Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility Internet Appendix

November 16, 2013

A Data Cleaning

Only transaction data were taken from the NYSE TAQ. All series were automatically cleaned according to a set of six rules:

1. Transactions outside of 9:30:00 and 16:00:00 were discarded.
2. Transactions with zero price or volume were discarded.
3. Each day the most active exchange was determined. Only transactions from this exchange were retained.
4. Only trades with conditions E, F or blank were retained.
5. Transaction prices outside of the CRSP high or low for the day were discarded.
6. Trade with immediate reversals more than 5 times a 50-sample moving window - excluding the transaction being tested - were discarded.

These rules are similar to those of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009), and prices were not manually cleaned for problems not addressed by these rules.

B Completely Decomposing Quadratic Variation

The specification in eq. (16) of the paper restricts the coefficients on the weekly and monthly realized semivariance to be identical. This restriction can be relaxed by decomposing the RV terms at all lags.[A.1] With this modification we obtain:

$$\overline{RV}_{h,t+h} = \mu + \phi_d^+ RS_t^+ + \phi_d^- RS_t^- + \phi_w^+ \overline{RS}_{w,t}^+ + \phi_w^- \overline{RS}_{w,t}^- + \phi_m^+ \overline{RS}_{m,t}^+ + \phi_m^- \overline{RS}_{m,t}^- + \epsilon_{t+h}. \tag{A.1}$$

Results from this extended specification are presented in the first row of each column of Table A.1a. In both sets of results, those using the SPDR and those based on the panel, the negative semivariance dominates the positive semivariance. In the models using the SPDR, the coefficients on positive semivariance are always either significantly negative or insignificantly different from zero. The coefficient on the terms involving negative semivariance are uniformly positive and significant. In the panel estimation the same general pattern appears, although some of the coefficients on positive semivariance, especially at short horizons and lag 1, are significantly positive. Interestingly, as the horizon increases, the persistence of the volatility in the panel shifts to the negative semivariance, particularly at the longer lags.

Decomposing realized variance at all lags allows us to consider a "Vector HAR" (VHAR) for the two semivariances. Such a model allows us to determine whether lagged realized semivariances of the same sign as the dependent variable are more useful than lagged semivariances of the opposite sign.

$$\begin{bmatrix} \overline{RS}_{h,t+h}^+ \\ \overline{RS}_{h,t+h}^- \end{bmatrix} = \begin{bmatrix} \mu^+ \\ \mu^- \end{bmatrix} + \boldsymbol{\phi}_d \begin{bmatrix} RS_t^+ \\ RS_t^- \end{bmatrix} + \boldsymbol{\phi}_w \begin{bmatrix} \overline{RS}_{w,t}^+ \\ \overline{RS}_{w,t}^- \end{bmatrix} + \boldsymbol{\phi}_m \begin{bmatrix} \overline{RS}_{m,t}^+ \\ \overline{RS}_{m,t}^- \end{bmatrix} + \begin{bmatrix} \epsilon_{t+h}^+ \\ \epsilon_{t+h}^- \end{bmatrix} \tag{A.2}$$

where

$$\boldsymbol{\phi}_j = \begin{bmatrix} \phi_{j+}^+ & \phi_{j+}^- \\ \phi_{j-}^+ & \phi_{j-}^- \end{bmatrix}, \quad j \in \{d, w, m\}$$

Results of the VHAR are presented in the two lower rows of each panel of Tables A.1a and A.1b. The estimates for each equation of the VHAR are virtually identical, with small or negative coefficients on lagged

[A.1] When a jump variable is also included, Chen and Ghysels (2011) call this model the "HAR-S-RV-J" model, and use it as one of the benchmarks in their study.

positive semivariance and large, significant coefficients on lagged negative semivariance. The results in the panel are similar with parameter estimates on negative semivariance uniformly large and highly significant. Thus negative semivariance is useful for predicting *both* positive and negative future semivariance. This is a novel and somewhat surprising result.

This leads us to test whether positive semivariance is actually needed in the VHAR models. We perform these tests on the individual models for the SPDR and the 105 constituent volatility series. The first null hypothesis is that positive semivariance can be excluded, $H_0 : \phi_d^+ = \phi_w^+ = \phi_m^+ = 0$, and the other tests whether negative semivariance can be excluded. We find that positive semivariance can be excluded from 19.1.% of the joint (bivariate VHAR) models, 19.1% of the models for positive semivariance, and 25.7% of the models for negative semivariance. Negative semivariance can only be excluded in 1 of the 105 volatility series. Thus, while most of the predictability for future semivariance appears to come from lagged negative semivariance, the lagged positive semivariance also carries some information.

We next test whether the sum of the coefficients on the positive semivariance is equal to the sum of the coefficients on the negative semivariance, $H_0 : \phi_d^+ + \phi_w^+ + \phi_m^+ = \phi_d^- + \phi_w^- + \phi_m^-$ in each of the two semivariance models. This null can be rejected for all but 3 (25) of the 105 positive (negative) semivariance models, and, when rejected, the sum of the coefficients on the negative semivariance is larger than the sum of the coefficient on the positive semivariance in all but two cases. We next test for equality only at the first lag, $H_0 : \phi_d^+ = \phi_d^-$. This null is rejected in 61% of the positive semivariance models and 78% of the negative semivariance models, and when rejected typically indicates that the coefficient on negative semivariance is larger than the coefficient on positive semivariance. Thus both of these sets of hypothesis tests reveal that negative semivariances have greater weight in these predictive models for almost all assets considered here.

Finally, we test whether the persistence of each series, as measured by the maximum eigenvalue of the companion form of a HAR, is equal for the two semivariances. This is done by restricting the off-diagonal elements in eq. A.2 to be zero, and estimating the remaining parameters and the (joint) asymptotic covariance matrix. The asymptotic distribution is used to simulate 1,000 draws of the parameters, each one is then transformed into companion form, and the maximum eigenvalue of the companion matrix is com-

Forecasting measures of future volatility using realized semivariances, results for the SPDR

$$\overline{RM}_{h,t+h} = \mu + \phi_d^+ RS_t^+ + \phi_w^+ \overline{RS}_{w,t}^+ + \phi_m^+ \overline{RS}_{m,t}^+ + \phi_d^- RS_t^- + \phi_w^- \overline{RS}_{w,t}^- + \phi_m^- \overline{RS}_{m,t}^- + \epsilon_{t+h}$$

	RM	ϕ_d^+	ϕ_w^+	ϕ_m^+	ϕ_d^-	ϕ_w^-	ϕ_m^-	R^2
$h=1$	RV	−0.087	−0.179	−0.233	1.162	0.788	0.536	0.621
		(−1.3)	(−2.5)	(−1.9)	(12.7)	(9.1)	(3.4)	
	RS^+	−0.084	−0.110	−0.069	0.644	0.404	0.219	0.539
		(−1.9)	(−2.9)	(−1.0)	(10.0)	(8.1)	(2.3)	
	RS^-	−0.003	−0.069	−0.163	0.518	0.384	0.317	0.621
		(−0.1)	(−1.7)	(−2.3)	(14.9)	(8.0)	(3.6)	
$h=5$	RV	−0.070	0.011	−0.474	0.836	0.815	0.880	0.642
		(−1.5)	(0.1)	(−2.4)	(13.4)	(5.9)	(3.7)	
	RS^+	−0.049	0.015	−0.207	0.447	0.405	0.403	0.635
		(−1.9)	(0.2)	(−2.0)	(12.6)	(5.8)	(3.2)	
	RS^-	−0.022	−0.004	−0.267	0.389	0.410	0.477	0.623
		(−0.9)	(−0.1)	(−2.6)	(13.6)	(5.9)	(3.9)	
$h=22$	RV	−0.058	−0.217	−0.920	0.583	0.907	1.607	0.565
		(−1.3)	(−1.3)	(−1.8)	(10.3)	(5.9)	(2.6)	
	RS^+	−0.031	−0.107	−0.494	0.301	0.451	0.851	0.583
		(−1.3)	(−1.2)	(−1.9)	(10.2)	(5.5)	(2.6)	
	RS^-	−0.027	−0.110	−0.425	0.282	0.457	0.756	0.538
		(−1.3)	(−1.4)	(−1.7)	(10.3)	(6.1)	(2.5)	
$h=66$	RV	−0.119	−0.411	−0.835	0.447	0.932	1.580	0.364
		(−2.7)	(−2.5)	(−1.3)	(7.9)	(5.0)	(2.3)	
	RS^+	−0.064	−0.218	−0.444	0.235	0.488	0.819	0.378
		(−2.8)	(−2.5)	(−1.3)	(8.0)	(5.0)	(2.3)	
	RS^-	−0.055	−0.193	−0.392	0.212	0.444	0.761	0.347
		(−2.7)	(−2.4)	(−1.3)	(7.8)	(5.0)	(2.4)	

Table A.1a: Extended model where RV at all lags is decomposed into positive and negative semivariance (robust t-statistics in parentheses). Each of the four panels contains results for the forecast horizon indicated at the left. RM indicates the dependent variable, realized variance (RV), positive realized semivariance (RS^+) or negative realized semivariance (RS^-). The R^2 measure is constructed using the WLS parameter estimates and the original, unmodified data.

puted. The null is tested using the percentage of times where $\lambda^- > \lambda^+$, were λ^+ is the largest eigenvalue from the companion matrix for HAR for positive semivariance. The null is rejected if λ^- is greater than λ^+ in more than 97.5% or in less than 2.5% of the simulations. quality is rejected in 89.6% of the series, and negative semivariance is found to be more persistent in 91.6% of the rejections.

These results indicate that negative semivariance is more useful for predicting realized variance and *both* realized semivariances, and that negative semivariance is more persistent than positive semivariance.

Forecasting measures of future volatility using realized semivariances, results for the panel of 105 individual stocks.

$$\overline{RM}_{h,i,t+h} = \mu_i + \phi_d^+ RS_{i,t}^+ + \phi_w^+ \overline{RS}_{w,i,t}^+ + \phi_m^+ \overline{RS}_{m,i,t}^+ + \phi_d^- RS_{i,t}^- + \phi_w^- \overline{RS}_{w,i,t}^- + \phi_m^- \overline{RS}_{m,i,t}^- + \epsilon_{i,t+h}$$

	RM	ϕ_d^+	ϕ_w^+	ϕ_m^+	ϕ_d^-	ϕ_w^-	ϕ_m^-	R^2
$h=1$	RV	0.262	0.105	0.065	0.696	0.529	0.290	0.399
		(15.3)	(5.5)	(2.5)	(24.3)	(25.1)	(11.0)	
	RS^+	0.144	0.061	0.058	0.339	0.253	0.118	0.380
		(14.5)	(6.0)	(4.1)	(18.3)	(22.6)	(8.3)	
	RS^-	0.118	0.045	0.007	0.357	0.277	0.172	0.309
		(13.5)	(4.1)	(0.5)	(28.7)	(23.4)	(11.1)	
$h=5$	RV	0.152	0.136	0.033	0.542	0.577	0.479	0.531
		(10.7)	(3.7)	(0.5)	(19.2)	(16.5)	(7.9)	
	RS^+	0.079	0.079	0.038	0.272	0.279	0.214	0.536
		(11.0)	(4.3)	(1.3)	(18.0)	(16.4)	(7.6)	
	RS^-	0.073	0.058	−0.005	0.270	0.298	0.265	0.472
		(10.0)	(3.1)	(−0.2)	(20.0)	(15.9)	(7.8)	
$h=22$	RV	0.085	0.057	0.022	0.377	0.561	0.724	0.517
		(6.4)	(1.3)	(0.2)	(13.2)	(11.9)	(5.2)	
	RS^+	0.045	0.035	0.015	0.189	0.275	0.358	0.524
		(6.8)	(1.6)	(0.3)	(12.8)	(11.9)	(5.2)	
	RS^-	0.039	0.022	0.008	0.188	0.286	0.366	0.494
		(5.9)	(1.0)	(0.1)	(13.5)	(11.8)	(5.2)	
$h=66$	RV	0.038	−0.001	0.119	0.266	0.467	0.770	0.455
		(2.9)	(−0.0)	(0.7)	(9.4)	(7.1)	(4.7)	
	RS^+	0.020	0.004	0.066	0.134	0.233	0.378	0.460
		(3.1)	(0.1)	(0.8)	(9.2)	(7.0)	(4.5)	
	RS^-	0.017	−0.005	0.053	0.132	0.235	0.392	0.444
		(2.6)	(−0.2)	(0.7)	(9.5)	(7.2)	(4.9)	

Table A.1b: Extended model where RV at all lags is decomposed into positive and negative semivariance (robust t-statistics in parentheses). Each of the four panels contains results for the forecast horizon indicated at the left. RM indicates the dependent variable, realized variance (RV), positive realized semivariance (RS^+) or negative realized semivariance (RS^-). The final column reports the average of the 105 R^2s for the individual assets constructed using the WLS parameter estimates and the original, unmodified data.

C Sign of Past Returns

We consider an extended model similar to the model presented in table 2a and 2b in the main paper where an indicator variable for a negative return on the previous day was included as a level-shift in addition to the interaction with lagged realized variance. Tables A.2a and A.2b contain the results from this additional specification applied to the SPDR and the panel of firm volatilities, respectively. The indicator variable typically has a negative sign indicating that level of volatility is lower subsequent to a negative shock when compared to a model without this term, although the response to the interaction term with lagged realized variance increases slightly and so the net effect of a negative return is still likely positive. The coefficients on the realized semivariance, RS^{+} and RS^{-}, are virtually unchanged by this modification.

Extended HAR estimation results for the SPDR, cumulative volatility

$$\overline{RV}_{h,t+h} = \mu + \delta I_{[r_{t-1}<0]} + \phi_d RV_t + \phi_d^+ RS_t^+ + \phi_d^- RS_t^- + \gamma RV_t I_{[r_t<0]} + \phi_w \overline{RV}_{w,t} + \phi_m \overline{RV}_{m,t} + \epsilon_{t+h}$$

	δ	ϕ_d	ϕ_d^+	ϕ_d^-	γ	ϕ_w	ϕ_m	R^2
$h=1$		0.607 (17.0)				0.268 (8.1)	0.120 (4.9)	0.532
			−0.024 (−0.3)	1.182 (13.0)		0.291 (9.3)	0.120 (4.9)	0.611
			0.037 (0.4)	1.064 (7.4)	0.050 (1.4)	0.293 (9.3)	0.121 (5.0)	0.611
	−4.185 (−2.0)		0.004 (0.0)	1.052 (7.4)	0.100 (2.3)	0.291 (9.2)	0.121 (5.0)	0.617
$h=5$		0.425 (14.7)				0.409 (8.6)	0.158 (4.0)	0.563
			−0.030 (−0.7)	0.862 (13.4)		0.421 (8.8)	0.155 (4.0)	0.620
			0.073 (1.1)	0.650 (6.6)	0.092 (2.5)	0.424 (8.8)	0.157 (4.0)	0.619
	−4.034 (−2.1)		0.049 (0.8)	0.639 (6.5)	0.132 (2.8)	0.422 (8.7)	0.157 (4.1)	0.621
$h=22$		0.305 (11.8)				0.357 (7.7)	0.265 (4.8)	0.468
			−0.009 (−0.3)	0.628 (9.9)		0.359 (7.5)	0.261 (4.8)	0.508
			−0.012 (−0.2)	0.635 (5.8)	−0.003 (−0.1)	0.359 (7.6)	0.261 (4.8)	0.508
	−0.656 (−0.4)		−0.015 (−0.3)	0.633 (5.7)	0.003 (0.1)	0.359 (7.6)	0.261 (4.8)	0.508
$h=66$		0.203 (8.4)				0.256 (7.4)	0.299 (5.3)	0.282
			−0.067 (−2.2)	0.501 (7.3)		0.253 (6.8)	0.294 (5.3)	0.313
			−0.121 (−1.8)	0.622 (4.1)	−0.054 (−1.3)	0.251 (6.8)	0.292 (5.2)	0.315
	1.666 (0.6)		−0.116 (−1.9)	0.627 (4.0)	−0.065 (−1.2)	0.252 (6.8)	0.292 (5.2)	0.315

Table A.2a: All models use the h-day cumulative variance as the dependent variable (robust t-statistics in parentheses). Each of the four panels contains results for the forecast horizon indicated in the left most column. Each panel contains 4 models: the first model is a standard RV HAR, the second decomposes realized variance into positive and negative realized semivariance at the first lag, the third specification adds an asymmetric term where the sign of the most recent daily return is used (these are repeated from Table 2a of the main paper) and the final model augments the third with an indicator variable as a level-shift. The R^2 measure is constructed using the WLS parameter estimates and the original, unmodified data.

Extended HAR estimation results for the panel of 105 individual stocks, cumulative volatility

$$\overline{RV}_{h,i,t+h} = \mu_i + \delta I_{[r_{i,t-1}<0]} + \phi_d RV_{t,i} + \phi_d^+ RS^+_{i,t} + \phi_d^- RS^-_{i,t} + \gamma RV_{i,t} I_{[r_{i,t}<0]} + \phi_w \overline{RV}_{w,i,t} + \phi_m \overline{RV}_{m,i,t} + \epsilon_{i,t+h}$$

	δ	ϕ_d	ϕ_d^+	ϕ_d^-	γ	ϕ_w	ϕ_m	R^2
$h=1$		0.488				0.315	0.172	0.395
		(39.7)				(28.1)	(16.2)	
			0.268	0.704		0.317	0.172	0.398
			(15.6)	(24.5)		(28.7)	(16.4)	
			0.316	0.607	0.046	0.317	0.172	0.398
			(17.1)	(18.9)	(6.8)	(28.7)	(16.5)	
	−3.306		0.311	0.597	0.061	0.317	0.173	0.399
	(−2.8)		(16.4)	(19.3)	(6.8)	(28.8)	(16.5)	
$h=5$		0.357				0.357	0.247	0.525
		(23.2)				(16.3)	(10.2)	
			0.158	0.551		0.359	0.247	0.529
			(11.4)	(19.0)		(16.4)	(10.3)	
			0.210	0.444	0.052	0.359	0.248	0.529
			(14.5)	(19.6)	(7.7)	(16.5)	(10.3)	
	−3.382		0.206	0.435	0.065	0.359	0.248	0.529
	(−2.5)		(14.4)	(20.1)	(6.3)	(16.4)	(10.4)	
$h=22$		0.241				0.312	0.360	0.511
		(14.8)				(11.9)	(10.3)	
			0.091	0.388		0.314	0.360	0.513
			(7.3)	(12.8)		(11.9)	(10.4)	
			0.126	0.314	0.036	0.314	0.360	0.514
			(9.4)	(13.8)	(5.1)	(11.9)	(10.4)	
	−0.780		0.125	0.313	0.039	0.314	0.360	0.514
	(−0.6)		(9.4)	(13.8)	(4.5)	(11.9)	(10.4)	
$h=66$		0.161				0.236	0.431	0.450
		(12.2)				(10.7)	(12.0)	
			0.044	0.275		0.238	0.431	0.452
			(3.9)	(8.8)		(10.6)	(12.1)	
			0.062	0.235	0.020	0.238	0.432	0.452
			(5.9)	(8.5)	(3.8)	(10.6)	(12.1)	
	0.731		0.063	0.237	0.018	0.238	0.431	0.452
	(0.5)		(6.0)	(8.3)	(3.2)	(10.6)	(12.1)	

Table A.2b: All models use the h-day cumulative variance as the dependent variable (robust t-statistics in parentheses). Each of the four panels contains results for the forecast horizon indicated in the left most column. Each panel contains 4 models: the first model is a standard RV HAR, the second decomposes realized variance into positive and negative realized semivariance at the first lag, the third specification adds an asymmetric term where the sign of the most recent daily return is used (these are repeated from Table 2b of the main paper) and the final model augments the third with an indicator variable as a level-shift. The final column reports the average of the 105 R^2s for the individual assets constructed using the WLS parameter estimates and the original, unmodified data.

D Model in Logs

The models in the paper, as well as the VHAR in this appendix, are all estimated in levels using weighted least squares. As an alternative, we estimate the same models using the log of realized variance and realized semivariance with ordinary least squares. The log transformation naturally eliminates the primary source of heteroskedasticity in realized variance and related estimators and so there is no need for weighting. The signed jump variation measures are not always positive, and so these were replaced with percentage jump variation measures defined below.

Level Parameterization	Log Parameterization
ΔJ_t^2	$\%\Delta J_t^2 = \ln\left(1 + \Delta J_t^2 / RV_t\right)$
ΔJ_t^{2+}	$\%\Delta J_t^{2+} = \ln\left(1 + \Delta J_t^2 / RV_t\right) I_{[\Delta J_t^2 > 0]}$
ΔJ_t^{2-}	$\%\Delta J_t^{2-} = \ln\left(1 + \Delta J_t^2 / RV_t\right) I_{[\Delta J_t^2 < 0]}$

Results for SPDR

Tables A.3a, A.4a and A.5a correspond to Tables 2a and 3a from the main paper, and Table A.1a from this internet appendix. The results in terms of the sign, relative magnitude and statistical significance are remarkably similar. The magnitudes of the coefficients differ since the exact decomposition of RV into RS^+ and RS^- no longer holds, and the log eliminates scale differences between the realized variance and the realized semivariance. Table A.4a contains the models built using the modified ΔJ^2 measures, which continue to have uniformly negative signs indicating that negative jumps increase volatility and positive jumps decrease volatility. The main conclusions, that negative semivariance is extremely informative about future volatility, are strengthened with these results.

Results for the Panel

Tables A.3b, A.4b and A.5b corresponds to Tables 2b and 3b from the main paper, and Table A.1b from this internet appendix. Like the results for the SPDR, these are virtually identical in terms of the relative magnitude of the coefficients and statistical significance, and serve to reinforce the importance of decomposed volatility estimators in forecasting volatility. In particular, Table A.4b confirms that the sign on $\%\Delta J^2$ is negative and statistically significant at all horizons.

HAR estimation results for the SPDR, log cumulative volatility

$$\ln \overline{RV}_{h,t+h} = \mu + \phi_d \ln RV_t + \phi_d^+ \ln RS_t^+ + \phi_d^- \ln RS_t^- + \gamma \ln RV_t I_{[r_t<0]} + \phi_w \ln \overline{RV}_{w,t} + \phi_m \ln \overline{RV}_{m,t} + \epsilon_{t+h}$$

	ϕ_d	ϕ_d^+	ϕ_d^-	γ	ϕ_w	ϕ_m	R^2
$h=1$	0.520 (24.0)				0.309 (11.9)	0.123 (5.9)	0.778
		0.044 (2.0)	0.446 (20.9)		0.335 (13.4)	0.124 (6.0)	0.786
		0.116 (4.0)	0.361 (11.8)	0.021 (4.0)	0.336 (13.4)	0.125 (6.1)	0.787
$h=5$	0.390 (19.6)				0.356 (10.2)	0.176 (5.3)	0.798
		0.033 (1.9)	0.336 (17.1)		0.374 (10.8)	0.176 (5.4)	0.803
		0.078 (3.0)	0.283 (10.1)	0.013 (2.7)	0.375 (10.7)	0.177 (5.4)	0.804
$h=22$	0.291 (13.6)				0.311 (8.8)	0.241 (5.3)	0.717
		0.036 (2.2)	0.243 (11.3)		0.322 (9.1)	0.240 (5.4)	0.720
		0.039 (1.5)	0.238 (7.3)	0.001 (0.2)	0.322 (9.1)	0.240 (5.4)	0.720
$h=66$	0.212 (11.0)				0.232 (8.7)	0.299 (5.5)	0.618
		0.012 (0.6)	0.190 (7.1)		0.241 (8.5)	0.299 (5.5)	0.621
		−0.007 (−0.2)	0.213 (4.5)	−0.006 (−0.9)	0.241 (8.5)	0.298 (5.5)	0.621

Table A.3a: Reference, base and asymmetric model parameter estimates using log h-day cumulative variance as the dependent variable (robust t-statistics in parentheses). Each of the four panels contains results for the forecast horizon indicated in the left most column. Each panel contains 3 models: the first model corresponds to the reference model using only realized variance, the second decomposes realized variance into positive and negative realized semivariance at the first lag, and the third specification adds an asymmetric term where the sign of the most recent daily return is used.

HAR estimation results for the panel of 105 individual stocks, log cumulative volatility

$$\ln \overline{RV}_{h,i,t+h} = \mu_i + \phi_d \ln RV_{i,t} + \phi_d^+ \ln RS_{i,t}^+ + \phi_d^- \ln RS_{i,t}^- + \gamma \ln RV_{i,t} I_{[r_{i,t}<0]} + \phi_w \ln \overline{RV}_{w,i,t} + \phi_m \ln \overline{RV}_{m,i,t} + \epsilon_{i,t+h}$$

	ϕ_d	ϕ_d^+	ϕ_d^-	γ	ϕ_w	ϕ_m	R^2
$h=1$	0.438				0.320	0.195	0.724
	(59.1)				(34.9)	(19.5)	
		0.140	0.296		0.321	0.195	0.725
		(23.1)	(37.3)		(35.4)	(19.6)	
		0.174	0.257	0.008	0.321	0.195	0.726
		(28.7)	(33.4)	(10.7)	(35.5)	(19.8)	
$h=5$	0.323				0.346	0.257	0.779
	(32.4)				(20.8)	(12.6)	
		0.085	0.238		0.346	0.256	0.780
		(16.4)	(25.9)		(21.0)	(12.7)	
		0.116	0.203	0.008	0.346	0.257	0.781
		(19.3)	(26.2)	(9.4)	(21.1)	(12.7)	
$h=22$	0.232				0.300	0.329	0.735
	(18.8)				(16.3)	(11.1)	
		0.053	0.179		0.300	0.329	0.737
		(8.4)	(16.1)		(16.4)	(11.1)	
		0.075	0.154	0.005	0.300	0.329	0.737
		(10.2)	(16.8)	(6.0)	(16.4)	(11.2)	
$h=66$	0.166				0.237	0.387	0.676
	(18.4)				(16.7)	(12.5)	
		0.031	0.134		0.237	0.387	0.677
		(4.8)	(11.1)		(16.6)	(12.6)	
		0.045	0.117	0.004	0.238	0.387	0.677
		(7.0)	(11.8)	(3.7)	(16.6)	(12.6)	

Table A.3b: Reference, base and asymmetric model parameter estimates using log h-day cumulative variance as the dependent variance (robust t-statistics in parentheses). Each of the four panels contains results for the forecast horizon indicated in the left most column. Each panel contains 3 models: the first model corresponds to the reference model using only realized variance, the second decomposes realized variance into positive and negative realized semivariance at the first lag, and the third specification adds an asymmetric term where the sign of the most recent daily return is used. The reported R^2 measure is the average of the 105 R^2s for the individual series.

The impact of signed jump variation on future log volatility, results for the SPDR

$$\ln \overline{RM}_{h,t+h} = \mu + \phi_J \%\Delta J_t^2 + \phi_{J^+} \%\Delta J_t^{2+} + \phi_{J^-} \%\Delta J_t^{2-} + \phi_C \ln BV_t + \phi_w \ln \overline{RV}_{w,t} + \phi_m \ln \overline{RV}_{m,t} + \epsilon_{t+h}$$

	RM	ϕ_J	ϕ_{J^+}	ϕ_{J^-}	ϕ_C	ϕ_w	ϕ_m	R^2
$h=1$	RV				0.523 (25.2)	0.297 (11.6)	0.124 (6.0)	0.780
	RV	−0.412 (−10.5)			0.489 (24.5)	0.332 (13.3)	0.124 (6.1)	0.788
	RV		−0.215 (−2.6)	−0.539 (−7.9)	0.488 (24.5)	0.335 (13.5)	0.124 (6.1)	0.789
	BV	−0.419 (−10.7)			0.505 (25.4)	0.335 (13.5)	0.120 (6.0)	0.794
$h=5$	RV				0.393 (19.4)	0.346 (9.7)	0.177 (5.4)	0.799
	RV	−0.303 (−9.7)			0.368 (18.3)	0.372 (10.6)	0.176 (5.4)	0.804
	RV		−0.234 (−3.4)	−0.347 (−6.7)	0.368 (18.2)	0.373 (10.6)	0.177 (5.4)	0.804
	BV	−0.306 (−9.6)			0.381 (19.2)	0.377 (10.8)	0.171 (5.3)	0.807
$h=22$	RV				0.296 (12.6)	0.301 (8.6)	0.241 (5.3)	0.718
	RV	−0.197 (−6.3)			0.279 (11.8)	0.318 (9.0)	0.241 (5.4)	0.721
	RV		−0.234 (−3.2)	−0.174 (−2.7)	0.280 (11.8)	0.318 (8.9)	0.240 (5.4)	0.721
	BV	−0.199 (−6.2)			0.287 (11.8)	0.322 (8.8)	0.239 (5.2)	0.718
$h=66$	RV				0.218 (10.0)	0.223 (8.2)	0.299 (5.4)	0.619
	RV	−0.167 (−3.7)			0.204 (9.7)	0.237 (8.3)	0.299 (5.5)	0.621
	RV		−0.252 (−2.8)	−0.111 (−1.1)	0.205 (9.6)	0.236 (8.1)	0.298 (5.5)	0.621
	BV	−0.169 (−3.6)			0.208 (9.7)	0.239 (8.2)	0.301 (5.5)	0.619

Table A.4a: Models estimated in logs that includes signed jump information where quadratic variation has been decomposed into signed jump variation, $\%\Delta J^2 = \ln\left(1 + \Delta J_t^2 / RV_t\right)$, and its continuous component using bipower variation, BV (robust t-statistics in parentheses). Each of the four panels contains results for the forecast horizon indicated at the left. RM indicates the dependent variable, realized variance (RV) or bipower variation (BV). $\%\Delta J_t^{2+}$ and $\%\Delta J_t^{2-}$ decompose $\%\Delta J_t^2$ using an indicator variable for the sign of the difference where $\Delta J_{i,t}^{2+} = \%\Delta J_{i,t}^2 I_{[RS_t^+ - RS_t^- > 0]}$.

The impact of signed jump variation on future log volatility, results for the panel of 105 individual stocks

$$\ln \overline{RM}_{h,i,t+h} = \mu_i + \phi_J \%\Delta J^2_{i,t} + \phi_{J^+} \%\Delta J^{2+}_{i,t} + \phi_{J^-} \%\Delta J^{2-}_{i,t} + \phi_C \ln BV_{i,t} + \phi_w \ln \overline{RV}_{w,i,t} + \phi_m \ln \overline{RV}_{m,i,t} + \epsilon_{i,t+h}$$

	RM	ϕ_J	ϕ_{J^+}	ϕ_{J^-}	ϕ_C	ϕ_w	ϕ_m	R^2
$h=1$	RV				0.421 (58.0)	0.330 (35.6)	0.206 (20.5)	0.723
	RV	−0.177 (−14.6)			0.418 (59.3)	0.332 (36.4)	0.206 (20.8)	0.725
	RV		0.036 (2.0)	−0.308 (−18.0)	0.415 (59.0)	0.334 (36.7)	0.206 (20.9)	0.726
	BV	−0.171 (−13.8)			0.437 (60.8)	0.322 (34.8)	0.186 (18.7)	0.712
$h=5$	RV				0.305 (31.3)	0.357 (21.1)	0.266 (12.9)	0.777
	RV	−0.162 (−14.9)			0.303 (31.5)	0.359 (21.4)	0.266 (13.0)	0.779
	RV		−0.104 (−5.3)	−0.197 (−15.3)	0.302 (31.3)	0.360 (21.4)	0.266 (13.0)	0.779
	BV	−0.159 (−14.5)			0.317 (32.8)	0.351 (21.1)	0.249 (12.4)	0.772
$h=22$	RV				0.211 (17.3)	0.313 (16.8)	0.338 (11.2)	0.733
	RV	−0.133 (−10.2)			0.209 (17.4)	0.315 (16.9)	0.338 (11.3)	0.734
	RV		−0.088 (−4.0)	−0.161 (−10.1)	0.208 (17.3)	0.315 (16.9)	0.338 (11.3)	0.734
	BV	−0.130 (−10.1)			0.220 (17.9)	0.308 (16.2)	0.321 (10.8)	0.728
$h=66$	RV				0.140 (14.9)	0.254 (16.8)	0.395 (12.6)	0.673
	RV	−0.110 (−6.4)			0.138 (15.1)	0.255 (16.8)	0.395 (12.6)	0.674
	RV		−0.058 (−2.3)	−0.143 (−6.5)	0.138 (15.1)	0.256 (16.7)	0.395 (12.6)	0.674
	BV	−0.107 (−6.2)			0.146 (15.3)	0.247 (15.9)	0.382 (12.6)	0.668

Table A.4b: Models estimated in logs that includes signed jump information where quadratic variation has been decomposed into signed jump variation, $\%\Delta J^2 = \ln\left(1 + \Delta J^2_t / RV_t\right)$, and its continuous component using bipower variation, BV (robust t-statistics in parentheses). Each of the four panels contains results for the forecast horizon indicated at the left. RM indicates the dependent variable, realized variance (RV) or bipower variation (BV). $\%\Delta J^{2+}_{i,t}$ and $\%\Delta J^{2-}_{i,t}$ decompose $\%\Delta J^2_{i,t}$ using an indicator variable for the sign of the difference where $\%\Delta J^{2+}_{i,t} = \%\Delta J^2_{i,t} I_{[RS^+_{i,t} - RS^-_{i,t} > 0]}$. R^2 values in the final column are the average across the R^2s of the 105 individual series.

Forecasting measures of future volatility using log realized semivariances, results for the SPDR

$$\ln \overline{RM}_{h,t+h} = \mu + \phi_d^+ \ln RS_t^+ + \phi_w^+ \ln \overline{RS}_{w,t}^+ + \phi_m^+ \ln \overline{RS}_{m,t}^+ + \phi_d^- \ln RS_t^- + \phi_w^- \ln \overline{RS}_{w,t}^- + \phi_m^- \ln \overline{RS}_{m,t}^- + \epsilon_{t+h}$$

	RM	ϕ_d^+	ϕ_w^+	ϕ_m^+	ϕ_d^-	ϕ_w^-	ϕ_m^-	R^2
$h=1$	RV	0.014	−0.070	−0.067	0.442	0.410	0.220	0.789
		(0.6)	(−1.8)	(−0.9)	(20.8)	(10.2)	(2.8)	
	RS^+	−0.015	−0.114	0.014	0.476	0.468	0.122	0.786
		(−0.6)	(−2.9)	(0.2)	(21.4)	(11.4)	(1.5)	
	RS^-	0.050	−0.003	−0.132	0.409	0.332	0.295	0.722
		(1.9)	(−0.1)	(−1.5)	(17.1)	(6.9)	(3.1)	
$h=5$	RV	0.011	0.010	−0.163	0.332	0.358	0.371	0.806
		(0.5)	(0.2)	(−1.3)	(17.4)	(5.7)	(2.7)	
	RS^+	−0.005	−0.008	−0.122	0.355	0.382	0.324	0.817
		(−0.3)	(−0.1)	(−1.0)	(18.5)	(6.3)	(2.5)	
	RS^-	0.029	0.030	−0.197	0.310	0.333	0.406	0.779
		(1.3)	(0.5)	(−1.4)	(15.5)	(4.8)	(2.7)	
$h=22$	RV	0.010	−0.054	−0.270	0.236	0.362	0.555	0.726
		(0.4)	(−0.7)	(−0.9)	(11.8)	(5.2)	(1.7)	
	RS^+	0.008	−0.055	−0.280	0.241	0.365	0.571	0.736
		(0.3)	(−0.7)	(−1.0)	(12.1)	(5.3)	(1.8)	
	RS^-	0.012	−0.051	−0.255	0.232	0.359	0.531	0.710
		(0.5)	(−0.7)	(−0.9)	(11.2)	(5.2)	(1.6)	
$h=66$	RV	−0.016	−0.116	−0.192	0.184	0.347	0.534	0.627
		(−0.5)	(−1.3)	(−0.6)	(8.0)	(3.9)	(1.6)	
	RS^+	−0.019	−0.124	−0.216	0.190	0.361	0.555	0.631
		(−0.6)	(−1.3)	(−0.7)	(8.1)	(4.0)	(1.7)	
	RS^-	−0.013	−0.109	−0.167	0.178	0.334	0.512	0.622
		(−0.4)	(−1.2)	(−0.5)	(7.9)	(3.9)	(1.6)	

Table A.5a: Extended model estimated in logs where RV at all lags is decomposed into positive and negative semivariance (robust t-statistics in parentheses). Each of the four panels contains results for the forecast horizon indicated at the left. RM indicates the dependent variable, realized variance (RV), positive realized semivariance (RS^+) or negative realized semivariance (RS^-).

Forecasting measures of future volatility using log realized semivariances, results for the panel of 105 individual stocks.

$$\ln \overline{RM}_{h,i,t+h} = \mu_i + \phi_d^+ \ln RS_{i,t}^+ + \phi_w^+ \ln \overline{RS}_{w,i,t}^+ + \phi_m^+ \ln \overline{RS}_{m,i,t}^+ + \phi_d^- \ln RS_{i,t}^- + \phi_w^- \ln \overline{RS}_{w,i,t}^- + \phi_m^- \ln \overline{RS}_{m,i,t}^- + \epsilon_{i,t+h}$$

	RM	ϕ_d^+	ϕ_w^+	ϕ_m^+	ϕ_d^-	ϕ_w^-	ϕ_m^-	R^2
$h=1$	RV	0.137	0.051	0.038	0.291	0.271	0.163	0.726
		(22.7)	(5.6)	(2.6)	(37.0)	(30.4)	(11.7)	
	RS^+	0.146	0.053	0.067	0.285	0.270	0.128	0.689
		(22.6)	(5.5)	(4.4)	(33.6)	(28.8)	(8.9)	
	RS^-	0.130	0.051	0.009	0.296	0.270	0.196	0.681
		(19.7)	(4.8)	(0.5)	(36.9)	(26.7)	(12.2)	
$h=5$	RV	0.082	0.062	0.036	0.233	0.284	0.227	0.782
		(15.4)	(4.1)	(1.2)	(26.2)	(19.7)	(8.3)	
	RS^+	0.085	0.065	0.062	0.234	0.280	0.199	0.778
		(16.1)	(4.5)	(2.1)	(25.6)	(19.9)	(7.7)	
	RS^-	0.080	0.059	0.009	0.232	0.287	0.255	0.760
		(14.1)	(3.7)	(0.3)	(26.5)	(18.9)	(8.5)	
$h=22$	RV	0.050	0.033	0.051	0.174	0.265	0.286	0.739
		(7.5)	(1.7)	(0.9)	(16.6)	(13.6)	(4.6)	
	RS^+	0.052	0.039	0.061	0.173	0.260	0.276	0.744
		(8.1)	(2.0)	(1.1)	(16.3)	(13.5)	(4.5)	
	RS^-	0.048	0.028	0.041	0.175	0.270	0.295	0.725
		(7.0)	(1.4)	(0.7)	(16.9)	(13.7)	(4.7)	
$h=66$	RV	0.029	0.015	0.116	0.129	0.221	0.278	0.679
		(4.0)	(0.6)	(1.6)	(12.1)	(8.3)	(4.5)	
	RS^+	0.030	0.019	0.123	0.129	0.218	0.270	0.683
		(4.3)	(0.8)	(1.7)	(11.8)	(8.2)	(4.3)	
	RS^-	0.027	0.011	0.109	0.130	0.223	0.286	0.670
		(3.8)	(0.5)	(1.5)	(12.3)	(8.5)	(4.6)	

Table A.5b: Extended model estimated in logs where RV at all lags is decomposed into positive and negative semivariance (robust t-statistics in parentheses). Each of the four panels contains results for the forecast horizon indicated at the left. RM indicates the dependent variable, realized variance (RV), positive realized semivariance (RS^+) or negative realized semivariance (RS^-). The reported R^2 are the average across the R^2s of the 105 individual series.

E Model using Spot RV

The main results in the paper use the h-day cumulative RV, and show that negative semivariance is useful for horizons out to 66 days. However, some of the longer term predictability may be simply due to short term predictability, and so we also estimate the models on "spot" volatility, that is, the volatility on day-$t+h$, rather than the cumulative volatility. Aside from this simple change in the dependent variable, the models are unmodified and all estimation uses weighted least squares. We do not report results for $h = 1$ case since this is identical to the results reported in the paper.

Results for SPDR

Tables A.6a, A.7a and A.8a correspond to tables 2a and 3a in the main paper, and Table A.1a of this web appendix. The results in terms of the sign, relative magnitude and statistical significance are preserved although the t-stats are noticeably smaller. Table A.6a shows that negative semivariance is useful for all prediction horizons even when only predicting the variance on a single day. Table A.8a shows that positive semivariance typically produces a reduction in both positive and negative semivariance (and total realized variance), and that the coefficients on negative semivariance are large and statistically significant even at long horizons. Finally, Table A.7a shows that the effect of jumps can be detected even at the longest horizons.

Results for Panel

Tables A.6b, A.7b and A.8b correspond to tables 2b and 3b in the main paper, and Table A.1b in this web appendix. The results on the panel are broadly similar to those in the paper and show that negative semivariance plays a larger role than positive semivariance, especially at longer horizons.

HAR estimation results for the SPDR, spot volatility

$$RV_{h,t+h} = \mu + \phi_d RV_t + \phi_d^+ RS_t^+ + \phi_d^- RS_t^- + \gamma RV_t I_{[r_{t-1}<0]} + \phi_w \overline{RV}_{w,t} + \phi_m \overline{RV}_{m,t} + \epsilon_{t+h}$$

	ϕ_d	ϕ_d^+	ϕ_d^-	γ	ϕ_w	ϕ_m	R^2
$h = 5$	0.270				0.526	0.177	0.327
	(6.5)				(6.0)	(3.3)	
		−0.014	0.559		0.528	0.173	0.341
		(−0.3)	(6.1)		(6.1)	(3.3)	
		0.072	0.370	0.084	0.532	0.175	0.340
		(1.0)	(2.7)	(1.9)	(6.1)	(3.3)	
$h = 22$	0.148				0.410	0.250	0.128
	(3.2)				(4.1)	(2.5)	
		−0.105	0.439		0.401	0.245	0.138
		(−1.5)	(4.0)		(4.0)	(2.5)	
		−0.109	0.450	−0.005	0.401	0.245	0.138
		(−1.1)	(2.5)	(−0.1)	(4.0)	(2.5)	
$h = 66$	0.038				0.159	0.257	0.021
	(1.9)				(2.7)	(3.2)	
		−0.105	0.220		0.150	0.253	0.022
		(−1.7)	(2.6)		(2.6)	(3.2)	
		−0.133	0.287	−0.031	0.149	0.252	0.022
		(−1.8)	(2.5)	(−0.8)	(2.6)	(3.1)	

Table A.6a: Reference, base and asymmetric model parameter estimates using day$-t+h$ realized variance as the dependent variance (t-statistics in parentheses). Each of the three panels contains results for the forecast horizon indicated in the left most column. Each panel contains 3 models: the first model corresponds to the reference model using only realized variance, the second decomposes realized variance into positive and negative realized semivariance at the first lag, and the third specification adds an asymmetric term where the sign of the most recent daily return is used. The R^2 measure is constructed using the WLS parameter estimates and the original, unmodified data.

HAR estimation results for the panel of 105 individual stocks, spot volatility

$$RV_{h,i,t+h} = \mu_i + \phi_d RV_{i,t} + \phi_d^+ RS_{i,t}^+ + \phi_d^- RS_{i,t}^- + \gamma RV_{i,t} I_{[r_{i,t-1}<0]} + \phi_w \overline{RV}_{w,i,t} + \phi_m \overline{RV}_{m,i,t} + \epsilon_{i,t+h}$$

	ϕ_d	ϕ_d^+	ϕ_d^-	γ	ϕ_w	ϕ_m	R^2
$h=5$	0.273				0.376	0.299	0.258
	(14.3)				(11.4)	(8.9)	
		0.119	0.424		0.377	0.299	0.260
		(5.2)	(13.1)		(11.4)	(8.9)	
		0.162	0.334	0.044	0.378	0.299	0.260
		(7.1)	(13.1)	(4.6)	(11.4)	(9.0)	
$h=22$	0.161				0.242	0.453	0.165
	(7.2)				(5.6)	(8.2)	
		0.019	0.300		0.243	0.453	0.165
		(0.8)	(7.6)		(5.7)	(8.2)	
		0.048	0.237	0.032	0.244	0.454	0.166
		(2.2)	(6.3)	(3.7)	(5.7)	(8.2)	
$h=66$	0.109				0.173	0.424	0.084
	(7.7)				(5.0)	(6.7)	
		0.006	0.210		0.174	0.423	0.084
		(0.2)	(5.5)		(5.0)	(6.7)	
		0.022	0.173	0.019	0.174	0.424	0.084
		(1.1)	(6.3)	(1.5)	(5.0)	(6.7)	

Table A.6b: Reference, base and asymmetric model parameter estimates using day$-t+h$ realized variance as the dependent variance (robust t-statistics in parentheses). Each of the three panels contains results for the forecast horizon indicated in the left most column. Each panel contains 3 models: the first model corresponds to the reference model using only realized variance, the second decomposes realized variance into positive and negative realized semivariance at the first lag, and the third specification adds an asymmetric term where the sign of the most recent daily return is used. The final column reports the average of the 105 R^2s for the individual assets constructed using the WLS parameter estimates and the original, unmodified data.

The impact of signed jump variation on future spot volatility, results for the SPDR

$$RM_{h,t+h} = \mu + \phi_J \Delta J_t^2 + \phi_{J^+} \Delta J_t^{2+} + \phi_{J^-} \Delta J_t^{2-} + \phi_C BV_t + \phi_w \overline{RV}_{w,t} + \phi_m \overline{RV}_{m,t} + \epsilon_{t+h}$$

	RM	ϕ_J	ϕ_{J^+}	ϕ_{J^-}	ϕ_C	ϕ_w	ϕ_m	R^2
$h=5$	RV				0.307 (7.2)	0.504 (5.8)	0.175 (3.3)	0.334
	RV	−0.248 (−4.0)			0.302 (7.2)	0.512 (5.9)	0.173 (3.3)	0.343
	RV		−0.252 (−2.7)	−0.243 (−1.7)	0.302 (6.6)	0.511 (5.9)	0.173 (3.2)	0.343
	BV	−0.262 (−4.5)			0.303 (8.0)	0.479 (6.3)	0.164 (3.2)	0.370
$h=22$	RV				0.194 (3.7)	0.378 (3.9)	0.246 (2.5)	0.130
	RV	−0.236 (−3.2)			0.200 (3.8)	0.379 (3.8)	0.244 (2.5)	0.138
	RV		−0.348 (−3.3)	−0.057 (−0.3)	0.230 (3.8)	0.371 (3.6)	0.244 (2.5)	0.138
	BV	−0.239 (−3.4)			0.186 (3.7)	0.367 (3.7)	0.235 (2.4)	0.148
$h=66$	RV				0.042 (1.8)	0.156 (2.6)	0.257 (3.2)	0.021
	RV	−0.144 (−2.1)			0.053 (2.0)	0.153 (2.6)	0.255 (3.2)	0.022
	RV		−0.029 (−0.4)	−0.367 (−2.4)	0.018 (0.6)	0.163 (2.6)	0.255 (3.2)	0.023
	BV	−0.150 (−2.5)			0.061 (2.4)	0.140 (2.5)	0.242 (3.2)	0.023

Table A.7a: Models that include signed jump information where quadratic variation has been decomposed into signed jump variation, ΔJ^2, and its continuous component using bipower variation, BV (robust t-statistics in parentheses). Each of the three panels contains results for the forecast horizon indicated at the left. RM indicates the dependent variable, realized variance (RV) or bipower variation (BV). $\Delta J_{i,t}^{2+}$ and $\Delta J_{i,t}^{2-}$ decompose $\Delta J_{i,t}^2$ using an indicator variable for the sign of the difference where $\Delta J_{i,t}^{2+} = \Delta J_{i,t}^2 I_{[RS^+ - RS^- > 0]}$. The R^2 measure is constructed using the WLS parameter estimates and the original, unmodified data.

The impact of signed jump variation on future spot volatility, results for the panel of 105 individual stocks

$$RM_{h,i,t+h} = \mu_i + \phi_J \Delta J^2_{i,t} + \phi_{J^+} \Delta J^{2+}_{i,t} + \phi_{J^-} \Delta J^{2-}_{i,t} + \phi_C BV_{i,t} + \phi_w \overline{RV}_{w,i,t} + \phi_m \overline{RV}_{m,i,t} + \epsilon_{i,t+h}$$

	RM	ϕ_J	ϕ_{J^+}	ϕ_{J^-}	ϕ_C	ϕ_w	ϕ_m	R^2
$h=5$	RV				0.320 (13.7)	0.380 (11.5)	0.305 (9.0)	0.258
	RV	−0.151 (−7.3)			0.317 (13.8)	0.382 (11.6)	0.305 (9.1)	0.259
	RV		−0.138 (−4.2)	−0.165 (−6.1)	0.314 (12.4)	0.382 (11.6)	0.305 (9.1)	0.259
	BV	−0.129 (−7.5)			0.278 (14.5)	0.308 (11.1)	0.236 (8.6)	0.281
$h=22$	RV				0.185 (6.6)	0.246 (5.8)	0.458 (8.2)	0.164
	RV	−0.139 (−5.8)			0.183 (6.7)	0.248 (5.8)	0.458 (8.3)	0.165
	RV		−0.111 (−3.6)	−0.168 (−4.4)	0.175 (6.1)	0.249 (5.8)	0.458 (8.3)	0.165
	BV	−0.109 (−5.5)			0.160 (6.7)	0.201 (5.5)	0.363 (7.8)	0.169
$h=66$	RV				0.115 (5.9)	0.180 (5.2)	0.429 (6.7)	0.083
	RV	−0.101 (−3.5)			0.114 (6.0)	0.183 (5.2)	0.429 (6.8)	0.084
	RV		−0.049 (−1.6)	−0.156 (−3.4)	0.099 (4.6)	0.184 (5.2)	0.429 (6.8)	0.084
	BV	−0.081 (−3.2)			0.098 (5.8)	0.144 (4.9)	0.345 (6.6)	0.087

Table A.7b: Models that include signed jump information where quadratic variation has been decomposed into signed jump variation, ΔJ^2, and its continuous component using bipower variation, BV (robust t-statistics in parentheses). Each of the three panels contains results for the forecast horizon indicated at the left. RM indicates the dependent variable, realized variance (RV) or bipower variation (BV). $\Delta J^{2+}_{i,t}$ and $\Delta J^{2-}_{i,t}$ decompose $\Delta J^2_{i,t}$ using an indicator variable for the sign of the difference where $\Delta J^{2+}_{i,t} = \Delta J^2_{i,t} I_{[RS^+ - RS^- > 0]}$. The final column reports the average of the 105 R^2s for the individual assets constructed using the WLS parameter estimates and the original, unmodified data.

Forecasting measures of future volatility using spot realized semivariances, results for the SPDR

$$RM_{h,t+h} = \mu + \phi_d^+ RS_t^+ + \phi_w^+ \overline{RS}_{w,t}^+ + \phi_m^+ \overline{RS}_{m,t}^+ + \phi_d^- RS_t^- + \phi_w^- \overline{RS}_{w,t}^- + \phi_m^- \overline{RS}_{m,t}^- + \epsilon_{t+h}$$

	RM	ϕ_d^+	ϕ_w^+	ϕ_m^+	ϕ_d^-	ϕ_w^-	ϕ_m^-	R^2
$h=5$	RV	−0.048	0.068	−0.694	0.526	0.948	1.171	0.363
		(−0.9)	(0.3)	(−2.4)	(5.8)	(3.6)	(3.3)	
	RS^+	−0.039	0.032	−0.322	0.263	0.519	0.546	0.302
		(−1.2)	(0.3)	(−2.1)	(4.9)	(3.3)	(3.1)	
	RS^-	−0.009	0.035	−0.372	0.263	0.429	0.624	0.377
		(−0.3)	(0.4)	(−2.5)	(6.2)	(3.6)	(3.4)	
$h=22$	RV	−0.145	−0.264	−0.721	0.391	1.053	1.366	0.165
		(−1.7)	(−0.9)	(−1.1)	(3.5)	(3.4)	(1.8)	
	RS^+	−0.071	−0.155	−0.437	0.203	0.544	0.781	0.137
		(−1.5)	(−1.0)	(−1.3)	(3.2)	(3.3)	(1.9)	
	RS^-	−0.074	−0.109	−0.284	0.188	0.509	0.585	0.169
		(−2.0)	(−0.8)	(−0.9)	(3.7)	(3.2)	(1.6)	
$h=66$	RV	−0.127	−0.228	−0.596	0.183	0.536	1.216	0.026
		(−1.9)	(−1.4)	(−0.9)	(2.4)	(2.8)	(1.7)	
	RS^+	−0.061	−0.115	−0.221	0.074	0.286	0.523	0.018
		(−1.5)	(−1.4)	(−0.6)	(1.5)	(2.7)	(1.3)	
	RS^-	−0.066	−0.114	−0.376	0.109	0.251	0.693	0.033
		(−2.2)	(−1.5)	(−1.2)	(3.3)	(2.7)	(2.1)	

Table A.8a: Extended model where RV at all lags is decomposed into positive and negative semivariance (robust t-statistics in parentheses). Each of the three panels contains results for the forecast horizon indicated at the left. RM indicates the dependent variable, realized variance (RV), positive realized semivariance (RS^+) or negative realized semivariance (RS^-). The R^2 measure is constructed using the WLS parameter estimates and the original, unmodified data.

Forecasting measures of future volatility using spot realized semivariances, results for the panel of 105 individual stocks.

$$RM_{h,i,t+h} = \mu_i + \phi_d^+ RS_{i,t}^+ + \phi_w^+ \overline{RS}_{w,i,t}^+ + \phi_m^+ \overline{RS}_{m,i,t}^+ + \phi_d^- RS_{i,t}^- + \phi_w^- \overline{RS}_{w,i,t}^- + \phi_m^- \overline{RS}_{m,i,t}^- + \epsilon_{i,t+h}$$

	RM	ϕ_d^+	ϕ_w^+	ϕ_m^+	ϕ_d^-	ϕ_w^-	ϕ_m^-	R^2
$h=5$	RV	0.113	0.155	−0.000	0.415	0.591	0.621	0.261
		(5.0)	(2.7)	(−0.0)	(13.0)	(11.9)	(6.9)	
	RS^+	0.059	0.093	0.012	0.207	0.287	0.293	0.241
		(4.4)	(2.9)	(0.3)	(12.4)	(11.3)	(6.8)	
	RS^-	0.055	0.063	−0.012	0.208	0.304	0.328	0.209
		(5.2)	(2.4)	(−0.3)	(12.5)	(12.1)	(6.8)	
$h=22$	RV	0.013	0.013	0.101	0.291	0.463	0.834	0.167
		(0.5)	(0.2)	(0.5)	(7.7)	(7.0)	(4.2)	
	RS^+	0.008	0.002	0.051	0.150	0.233	0.419	0.150
		(0.6)	(0.1)	(0.5)	(7.5)	(7.3)	(4.0)	
	RS^-	0.005	0.010	0.050	0.141	0.230	0.414	0.144
		(0.5)	(0.3)	(0.5)	(7.8)	(6.6)	(4.2)	
$h=66$	RV	0.001	−0.007	0.257	0.204	0.350	0.608	0.085
		(0.0)	(−0.1)	(1.7)	(5.7)	(5.3)	(3.6)	
	RS^+	0.001	−0.001	0.140	0.100	0.175	0.293	0.075
		(0.1)	(−0.0)	(1.8)	(4.9)	(5.5)	(3.3)	
	RS^-	−0.000	−0.007	0.117	0.103	0.176	0.315	0.073
		(−0.0)	(−0.2)	(1.6)	(6.5)	(5.1)	(4.0)	

Table A.8b: Extended model where RV at all lags is decomposed into positive and negative semivariance (robust t-statistics in parentheses). Each of the three panels contains results for the forecast horizon indicated at the left. RM indicates the dependent variable, realized variance (RV), positive realized semivariance (RS^+) or negative realized semivariance (RS^-). The final column reports the average of the 105 R^2s for the individual assets constructed using the WLS parameter estimates and the original, unmodified data.

F Autocorrelations of RV, RS^+ and RS^-

Figure A.1 contains a plot of the first 66 autocorrelations of realized variance and positive and negative realized semivariance for the S&P 500 ETF. While the long-run behavior of the three autocorrelation series is similar, negative semivariance has uniformly larger autocorrelations than positive semivariance. The difference at the first lag is 0.19.

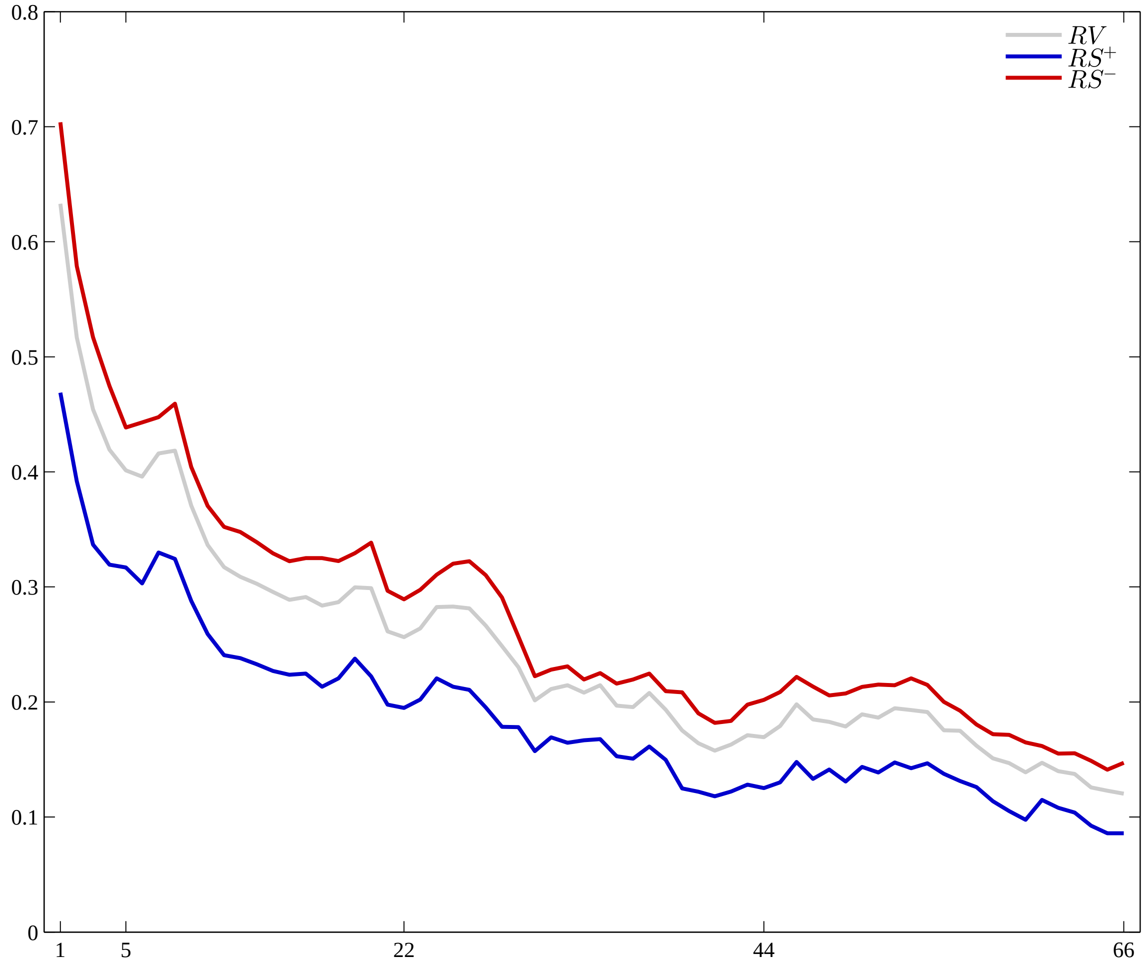

Figure A.1: The first 66 autocorrelations of realized variance and positive and negative realized semi variance for the S&P 500 ETF (SPDR).