

Cambridge Working Papers in Economics

Cambridge Working Papers in Economics: 1950

SCORE-DRIVEN MODELS FOR REALIZED VOLATILITY

Andrew Harvey

Dario Palumbo

30 May 2019

This paper sets up a statistical framework for modeling realised volatility (RV) using a Dynamic Conditional Score (DCS) model. It first shows how a preliminary analysis of RV, based on fitting a linear Gaussian model to its logarithm, confirms the presence of long memory effects and suggests a two component dynamic specification. It also indicates a weekly pattern in the data and an analysis of squared residuals suggests the presence of heteroscedasticity. Furthermore working with a Gaussian model in logarithms facilitates a comparison with the popular Heterogeneous Autoregression (HAR), which is a simple way of accounting for long memory in RV. Fitting the two component specification with leverage and a day of the week effect is then carried out directly on RV with a Generalised Beta of the second kind (GB2) conditional distribution. Estimating $\log RV$ with an Exponential Generalised Beta of the second kind (EGB2) distribution gives the same result. The EGB2 model is then fitted with heteroscedasticity and its forecasting performance compared with that of HAR. There is a small gain from using the DCS model. However, its main attraction is that it gives a comprehensive description of the properties of the data and yields multi-step forecasts of the conditional distribution of RV.

Score-Driven Models for Realized Volatility

Andrew Harvey and Dario Palumbo
Faculty of Economics, Cambridge University

May 30, 2019

Abstract

This paper sets up a statistical framework for modeling realised volatility (RV) using a Dynamic Conditional Score (DCS) model. It first shows how a preliminary analysis of RV, based on fitting a linear Gaussian model to its logarithm, confirms the presence of long memory effects and suggests a two component dynamic specification. It also indicates a weekly pattern in the data and an analysis of squared residuals suggests the presence of heteroscedasticity. Furthermore working with a Gaussian model in logarithms facilitates a comparison with the popular Heterogeneous Autoregression (HAR), which is a simple way of accounting for long memory in RV. Fitting the two component specification with leverage and a day of the week effect is then carried out directly on RV with a Generalised Beta of the second kind (GB2) conditional distribution. Estimating logRV with an Exponential Generalised Beta of the second kind (EGB2) distribution gives the same result. The EGB2 model is then fitted with heteroscedasticity and its forecasting performance compared with that of HAR. There is a small gain from using the DCS model. However, its main attraction is that it gives a comprehensive description of the properties of the data and yields multi-step forecasts of the conditional distribution of RV.

Keywords: EGARCH; GB2 distribution; HAR model; heteroscedasticity; long memory; weekly volatility pattern.

1 Introduction

This paper sets up a general model for realized variance (RV) based on a distribution that is coherent for non-negative variables. The dynamics de-

pend on the score of the conditional distribution. Score-driven models were developed in Harvey (2013) and Creal et al (2013) where they were called DCS and GAS models respectively. The contribution of this paper is to show how the score-driven approach provides an integrated framework for volatility modeling and how its application to RV yields new insights, as well as giving forecasts which are as good as, if not better, than those obtained from existing techniques. As well as giving a comprehensive description of the properties of the data, it is able to provide multi-step forecasts of the conditional distribution of RV.

A preliminary analysis of RV can be carried out by fitting a linear Gaussian model to its logarithm. (This has an additional attraction of yielding smoothed estimates of the underlying level of the series.) In our example, this analysis confirms the two component specification that has often been found in GARCH and EGARCH models and at the same time reveals a weekly pattern in RV. It also clarifies the relationship with the popular Heterogeneous Autoregression (HAR) model, devised by Corsi (2009) and Müller et al (1997), which is a simple approximation to the high-order autoregression implied by long memory. Fitting the two component specification is then carried out directly on RV with a Generalised Beta distribution of the second kind (GB2) conditional distribution. The Burr, Pareto and log-logistic are all special cases of GB2 and the F distribution is closely related. These distributions have fat tails but the DCS model handles extreme values robustly. Estimating a GB2 for RV is equivalent to estimating a model for its logarithm with an Exponential GB2 (EGB2) distribution. The EGB2 provides further insights into the overall picture because it has the normal distribution as a limiting case. Furthermore it is easily extended to allow for heteroscedasticity.

Finally we comment on the relevance of our findings for the multivariate score-driven models for RV covariance matrices proposed by Opshoor et al (2016).

2 Statistical framework

This section sets out a coherent statistical framework for modeling RV that respects the non-negativity of the observations and the tendency of variance to increase with the level; see Corsi et al (2008) and Taylor (2005, p 335-7).

When volatility is estimated from RV, it may be calibrated for returns by re-estimating the constant from the standardized returns assuming a distri-

bution, such as generalized- t . The resulting estimates of scale, or standard deviation, are then used to give forecasts based on the conditional distribution of returns; compare Brownlees and Gallo (2006) or Harvey (2013, p 180-1).

2.1 Dynamic location/scale model

A dynamic location/scale model for a non-negative variable is

$$y_t = \epsilon_t \mu_{t|t-1}, \quad t = 1, \dots, T, \quad (1)$$

where the random variable ϵ_t has unit mean. In the MEM model of Engle and Gallo (2006), the conditional mean, $\mu_{t|t-1}$, follows a GARCH-type process and has to be constrained to be positive. Estimation is by maximum likelihood (ML), based on the conditional distribution, $f(y_t | Y_{t-1})$.

An equivalent model can be formulated in terms of scale, $\alpha_{t|t-1}$, in which case $y_t = \epsilon_t \alpha_{t|t-1}$, where $E(\epsilon_t)$, unlike $E(\epsilon_t)$, is not equal to one, unless the distribution is exponential. However, $E(\epsilon_t)$ is easily found. With an exponential link function for the scale,

$$y_t = \epsilon_t \exp(\lambda_{t|t-1}), \quad t = 1, \dots, T, \quad (2)$$

where the logarithm of scale, $\lambda_{t|t-1}$, is unconstrained. The structure is the same as that of EGARCH, as in Harvey (2013, chs 4 and 5) and Harvey and Lange (2017), and much of the theory is the same: hence the name Beta-GB2-EGARCH. First-order dynamics for $\lambda_{t|t-1}$ take the form

$$\lambda_{t+1|t} = \omega(1 - \phi) + \phi \lambda_{t|t-1} + \kappa u_t, \quad (3)$$

where u_t is the score of the conditional distribution of y_t . As with EGARCH, $|\phi| < 1$ ensures stationarity and there is no need to restrict parameters to be non-negative. When the distribution is fat-tailed, the score is bounded. This makes it possible to show the model is invertibility subject to restrictions on κ .

Taking logarithms in (2) gives

$$\ln y_t = x_t = \lambda_{t|t-1} + \ln \epsilon_t \quad (4)$$

so $\lambda_{t|t-1}$ now becomes location. A distribution for ϵ_t implies a distribution for $\ln \epsilon_t$, so the model can, in principle, be estimated in levels or logs. Comparison of the likelihoods requires a Jacobian term to account for the transformation.

2.2 GB2 conditional distribution

The usual form of the GB2 density, as given in Kleiber and Kotz (2003, p.187) and Harvey (2013, ch 5), is

$$f(y) = \frac{v(y/\alpha)^{v\xi-1}}{\alpha B(\xi, \varsigma) [(y/\alpha)^v + 1]^{\xi+\varsigma}}, \quad \alpha, v, \xi, \varsigma > 0, \quad (5)$$

where α is the scale parameter, v, ξ and ς are shape parameters and $B(\xi, \varsigma)$ is the beta function. The standardized GB2, as it appears in (2), is

$$f(\varepsilon_t) = \frac{v\varepsilon_t^{v\xi-1}}{B(\xi, \varsigma) [\varepsilon_t^v + 1]^{\xi+\varsigma}}, \quad v, \xi, \varsigma > 0.$$

The score wrt the logarithm of scale is

$$\frac{\partial \ln f_t}{\partial \lambda_{t-1}} = u_t = v(\xi + \varsigma)b_t(\xi, \varsigma) - v\xi, \quad (6)$$

where

$$b_t(\xi, \varsigma) = \frac{(y_t e^{-\lambda_{t-1}})^v}{(y_t e^{-\lambda_{t-1}})^v + 1}, \quad t = 1, \dots, T, \quad (7)$$

is distributed as $beta(\xi, \varsigma)$; see Harvey (2013, ch 5). As $y \rightarrow \infty$, the score approaches an upper bound of $\eta = v\varsigma$, whereas the lower bound of $-v\xi$ is obtained when $y = 0$.

Remark 1 *The GB2 distribution can be reparameterized so that the (upper) tail index, η , replaces ς , that is we define $\eta = v\varsigma$. The generalized gamma (GG) is then obtained as a limiting case as $\eta \rightarrow \infty$ by replacing α by $\varphi\eta^{1/v}$. The Weibull is a special case of GG in which $\xi = 1$ (and a special case of Burr as $\eta \rightarrow \infty$) while setting $v = 1$ gives the gamma distribution. The chi-square is a special case of standardized gamma and a limiting case of the F distribution. The score for GG is unbounded and for gamma it is $y_t e^{-\lambda_{t-1}} - \xi$ so the response is linear.*

The asymptotic distribution of the ML estimator is given in Harvey (2013, ch 5). When, as is usually the case, κ is positive, a sufficient condition for the model to be invertible is that

$$\kappa < \frac{4(1 + \phi)}{v^2(\xi + \varsigma)} \quad (8)$$

for all values in the parameter space; see Appendix.

Taking the logarithm of a GB2 gives the exponential GB2 (EGB2) distribution. The EGB2 is symmetric when $\xi = \varsigma$. It includes normal, when $\xi = \varsigma \rightarrow \infty$, Laplace, when $\xi = \varsigma \rightarrow 0$, and logistic, when $\xi = \varsigma = 1$; see McDonald and Xu (1995) and Caivano and Harvey (2014).

If y is distributed as $GB2(\alpha, v, \xi, \varsigma)$ and $x = \ln y$, the PDF of the EGB2 variate x is

$$f(x; \mu, v, \xi, \varsigma) = \frac{v \exp\{\xi(x - \lambda)v\}}{B(\xi, \varsigma)(1 + \exp\{(x - \lambda)v\})^{\xi + \varsigma}}. \quad (9)$$

The logarithm of scale in GB2, that is λ , becomes a location parameter in EGB2. Furthermore v is now a scale parameter, but ξ and ς are still shape parameters and they determine skewness and kurtosis. The distribution has light (exponential) tails. All moments exist with the mean equal to $\lambda + v^{-1}[\psi(\xi) - \psi(\varsigma)]$ and the standard deviation given by $\sigma = h/v$, where $h^2 = \psi'(\xi) + \psi'(\varsigma)$, with ψ and ψ' denoting the digamma and trigamma functions respectively. The distribution is positively (negatively) skewed when $\xi > \varsigma$ ($\xi < \varsigma$) and its kurtosis decreases as ξ and ς increase; excess kurtosis does not exceed six.

Although v is a scale parameter in (9), it is the inverse of what would be considered a more conventional measure of scale. Scale is better defined as the standard deviation

$$\sigma = h/v, \quad \text{where} \quad h = \sqrt{\psi'(\xi) + \psi'(\varsigma)}, \quad (10)$$

leading to

$$u_t = \sigma^2 \frac{\partial \ln f_t}{\partial \lambda} = \sigma h [(\xi + \varsigma)b_t(\xi, \varsigma) - \xi], \quad (11)$$

where

$$b_t(\xi, \varsigma) = \frac{e^{(x_t - \lambda_{t,t-1})h/\sigma}}{e^{(x_t - \lambda_{t,t-1})h/\sigma} + 1}, \quad t = 1, \dots, T. \quad (12)$$

Note that $b_t(\xi, \varsigma)$ is as in the GB2, that is (7), but with v replaced by h/σ and $y_t \exp(-\lambda_{t,t-1})$ replaced by $\exp(x_t - \lambda_{t,t-1})$. Hence the score is bounded as $|x| \rightarrow \infty$.

The inclusion of σ^2 in (11) means that letting $\varsigma = \xi \rightarrow \infty$ yields $u_t \rightarrow x_t - \lambda_{t,t-1}$; see Caivano and Harvey (2014). The DCS model is then the innovations form of the Kalman filter for a Gaussian unobserved components

model, which in the first-order case, (3), is

$$x_t = \mu_t + \varepsilon_t, \quad \xi_t \sim NID(0, \sigma_\xi^2), \quad t = 1, \dots, T,$$

$$\mu_{t+1} = \phi\mu_t + \eta_t, \quad \eta_t \sim NID(0, \sigma_\eta^2),$$

where $NID(0, \sigma^2)$ denotes normally and independently distributed with mean zero and variance σ^2 and $E(\xi_t \eta_s) = 0$ for all t, s .

Although the invertibility constraint, (8), is often not restrictive in practice, it clearly creates difficulties as ξ and ς become bigger. However, when u_t is defined as in (11), the invertibility condition is

$$|\phi - \kappa h^2(\xi + \varsigma)/4| < 1.$$

For a (log) logistic distribution, when $\xi = \varsigma = 1$, the invertibility condition is $|\phi - 1.645\kappa| < 1$. As $\varsigma = \xi \rightarrow \infty$, it is shown¹ in the appendix that the standard invertibility condition for a Gaussian model is obtained, that is $|\phi - \kappa| < 1$; see Harvey (2013, p 67). When $\kappa > 0$

$$\kappa < \frac{4(1 + \phi)}{h^2(\xi + \varsigma)} \quad (13)$$

which yields $\kappa < 1 + \phi$ for the Gaussian model.

2.3 Dynamic components

Instead of capturing long memory by a fractionally integrated process, two components may be used. For RV (and range), the leverage term is governed by $sgn(-r_t)$, where r_t denotes mean-adjusted returns; see Harvey (2013, pp 178-9). Thus

$$\begin{aligned} \lambda_{t|t-1} &= \omega + \lambda_{1,t|t-1} + \lambda_{2,t|t-1}, \\ \lambda_{i,t+1|t} &= \phi_i \lambda_{i,t|t-1} + \kappa_i u_{i,t} + \kappa_i^* sgn(-r_t)(u_t + 1), \quad i = 1, 2. \end{aligned} \quad (14)$$

The score curve from the RV model is rotated and, when the distribution of (mean-adjusted) returns is symmetric, the expectation of the composite variable driving the dynamic equation is zero. Identifiability requires $\phi_1 \neq \phi_2$, together with $\kappa_1 \neq 0$ or $\kappa_1^* \neq 0$ and $\kappa_2 \neq 0$ or $\kappa_2^* \neq 0$. An attraction of the

¹If $\varsigma \rightarrow \infty$ while ξ remains fixed at a finite value, establishing whether the model is invertible appears not to be straightforward.

two-component DCS model is that the leverage effect can differ in the long and short run.

Other components may be added. Here we employ a seasonal component $\gamma_{t|t-1}$ so

$$\lambda_{t|t-1} = \omega + \lambda_{1,t|t-1} + \lambda_{2,t|t-1} + \gamma_{t|t-1}, \quad (15)$$

where $\gamma_{t|t-1}$ is as in Harvey (2013, p 79-80), that is

$$\begin{aligned} \gamma_{t|t-1} &= \mathbf{z}'_t \boldsymbol{\alpha}_{t|t-1} \\ \boldsymbol{\alpha}_{t+1|t} &= \boldsymbol{\alpha}_{t|t-1} + \boldsymbol{\kappa}_t u_t, \end{aligned} \quad (16)$$

where \mathbf{z}_t picks out the current season. In season j we set $\kappa_j = \kappa$, where κ is a non-negative unknown parameter, whereas $\kappa_i = -\kappa/(s-1)$ for $i \neq j$. The amounts by which the seasonal effects change therefore sum to zero. The filter treating the elements of $\boldsymbol{\alpha}_{1|0}$ as parameters. (There are $s-1$ free parameters because they sum to zero.) In the present application the ‘seasons’ are days of the (working) week, so $s = 5$. The exponential link function prevents the scale from becoming negative.

A random walk level can be extended so as to mirror an integrated random walk (IRW) in a UC model, that is

$$\begin{aligned} \lambda_{1,t+1|t} &= \lambda_{1,t|t-1} + \lambda_{1,t|t-1}^\beta \\ \lambda_{1,t+1|t}^\beta &= \lambda_{1,t|t-1}^\beta + \kappa u_t. \end{aligned}$$

In a Gaussian UC model the IRW yields a smooth trend when the Kalman filter and smoother (KFS) is applied.

2.4 Long memory, Fractional integration and HAR

The standard HAR model, as in Corsi (2009), takes the form

$$\widehat{y}_{h,t+h} = \mu + \beta_d y_t + \beta_w \bar{y}_{w,t} + \beta_m \bar{y}_{m,t} + \zeta_{h,t+h}, \quad (17)$$

where y_t is RV, $\bar{y}_{w,t} = (\sum_{i=1}^5 y_{t-i}/5)$, $\bar{y}_{m,t} = (\sum_{i=1}^{22} y_{t-i}/22)$ and $\widehat{y}_{h,t+h}$ is the h -day cumulative average for $h = 1, 2, \dots$. Usually $h = 1$ in which case $\widehat{y}_{h,t+h} = y_{t+1}$. The disturbance term is $\zeta_{h,t+h}$. The coefficients are estimated by regression. The weekly and monthly averages approximate the kind of lag structure found with long memory and are seen a simple alternative to a fractionally integrated model.

Remark 2 *In the reparametrization of Patton and Sheppard (2015) the model specification has the second term consisting of only the realized volatilities between lags 2 and 5, and the third term consisting of only the realized volatilities between lag 6 and 22. If the Patton and Sheppard parameters are denoted with a star, then $\beta_d^* = \beta_d + (1/5)\beta_w + (1/22)\beta_m$, $\beta_w^* = (4/5)\beta_w + (4/22)\beta_m$ and $\beta_m^* = (17/22)\beta_m$. The sum of the coefficients is the same.*

As set up in (17), the HAR is a regression model for a variable that is intrinsically positive, with a standard deviation that, in a GB2 framework, is proportional to the location. Working in logarithms, particularly when $h = 1$, this stabilizes the variance²; see also Corsi et al (2008) and Corsi and Reno (2012). In this case the y'_t s in the summations of (17) are replaced by their logarithms. The fact that the model for logarithms, (4), is additive makes aggregation more appealing because the (conditional) variance is constant.

3 Data Description and preliminary analysis

The simplest estimator of RV in day t is

$$RV_t = \sum_{i=1}^m r_{t,i}^2, \quad t = 1, \dots, T,$$

where $r_{t,i}$ is the return in the i -th subperiod in the day; see, for example, Andersen et al (2003). The current study uses RV based on high-frequency index transaction prices from FTSE 100, DAX, S&P 500 and NASDAQ. The dataset is the one used by Patton and Sheppard (2015) and is constructed from five-minute intraday returns in the Oxford Man Institute library³. The observations are for weekdays, with public holidays assumed to be the same as the previous day. The starting date, 3rd January 2000, corresponds to the first non-weekend/holiday date in the U.S. The end is May 23, 2017. The sample size is therefore $T = 4536$.

When $\ln \varepsilon_t$ in (4) is not too far from normality⁴, fitting a linear state space model may be a convenient way of finding a suitable specification

²Weighted least squares (WLS), as used by Patton and Sheppard (2015), is better than OLS, but it still does not acknowledge that the conditional distribution cannot possibly be normal.

³We use the variable 'rv5' from <https://realized.oxford-man.ox.ac.uk>.

⁴When the GB2 distribution reduces to GG, that is $\varsigma = \infty$, but ξ is finite, the EGB2

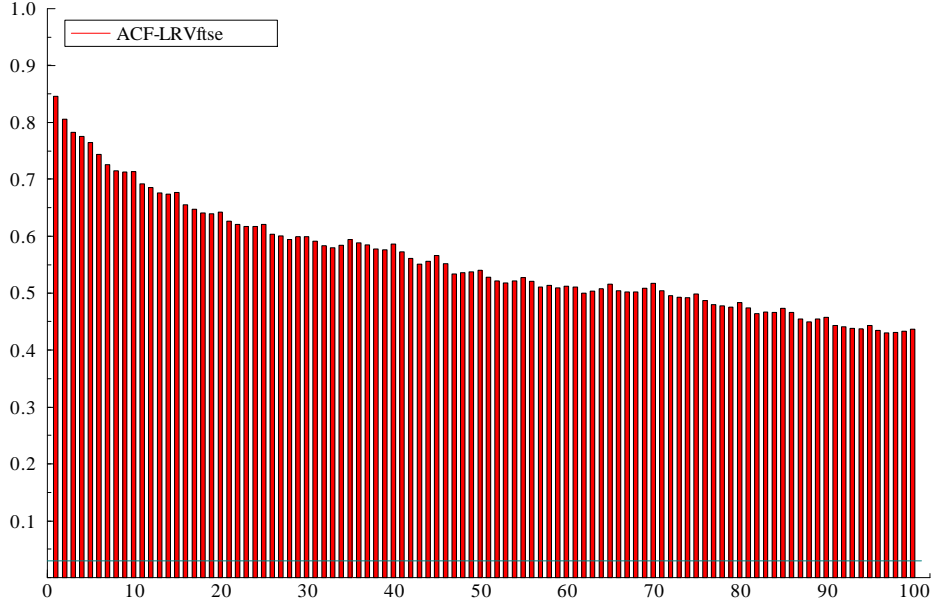


Figure 1: Correlogram of log of FTSE RV

for a dynamic model based on the conditional EGB2 distribution. This is easily done using the STAMP package of Koopman et al (2008). Smoothed estimates of volatility, that is estimates constructed from two-sided filters, are readily available. Taking logarithms also facilitates a comparison with HAR.

The preliminary analysis is carried out with FTSE RV, but the analysis of the other indices leads to similar conclusions. The first 23 observations were dropped so as to enable a comparison to be made with HAR. The long memory features are apparent in the ACF of its logarithm shown in Figure 1. Hence the decision to fit two components. In any case, fitting only one AR(1) component gives a very high portmanteau statistic, $Q(67)$, at 410 and the first-order residual autocorrelation is also high at $r(1) = 0.078$. When two AR(1) components are included, $r(1)$ reduces to 0.033, but Q is still very high and correlations at lags of ten, fifteen and twenty (but not five) are apparent

is negatively skewed and the excess kurtosis goes to six as $\xi \rightarrow 0$. Taking logarithms may be less appealing in such cases. A better option for QML is to model y_t with a gamma distribution and a linear response function.

in the residual ACF shown in Figure 2. Furthermore the associated spectrum shows a clear peak at $2/5$ and at its harmonic $4/5$. Thus there is strong evidence for a day of the week effect. The model is therefore augmented by including a weekly seasonal component as in (16). The seasonals appear to be fixed and the seasonal test statistic, which is asymptotically χ_5^2 under the null of no seasonal effect, is 203.2. The day of the week coefficients are -0.132, -0.052, -0.002, 0.088 and 0.098. Thus Thursday and Friday exhibit much higher volatility than the other days, with Monday being particularly low⁵. The short-term component has a coefficient of 0.75 which when combined with a long-run coefficient close to unity implies a long memory pattern for the ACF. With the seasonal included, the implied weights in the filter are as in Figure 3.

There is no evidence of first-order serial correlation as $r(1) = 0.007$ and although the 117.5 value for $Q(70)$ is statistically significant at conventional levels its value is not excessive given the large sample size. The only real failure of the model is in the distribution of the residuals: there is positive skewness and excess kurtosis of about 2.6.

The HAR model does rather well given its simplicity but $Q(67)$ is high⁶ at 230.9 and the prediction error variance is 0.240 compared with 0.230 for the unobserved component model. The sum of the coefficients is 0.96 and in the Patton-Sheppard form, the estimates of β_d^* , β_w^* and β_m^* are 0.46, 0.38 and 0.12 respectively; compare the weights in Figure 3.

The Gaussian model allows smoothing. If the first component is modelled as an integrated random walk (IRW), that is

$$\lambda_{1,t+1} = \lambda_{1,t} + \lambda_{1,t}^\beta, \quad \lambda_{1,t+1}^\beta = \lambda_{1,t}^\beta + \zeta_t,$$

the smoothed estimates are as shown in Figure 4.

4 Location/scale DCS model

Estimating a model with a conditional GB2 distribution with the parameters ν , ζ and ξ can be problematic. We therefore focus on the Burr distribution,

⁵The realised measures in the Oxford library are explicitly constructed so as to ignore any extreme movements in the first few minutes of the trading day, which may be caused by large overnight volatility.

⁶This is only partly explained by the failure to take account of the day of the week effect.

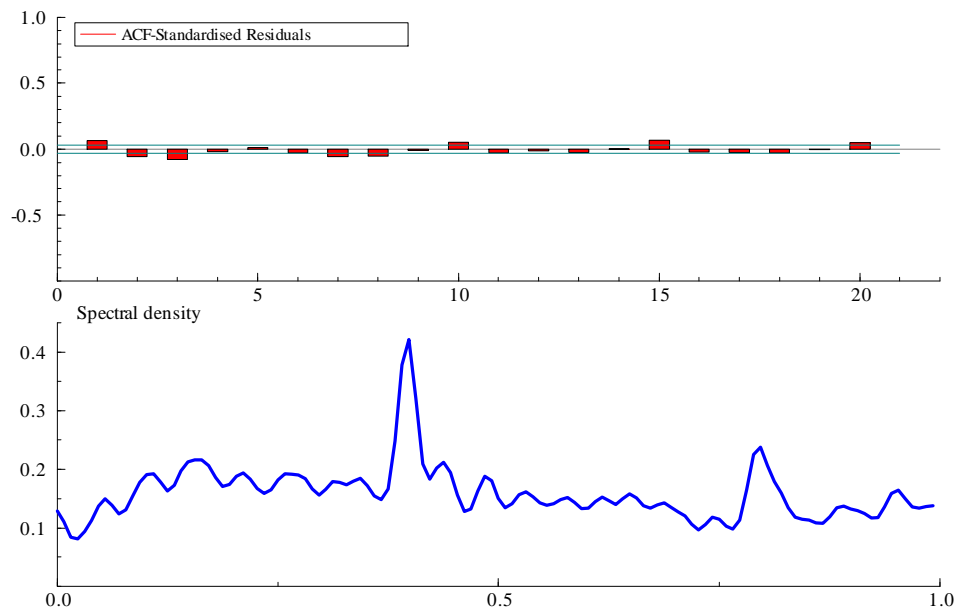


Figure 2: ACF and spectrum of residuals from fitting a RW and AR1 to $\ln RV$

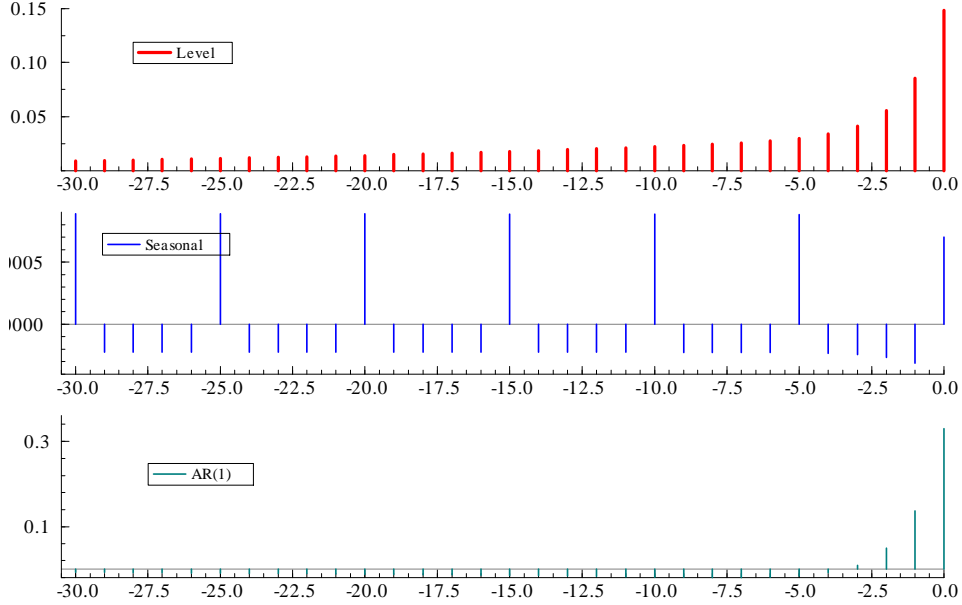


Figure 3: Filter weights for linear Gaussian model for $\ln FTSErv$

in which $\xi = 1$, and what we call the *balanced* GB2 where $\xi = \varsigma$. The log-logistic distribution is a special case of both Burr and balanced GB2 in which $\varsigma = \xi = 1$, whereas the generalized Pareto distribution is a special case of Burr with $\nu = 1$. The logarithm of the balanced GB2 has a symmetric EGB2 distribution which, as noted earlier, goes to a normal distribution as $\xi, \varsigma \rightarrow \infty$ meaning that the GB2 becomes lognormal. Finally the F -distribution with (ν_1, ν_2) degrees of freedom, is a special case of balanced GB2 in which $\nu = 1$ and $\xi = \varsigma = \nu_1/2 = \nu_2/2$, but more generally, $(\nu_1/\nu_2)F$ is $GB2(\nu_1/2, \nu_2/2)$.

Based on the specification suggested by the preliminary analysis, we fitted these distributions with two autoregressive components. The Weibull is included for comparison. It is not a fat-tailed distribution, but like the lognormal, it can be obtained from GB2 as a limiting case. Table 1 shows the FTSE results, where, according to the AIC and BIC information criteria, Burr is best, whereas Table 2 for S & P 500 shows balanced GB2 to be better. The results for DAX and NASDAQ, which are available on request, show DAX leaning towards Burr and NASDAQ preferring balanced GB2. The estimate of ς for FTSE is around 0.8 which implies positive skewness

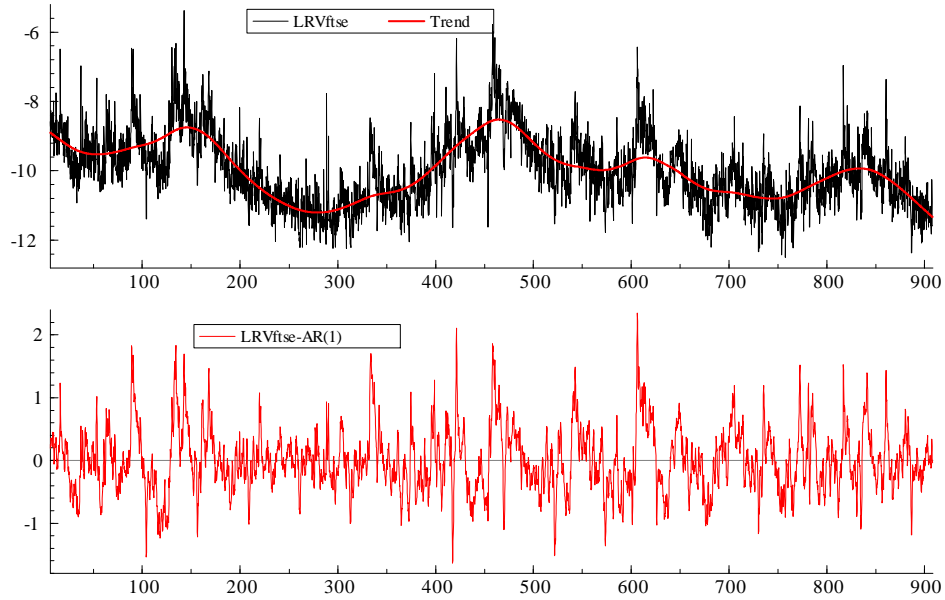
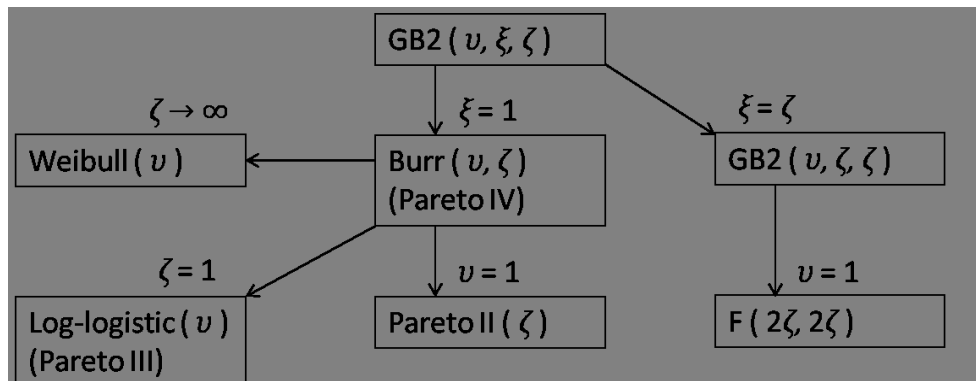


Figure 4: Smoothed estimates when long-run component is an IRW and short-run component is an AR(1).



and excess kurtosis for the EGB2. The implied EGB2 values for excess kurtosis and skewness are 2.76 and 0.54, which are close to the values for the residuals reported in the previous section.

The values of the tail indices are similar for all four markets, but tend to be lower for the Burr distribution. For example with FTSE the upper (lower) indices for Burr and balanced are 3.33 (3.89) and 4.48 (4.48) for balanced. The F-distribution does well, particularly when the balanced GB2 is best overall. The estimated ν in balanced GB2 is typically greater than one and the estimates of ξ are ζ correspondingly smaller than those reported for F ; as a result the tail indices are of a similar magnitude. In the log-logistic, where ξ and ζ are constrained to be unity, it is ν that adjusts to give comparable tail indices. Although the fit of the log-logistic is good, the hypothesis that $\nu = 1$ is rejected for all markets by Wald tests on the balanced GB2 estimates. The fit with distributions from the generalized gamma class, such as Weibull, is poor.

The Kolmogorov-Smirnov test is informative for comparing the distributional fit. When a formal test is carried out at a nominal 5% significance level, it tends to reject most of the distributions, an exception being the balanced GB2 for S&P. However, such rejections may not be surprising given the large sample size.

The low tail indices found in the best GB2 models indicate that the lognormal distribution gives an inferior fit and this is confirmed by the goodness of fit statistics; for FTSE the log-likelihood is 42,067 which is well below the log-likelihood of 42,136 obtained for the balanced GB2 where $\xi = \zeta = 1.76$ rather than infinity. For S&P, the lognormal log-likelihood is 40,312, again well below balanced GB2. The fit for Pareto is much worse than lognormal and like Weibull it is not considered further.

The seasonality does not show up clearly in residuals from fitting a non-seasonal location/scale model to levels, that is $y_t / \exp(\lambda_{t|t-1})$. However, the ACFs of scores from the models reported in Tables 1 and 2 yield results similar to those for the ACF of the Gaussian UC model in Figure 2. This is because the scores for GB2 and EGB2 are the same and the Gaussian model is not too far away from a balanced EGB2 with moderate size tail indices.

The results in Table 3 show the parameter estimates for each index with the best-fitting distribution. The dynamic equation now includes a seasonal component and the two autoregressive components incorporate leverage effects. Taking logarithms and fitting a corresponding EGB2 gives the same estimates and score residuals. The seasonal pattern is similar to that re-

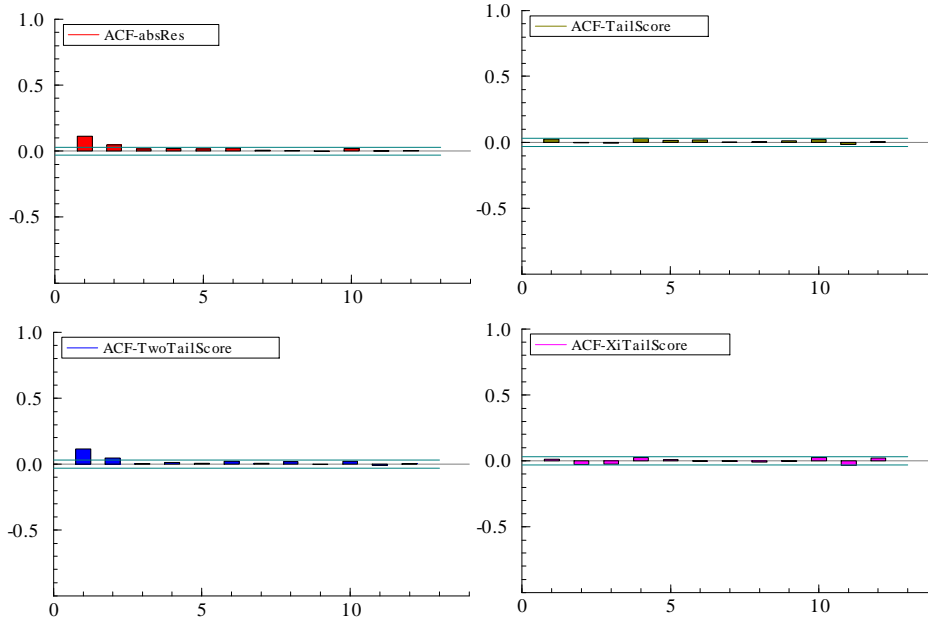


Figure 5: FTSE: ACFs of scores for absolute values of residuals from preferred $\ln RV$ model (top left hand) plus ACFs of scores for tail parameters.

vealed by the preliminary modelling and the leverage effect is stronger in the short-term component, a finding that is not unusual for returns or range; see Harvey and Lange (2018) and Harvey (2013, pp 176-81).

Table 4 shows the day of the week effect. In all cases there is very little movement with the estimate of κ_s being zero in three cases. For the US indices the most volatile days are Wednesday and Thursday, whereas for the European indices it is Thursday and Friday.

5 Heteroscedasticity and a Changing Tail Index

Although the residuals from the preferred Gaussian model have relatively little serial correlation, this is not true for their squares or absolute values: Figure 5 clearly indicates heteroscedasticity.

For fixed ξ and ς , the scale of the EGB2 is $1/v$. An exponential link

function, that is $v = \exp(-\bar{v})$, yields

$$\partial \ln f_t / \partial \bar{v} = (\xi + \varsigma) \varepsilon_t b_t - \xi \varepsilon_t - 1, \quad (18)$$

where $\varepsilon_t = (x_t - \lambda_{t|t-1}) / \exp(\bar{v})$ and

$$b_t = \frac{\exp \varepsilon_t}{1 + \exp \varepsilon_t}.$$

The score is symmetric but unbounded⁷; see Figure 2 in Caivano and Harvey (2014).

Parameterizing in terms of $\ln \sigma$, gives

$$\partial \ln f_t / \partial \ln \sigma = u_t = h(\xi + \varsigma) \varepsilon_t b_t - h \xi \varepsilon_t - 1, \quad (19)$$

where ε_t is $(x_t - \lambda_{t|t-1}) / \sigma = \varepsilon_t / h$. Caivano and Harvey (2014, p 566) show that in the limit as $\xi = \varsigma \rightarrow \infty$, $u_t = (x_t - \lambda_{t|t-1})^2 / \sigma^2 - 1$, which is the score for a Gaussian EGARCH model. At the other extreme, when $\xi = \varsigma = 0$, the score is $\sqrt{2} |x_t - \lambda_{t|t-1}| / \sigma - 1$.

How should heteroscedasticity in the logarithm of a variable that is already subject to changing variance be interpreted? As was noted earlier, the (upper) tail index for a GB2 is $\eta = v\varsigma$. Thus for a fixed value of ς , a dynamic v implies that η is dynamic. It also might imply that the lower tail index, $v\xi$, is dynamic. We could let the tail indices be dynamic and keep v fixed. Setting $\xi = \varsigma$ and letting $\varsigma = \exp(-\bar{\varsigma})$ so $\bar{\varsigma} = -\ln \varsigma$, gives

$$\partial \ln f_t / \partial \bar{\varsigma} = 2\varsigma\psi(\varsigma) - 2\varsigma\psi(2\varsigma) - \varsigma\varepsilon_t - 2\varsigma \ln(1 - b_t). \quad (20)$$

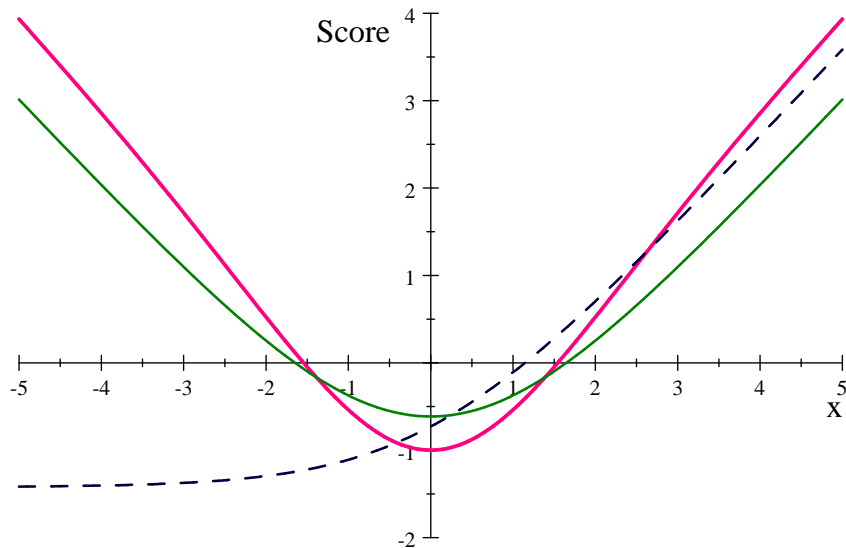
This score is symmetric. Figure ?? compares $\partial \ln f_t / \partial \bar{\varsigma}$ and $\partial \ln f_t / \partial \bar{v}$ for $\varsigma = \xi = 1$. The response is similar in this case. If $\xi = \varsigma$ is not assumed,

$$\partial \ln f_t / \partial \bar{\varsigma} = \varsigma\psi(\varsigma) - \varsigma\psi(\varsigma + \xi) - \varsigma \ln(1 - b_t) \quad (21)$$

and, as can be seen from the graph, the (dashed) curve is asymmetric with large positive (negative) observations lowering (raising) the value of ς . Figure 5 shows that the ACF of this score, like that of the lower tail, gives little indication of any serial correlation⁸.

⁷Gen-t parameterized as in Harvey and Lange (2017) does have a bounded score for this parameter but tail index is fixed.

⁸This is somewhat surprising (??) as the sum of the scores equals the score for both.



Estimation of the GB2/EGB2 with a two component location/scale, seasonal dummies and dynamic v , captured by a first-order equation for $\bar{v}_{t|t-1}$, gives the results shown in Table 5. In all cases apart from DAX, the dynamic effects in $v_{t|t-1}$ are short-lived with the AR coefficient only slightly bigger than 0.5. Nevertheless, the information criteria in Table 6 indicate that their inclusion improves the fit. Furthermore there is a big reduction in the value of the portmanteau statistic for v , but not by enough to completely eliminate the residual heteroscedasticity.

Remark 3 *Corsi et al (2008) introduce heteroscedasticity into a model for $\log RV$ using GARCH combined with a normal inverse Gaussian distribution (NIG). The score-driven approach is more easily developed with EGB2 distribution.*

6 Forecasting

In order to compare the forecasting performance of the DCS models with the HAR model, we identified two time periods of five months, one with large market fluctuations, labelled High Volatility Period, and one of relatively quiet market conditions, the Low Volatility Period. The High Volatility Period, between the 01/10/2008 and the 31/03/2009, covers the 2008 financial

crisis and includes the announcement of the bankruptcy of Lehman Brothers. The Low Volatility period is between the 01/02/2005 and the 29/07/2005. Models for each period were estimated using all the observations up to the beginning of the the period. Then one step ahead forecasts were computed for all the forecasting period.

The models compared are the location/scale DCS Seasonal EGARCH model, with and without heteroscedasticity (dynamic $v_{t|t-1}$), the HAR model in levels, as defined⁹ by Patton and Sheppard (2015), and the HAR model in logarithms, as defined by Corsi (2009). The one step ahead forecasts for the mean of the GB2 model are obtained by making a modification to the forecast of the scale; see Harvey (2013, p163). In the general case with heteroscedasticity

$$\tilde{y}_{t+1|t} = \frac{\Gamma(\xi + 1/v_{T+1|T}) \Gamma(\zeta + 1/v_{T+1|T})}{\Gamma(\xi) \Gamma(\zeta)} e^{\lambda_{t+1|t}}, \quad t = T + 1, \dots, T + T^*,$$

where T^* denoted the number of observations in the forecasting period. When the HAR model is estimated in logarithms. In the latter case, the standard lognormal correction factor is used to give $\hat{y}_{T+1|T} = \exp(\hat{x}_{T+1|T} + \hat{\sigma}^2/2)$.

In Table 7 the forecasting accuracy is measured in terms of Root Mean Squared Forecast Errors (RMSFE) and Mean Absolute Forecast Errors (MAFE) which are presented as percentages with respect to the standard DCS Seasonal EGARCH model. The significance of differences in forecasting accuracy is assessed from the p-values of the Diebold and Mariano (1995) test for equal expected loss of forecasts based on the proposed model versus each competitor. Finally, we have included the results for the predictive likelihood; see Mitchell and Wallis (2011).

The results in Table 7 show that the DCS model easily beats the HAR models according to predictive likelihood. This is what matters if we are concerned with the full distribution; see, for example, Corsi et al (2008, p 73). Comparisons based on RMFSE and MAFE are less clear cut: although DCS beats HAR in most cases, the differences are rarely statistically significant. In periods of Low Volatility the DCS forecasts better than the HAR on the European indices whereas for High Volatility regimes it forecasts better on the American indexes. In summary although HAR is generally difficult to beat when the criterion is RMSFE or MAFE, the DCS model is certainly not

⁹They allow for heteroscedasticity by Weighted Least Squares, with weights given by the residuals from an initial OLS estimation.

worse and may be slightly better. What is clear, however, is that the HAR performs better when formulated in logarithms.

In all the cases the predictive likelihood is the highest for the DCS models, always slightly in favour of the heteroscedastic EGB2 (DCS- $v_{t|t-1}$). As might be expected, HAR is much better in logarithms. This is because in levels the assumption of normality, even after removing the heteroscedasticity by weighted least squares, is still not a good one for observations which are strictly positive and skewed to the right (upwards) with a heavy upper tail. As was observed in Section 3, assuming normality for $\ln RV$, thereby implicitly assuming that RV is lognormal, is not unreasonable as an approximation.

In order to assess the relevance of accurately describing the predictive density, we looked at the upper quantiles. This is analogous to computing Value at Risk (VaR) for a return. For the DAX index, the Burr distribution gave the best fit and in this case the quantile function takes a particularly simple form, with the probability that RV is in the upper $100\alpha\%$ given by $\xi_\alpha^u(\tilde{y}_{T+1|T}) = 1 - \exp(\lambda_{T+1|T}) (\alpha^{-1/\zeta} - 1)^{1/v_{T+1|T}}$. Comparing these quantile for $\alpha = 0.10$ and 0.25 with the corresponding quantiles for HAR (in logs), it seems that the HAR quantiles are consistently higher. This suggests an unreasonable hedge in a portfolio to protect against volatility risk and a tendency to overestimate a volatility spillover effect in a systemic risk assessment. These are implications that go beyond the simple point forecast which has been extensively highlighted in the literature, and should be considered carefully by risk managers and policy makers.

Table 7 also includes results for the HAR-with-Jumps (HARJ) and Continuous-HAR (CHAR) models proposed by Andersen et al. (2007). These are based on the idea that the total variation of RV can be decomposed into a continuous and discontinuous component, that is a ‘jump’ component. Following Barndorff-Nielsen and Shephard (2004), this decomposition can be implemented through the Bi-Power Variation (BPV), defined as

$$BPV_t = z_t = \frac{\pi}{2} \sum_{i=1}^{M-1} |r_{t_i}| |r_{t_{i-1}}|.$$

The series on BPV for the four indices come from the same Oxford Man Realized Library dataset as was used before. The CHAR model replaces RV by BPV in the right hand side of the levels HAR specification, so that

$$\bar{y}_{h,t+h} = \mu + \beta_d z_t + \beta_w \bar{z}_{w,t} + \beta_m \bar{z}_{m,t} + \zeta_{h,t+h}. \quad (22)$$

The HARJ model adds a jump variable, $J_t = \max\{y_t - z_t, 0\}$, to the standard HAR giving

$$\bar{y}_{h,t+h} = \mu + \beta_d y_t + \beta_w \bar{y}_{w,t} + \beta_m \bar{y}_{m,t} + \beta_J J_t + \zeta_{h,t+h}. \quad (23)$$

The performance of both CHAR and HARJ is disappointing. Perhaps formulations designed for the logarithm of RV will fare better.

7 Conclusion

A location/scale DCS model for RV, formulated with a GB2 distribution in levels, or equivalently an EGB2 distribution in logarithms, is able to capture the main features of the data. It is statistically coherent and can model long memory, through a two component structure, and asymmetric response (leverage) in a straightforward and transparent manner. The asymptotic theory for the ML estimator is as in Harvey (2013, ch 5), but here we show an additional result which indicates that invertibility for models with parameters that tend to arise in practice is unlikely to be a problem.

Working in logarithms is useful in that the normal distribution is a limiting case of EGB2 and preliminary investigation is easily carried out using linear state space models. For RV data on the four indices FTSE, DAX, S&P and Nasdaq, preliminary analysis reveals a day of the week pattern, as well as long memory effects. The more general score-driven models can handle these effects and are able to introduce leverage into both short and long run components. The best fitting distributions are found to be the Burr and the balanced GB2, where the two shape parameters are set equal, resulting in a symmetric distribution for $\ln RV$ and a normal distribution as a limiting case. Both Burr and balanced GB2 both have only two parameters in addition to those required to model dynamic location/scale. The F -distribution also gives a relatively good fit to RV , even though it is not the best. However, it does generalize the most easily to the modelling of an $N \times N$ realized volatility covariance matrix as in Opschoor et al (2017).

Heteroscedasticity, or Volatility of Volatility ('Vol of Vol') as described by Corsi (2008), is detected in $\ln RV$ and can be modeled within the framework of an EGB2 distribution. We note that such heteroscedasticity implies dynamic tail indices in the corresponding GB2 distributions.

The forecasting performance of the models was assessed using HAR as a benchmark. HAR performs much better in logarithms than in levels, because in levels it fails to match the empirical distribution of the residuals. The point is worth stressing because some researchers, such as Patton and Sheppard (2015) and Bollerslev, Patton and Quaadvlieg (2016), continue to make a case for modeling in levels. When formulated in logarithms, the HAR model forecasts well, but it rarely beats the the DCS model in terms of RMSE or MAE. Furthermore the DCS model captures the conditional distribution more accurately. The difference is reflected in the upper quantiles, and we argue that this may have implications for risk management; compare the use of score-driven models proposed by Lucas et al. (2016) for forecasting implied default probabilities in Credit Default Swap (CDS) prices.

References

- Andersen, T.G., Bollerslev, T. and Diebold, F.X. (2010), "Parametric and Nonparametric Volatility Measurement," in L.P. Hansen and Y. Ait-Sahalia (eds.), *Handbook of Financial Econometrics*. Amsterdam: North-Holland, 67-138.
- Andersen, T.G., Bollerslev, T., Diebold, F., Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, 71, 579-625.
- Audrino, F. and F. Corsi (2010) Modeling tick-by-tick realized correlations. *Computational Statistics and Data Analysis* 54, 2372-2382.
- Barndorff-Nielsen, O.E. and Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B*, 64, 253–280.
- Blasques, F., Gorgi, P., Koopman, S.J. and O. Wintenberger (2018). Feasible invertibility conditions and maximum likelihood estimation for observation-driven models. *Electronic Journal of Statistics*, 12, 1019-1052.
- Bollerslev, T., Patton, A.J., and Quaadvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 192, 1-18.
- Caivano, M. and A.C. Harvey (2014). Time series models with an EGB2 conditional distribution. *Journal of Time Series Analysis*, 34, 558-71.
- Corsi, F., Mittnik, S., Pigorsch, C., and U. Pigorsch (2008). The volatility of realized volatility. *Econometric Reviews*, 27, 46-78.
- Corsi, F. and R. Reno (2012). Discrete-Time Volatility Forecasting With Persistent Leverage Effect and the Link With Continuous-Time Volatility Modeling, *Journal of Business and Economic Statistics*, 30, 368-80.

Creal, D., Koopman, S. J. and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28, 777-95.

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

Harvey, A.C. (2013). *Dynamic Models for Volatility and Heavy Tails: with applications to financial and economic time series*. Econometric Society Monograph, Cambridge University Press.

Harvey, A.C. and R-J. Lange (2017). Volatility Modelling with a Generalized t-distribution, *Journal of Time Series Analysis*, 38, 175-90

Harvey, A.C. and R-J. Lange (2018). Modelling the Interactions between Volatility and Returns. *Journal of Time Series Analysis*, 39, 909–919.

Kleiber, C. and S. Kotz (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*. New York: Wiley.

Koopman, S. J. Harvey, A. C., Doornik, J. A. and N. Shephard (2009). *STAMP 8.2 Structural Time Series Analysis Modeller and Predictor*. London: Timberlake Consultants Ltd.

Lucas, A., Schwaab, B. and X. Zhang (2014). Conditional Euro Area sovereign default risk. *Journal of Business and Economic Statistics*, 32, 271-84.

McDonald J. B. and Y. J. Xu (1995) A generalization of the beta distribution with applications. *Journal of Econometrics* 66, 133-152.

Mitchell, J. and K. F. Wallis (2011). Evaluating density forecasts: forecast combinations, model mixtures, calibration and sharpness. *Journal of Applied Econometrics* 26, 1023-1040.

Muller, U., M. Dacorogna, R. Dav, R. Olsen, O. Pictet, and J. von Weizsacker (1997). Volatilities of different time resolutions – Analysing the dynamics of market components. *Journal of Empirical Finance* 4, 213–239.

Opschoor, A, Janus, P., Lucas, A. and D J. van Dijk (2018). New HEAVY Models for Fat-Tailed Realized Covariances and Returns, *Journal of Business and Economic Statistics*, 36, 643-7.

Patton, A. J. and K. Sheppard (2015). Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility. *Review of Economics and Statistics*, 97, 683–697.

Taylor, S. (2005). *Asset Price Dynamics, Volatility, and Prediction*. Princeton University Press.

Appendix on invertibility

Let

$$\Lambda_t(\boldsymbol{\psi}) := \sup |x_t|,$$

where $x_t = d\lambda_{t+1|t}/d\lambda_{t|t-1} = \phi + \kappa(\partial u_t/\partial \lambda_{t|t-1})$. Blasques et al (2018) show that a sufficient condition for invertibility is $E \ln \Lambda_0(\boldsymbol{\psi}) < 0$ over all admissible $\boldsymbol{\psi}$.

For GB2

$$\frac{\partial u_t}{\partial \lambda_{t|t-1}} = u'_t = -v^2(\xi + \varsigma)b_t(1 - b_t).$$

Thus

$$x_t = \phi + \kappa u'_t = \phi - \kappa v^2(\xi + \varsigma)b_t(1 - b_t)$$

The maximum of $b_t(1 - b_t)$ is 1/4 and so the model is invertible if we ensure

$$|\phi - \kappa v^2(\xi + \varsigma)/4| < 1 \tag{24}$$

Hence

$$-\frac{4(1 - \phi)}{v^2(\xi + \varsigma)} < \kappa < \frac{4(1 + \phi)}{v^2(\xi + \varsigma)}$$

These bounds appear to place a constraint on the magnitude of ξ and ς . However, this difficulty can be avoided by analysing the EGB2 with the forcing variable defined as in (11). Then

$$-\sigma^2 \frac{\partial_t^2 \ln f_t}{\partial \lambda_{t|t-1}^2} = h^2(\xi + \varsigma)b_t(1 - b_t)$$

and since the maximum of $b_t(1 - b_t)$ is again 1/4, the invertibility condition is

$$|\phi - \kappa h^2(\xi + \varsigma)/4| < 1;$$

compare (24). For a logistic distribution, when $\xi = \varsigma = 1$, the invertibility condition is $|\phi - 1.645\kappa| < 1$. As $\varsigma = \xi \rightarrow \infty$, the distribution becomes normal and since, as noted in Caivano and Harvey (2014, Lemma 1, p 560), $\varsigma h^2 = \xi h^2 \rightarrow 2$, it follows that $h^2(\xi + \varsigma)/4 \rightarrow 1$ and the standard invertibility condition for a Gaussian model is obtained, that is $|\phi - \kappa| < 1$; see Harvey (2013, p 67).

1. Tables

Table 1: Realized Volatility of FTSE 100 Index

Distribution	Shape			Components					Fit			
	ξ	ν	ς	ω	ϕ_1	κ_1	ϕ_2	κ_2	Logl	AIC	BIC	K-S Test
GB2 $\xi = \varsigma$	1.763 (0.136)	2.543 (0.085)	1.763 (0.136)	-8.827 (0.415)	0.999 (0.001)	0.036 (0.007)	0.842 (0.040)	0.058 (0.006)	42,135.76	-84,257.52	-84,212.58	0.027 (0.002)
Burr	1 -	3.893 (0.022)	0.856 (0.048)	-8.796 (0.383)	0.999 (0.001)	0.033 (0.006)	0.867 (0.030)	0.056 (0.006)	42,130.52	-84,247.03	-84,202.09	0.020 (0.052)
Log-Logistic	1 -	3.676 (0.012)	1 -	-8.764 (0.387)	0.999 (0.001)	0.034 (0.006)	0.862 (0.033)	0.056 (0.006)	42,125.46	-84,238.92	-84,200.40	0.023 (0.020)
F $\nu_1 = \nu_2$	4.494 (0.020)	1 -	4.494 (0.020)	-8.919 (0.469)	0.999 (0.002)	0.039 (0.007)	0.814 (0.049)	0.062 (0.007)	42,099.74	-84,187.48	-84,148.96	0.041 (0.000)
Pareto	1 -	1 -	39.499 (0.187)	-9.789 (0.236)	0.991 (0.004)	0.141 (0.030)	0.580 (0.174)	0.207 (0.033)	40,012.12	-80,012.24	-79,973.72	0.284 (0.000)
Weibull	1 -	1.330 (0.008)	∞ -	-9.688 (0.073)	0.992 (0.001)	0.023 (0.003)	0.667 (0.050)	0.131 (0.011)	40,433.65	-80,855.29	-80,816.77	0.167 (0.000)

Table 2: Realized Volatility of S&P 500 Index

Distribution	Shape			Components					Fit			
	ξ	ν	ς	ω	ϕ_1	κ_1	ϕ_2	κ_2	Logl	AIC	BIC	K-S Test
GB2 $\xi = \varsigma$	3.234 (0.220)	1.450 (0.126)	3.234 (0.220)	-9.546 (0.277)	0.994 (0.003)	0.058 (0.010)	0.781 (0.045)	0.091 (0.010)	40,337.42	-80,660.84	-80,615.90	0.016 (0.176)
Burr	1 -	3.271 (0.023)	0.798 (0.048)	-9.617 (0.287)	0.994 (0.003)	0.061 (0.009)	0.762 (0.047)	0.082 (0.009)	40,320.62	-80,627.25	-80,582.31	0.013 (0.432)
Log-Logistic	1 -	2.997 (0.012)	1 -	-9.567 (0.298)	0.994 (0.003)	0.059 (0.010)	0.779 (0.046)	0.086 (0.010)	40,310.28	-80,608.56	-80,570.04	0.026 (0.004)
F $\nu_1 = \nu_2$	3.153 (0.020)	1 -	3.153 (0.020)	-9.544 (0.274)	0.994 (0.003)	0.058 (0.010)	0.782 (0.045)	0.093 (0.010)	40,334.40	-80,656.81	-80,618.29	0.029 (0.001)
Pareto	1 -	1 -	45.402 (0.235)	-9.518 (0.319)	0.996 (0.003)	0.090 (0.025)	0.821 (0.048)	0.238 (0.028)	38,880.68	-77,749.37	-77,710.85	0.223 (0.000)
Weibull	1 -	1.292 (1.292)	∞ -	-9.388 (0.165)	0.997 (0.001)	0.022 (0.003)	0.793 (0.020)	0.170 (0.009)	39,203.06	-78,394.11	-78,355.59	0.109 (0.000)

Table 3: Estimation results from location/scale DCS Seasonal EGARCH with Leverage Effect

Index	Shape			Components							Logl	Fit AIC	BIC
	ξ	ν	ς	ω	ϕ_1	κ_1	κ_{1l}	ϕ_2	κ_2	κ_{2l}			
FTSE 100	1	3.930	0.841	-8.754	0.999	0.030	0.000	0.886	0.052	0.019	42,158.09	- 84,298.19	- 84,240.41
	-	(0.022)	(0.047)	(0.385)	(0.001)	(0.004)	(0.003)	(0.020)	(0.005)	(0.004)			
	1	4.038	0.825	-8.740	0.999	0.029	0.000	0.888	0.052	0.018	42,250.85	- 84,473.70	- 84,383.82
	-	(0.022)	(0.047)	(0.396)	(0.001)	(0.004)	(0.003)	(0.019)	(0.005)	(0.004)			
DAX	1	3.695	0.877	-7.788	1.000	0.037	0.004	0.838	0.056	0.027	38,632.90	- 77,283.06	- 77,225.28
	-	(0.056)	(0.081)	(0.591)	(0.001)	(0.005)	(0.004)	(0.036)	(0.006)	(0.005)			
	1	3.831	0.828	-8.472	0.998	-0.033	-0.011	0.986	0.114	0.033	38,695.46	- 77,362.92	- 77,273.04
	-	(0.018)	(0.032)	(0.066)	(0.000)	(0.002)	(0.000)	(0.000)	(0.005)	(0.001)			
S&P	2.605	1.652	2.605	-8.935	1.000	0.040	0.000	0.848	0.092	0.051	40,418.58	- 80,819.16	- 80,761.38
	(0.184)	(0.108)	(0.184)	(3.574)	(0.011)	(0.010)	(0.020)	(0.022)	(0.010)	(0.018)			
	2.143	1.878	2.143	-8.842	1.000	0.041	0.000	0.848	0.090	0.052	40,466.58	- 80,905.17	- 80,815.29
	(0.168)	(0.102)	(0.168)	(1.603)	(0.005)	(0.007)	(0.010)	(0.022)	(0.008)	(0.010)			
Nasdaq	2.497	1.978	2.497	-8.195	1.000	0.035	0.001	0.779	0.089	0.026	40,440.10	- 80,862.21	- 80,804.43
	(0.178)	(0.105)	(0.178)	(0.405)	(0.001)	(0.006)	(0.003)	(0.032)	(0.006)	(0.004)			
	2.361	2.076	2.361	-8.130	1.000	0.032	0.001	0.807	0.086	0.026	40,507.59	- 80,987.18	- 80,897.30
	(0.169)	(0.101)	(0.169)	(0.412)	(0.001)	(0.005)	(0.003)	(0.026)	(0.006)	(0.004)			

Table 4: Estimation results from location/scale DCS Seasonal EGARCH with Leverage Effect - Seasonal Component

Index	Seasonal					
	κ_s	γ_1	γ_2	γ_3	γ_4	γ_5
FTSE 100	0.001 (0.001)	-0.119 (0.020)	-0.029 (0.015)	0.031 (0.021)	0.043 (0.026)	0.075 -
DAX	0.000 (0.000)	-0.113 (0.007)	-0.044 (0.016)	0.004 (0.012)	0.081 (0.007)	0.072 -
S&P	0.000 (0.002)	-0.129 (0.075)	-0.006 (0.015)	0.063 (0.014)	0.069 (0.068)	0.003 -
Nasdaq	0.001 (0.000)	-0.114 (0.031)	-0.045 (0.034)	0.122 (0.028)	0.045 (0.028)	-0.007 -

Table 5: Estimation results from location/scale DCS-H Seasonal EGARCH

Index	Shape						Components						
	ξ	v	ζ	ω_v	ϕ_v	κ_v	ω_λ	$\phi_{\lambda 1}$	$\kappa_{\lambda 1}$	$\kappa_{\lambda 1l}$	$\phi_{\lambda 2}$	$\kappa_{\lambda 2}$	$\kappa_{\lambda 2l}$
FTSE 100	1	4.034	0.825				-8.699	1.000	0.029	0.000	0.887	0.052	0.018
	-	(0.022)	(0.047)				(0.386)	(0.001)	(0.004)	(0.003)	(0.019)	(0.005)	(0.004)
	1		0.817	-1.400	0.510	0.045	-8.620	1.000	0.028	0.000	0.892	0.049	0.018
DAX	-		(0.047)	(0.023)	(0.139)	(0.009)	(0.366)	(0.001)	(0.004)	(0.003)	(0.017)	(0.005)	(0.004)
	1	3.800	0.851				-7.792	1.000	0.036	0.004	0.846	0.056	0.025
	-	(0.022)	(0.049)				(0.523)	(0.001)	(0.005)	(0.004)	(0.028)	(0.005)	(0.004)
S&P 500	1		0.851	-1.322	0.985	0.008	-7.810	1.000	0.035	0.004	0.857	0.054	0.023
	-		(0.049)	(0.029)	(0.009)	(0.003)	(0.473)	(0.001)	(0.004)	(0.003)	(0.026)	(0.005)	(0.004)
	2.017	1.947	2.017				-8.408	1.000	0.039	-0.001	0.866	0.090	0.052
Nasdaq 100	(0.153)	(0.094)	(0.153)				(0.723)	(0.002)	(0.006)	(0.006)	(0.018)	(0.007)	(0.007)
	2.161		2.161	-0.627	0.531	0.037	-8.703	1.000	0.039	0.001	0.860	0.085	0.049
	(0.157)		(0.157)	(0.095)	(0.132)	(0.008)	(0.813)	(0.003)	(0.005)	(0.006)	(0.020)	(0.006)	(0.004)
Nasdaq 100	2.443	2.032	2.443				-7.716	1.000	0.029	0.001	0.837	0.089	0.026
	(0.173)	(0.103)	(0.173)				(0.378)	(0.020)	(0.006)	(0.004)	(0.001)	(0.005)	(0.003)
	2.742		2.742	-0.646	0.556	0.044	-8.139	1.000	0.030	0.001	0.820	0.081	0.026
	(0.175)		(0.175)	(0.102)	(0.092)	(0.007)	(0.367)	(0.001)	(0.005)	(0.003)	(0.024)	(0.005)	(0.004)

Table 6: Goodness of Fit

Index	Model	Logl	Fit		$Q_v(5)$	$Q_v(20)$
			AIC	BIC		
FTSE 100	DCS	42,236.70	- 84,447.39	- 84,363.93	122.455 (0.000)	157.988 (0.000)
	DCS-H	42,270.43	- 84,508.86	- 84,406.14	53.705 (0.000)	84.469 (0.000)
DAX	DCS	38,727.42	- 77,426.83	- 77,336.95	131.337 (0.000)	209.129 (0.000)
	DCS-H	38,748.30	- 77,464.60	- 77,361.88	73.715 (0.000)	95.681 (0.000)
S&P 500	DCS	40,466.07	- 80,904.14	- 80,814.26	112.195 (0.000)	188.159 (0.000)
	DCS-H	40,481.79	- 80,931.59	- 80,828.87	22.111 (0.000)	78.503 (0.000)
Nasdaq 100	DCS	40,505.39	- 80,982.79	- 80,892.91	157.272 (0.000)	246.286 (0.000)
	DCS-H	40,539.83	- 81,047.66	- 80,944.93	24.458 (0.000)	86.623 (0.000)

Table 7: Forecasting results on RV for location/scale DCS and HAR.

Index		01/02/2005-29/07/2005						01/10/2008-31/03/2009					
		Low Vol						High Vol					
		DCS	DCS-H	HAR Levels	HAR Log	CHAR	HARJ	DCS	DCS-H	HAR Levels	HAR Log	CHAR	HARJ
FTSE	RMSFE	1.000	1.002	1.094	1.010	1.111	1.128	1.000	1.000	1.022	0.992	1.045	1.054
	MAFE	1.000	0.996	1.560	1.098	1.795	1.585	1.000	1.000	1.050	0.995	1.148	1.096
	D-M Test RMSE Vs DCS	-	1.345 (0.179)	-1.605 (0.109)	-2.553 (0.011)	-1.752 (0.080)	-1.413 (0.158)	-	0.034 (0.973)	-0.391 (0.696)	0.364 (0.716)	-0.631 (0.528)	-0.622 (0.534)
	D-M Test RMSE Vs DCS-H	-	-	-1.569 (0.117)	-2.036 (0.042)	-1.717 (0.086)	-1.392 (0.164)	-	-	-0.388 (0.698)	0.399 (0.690)	-0.625 (0.532)	-0.621 (0.535)
	Predictive Likelihood	1,925.10	1,922.38	1,404.07	1,890.31	1,396.93	1,402.01	1,375.22	1,376.00	1,038.24	1,361.11	1,045.56	1,023.70
DAX	RMSFE	1.000	1.001	1.094	1.004	1.068	1.319	1.000	1.024	1.066	1.032	1.016	1.080
	MAFE	1.000	1.005	1.322	1.080	1.408	1.624	1.000	1.031	1.087	1.006	1.069	1.133
	D-M Test RMSE Vs DCS	-	0.281 (0.779)	-1.127 (0.260)	-0.176 (0.860)	-1.749 (0.080)	-2.246 (0.025)	-	1.880 (0.060)	-1.299 (0.194)	-1.460 (0.144)	-0.279 (0.780)	-1.328 (0.184)
	D-M Test RMSE Vs DCS-H	-	-	-1.125 (0.261)	-0.138 (0.890)	-1.741 (0.082)	-2.246 (0.025)	-	-	-0.794 (0.427)	-0.427 (0.669)	0.130 (0.897)	-0.884 (0.377)
	Predictive Likelihood	1,710.92	1,711.87	1,317.49	1,698.67	1,317.39	1,315.22	1,310.37	1,310.95	617.40	1,299.38	670.13	599.49
S&P	RMSFE	1.000	1.007	1.057	1.029	1.007	1.216	1.000	1.014	1.049	1.009	1.001	1.100
	MAFE	1.000	1.011	1.065	1.024	1.029	1.231	1.000	0.996	1.067	1.014	1.048	1.133
	D-M Test RMSE Vs DCS	-	1.375 (0.169)	-1.677 (0.094)	-1.031 (0.303)	-0.212 (0.832)	-4.200 (0.000)	-	1.120 (0.263)	-0.759 (0.448)	-0.396 (0.692)	0.017 (0.986)	-0.839 (0.402)
	D-M Test RMSE Vs DCS-H	-	-	-1.545 (0.122)	-0.804 (0.421)	0.002 (0.998)	-4.151 (0.000)	-	-	-0.538 (0.591)	0.263 (0.793)	0.218 (0.828)	-0.713 (0.476)
	Predictive Likelihood	1,710.92	1,711.87	1,317.49	1,698.67	1,317.39	1,315.22	1,259.00	1,261.87	-1,096.13	1,257.26	- 956.97	-1,259.93
Nasdaq	RMSFE	1.000	1.011	1.044	1.030	1.100	1.056	1.000	1.033	1.048	1.007	1.019	1.052
	MAFE	1.000	1.020	1.078	1.048	1.179	1.084	1.000	1.010	1.086	1.042	1.099	1.097
	D-M Test RMSE Vs DCS	-	1.028 (0.304)	-1.242 (0.214)	-1.087 (0.277)	-2.203 (0.028)	-1.638 (0.102)	-	1.508 (0.132)	-0.584 (0.559)	-0.170 (0.865)	-0.294 (0.769)	-0.597 (0.551)
	D-M Test RMSE Vs DCS-H	-	-	-1.035 (0.301)	-0.753 (0.451)	-2.150 (0.032)	-1.462 (0.144)	-	-	-0.171 (0.865)	0.494 (0.622)	0.172 (0.864)	-0.203 (0.839)
	Predictive Likelihood	1,713.32	1,717.33	1,340.90	1,712.65	1.337.11	1.340.69	1.336.05	1,337.74	911.94	1,328.90	950.74	910.06