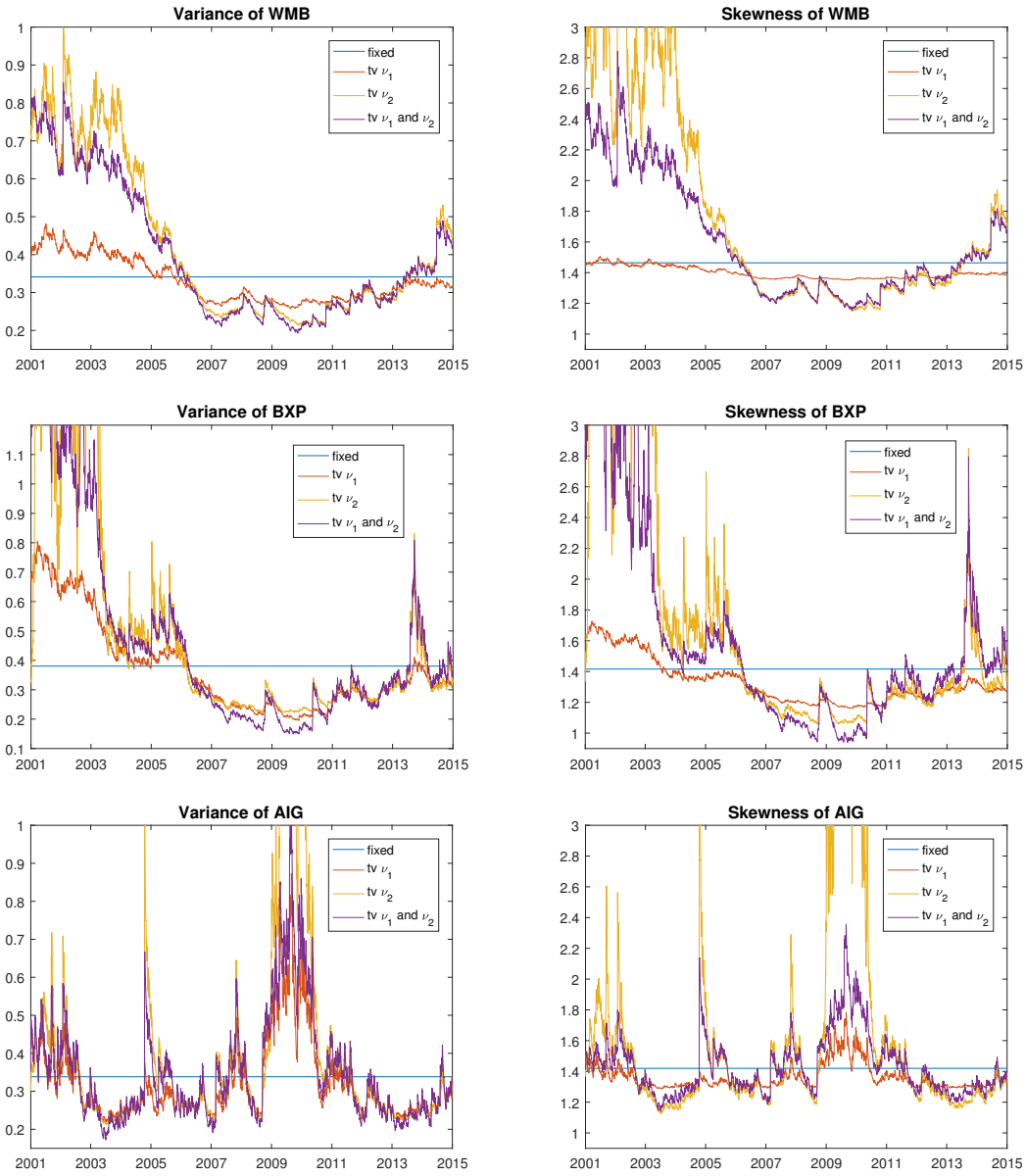

Figure 6: Estimated Vol-of-Vol ($\nu_{1,t}$) and skewness ($\nu_{2,t}$)
This figure shows time series of the estimated Vol-of-Vol (variance) and skewness of ϵ_t in (4) associated with the stocks WMB, BXP and AIG according to four models: the GAS HAR (blue line), GAS HAR ν_1 (red line), GAS HAR ν_2 (yellow line), and the GAS HAR ν_{12} model (purple line). The sample covers the period January 2, 2001 until December 31 2014 and contains 3521 observations.

ν_1 varies. Allowing both parameters to vary thus appears empirically important. The next section investigates the possible implications of these patterns in an out-of-sample analysis.

Based on the full-sample analysis, we conclude that the Vol-of-Vol and tail-shape parameters of the distribution of the realized kernel series vary through time. Allowing for this time-variation improves the fit of the conditional distribution. The following subsection investigates whether allowing for time-varying DoF parameters also improves the forecasting

power of the distribution of RK_t.

4.3 Out-of-sample results

We assess the short-term forecasting performance of the models in an economic setting by predicting the 1-step ahead Volatility-at-Risk. As indicated by Caporin *et al.* (2017), the VolaR is important for volatility traders and hedgers. Similar to the in-sample analysis of the previous subsection, we consider the GAS HAR type models and the MEM model. In addition to the MEM benchmark, we also consider the GAS t model for the logarithm of the realized kernel in (11a)–(11d) as a further benchmark model. We use a recursive estimation approach starting from an initial sample period of 1500 observations. This covers the period 2001–2006, just before the start of the global financial crisis, and therefore constitutes a strong test on the forecasting performance of all models. We update the recursive estimates roughly every two months (50 observations).

The VolaR computes the risk of extremely high volatility. The q% VolaR is defined as the q-quantile of the distribution of the realized kernel,

$$P[RK_{t+1} > VolaR^q_{t+1}|\mathcal{F}_t] = q, \tag{12}$$

where q is set to 5 percent. Note that the probability depends on the shape of the distribution modeled by the (time-varying) degrees of freedom parameters, as well as the forecast of μ_t. We backtest our predicted $VolaR$ using the unconditional coverage (UC) test proposed by Christoffersen (1998), which tests whether the number of violations, i.e. the number of times $RK_{t+1} > \mu_{t+1}$, is equal to the unconditional coverage probability q.

Table 3 shows the results. The upper part of the table gives the mean violation rates. Both the MEM and the GAS models with fixed tail shape fail in at least 12 out of the 30 cases at a 1% significance level (bolded entries). The GAS HAR ν_2 model, by contrast, only fails in 3 out of 30 cases. Hence accounting for a time-varying shape results on average in considerably better Volatility-at-Risk forecasts, even if one accounts for a time-varying

Table 3: 1-step ahead Volatility-at-Risk predictions

This table reports results on 1-step ahead 95% Volatility-at-Risk predictions, using the MEM, GAS HAR, GAS HAR ν_1, GAS HAR ν_2 and the GAS HAR ν_{12} models, applied to daily realized kernels of 30 stocks from the S&P 500 index. In addition, we consider the GAS t (denoted as GAS t_1), GAS t_2 Vol-of-Vol (denoted as GAS t_2) applied on the logarithm of the realized kernel. We use an expanding window where the first window contains 1500 observations. The first part of the table lists the violation rate, i.e. the fraction of predictedd VolaR violations by the realized RK_t at time t. A bold number indicates that the null-hypothesis of the unconditional coverage of Christoffersen (1998) is rejected at a 1% significance level. The second part of the table summarizes the number of times the p-value of the test is below 10, 5 or 1%. The out-of-sample period is from December 20, 2006 until December 31 2014 and contains 2021 observations.

Ticker	$\log RK_t$		RK_t				
	GAS t_1	GAS t_2	MEM	GAS	GAS ν_1	GAS ν_2	GAS ν_{12}
AES	**0.036**	0.042	**0.035**	0.040	0.043	0.048	0.048
AIG	**0.072**	**0.074**	0.052	**0.063**	**0.065**	0.063	**0.064**
BAC	**0.092**	**0.084**	**0.082**	**0.091**	**0.083**	**0.085**	**0.083**
BXP	**0.026**	0.053	**0.032**	**0.024**	0.052	0.050	0.047
C	**0.090**	**0.082**	**0.079**	**0.087**	**0.074**	**0.081**	**0.077**
CNX	**0.031**	**0.064**	**0.027**	**0.025**	0.053	0.051	0.055
CSC	0.054	0.055	0.039	0.054	0.053	0.045	0.046
EQR	0.040	0.056	0.043	0.040	0.055	0.054	0.060
FLR	0.045	0.054	**0.038**	0.040	0.051	0.047	0.051
GS	**0.069**	**0.066**	0.050	0.061	**0.065**	0.052	0.052
HAL	0.043	0.047	**0.032**	0.043	0.047	0.045	0.048
HON	0.044	0.049	**0.037**	0.042	0.049	0.047	0.048
JCP	**0.071**	**0.068**	**0.069**	**0.066**	**0.068**	0.062	**0.068**
KEY	**0.069**	**0.066**	0.059	0.058	0.059	0.057	0.057
KO	**0.068**	**0.068**	0.061	**0.064**	**0.066**	0.063	**0.064**
LLY	0.054	0.051	0.047	0.054	0.053	0.051	0.051
MCO	0.062	**0.067**	0.052	0.054	0.062	0.051	0.056
MDT	0.052	0.059	0.046	0.051	0.053	0.048	0.052
MMC	0.045	0.045	**0.032**	0.040	0.044	0.042	0.044
MO	**0.035**	**0.037**	**0.017**	**0.031**	**0.036**	**0.037**	0.042
MS	**0.065**	**0.063**	0.062	0.059	**0.063**	0.061	0.061
MTB	0.058	**0.063**	**0.067**	0.058	**0.065**	0.058	0.061
NLY	**0.035**	0.056	**0.026**	**0.032**	0.042	0.044	0.048
POT	**0.034**	**0.070**	0.040	**0.032**	0.051	0.057	0.057
PXD	**0.037**	0.054	0.043	0.040	0.057	0.055	0.056
SU	**0.023**	0.056	**0.029**	**0.024**	0.050	0.053	0.053
TV	0.039	0.044	0.045	0.042	0.046	0.042	0.047
USB	0.045	0.047	0.049	0.048	0.056	0.048	0.055
WMB	**0.035**	0.043	**0.031**	**0.035**	0.045	0.048	0.048
XL	0.051	0.056	0.048	0.050	0.054	0.049	0.054
Summary of UC-test							
♯ p-val < 0.10	21	15	20	23	12	9	9
♯ p-val < 0.05	20	13	19	18	10	7	8
♯ p-val < 0.01	17	13	15	12	9	3	5

Vol-of-Vol. Interestingly, the GAS models with F distribution and time-varying tail shape for RK_t also outperform the time-varying location and scale model for $\log RK_t$: the UC test for the $\log RK_t$ models fails too many times, namely 17 and 13 times out of 30 cases, respectively.

We also compare the forecasting power of our models in a statistical setting, by considering 1-step ahead density forecasts of the realized kernel. We use the log score (see Mitchell and Hall, 2005; Amisano and Giacomini, 2007) as a scoring rule to differentiate between the density forecasts of the models. Define the difference in the log score between two density forecasts M_1 and M_2 as

$$d_{ls,t} = S_{ls,t}(RK_t, M_1) - S_{ls,t}(RK_t, M_2) \tag{13}$$

for $t = 1501, 1502, \ldots, T-1$ with $S_{ls,t}(RK_t, M_j)$ $(j = 1, 2)$ the log score at time t of the density forecast corresponding to model M_j,

$$S_{ls,t}(RK_t, M_j) = \log p_t(RK_t|\mu_t, \mathcal{F}_{t-1}, M_j) \tag{14}$$

where $p_t(\cdot)$ is the χ^2 or F density. The null and alternative hypothesis of equal predictive ability are now given by

$$H_0 : \mathbb{E}[d_{ls}] = 0, \qquad H_A : \mathbb{E}[d_{ls}] \neq 0, \tag{15}$$

for all P out-of-sample (OOS) forecasts. This hypothesis can be tested by means of a Diebold and Mariano (1995) (DM) test statistic

$$DM_{ls} = \bar{d}/\sqrt{\hat{\sigma}^2/P}, \tag{16}$$

with $\bar{d}$ the out-of-sample average of the log score differences and $\hat{\sigma}^2$ a HAC-estimator of the true variance σ^2 of d_{ls}. A significantly negative value of DM_{ls} means that model M_2 has a

Table 4: 1-step ahead density forecasts of the realized kernel

This table reports results on 1-step ahead density forecasts of the realized kernel, using the MEM model assuming a Γ distribution, the GAS HAR F model assuming an F distribution with fixed degrees of freedom parameters and three GAS HAR models with time varying tail shape and/or dispersion parameter(s). The models are applied to daily realized kernels of 30 stocks from the S&P 500 index. We use an expanding window where the first window contains 1500 observations. The table summarizes the number of times the test-statistic is negative (implying relatively better density forecasts produced by model M_2) and the number of negative significant test-statistics at a 10, 5 or 1% level. The out-of-sample period covers December 20, 2006 until December 31 2014 and contains 2021 observations.

Summary of DM-test				
Model M_1	MEM	GAS HAR	GAS HAR	GAS HAR
Model M_2	GAS HAR	GAS HAR ν_1	GAS HAR ν_2	GAS HAR ν_{12}
♯ neg	30	21	24	22
♯ neg sign. at 10%	29	12	14	17
♯ neg sign. at 5%	29	11	12	13
♯ neg sign. at 1%	29	6	10	9

superior forecast performance over model M_1.

Figure 7 and Table 4 present the results on the density forecasts. The figure present the histogram of the Diebold-Mariano tests for four different pairs of models. The dark blue bar corresponds to the comparison of the log score performance of the MEM model (Gamma distribution) to the GAS HAR model (F distribution). Clearly, the F distribution produces considerably better density forecasts than the Gamma distribution: all Diebold-Mariano tests for the 30 stocks are negative, and 29 significantly so. The light blue, green, and yellow bars give the DM statistics across all 30 stocks for the GAS HAR model with static ν_1 and ν_2 compared to its dynamic counterparts. Again, most DM statistics are negative. As expected, the effect is less strong than the switch from a Gamma to an F distribution. Still, however, it also appears relevant for out-of-sample density forecasts to allow for a time-varying shape and/or Vol-of-Vol parameter. Table 4 confirms this by showing summary statistics of the 30 DM test-statistics: the t-statistic is negative and statistically significant in 12, 14 and 17 out of 30 cases at a 10% level. Even at a 1% significance level, the model with a time-varying shape parameter $\nu_{2,t}$ still significantly outperforms the GAS HAR model with static shape parameter in a third of the cases.

We conclude that density forecasts improve when accounting for a time-varying shape

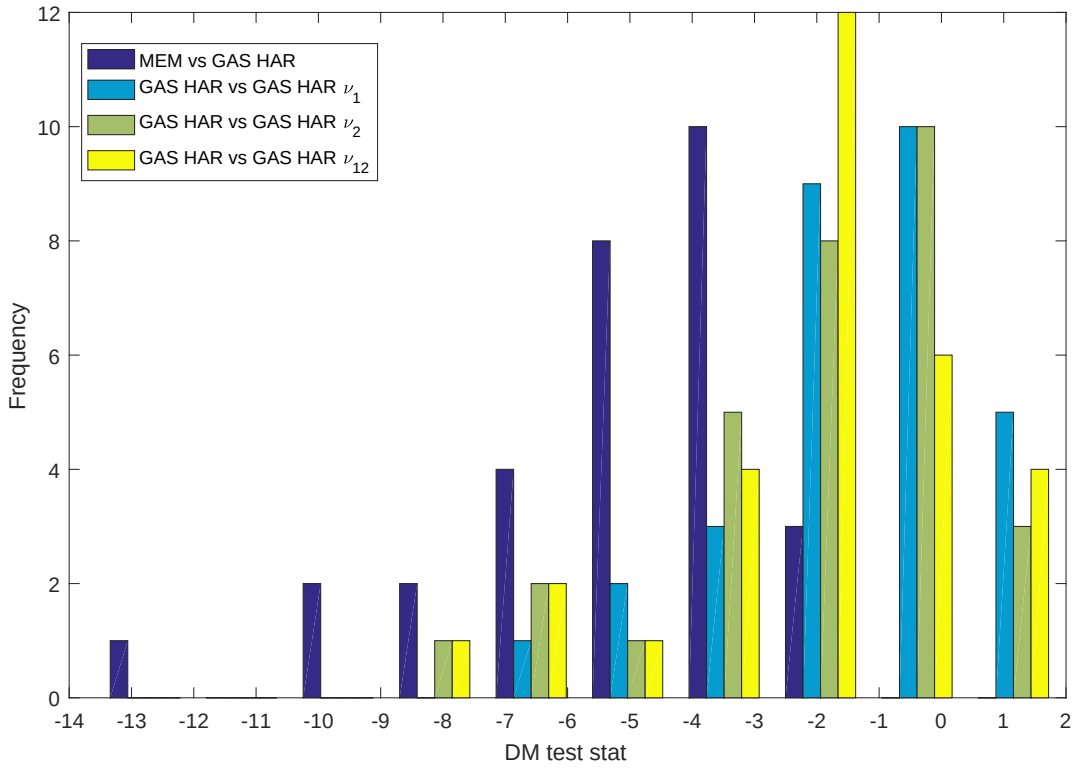

Figure 7: Histogram of test-statistics on equal predictive accuracy
This figure shows a histogram of Diebold-Mariano test-statistics on equal predictive accuracy between a pair of two models M_1 and M_2 using the log score. The four different colours present four different pairs of models: the dark blue bar corresponds with differences in the log score between the MEM model (Gamma distribution) and the GAS HAR model (F distribution). The light blue, green and yellow bars show test statistics of predictive accuracy between the GAS HAR ν_1, GAS HAR ν_2 and GAS HAR ν_{12} model vis-á-vis the benchmark GAS HAR model. A negative DM statistics means that the model M_2 has superior predictive accuracy. The out-of-sample period covers December 20, 2006 until December 31 2014 and contains 2021 observations.

and/or Vol-of-Vol parameter. In addition, we find that VolaR forecasts improve when the realized kernel is modeled directly with the F distribution compared to modeling the logarithm of RK_t with a Student's t distribution.

5 Conclusions

We introduced a new dynamic score driven model for the Vol-of-Vol and the tail shape of realized kernels. The proposed model explicitly acknowledges that realized kernels are fat-tailed. The proposed set-up is particularly suitable for cases where no explicit robustification methods are applied while estimating realized measures. Using the GAS dynamics of Creal

et al. (2013) and the recent work of Opschoor *et al.* (2018) based on the F distribution for the realized kernel, we derived an observation driven model for the unobserved tail shape and Vol-of-Vol parameters with robust propagation dynamics.

Using realized kernels of 30 U.S. stocks of the S&P 500 Index over 2001–2014, the new model improves both the in-sample and out-of-sample fit of the realized kernel dynamics vis-á-vis the MEM model of Engle and Gallo (2006) and the GAS F model with fixed tail shape and Vol-of-Vol of Opschoor *et al.* (2018). Particularly Volatility-at-Risk predictions are significantly better if one accounts for time-varying tail shape parameters. The model also outperforms models based on the logarithm of the realized kernel. We conclude that our new model provides a valuable tool when modeling and forecasting realized kernels, and time-varying tail shapes are empirically important.

References

Amisano, Gianni and Raffaella Giacomini (2007), Comparing density forecasts via weighted likelihood ratio tests, *Journal of Business & Economic Statistics* **25**, 177–190.

Andersen, T.G. and T. Bollerslev (1998), Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review* pp. 885–905.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2003), Modeling and forecasting realized volatility, *Econometrica* **71**, 529–626.

Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde and N. Shephard (2008), Designing realized kernels to measure the ex post variation of equity prices in the presence of noise, *Econometrica* **76**, 1481–1536.

Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde and N. Shephard (2009), Realized kernels in practice: trades and quotes, *Econometrics Journal* **12**, 1–32.

Blasques, C., S.J. Koopman and A. Lucas (2015), Information theoretic optimality of observation driven time series models for continuous responses, *Biometrika* **102**, 325–343.

Brownlees, C.T. and G.M. Gallo (2006), Financial econometric analysis at ultra-high frequency: Data handling concerns, *Computational Statistics and Data Analysis* **51**, 2232–2245.

Caporin, M., E. Rossi and P.S. de Magistris (2017), Chasing volatility: A persistent multiplicative error model with jumps, *Journal of Econometrics* **198**, 122–145.

Christoffersen, P. (1998), Evaluating interval forecasts, *International Economic Review* **39**, 841–862.

Corsi, F. (2009), A simple approximate long-memory model of realized volatility, *Journal of Financial Econometrics* **7**, 174–196.

Corsi, F., S. Mittnik, C. Pigorsch and U. Pigorsch (2008), The volatility of realized volatility, *Econometric Reviews* **27**, 46–78.

Creal, D., B. Schwaab, S.J. Koopman and A. Lucas (2014), Observation Driven Mixed-Measurement Dynamic Factor Models with an Application to Credit Risk, *Review of Economics and Statistics* **96**, 898–915.

Creal, D., S.J. Koopman and A. Lucas (2013), Generalized Autoregressive Score Models with Applications, *Journal of Applied Econometrics* **28**, 777–795.

Diebold, Francis X and Robert S Mariano (1995), Comparing predictive accuracy, *Journal of Business & economic statistics* **13**, 253–263.

Engle, R.F. and G.M. Gallo (2006), A multiple indicators model for volatility using intra-daily data, *Journal of Econometrics* **131**, 3–27.

Gerlach, R., Z. Lu and H. Huang (2013), Exponentially Smoothing the Skewed Laplace Distribution for Value-at-Risk Forecasting, *Journal of Forecasting* **32**, 534–550.

Golosnoy, V., B. Gribisch and R. Liesenfeld (2012), The conditional autoregressive Wishart model for multivariate stock market volatility, *Journal of Econometrics* **167**, 211–223.

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

Harvey, A.C. and A. Luati (2014), Filtering with heavy tails, *Journal of the American Statistical Association* **109**, 1112–1122.

Lucas, A., B. Schwaab and X. Zhang (2014), Conditional Euro Area Sovereign Default Risk, *Journal of Business and Economic Statistics* **32**, 271–284.

Lucas, A. and X. Zhang (2016), Score-driven exponentially weighted moving averages and Value-at-Risk forecasting, *International Journal of Forecasting* **32**, 293–302.

Mitchell, James and Stephen G Hall (2005), Evaluating, Comparing and Combining Density Forecasts Using the KLIC with an Application to the Bank of England and NIESR Fan Charts of Inflation, *Oxford bulletin of economics and statistics* **67**, 995–1033.

Noureldin, D., N. Shephard and K. Sheppard (2012), Multivariate high-frequency-based volatility (HEAVY) models, *Journal of Applied Econometrics* **27**, 907–933.

Oh, D.H. and A.J. Patton (2018), Time-varying systemic risk: Evidence from a dynamic copula model of cds spreads, *Journal of Business & Economic Statistics* **36**, 181–195.

Opschoor, A., P. Janus, A. Lucas, and D. van Dijk (2018), New HEAVY models for fat-tailed realized covariances and returns, *Journal of Business & Economic Statistics* **36**, 642–657.

Shephard, N. and K. Sheppard (2010), Realising the future: forecasting with high-frequency-based volatility (heavy) models, *Journal of Applied Econometrics* **25**, 197–231.

Sinclair, Euan (2013), *Volatility trading*, John Wiley & Sons.