Estimating a Local Heston Model

Bryan Ellickson, Miao Sun, Duke Whang and Sibo Yan
Department of Economics, UCLA*

September 6, 2017

Abstract

We develop a new approach to modeling the volatility of stock prices that overcomes well-known problems with using realized volatility as a proxy for integrated volatility. Using high-frequency data for SPY, an exchange-traded fund that tracks the S&P 500, we estimate the parameters of a structural model of stochastic volatility for every trading day from 2007 to 2014. The estimation is successful for 46% of trading days and 71% of five-day pools. When used in conjunction with the structural model, realized volatilities are also quite effective in detecting jumps in the price or volatility process.

Realized volatility is in theory a natural proxy for the latent integrated volatility of a stock-price process. In practice, realized volatility has not worked well, a failure usually attributed to errors in stock prices.[1] However, this conclusion is based entirely on using realized volatilities computed over the entire trading day. The outcome is very different when we estimate the Heston (1993) model of stochastic volatility using realized volatilities computed over 100-second intervals within the trading day.

Figure 1 plots the sequence of 100-second realized volatilities on January 6, 2011, for SPY (an exchange-traded fund that tracks the S&P 500). Because prices are sampled once a second, each of these 234 realized volatilities is the sum of the 100 one-second log returns contained in the block.[2] We use the Heston model to derive a stochastic difference

*This paper owes a great debt to the research reported in the UCLA Economics Department dissertations of Benjamin Hood (2011), Tin Shing Liu (2011) and Peilan Zhou (2007). We are grateful to Zhipeng Liao for helpful comments and advice.

[1]See Aït-Sahalia, Mykland and Zhang (2005) and Aït-Sahalia and Jacod (2014), Chapter 7, for a discussion of the implications of market microstructure noise for volatility estimation and the rationale for sampling prices infrequently.

[2]In the figure "time" represents the fraction of the trading day expired since the market open.

equation that serves as a structural model of the sequence of realized volatilities for each trading day. In the Heston model, the volatility $c_t := \sigma_t^2$ at time t is itself a random variable with *asymptotic mean* $\bar{c}$ and *volatility of volatility* γ.[3] A third parameter, κ, measures the speed with which the volatility process reverts to the asymptotic mean. The parameters of the stochastic difference equation allow us to estimate the parameters κ and $\bar{c}$ of the continuous-time Heston model. Asymptotic autocovariances of the stochastic difference equation enable us to estimate γ.

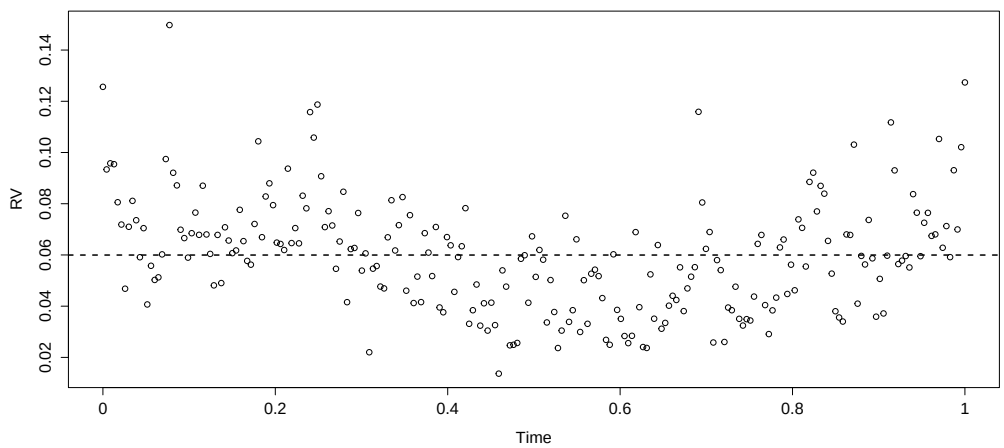

Figure 1: Realized volatilities on 1/6/2011

Adapting Bollerslev and Zhou (2002) to our setting, we use the general method of moments to estimate the parameters of the Heston model for every trading day from 2007 to 2014.[4] Our GMM estimation of the Heston model for *this trading day* yields an estimate $\widehat{\bar{c}} = 0.060$ of the asymptotic mean $\bar{c}$ of the volatility process, scaled to an annual rate. (In the figure, the asymptotic mean is indicated by the horizontal dashed line.) The annual log return of a geometric Brownian motion with this volatility would

[3] In the literature, σ_t rather than σ_t^2 is often referred to as the volatility.

[4] Bollerslev and Zhou apply their model to spot exchange rates rather than stock prices. They use daily realized volatilities that are computed with prices sampled every five minutes, reflecting the widespread belief that realized volatilities become unreliable gauges of volatility if prices are sampled too frequently. Three papers apply their methodology to stocks or the S&P 500 stock index, all using realized volatility over the entire trading day with prices sampled every five minutes. Corradi and Distaso (2006) test their model with General Electric, Intel and Microsoft stocks. Garcia, Lewis, Pastorello and Renault (2011) and Bergantini (2013) use the S&P 500 index. All of these papers assume that the parameters of the Heston model remain fixed throughout a sample period extending over many years.

have a standard deviation of $\sqrt{0.060} = 0.245$ (24.5%). The estimated speed of mean reversion for this trading day implies that on average 21% of the gap between c_t and $\bar{c}$ is eliminated within 100 seconds. In Figure 1 the realized volatility process has fallen well below the asymptotic mean by the seventh block, 700 seconds after the market opened.

We estimate this structural model separately for each of the 2,014 trading days in the sample and use the z-scores and J-statistics to set a uniform criterion for each daily estimation to be a success (a *good* day).[5] For our entire sample of 2,014 trading days, 45.8% are good, including the day portrayed in Figure 1.

No structure is imposed on the sequence of daily estimates. However, if we assume the daily parameters are slow moving over time, we can reduce the standard errors of our parameter estimates by pooling the data at the cost of introducing some bias. In constructing five-day pools, we are careful not to link the final realized volatility of one day to the initial realized volatility of the next, in keeping with our interpretation of the Heston model as a local structural model, applying only within a trading day. The only restriction we impose is that the parameters are the same for all of the trading days in the pool. When we estimate the local Heston model for five-day pools and apply the same criteria for good parameter estimates and good J-statistics we used for the daily estimates, the percentage of good pools is a remarkable 70.9%.

For these good pools, the 8-year time series for the mean reversion parameter narrows considerably relative to the time series of the daily estimates and appears not to be affected by the financial crisis in 2008–2009. In contrast, the time series for the asymptotic mean and volatility of volatility show wide variation over time, climbing to great heights during the financial crisis. The time series for the asymptotic mean resembles the VIX, the Chicago Board of Exchange (CBOE) volatility index that is constructed using an index of S&P 500 index options. The estimated asymptotic mean is a close match to the average of the 100-second realized volatilities for good pools, evidence that the estimate $\widehat{\bar{c}}$ is a good proxy for the average volatility in the pool. The median z score for $\widehat{\bar{c}}$ is an astonishing 15.1, which implies the local asymptotic mean is estimated with a median precision of $100/15.1 = 6.6\%$ of its estimated value.

[5]The Heston model requires that $\bar{c}$, κ and γ be strictly positive. We classify a parameter estimate as *good* if the ratio of the estimate to its standard error (its z score) is large enough to reject at a 5% significance level the null hypothesis that the estimate is 0 relative to the alternative that it is strictly positive. A J-statistic is *good* if the model is not rejected at a 10% significance level. The estimation for a trading day is *good* if the parameter estimates for the asymptotic mean and the speed of mean reversion are good, the estimate of γ^2 is strictly positive and the J-statistic is good.

The Heston model assumes there are no jumps in the log-price or volatility process. Fortunately, realized volatilities computed over 100-second intervals are quite effective in detecting the presence of jumps, which do exist. To adjust for the variation in mean realized volatility over the course of our 8-year sample period, we introduce *relative* realized volatilities (RRVs), 100-second realized volatilities divided by the median realized volatility of the pool to which they belong. RRVs capture the notion that what constitutes an atypically large realized volatility for a pool should depend on the average level of volatility for the pool.[6] We find that large RRVs are much more common in pools for which our estimation fails than in pools where the estimation is successful.

Incorporating jumps into our local Heston model is beyond the scope of this paper. Once jumps are taken into account, we will have a framework that allows for both the everyday processing of information and the disruption associated with large informational shocks.

Section 1 constructs our GMM framework. Section 2 describes the data. Section 3 presents the results of the estimation for SPY. We discuss jumps in Section 4. Section 5 offers some conclusions. Almost all formal proofs and several plots are in an appendix.

1 A GMM framework for the Heston model

Let $X = \log(S)$ denote a log-stock price process over the time interval $[0, \infty)$. A real-valued stochastic process $X = (X_t)_{t \geq 0}$ on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a semimartingale if it can be decomposed into the sum $A + M$ where A is an adapted cádlág process with finite variation and M is a local martingale.[7] Heuristically, the component A represents the trend in the log-price process over time and M represents random variation around the trend. A real-valued semimartingale with continuous paths takes the form

$$X_t = X_0 + \int_0^t b_s \, ds + \int_0^t \sigma_s \, dW_s \qquad (t \geq 0) \tag{1}$$

where $M_t = \int_0^t \sigma_s \, dW_s$ and W is a Wiener process. The processes $b = (b_t)_{t \geq 0}$ and $\sigma = (\sigma_t)_{t \geq 0}$ are progressively-measurable stochastic processes satisfying appropriate in-

[6] We use the median realized volatility rather than the mean because the median is less sensitive to outliers.

[7] See Aït-Sahalia and Jacod (2014), Definition 1.13, pp. 35–36. $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space and $\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}$ is the filtration.

tegrability conditions.[8]

Letting $c_t := \sigma_t^2$, in the Heston model the volatility process $c = (c_t)_{t\geq 0}$ is described by the stochastic differential equation

$$dc_t = \kappa(\bar{c} - c_t)\, dt + \gamma\sqrt{c_t}\, dB_t \qquad (t \geq 0) \tag{2}$$

The volatility process c is a mean-reverting, ergodic Markov process with *asymptotic mean* $\bar{c}$, *speed of mean reversion* κ and *volatility of volatility* γ. These parameters are assumed to be strictly positive and to satisfy the Feller condition $\gamma^2 < 2\kappa\bar{c}$, which guarantees that c_t is almost surely positive for all t.[9]

Bollerslev and Zhou (2002) distinguish between one-dimensional and multidimensional Heston volatility processes, but this is not necessary. Let

$$X_t = X_0 + \int_0^t b_s\, ds + \sum_{d=1}^{D} \int_0^t \sigma_s^d\, dW_s^d \qquad (t \geq 0) \tag{3}$$

where $(W^1, W^2, \ldots, W^D)$ is a standard D-dimensional Wiener process and assume that $c_t := \sum_{d=1}^{D} (\sigma_t^d)^2$ for all $t \geq 0$ is never zero. Assume that the stochastic process c is the solution to the stochastic differential equation

$$dc_t = \kappa(\bar{c} - c_t)\, dt + \sqrt{c_t} \sum_{d=1}^{D} \gamma^d\, dW_t^d \qquad (t \geq 0) \tag{4}$$

where the parameters γ^d are constant and $\gamma^2 := \sum_{d=1}^{D} (\gamma^d)^2$ is never zero. We claim that equations (3) and (4) imply equations (1) and (2), the usual representation of the Heston model. If we define

$$W_t = \sum_{d=1}^{D} \int_0^t \frac{\sigma_s^d}{\sqrt{c_s}}\, dW_s^d \qquad (t \geq 0) \tag{5}$$

and

$$B_t = \sum_{d=1}^{D} \int_0^t \frac{\gamma^d}{\gamma}\, dW_s^d \qquad (t \geq 0) \tag{6}$$

[8] See Aït-Sahalia and Jacod (2014), Definition 1.1, p. 10.

[9] Equation (2) is also called the Feller equation or the CIR equation (see Feller (1951), Cox, Ingersoll and Ross (1993) and Aït-Sahalia and Jacod (2014), p. 15 and p. 284.) As we verify later, $\lim_{t\to\infty} \mathbb{E}c_t = \bar{c}$ and $\lim_{t\to\infty} \mathrm{Var}(c_t) = \bar{c}\gamma^2/2\kappa$, which justifies the interpretation of $\bar{c}$ as the asymptotic mean of the volatility process and γ as the volatility of volatility.

then W and B are standard 1-dimensional Wiener processes, possibly correlated.[10] Equation (5) implies $\sqrt{c_t}\, dW_t = \sum_{d=1}^{D} \sigma_t^d\, dW_t^d$ for all $t \geq 0$. Substituting into the differential version of equation (3) yields $dX_t = b_t\, dt + \sqrt{c_t}\, dW_t$, which is the differential version of equation (1). Equation (6) implies $\gamma\, dB_t = \sum_{d=1}^{D} \gamma^d\, dW_t^d$, which when substituted into equation (4) yields equation (2).

Thus, we can reduce the D-dimensional Heston model represented by equations (3) and (4) to the one-dimensional Heston model represented by equations (1) and (2). Introducing multidimensionality could be useful in comparing the response of different assets to volatility factors, but it is irrelevant in estimating the parameters κ, $\bar{c}$ and γ of the Heston model for an asset considered in isolation.

1.1 Stochastic difference equations for QV and RV

Let $[t_0, t_N]$ denote a 6.5-hour trading day. We use the partition

$$\Pi^N := \{t_0, t_1, \ldots, t_N\} \qquad (t_0 < t_1 < \ldots < t_N)$$

to divide each trading day into $N = 23{,}400$ one-second intervals and the partition

$$\Pi^M := \{t'_0, t'_1, \ldots, t'_M\} \qquad (t_0 = t'_0 < t'_1 < \cdots < t'_M = t_N, \quad \Pi^M \subset \Pi^N)$$

to divide the trading day into $M = 234$ one-hundred-second blocks.

The integral

$$C_{t,t+h} = \frac{1}{h} \int_t^{t+h} c_s\, ds \tag{7}$$

is the (scaled) *quadratic variation (QV) over the block* $[t, t+h]$. Scaling by h gives the quadratic variation over an interval of length h the interpretation of a rate of change per unit time, just like c_t. We choose $h = 1/315576$, giving $C_{t,t+h}$ the dimensions of an annual rate.[11]

[10] See Shreve (2004), p. 226. The quadratic variation of W is the same as that of a standard Wiener process, which implies by Lévy's Theorem that W is a standard Wiener process. The same argument applies to the process B. The instantaneous correlation of the processes W and B at time t equals the differential of the quadratic covariation $[W, B]$ at time t, $d[W, B]_t = \left(\sum_{d=1}^{D} \sigma_t^d \gamma^d\right) / \sqrt{c_t}\, \gamma$.

[11] If $h = 1$, then QV over a 100-second interval is scaled to a rate per 100 seconds. Setting $h = 100/(3600 \times 24 \times 365.25)$ (the size of a 100-second interval relative to a year with 365.25 days) scales QV to an annual rate. Annual scaling helps avoid computational errors that can arise in GMM if RVs

Let

$$\widehat{C}_t = \sum_{t_i \le t} (\Delta X_{t_i})^2 \qquad (t_i \in \Pi^N)$$

denote the cumulative sum of squared log returns up to time t of the log-price process X where $\Delta X_{t_i} = X_{t_i} - X_{t_{i-1}}$ is the log return over the interval $[t_{i-1}, t_i]$. The (scaled) *realized volatility (RV)* over the block $[t, t+h]$ is

$$\widehat{C}_{t,t+h} = \frac{1}{h}(\widehat{C}_{t+h} - \widehat{C}_t) \tag{8}$$

The following proposition establishes a moment condition that is satisfied by the sequence of scaled quadratic variations. Apart from scaling, equation (9) is equivalent to equation (4) in Bollerslev and Zhou (2002).

Proposition 1. *For any pair of adjacent blocks* $[t, t+h]$ *and* $[t+h, t+2h]$

$$\mathbb{E}\left[C_{t+h,t+2h} - \alpha\, C_{t,t+h} - \beta\, \bar{c} \mid \mathcal{F}_t\right] = 0 \tag{9}$$

where $\alpha = e^{-\kappa h}$ *and* $\beta = 1 - e^{-\kappa h}$. *Because* $\kappa \in (0, \infty)$, $\alpha \in (0,1)$ *and* $\beta \in (0,1)$.

Proof. See Appendix A. □

Bollerslev and Zhou (2002) sample prices every five minutes to justify replacing QVs by RVs in this moment condition. In contrast, we assume $\widehat{C}_{t,t+h} = C_{t,t+h} + \nu_{t,t+h}$ for every block $[t, t+h]$ and make the following assumption about the measurement error.

Assumption 1. *The measurement error* $\nu_{t,t+h}$ *satisfies the condition*

$$\mathbb{E}\left[\nu_{t,t+h} \mid \mathcal{F}_s\right] = 0 \qquad \textit{for all} \quad s \le t \tag{10}$$

The interpretation is straightforward: the difference between RV and QV is mean zero and uncorrelated with what happens before the beginning of the block.

Proposition 2. *If the measurement errors satisfy Assumption 1, then for any pair of adjacent blocks* $[t, t+h]$ *and* $[t+h, t+2h]$

$$\mathbb{E}\left[\widehat{C}_{t+h,t+2h} - \alpha\, \widehat{C}_{t,t+h} - \beta\, \bar{c} \mid \mathcal{F}_t\right] = 0 \tag{11}$$

are too small.

where $\alpha = e^{-\kappa h}$ *and* $\beta = 1 - e^{-\kappa h}$.

Proof. Proposition 1 is equivalent to the assertion that the sequence of quadratic variations satisfies the stochastic difference equation

$$C_{t+h,t+2h} = \alpha\, C_{t,t+h} + \beta\, \bar{c} + \eta_{t,t+2h} \tag{12}$$

where the error term $\eta_{t,t+2h} := C_{t+h,t+2h} - \alpha C_{t,t+h} - \beta\bar{c}$ has conditional mean

$$\mathbb{E}\left[\eta_{t,t+2h} \mid \mathcal{F}_t\right] = 0 \tag{13}$$

Substituting $C_{t,t+h} = \widehat{C}_{t,t+h} - \nu_{t,t+h}$ and $C_{t+h,t+2h} = \widehat{C}_{t+h,t+2h} - \nu_{t+h,t+2h}$ into equation (12),

$$\widehat{C}_{t+h,t+2h} = \alpha\, \widehat{C}_{t,t+h} + \beta\, \bar{c} + \varepsilon_{t,t+2h} \tag{14}$$

where $\varepsilon_{t,t+2h} := \nu_{t+h,t+2h} - \alpha\, \nu_{t,t+h} + \eta_{t,t+2h}$. Equation (11) is equivalent to

$$\mathbb{E}\left[\varepsilon_{t,t+2h} \mid \mathcal{F}_t\right] = \mathbb{E}\left[\nu_{t+h,t+2h} - \alpha\, \nu_{t,t+h} + \eta_{t,t+2h} \mid \mathcal{F}_t\right] = 0$$

which holds because of Assumption 1 and equation (13).[12] □

Because $\alpha = 1 - \beta$, equation (11) can be rewritten in the form

$$\mathbb{E}\left[\widehat{C}_{t+h,t+2h} - \widehat{C}_{t,t+h} \mid \mathcal{F}_t\right] = \beta\left(\bar{c} - \mathbb{E}\left[\widehat{C}_{t,t+h} \mid \mathcal{F}_t\right]\right) \tag{15}$$

Conditioned on $\mathcal{F}_t$, β is the average fraction of the gap between the conditional expectation of $\widehat{C}_{t,t+h}$ and the asymptotic mean $\bar{c}$ eliminated over the interval $[t+h, t+2h]$. Thus, β has a natural interpretation as a rate of mean reversion.

In equation (14) the regressor $\widehat{C}_{t,t+h}$ is necessarily endogenous: $\widehat{C}_{t,t+h}$ is correlated with $\eta_{t,t+2h}$ which in turn is correlated with the error term $\varepsilon_{t,t+2h}$. To deal with the endogeneity problem, we introduce instrumental variables (IVs). IVs should be both valid and relevant: orthogonal to the error term and correlated with the endogenous regressors. The lagged realized volatility and its powers satisfy both conditions: they are orthogonal to the error term under Assumption 1 and equation (13), and realized volatility is serially autocorrelated. The number of IVs should be no less than two, the

[12] Note that $\mathcal{F}_t \in \mathbb{F}$ where $\mathbb{F}$ is the filtration generated by W and B.

number of parameters.

In the empirical analysis, we use the lagged RV, its square root and its fourth root as the instruments, yielding the orthogonality conditions

$$\mathbb{E}\left[\widehat{C}_{t+h,t+2h} - (1-\beta)\widehat{C}_{t,t+h} - \beta\bar{c}\right] = 0 \tag{16}$$

$$\mathbb{E}\left[\left(\widehat{C}_{t+h,t+2h} - (1-\beta)\widehat{C}_{t,t+h} - \beta\bar{c}\right)\widehat{C}_{t-h,t}\right] = 0 \tag{17}$$

$$\mathbb{E}\left[\left(\widehat{C}_{t+h,t+2h} - (1-\beta)\widehat{C}_{t,t+h} - \beta\bar{c}\right)(\widehat{C}_{t-h,t})^{1/2}\right] = 0 \tag{18}$$

$$\mathbb{E}\left[\left(\widehat{C}_{t+h,t+2h} - (1-\beta)\widehat{C}_{t,t+h} - \beta\bar{c}\right)(\widehat{C}_{t-h,t})^{1/4}\right] = 0. \tag{19}$$

These orthogonality conditions can be used to estimate the asymptotic mean $\bar{c}$ and speed of mean reversion β (or κ) of the Heston model using GMM. Second-moment conditions allow us to estimate the volatility-of-volatility parameter as well.

1.2 Second moments of QV and RV

The GMM framework of Section 1.1 yields estimates of κ (or β) and $\bar{c}$, but not γ. In the Heston model the volatility c_t has mean

$$\mathbb{E}c_t = e^{-\kappa t}\, c_0 + (1 - e^{-\kappa t})\, \bar{c} \tag{20}$$

and variance[13]

$$\mathrm{Var}(c_t) = \frac{\gamma^2}{\kappa}(e^{-\kappa t} - e^{-2\kappa t})\, c_0 + \frac{\gamma^2}{2\kappa}(1 - 2e^{-\kappa t} + e^{-2\kappa t})\, \bar{c} \tag{21}$$

As $t \to \infty$ the effects of the initial value wear off, and the mean and the variance approach limits that are independent of c_0: $\lim_{t\to\infty} \mathbb{E}c_t = \bar{c}$ and $\lim_{t\to\infty} \mathrm{Var}(c_t) = \gamma^2\bar{c}/2\kappa$. We will prove that the variance of QV for an interval of length h also converges as $t \to \infty$ to a limit that depends only on h and the parameters of the Heston model:

$$\lim_{t\to\infty} \mathrm{Var}(C_{t,t+h}) = \left(\frac{\gamma}{\kappa h}\right)^2 \left(h - \frac{\beta}{\kappa}\right) \bar{c}$$

This provides a candidate for a moment condition that will allow us to estimate γ^2:

[13] These equations correspond to equations (4.4.36) and (4.4.38) respectively in Shreve (2004), which are derived for the CIR model of interest rates.

if the speed of mean reversion is rapid, sample second moments should be close to their asymptotic limits. However, replacing QVs by RVs again raises the issue of error in variables. The asymptotic variance is not robust to measurement error. For that reason, we also derive formulas for asymptotic autocovariances of lag $j \geq 1$:

$$V_j := \lim_{t\to\infty} \mathbb{E}\left[\left(C_{t,t+h} - \bar{c}\right)\left(C_{t+jh,t+(j+1)h} - \bar{c}\right)\right]$$

Under a strengthened assumption about the measurement error, all asymptotic autocovariances of lag $j \geq 1$ are robust to measurement error, but the asymptotic variance is not. We use moment conditions for the asymptotic autocovariances of lag 1 and 2 to estimate γ^2 using GMM.

We begin by establishing that the conditional variance of $C_{t,t+h}$ is a linear function of c_t and $\bar{c}$. Apart from scaling, equation (22) is equivalent to equation (8) in Bollerslev and Zhou (2002).

Lemma 1. *For any block* $[t, t+h]$

$$Var(C_{t,t+h} \mid \mathcal{F}_t) = \left(\frac{\gamma}{\kappa h}\right)^2 (a_2\, c_t + b_2\, \bar{c}) \tag{22}$$

where a_2 *and* b_2 *are constants that depend only on* h *and the parameters of the Heston model.*

Proof. See Appendix B. □

Let $V_0 := \lim_{t\to\infty} \mathrm{Var}(C_{t,t+h})$ denote the asymptotic variance.

Proposition 3. *The asymptotic variance for blocks of length* h *is*

$$V_0 = \lim_{t\to\infty} Var(C_{t,t+h}) = \left(\frac{\gamma}{\kappa h}\right)^2 \left(h - \frac{\beta}{\kappa}\right) \bar{c} \tag{23}$$

Proof. See Appendix C. □

To compare sizes of the asymptotic variances of $C_{t,t+h}$ and c_t, note that

$$\begin{aligned}\lim_{t\to\infty} \mathrm{Var}(C_{t,t+h}) &= \frac{2}{\kappa h^2}\left(h - \frac{1}{\kappa} + \frac{1}{\kappa}e^{-\kappa h}\right)\frac{\gamma^2 \bar{c}}{2\kappa} \\ &= \left(\frac{2}{\kappa h} - \frac{2}{(\kappa h)^2} + \frac{2}{(\kappa h)^2}e^{-\kappa h}\right) \lim_{t\to\infty} \mathrm{Var}(c_t)\end{aligned}$$

In our case $h = 1/315576$ and the median estimate of κ for pools is 3.32×10^4 (see Table 3), which implies

$$\lim_{t\to\infty} \mathrm{Var}(C_{t,t+h}) \approx 0.97 \lim_{t\to\infty} \mathrm{Var}(c_t)$$

The asymptotic variance of $C_{t,t+h}$ accounts for a large proportion of the asymptotic variance of c_t.

Lemma 2. *For any interval $[t, t+2h]$ the conditional variance of the error $\eta_{t,t+2h} := C_{t+h,t+2h} - \alpha\, C_{t,t+h} - \beta\, \bar{c}$ is*

$$Var\left(\eta_{t,t+2h} \mid \mathcal{F}_t\right) = \left(\frac{\gamma}{\kappa h}\right)^2 (a_3\, c_t + b_3\, \bar{c}) \tag{24}$$

where a_3 and b_3 are constants that depend only on h and the parameters of the Heston model.

Proof. See Appendix D. □

Proposition 4. *The asymptotic variance of the error $\eta_{t,t+2h}$ is*

$$\lim_{t\to\infty} Var\left(\eta_{t,t+2h}\right) = \left(\frac{\gamma}{\kappa h}\right)^2 (a_3 + b_3)\, \bar{c} \tag{25}$$

where

$$a_3 + b_3 = h - \frac{1}{\kappa} + \left(h + \frac{1}{\kappa}\right) e^{-2\kappa h} \tag{26}$$

Proof. See Appendix E. □

To compare the size of the asymptotic variance of $\eta_{t,t+2h}$ to that of the asymptotic variance of c_t, rewrite equation (25) to express the former as a fraction of the latter:

$$\begin{aligned}\lim_{t\to\infty} \mathrm{Var}(\eta_{t,t+2h}) &= 2\left(\frac{a_3 + b_3}{\kappa h^2}\right)\frac{\gamma^2 \bar{c}}{2\kappa} \\ &= \left(\frac{2}{\kappa h} - \frac{2}{(\kappa h)^2} + \left(\frac{2}{\kappa h} + \frac{2}{(\kappa h)^2}\right) e^{-2\kappa h}\right) \lim_{t\to\infty} \mathrm{Var}(c_t)\end{aligned}$$

If $h = 1/315576$ and $\kappa = 3.32 \times 10^4$,

$$\lim_{t\to\infty} \mathrm{Var}(\eta_{t,t+2h}) \approx 0.18 \lim_{t\to\infty} \mathrm{Var}(c_t)$$

The asymptotic variance of the error term is about 18% of the asymptotic variance of c_t.

Proposition 5. *The asymptotic second moments of the quadratic variation $C_{t,t+h}$ in the Heston model with blocks of length h are*

$$\begin{aligned} V_0 &= \left(\frac{\gamma}{\kappa h}\right)^2 \left(h - \frac{\beta}{\kappa}\right) \bar{c} \\ V_1 &= \left(\frac{\gamma}{\kappa h}\right)^2 \frac{\beta^2}{2\kappa} \bar{c} \\ V_j &= (1-\beta)\, V_{j-1} \qquad (j \geq 2) \end{aligned}$$

Proof. See Appendix F. □

Corollary 1. *The asymptotic autocorrelations $R_j := V_j/V_0$ of $C_{t,t+h}$ for lags $j \geq 0$ are given by*

$$\begin{aligned} R_0 &= 1 \\ R_1 &= \frac{-\beta^2}{2\left(\log(1-\beta) + \beta\right)} \\ R_j &= (1-\beta)\, R_{j-1}, \qquad j \geq 2 \end{aligned}$$

Proof. See Appendix G. □

The asymptotic autocorrelations of $C_{t,t+h}$ for all lags depend only on β. In particular, the median estimate $\widehat{\beta} = 0.10$ (see Table 3, 5-day pools) implies that $R_1 = 0.93$, $R_2/R_1 = 0.9$ and $R_3/R_1 = 0.81$, implications we will compare with our data.

We now consider the effect of measurement error on the second moments when realized volatility is used in place of quadratic variation. As before we assume that $\widehat{C}_{t,t+h} = C_{t,t+h} + \nu_{t,t+h}$ where $\nu_{t,t+h}$ is measurement error, but we impose an assumption on the error stronger than Assumption 1.

Assumption 2. *The measurement errors $\nu_{t,t+h}$ are independent of the volatility process c and i.i.d. with mean 0 and variance $\nu^2 > 0$.*

Let

$$\widehat{V}_j := \lim_{t\to\infty} \mathbb{E}\left[\left(\widehat{C}_{t,t+h} - \bar{c}\right)\left(\widehat{C}_{t+jh,t+(j+1)h} - \bar{c}\right)\right]$$

denote the asymptotic autocovariance of $\widehat{C}_{t,t+h}$ for lag $j \geq 0$.

Proposition 6. *Under Assumption 2,* $\widehat{V}_0 = V_0 + \nu^2$ *and* $\widehat{V}_j = V_j$ *for* $j \geq 1$.

Proof. See Appendix H. □

To summarize, the asymptotic variance V_0 is vulnerable to measurement error, but the asymptotic autocovariances V_j for $j \geq 1$ are not. We will use V_1 and V_2 to estimate the volatility-of-volatility parameter γ^2. Substituting $\kappa h = -\log(1-\beta)$ into the formulas of Proposition 5 yields

$$V_1 = -\left(\frac{\beta^2 h\bar{c}}{2(\log(1-\beta))^3}\right)\gamma^2 \qquad \text{and} \qquad V_2 = -\left(\frac{(1-\beta)\beta^2 h\bar{c}}{2(\log(1-\beta))^3}\right)\gamma^2$$

and hence the moment conditions

$$\mathbb{E}\left[\left(\widehat{C}_{t,t+h} - \bar{c}\right)\left(\widehat{C}_{t+h,t+2h} - \bar{c}\right) + \left(\frac{\beta^2 h\bar{c}}{2(\log(1-\beta))^3}\right)\gamma^2\right] = 0 \tag{27}$$

$$\mathbb{E}\left[\left(\widehat{C}_{t,t+h} - \bar{c}\right)\left(\widehat{C}_{t+2h,t+3h} - \bar{c}\right) + \left(\frac{(1-\beta)\beta^2 h\bar{c}}{2(\log(1-\beta))^3}\right)\gamma^2\right] = 0 \tag{28}$$

We assume that the usual assumptions required for nonlinear GMM estimation are satisfied (see Hall (2005), Chapter 3.) The six moment conditions (16)–(19), (27) and (28) are rescaled to give all the same order of magnitude. Because the error terms are serially correlated, we use a heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimator with a Bartlett kernel (see Newey and West (1987)). The model was estimated using the R gmm package (see Chaussé (2010)).

Bollerslev and Zhou (2002) employ a different second-moment condition. Using our notation, their second-moment condition is

$$\mathbb{E}[C^2_{t+h,t+2h} - HC^2_{t,t+h} - IC_{t,t+h} - J \mid \mathcal{F}_t] = 0 \tag{29}$$

where H, I and J are constants that depend only on κ, $\bar{c}$ and γ.[14] They replace the daily quadratic variations with daily realized volatilities. Moment condition (16) and the unconditional version of moment-condition (29) (with QVs replaced by RVs) constitute the basic moment conditions they use for estimation. As do we, they use functions of

[14] In their setting of daily quadratic variations, $h = 1$ which scales quadratic variation to a day. The functions expressing H, I and J in terms of κ, $\bar{c}$ and γ, which they explicitly derive, are quite complicated.

RVs lagged one period as instruments, yielding a total of six moment conditions, three based on the first-moment condition and three based on the second-moment condition.

2 Constructing the data set

The data for SPY come from the NYSE Euronext Trades and Quotes (TAQ) database, available at the Wharton WRDS website. The data include price, number of shares and time of each transaction (to the nearest second). We begin at 9:30 AM ET and end at 4:00 PM ET, coinciding with the market open and close. We use the TAQ "condition codes" to remove trades from the dataset that were canceled or otherwise flagged as illegitimate, and we sort prices by time stamp.

SPY averaged 17.2 trades per second over our eight year sample period. Nevertheless, 23% of time stamps (*inactive time stamps*) have no trade. Because active time stamps have many trades, we need a procedure to select a single price for each active time stamp. We use the *median share-price*, the median price per share treating each share traded as a separate observation. For inactive time stamps, we assign the median share-price associated with the most recent time stamp with a trade. This yields prices for each second used to compute the realized volatility for each 100-second block.[15]

Apart from using the condition codes to remove illegitimate trades, the only data filtering is the elimination of "bouncebacks," which Aït-Sahalia and Jacod (2014) describe as follows:

> Bouncebacks are price observations that are either higher or lower than the sequence of prices that both immediately precede and follow them. Such prices generate a log-return from one transaction to the next that is large in magnitude and is followed immediately by a log-return of the same magnitude but of the opposite sign, so that the price returns to its starting level before that particular transaction.[16]

Figure 2.7 in Aït-Sahalia and Jacod (2014) provides a clear illustration of a bounceback for SPY on April 8, 2009.

[15] The data keep improving. The TAQ data with time measured in seconds runs only to 2014, the end of our sample period. In 2014 WRDS archived this data set, replaced by TAQ data that measures time to the nearest millisecond. Millisecond data will allow much more accurate sampling of transactions once a second.

[16] Aït-Sahalia and Jacod (2014), p. 74.

We eliminate bouncebacks using an influence statistic that measures the impact of removing a price on the realized volatility over an interval centered on the time stamp of the transaction. Let t_i, t_j and t_k be active time stamps with $t_i < t_j < t_k$. Define the *influence statistic* I_j to be the change in the unscaled realized volatility $\widehat{C}_{t_k} - \widehat{C}_{t_i}$ over $[t_i, t_k]$ if the median log share-price associated with the active second $t_j \in (t_i, t_k)$ is eliminated from the calculation.[17]

$$\begin{aligned} I_j &= (X_{t_{j+1}} - X_{t_j})^2 + (X_{t_j} - X_{t_{j-1}})^2 - (X_{j+1} - X_{j-1})^2 \\ &= -2(X_{t_{j+1}} - X_{t_j})(X_{t_j} - X_{t_{j-1}}) \end{aligned} \tag{30}$$

The influence statistic I_j is

- 0 if either $X_{t_j} = X_{t_{j-1}}$ or $X_{t_j} = X_{t_{t+1}}$
- strictly negative if $X_{t_{j-1}} < X_{t_j} < X_{t_{j+1}}$ or $X_{t_{j+1}} < X_{t_j} < X_{t_{j-1}}$
- strictly positive if $X_{t_j} < \min\{X_{t_{j-1}}, X_{t_{j+1}}\}$ or $X_{t_j} > \max\{X_{t_{j-1}}, X_{t_{j+1}}\}$

We declare X_{t_j} to be a *bounceback* if its influence statistic I_j is positive and "large" in the sense that (1) I_{t_j} is greater than 20% of the unscaled realized volatility for an interval $[t_i, t_k]$ containing 201 active time stamps centered on t_j or (2) I_{t_j} exceeds 5% of the unscaled realized volatility for the entire trading day. The procedure was applied iteratively until no median share-price met the criterion for a bounceback. For SPY the average number of bouncebacks eliminated per day ranged from a low of 11 in 2009 to a high of 17 in 2007, a tiny fraction of the total number of trades in a day. Eliminating bouncebacks caused little reduction in the number of active time stamps because we were usually able to replace the bounceback with the median share-price of the remaining trades with the same time stamp.[18]

We estimate the parameters of the Heston model for individual trading days and for five-day pools. For each daily estimate, the sample moments sum over the 232 pairs of blocks $\{[t, t+h], [t+h, t+2h]\}$, starting from blocks 2 and 3 (the first block $[t_0, t_1]$ is excluded because lagged RV is used as an instrument). For the five-day pools, we start each day from blocks 2 and 3, summing over a total of $232 \times 5 = 1160$ pairs of blocks.

[17] Letting $a = X_{t_{j+1}} - X_{t_j}$ and $b = X_{t_j} - X_{t_{j-1}}$, $a^2 + b^2 - (a+b)^2 = -2ab$.

[18] For a detailed description of the construction of the data set, see Whang (2012).

3 Estimating the model

We turn now to the estimation results for SPY. In reporting these results, it is useful to adopt a rough classification of trading days or five-day pools into two classes: good days or pools where we deem the estimation a success and bad days or pools where we regard the estimation as a failure. The classification is not meant to be definitive but rather to provide a useful way of organizing the reporting of the results. Good days or pools need to be examined closely to access whether the estimation is really good and similarly bad pools need to be examined closely to access the extent to which the estimation is really bad.

The Heston model requires that κ, $\bar{c}$ and γ be strictly positive and $\beta \in (0, 1)$. A parameter estimate is *good* if the ratio of the estimate to its standard error (its z score) is large enough to reject at a 5% significance level the null hypothesis that the estimate is 0 relative to the alternative it is strictly positive. A J-statistic is *good* if the model is not rejected at a 10% significance level. The estimation for a trading day or a pool is *good* if the parameter estimates for β and $\bar{c}$ are good, the estimate of γ^2 is strictly positive and the J-statistic is good.

Table 1 summarizes the overall performance of the estimation for individual trading days and for 5-day pools. Pooling increases the success rate dramatically from 45.8% of trading days to 70.9% of 5-day pools. The percentage of good parameter estimates increased for all three parameters. The percentage of good J-statistics declined slightly, which is not surprising: we are forcing the parameters to be the same for all 5 days of the pool.

Table 1: Overall performance: 2007–2014

Classification	days (%)	pools (%)
Good days/pools	923 (45.8%)	286 (70.9%)
Bad days/pools	1091 (54.2%)	117 (29.1%)
Good $\widehat{\beta}$	967 (48.0%)	304 (75.4%)
Good $\widehat{\bar{c}}$	1495 (74.2%)	352 (87.3%)
Good $\widehat{\gamma^2}$	527 (26.2%)	252 (62.5%)
Good J-statistic	1899 (94.3%)	360 (89.3%)

Table 2: Estimates of β, $\bar{c}$ and γ^2 (good days and good pools)

	days			pools		
	$\widehat{\beta}$	$\widehat{\bar{c}}$	$\widehat{\gamma^2}(\times 10^3)$	$\widehat{\beta}$	$\widehat{\bar{c}}$	$\widehat{\gamma^2}(\times 10^3)$
median	0.19	0.081	1.89	0.10	0.081	1.11
lower quartile	0.14	0.051	0.97	0.07	0.053	0.64
upper quartile	0.27	0.14	3.82	0.13	0.15	2.31
median s.e.	0.07	0.007	1.11	0.031	0.006	0.52
mean	0.22	0.14	4.07	0.11	0.14	1.97
median z score	2.62	11.7	1.68	3.16	15.1	2.33

Table 2 describes the parameter estimates for good days and good pools. Pooling has little impact on the estimates of $\bar{c}$: the mean, median, lower and upper quartiles are almost the same. Standard errors and z scores improve slightly. The impact of pooling on the estimates of β and γ^2 is much greater. The median speed of mean reversion β fell from 0.19 to 0.10, and the interquartile range (the spread between the upper and lower quartiles) fell from 0.13 to 0.06. The median estimate of γ^2 fell from 1.89×10^3 to 1.11×10^3 and the interquartile range fell from 2.85×10^3 to 1.67×10^3. Standard errors and z scores for all parameters improved substantially. Every good day and every good pool satisfied the Feller condition, $\gamma^2 < 2\kappa\bar{c}$.

Figures 2 and 3, which plot the estimates of β for good days and good pools respectively, dramatically portray the impact of pooling. The dates on the horizontal axis indicate the beginning of each calendar year. Neither plot shows any evidence of a trend over time, but the daily estimates are highly variable and sometimes unrealistically high. In contrast, the estimates for good pools are tightly clustered around a plausible speed of mean reversion.

Table 3 interprets the speed of mean reversion in various ways. The top row copies the median, lower quartile and upper quartile estimates of β from Table 2. From $\kappa = -\log(1-\beta)/h$ we obtain estimates for κ, reported on the bottom row. Using these estimates of κ, the top panel uses $\beta_h = 1 - e^{-\kappa h}$ to compute β_h for various lengths h. The median estimate implies that 10% of the gap between volatility and its asymptotic mean is eliminated in 100 seconds, 27% in 5 minutes, 47% in 10 minutes and 85% within

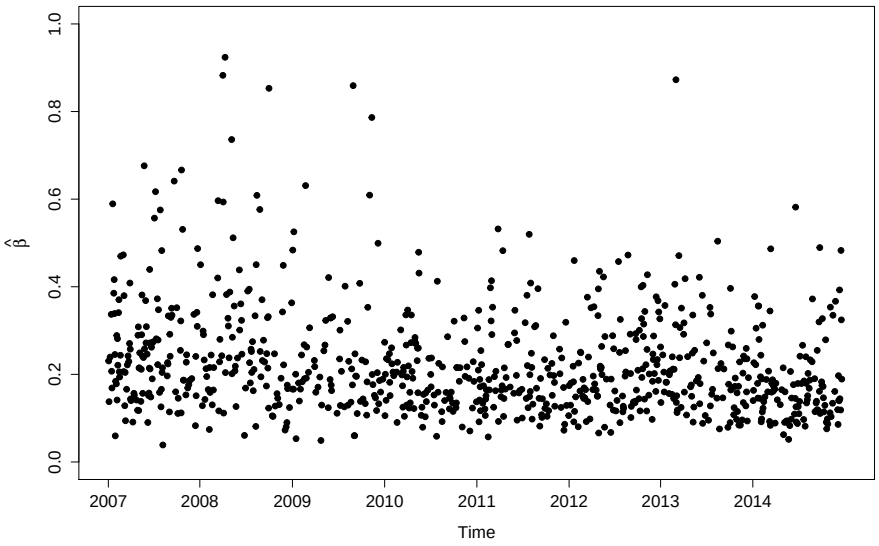

Figure 2: Time series for $\widehat{\beta}$ (good days)

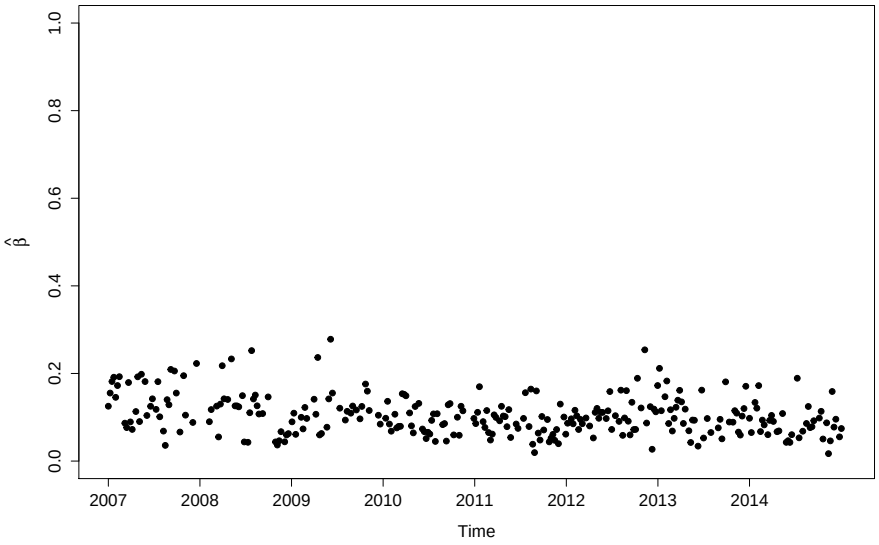

Figure 3: Time series for $\widehat{\beta}$ (good pools)

a half hour. Even the lower quartile estimates suggest mean reversion is rapid.

Table 3: Estimates of mean reversion for good pools

	median	lower quartile	upper quartile
$\widehat{\beta}$: 100 seconds	0.10	0.07	0.13
$\widehat{\beta}$: 5 minutes	0.27	0.20	0.34
$\widehat{\beta}$: 10 minutes	0.47	0.35	0.57
$\widehat{\beta}$: 30 minutes	0.85	0.73	0.92
$\widehat{\kappa}$ (rate per year)	3.32×10^4	2.29×10^4	4.39×10^4

Figures 4 and 5 plot the estimates of $\bar{c}$ and γ^2 for good pools. Three pairs of vertical lines mark periods of interest. The pair of lines straddling the beginning of 2009 highlight the mortgage-financing crisis, starting with the declaration of bankruptcy by Fannie Mae and Freddie Mac early in September 2008 and ending in late June 2009 when our estimate of $\bar{c}$ has returned to roughly the level it had before the crisis. The pair of vertical lines in the middle of 2010 mark a surge in volatility associated with the sovereign debt crisis in the Eurozone. The Flash Crash occurred near the beginning of this period and the Deepwater Horizon explosion near the end. The final pair of vertical lines identify a third period of high volatility stretching from early August to the end of September 2011. Standard and Poor's credit rating of the U.S. Government debt was lowered shortly before the start of this period.

The median estimates $\widehat{\bar{c}} = 0.081$ and $\widehat{\gamma^2} = 1.11 \times 10^3$ for good pools correspond to a standard deviation $\sqrt{0.081} = 0.29$ (29%) of annualized log returns with a volatility of volatility $\sqrt{1.11 \times 10^3} = 33.6$. At the peak of the financial crisis both parameters exceeded their medians by an order of magnitude, testimony to the severity of the crisis. Figure 6 plots the estimate of $\bar{c}$ versus the sample mean of the 100-second RVs for good pools, along with a 45° line: the estimate $\widehat{\bar{c}}$ matches closely the mean 100-second RV over a range that spans two orders of magnitude.[19] Figure 7 plots the well-known VIX index for our sample period, which resembles the time series for $\widehat{\bar{c}}$.[20]

[19]Note that both axes are scaled logarithmically.

[20]The VIX is the Chicago Board of Exchange (CBOE) volatility index, which is constructed using a range of S&P 500 index options.

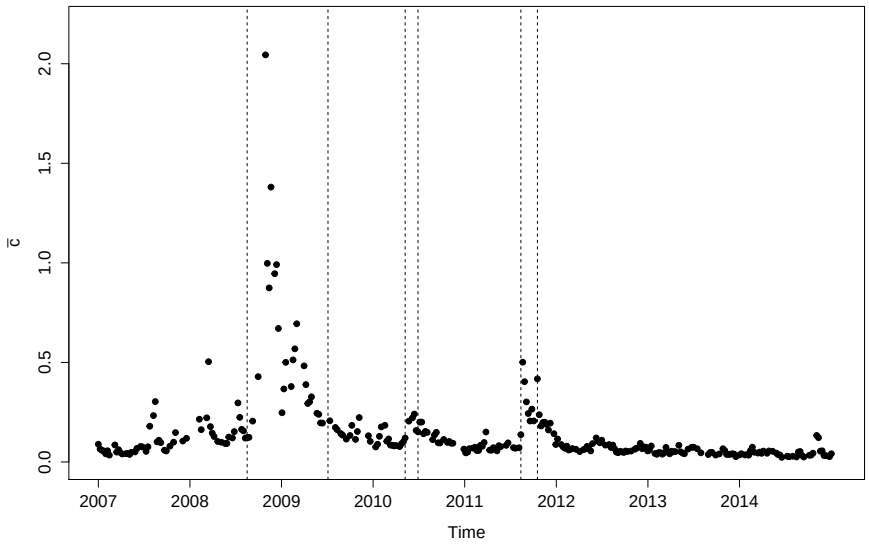

Figure 4: Time series for $\widehat{\overline{c}}$ (good pools)

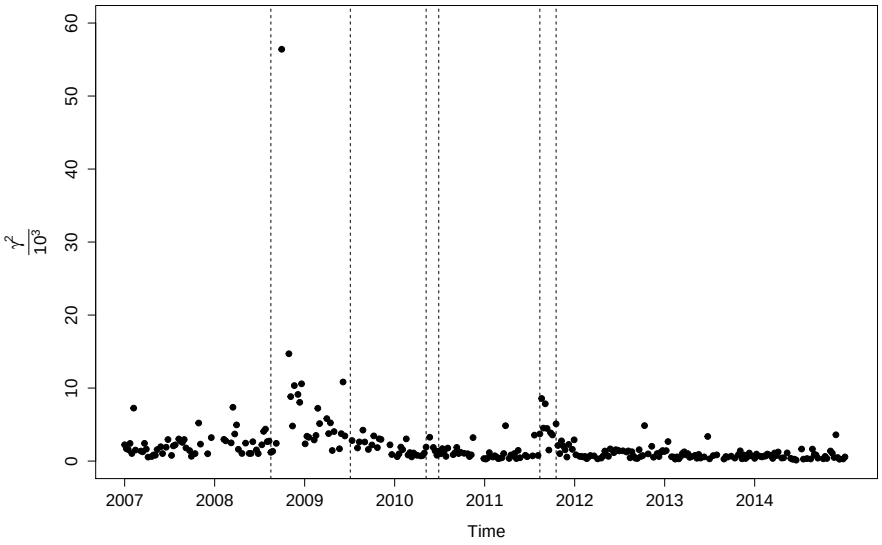

Figure 5: Time series for $\widehat{\gamma^2}$ (good pools)

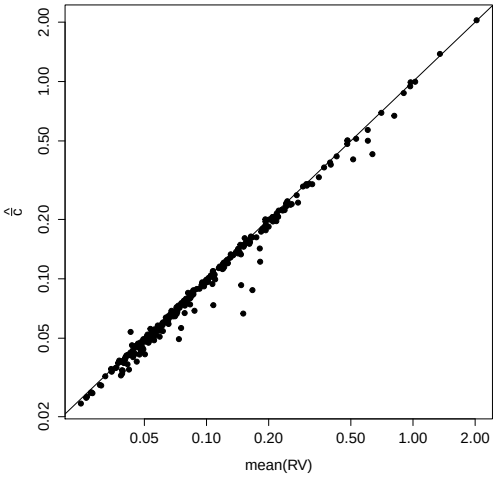

Figure 6: $\widehat{\bar{c}}$ versus mean realized volatility (good pools)

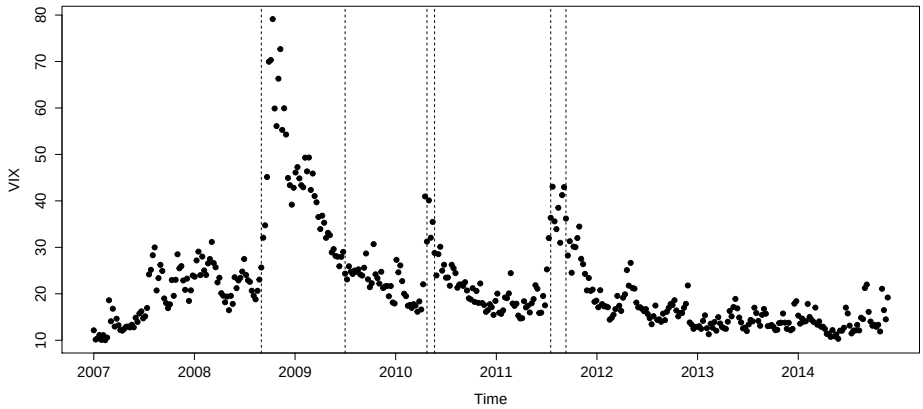

Figure 7: Volatility index for the S&P 500 (VIX)

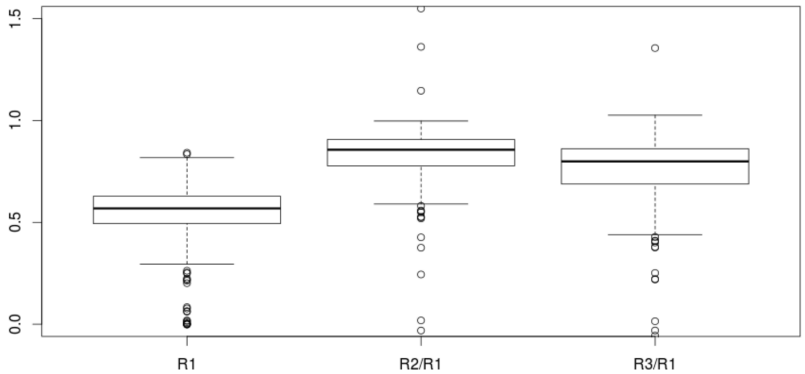

Figure 8: Box plots of $\widehat{R}_1$, $\widehat{R}_2/\widehat{R}_1$ and $\widehat{R}_3/\widehat{R}_1$ (good pools)

To check the validity of our claim that autocovariances of all lags are robust to measurement error, we examine the sample autocorrelations for good pools. Corollary 1 provides formulas for the asymptotic autocorrelations of $C_{t,t+h}$ that depend only on the parameter β. If $\beta = 0.10$ (the sample median for good pools), then the formulas imply $R_1 = 0.93$, $R_2/R_1 = 0.90$ and $R_3/R_1 = 0.81$. Figure 8 displays box plots of the sample autocorrelation $\widehat{R_1}$ and the ratios $\widehat{R}_2/\widehat{R}_1$ and $\widehat{R}_3/\widehat{R}_1$ for good pools. The horizontal line inside each box represents the median. The upper and lower edges of the box represent the upper and lower quartiles of the distribution. The isolated points above the "whiskers" are outliers. The median for $\widehat{R}_1$ is 0.57, far below the median 0.93 predicted by the formula, exactly what we expect because R_1 has the variance in the denominator. Proposition 6 asserts that measurement error increases the variance but not the autocovariance for lag 1, which lowers the sample autocorrelation below the prediction of the formula. On the other hand, the medians for $\widehat{R}_2/\widehat{R}_1$ and $\widehat{R}_3/\widehat{R}_1$ are 0.87 and 0.80, close to the predicted values 0.90 and 0.81, just as expected.

Moment conditions (27) and (28) match the sample autocovariances $\widehat{V}_1$ and $\widehat{V}_2$ to the estimates implied by the formulas we derived for the asymptotic covariances V_1 and

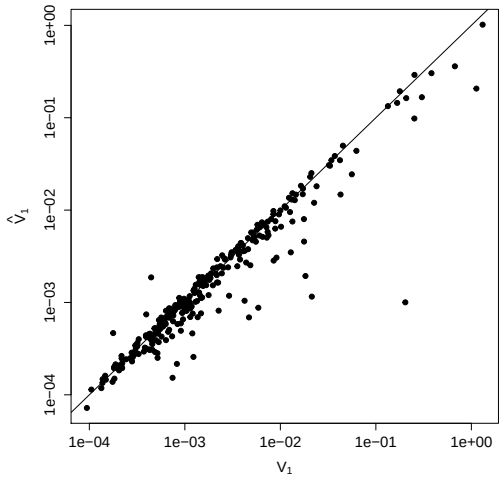

Figure 9: V_1 versus $\widehat{V}_1$ (good pools)

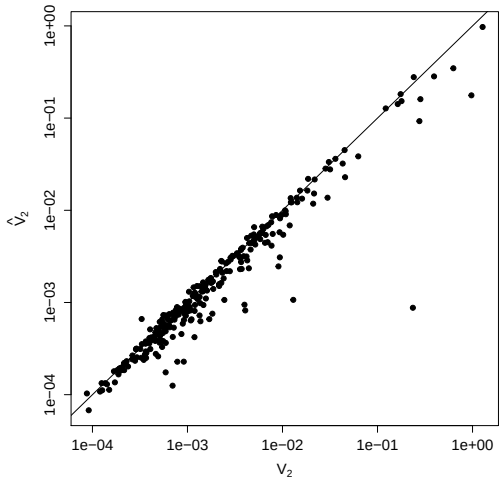

Figure 10: V_2 versus $\widehat{V}_2$ (good pools)

V_2 respectively, formulas that depend solely on h and the parameters β, $\bar{c}$ and γ of the Heston model. Figures 9 and 10 plot the estimated asymptotic autocovariances predicted by the parameter estimates versus the sample autocovariances for each good pool, along with a 45° line. Although there are a few outliers, the match is reasonably close over a range of four orders of magnitude.[21]

Finally, it is instructive to use our moment conditions to estimate the Heston model using daily RVs, with prices sampled either once every five minutes (the usual procedure in the literature) or every second. We apply the moment conditions (16)–(19), (27) and (28) to the sequence of daily QVs implied by the Heston model and estimate the model for the entire eight-year sequence of 2014 trading days, a sample roughly equal in size to two five-day pools for the intraday model. We set $h = 6.5/(24 \times 365.25)$, the length of a 6.5-hour day relative to a calendar year, rescaling daily QV and RV to an annual rate.

Table 4: Heston model using daily RV

	5-minute sampling		1-second sampling	
parameter	estimate	z score	estimate	z score
$\widehat{\beta}$	0.04	2.25	0.002	0.48
$\widehat{\bar{c}}$	0.13	3.87	0.12	0.78
$\widehat{\gamma^2}$	36.6	1.35	13.6	0.96

Table 4 reports the results of the estimation, giving the parameter estimates and their z scores. The model fails badly when prices are sampled once a second: z scores are all less than one. When prices are sampled every 5 minutes, the z scores of β and $\bar{c}$ are significantly different from 0 at the 2.5% level, the estimate of γ^2 is significantly different from 0 at the 10% level, the J-statistic does not reject the specification and the parameter estimates satisfy the Feller condition.

The estimate $\widehat{\bar{c}} = 0.13$ in Table 4 is close to the mean estimates reported in Table 2 for good days or good pools (0.14). However, the estimates of γ^2 and κ are two or three orders of magnitude smaller than the corresponding estimates for our intraday model. The estimate of γ^2 in Table 4 is about 0.17% of the median estimate of γ^2 for good pools in

[21] Note that both axes are scaled logarithmically.

Table 2. The estimate $\widehat{\beta} = 0.04$ in Table 4 implies that $\widehat{\kappa} = -\log(1-\widehat{\beta})/h = 55$, which is 3.3% of the median estimate of $\kappa = 3.32 \times 10^4$ reported in Table 3 for good pools.[22] The conclusion is obvious: Computing daily realized volatilities with prices sampled every five minutes destroys all evidence of high volatility and rapid mean reversion within the trading day and does not allow for variation of the asymptotic mean from one day to the next.

4 Relative RV and jumps

Our local Heston model assumes that the log-price process X and the volatility process c never jump within a trading day. In this section we briefly discuss how jumps fit within our framework.

To find evidence of jumps, it is natural to look for realized volatilities that are exceptionally large. However, there is a complication. The asymptotic mean of the Heston model varies widely over the course of our 8-year sample, which needs to be taken into account when setting a threshold for what constitutes an "exceptionally large" realized volatility. For this reason, we construct *relative realized volatilities* (RRVs) by dividing each RV_i by the median RV for the pool that contains block i: $\mathrm{RRV}_i = \mathrm{RV}_i/\mathrm{median}(\mathrm{RV})$.[23]

Figures 11 and 12 plot RRV for two good pools with dramatically different estimated asymptotic means, the second ($\widehat{\bar{c}} = 2.04$) over 33 times the first ($\widehat{\bar{c}} = 0.060$). Figure 11 plots RRV for pool 203, which starts on a Monday and contains the trading day (1/5/2011) featured in Figure 1.[24] Figure 12 plots RRV for Pool 92, which starts on a Tuesday. Pool 92 is the pool that attains the peak value in Figure 4, the time-series plot of $\widehat{\bar{c}}$. Dashed vertical lines separate trading days. The dashed horizontal line has height $\widehat{\bar{c}}/\mathrm{median}(RV)$. These two plots look remarkably alike, with most RRVs falling in the interval $[0, 2.5)$. If we plotted RV rather than RRV for these two pools, the scale for the second pool would be around 33 times that of the first.

[22] Bergantini (2013) estimated the Bollerslev and Zhou model for the period 8/1/1987–6/20/2011 using the S&P 500 index sampled every 5 minutes. The estimates obtained were $\widehat{\bar{c}} = 0.0103$, $\widehat{\kappa} = 0.1135$ and $\widehat{\gamma} = 0.0840$ with standard errors 0.0011, 0.0353 and 0.0104 respectively (see Bergantini (2013), Table 2.) The Feller condition was not satisfied.

[23] We use the median rather than mean because the median is less sensitive to outliers.

[24] We number the pools in our sample sequentially from 1 to 403.

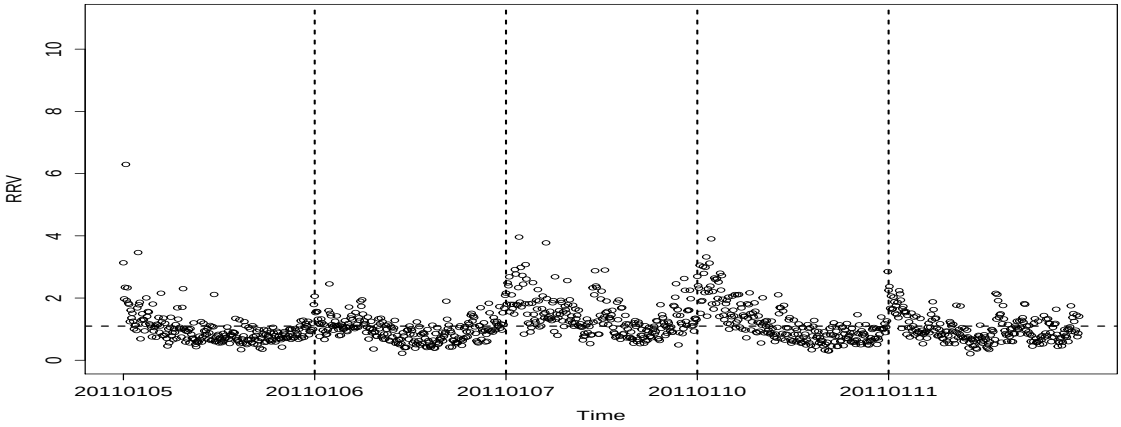

Figure 11: RRV for pool 203 (contains 1/6/2011)

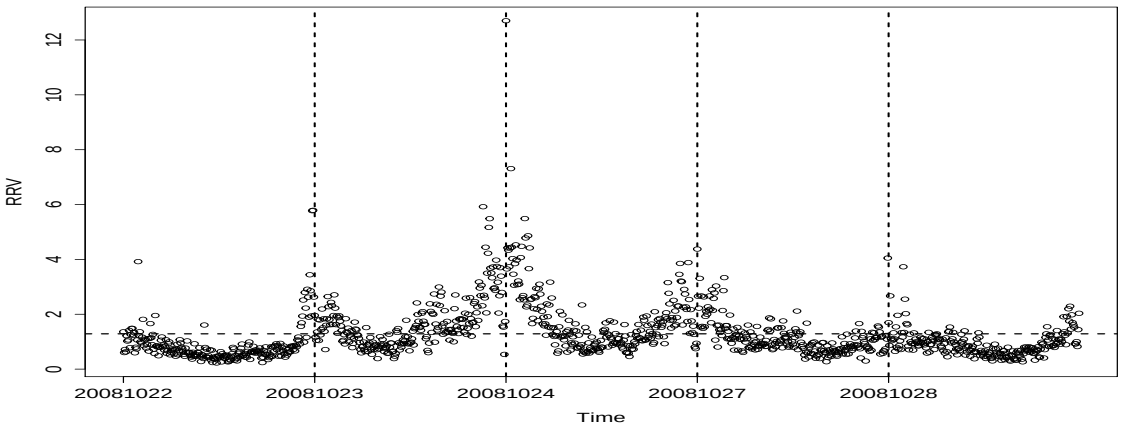

Figure 12: RRV for pool 92 (peak of financial crises)

We created RRV plots for 35 of our 403 pools, 23 good and 12 bad.[25] The plots for most of the good pools look like Figures 11 or 12. There is a clear tendency for RRVs to start the trading day over twice the median but rarely above 5. (Figure 12 provides an illustration of an initial RRV above 5 at the start of the third day.) From the point of view of the local Heston model, a high realized volatility in the first block of the trading day is not a jump but simply a high initial value for the stochastic difference equation, its influence dissipating once the market participants have had time to process the information revealed by trade. RRVs above 10 (which we call *bad blocks*) are quite rare. There are no bad blocks in Figure 11 and one in Figure 12, the initial RRV that starts the third day.

Table 5 describes the distribution of RRV for the blocks in our entire 8-year sample. Dividing the range $[0, \infty)$ into four intervals, the table gives the number of blocks falling into each interval with the percentage shown in parentheses. Bad pools differ very little from good in the percentage of RRVs in $[0, 2.5)$, $[2.5, 5)$ or $[5, 10)$. Approximately 95% of RRVs for good pools and bad lie in the interval $[0, 2.5)$ and 4% in the interval $[2.5, 5)$. A slightly higher percentage of bad pools than good have RRVs in the interval $[5, 10)$. In contrast, the percentage of RRVs in the interval $[10, \infty)$ (bad blocks) is 4 times greater for bad pools than good.

Table 5: Distribution of relative RV for SPY (pools): 2007–2014

Interval	Good pools	Bad pools	All pools
$[10, \infty)$	365 (0.11%)	582 (0.43%)	947 (0.20%)
$[5, 10)$	1,291 (0.39%)	659 (0.48%)	1,950 (0.41%)
$[2.5, 5)$	13,875 (4.14%)	5,649 (4.13%)	19,524 (4.14%)
$[0, 2.5)$	319,089 (95.36%)	129,766 (94.96%)	448,855 (95.24%)

To refine the classification of bad blocks, we broke the interval $[10, \infty)$ into three categories of increasing severity: $[10, 100)$, $[100, 1000)$ and $[1000, \infty)$, which we call category 1, 2 and 3. Good and bad pools have almost the same number of category 1 bad blocks: 353 for good pools and 352 for bad. For category 2, the difference widens greatly: 11 for good pools, 186 for bad. For category 3, the difference is stark: 1 for good pools, 44 for

[25] See Yan (2017), Chapter 3, for a detailed discussion of these RRV plots.

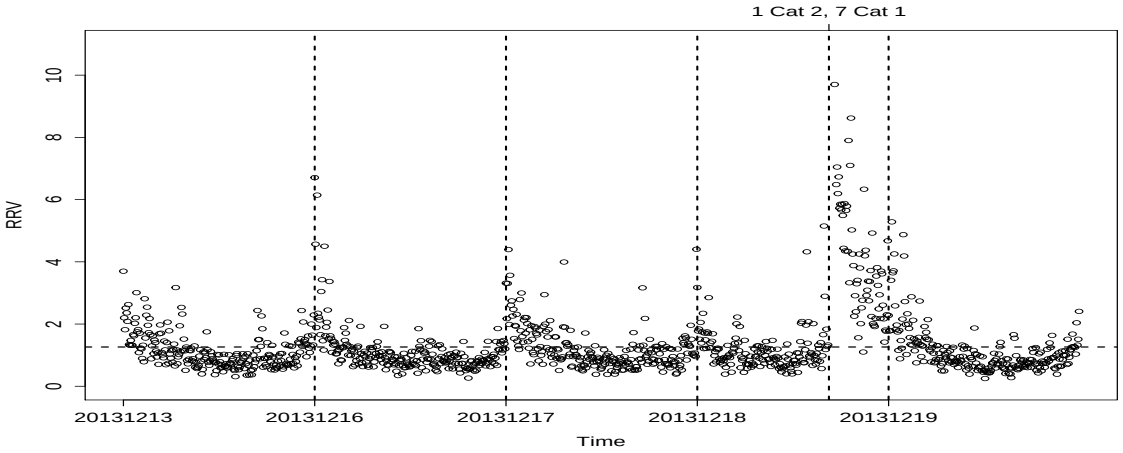

Figure 13: RRV for pool 351 (FOMC announcement on 12/18/2013)

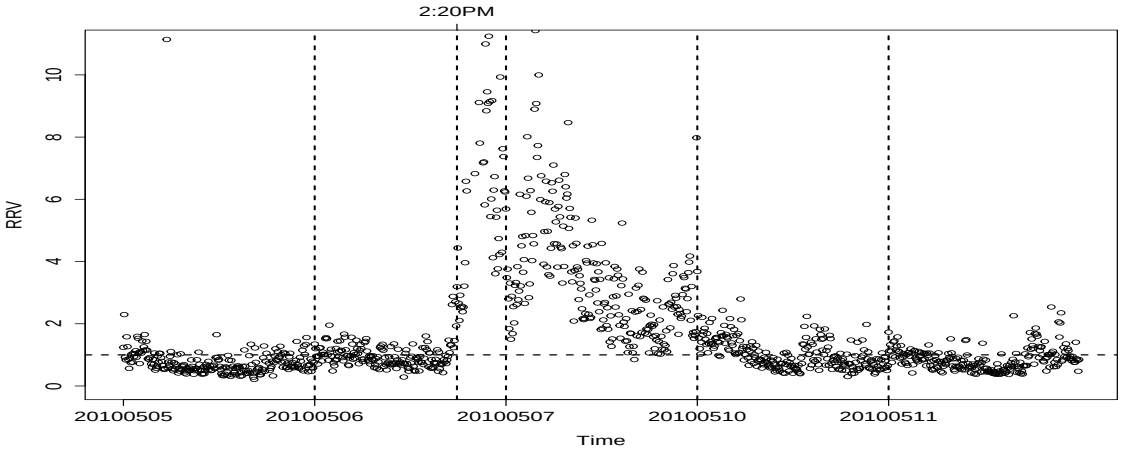

Figure 14: RRV for pool 169 (Flash Crash on 5/6/2010)

bad pools.

For the 35 pools we examined, a common pattern is for bad blocks to cluster and to leave a cascade of RRVs larger than 2 in their wake. Figure 13 gives an example of a cascade in a good pool triggered by an FOMC announcement on Wednesday at 2 PM. The vertical line on the fourth day labeled "1 Cat 2, 7 Cat 1" coincides with the announcement. The vertical axis is truncated from above at 11. As indicated by the label above the vertical line, 1 category 2 and 7 category 1 bad blocks (lying outside the plot because of the truncation) occurred in rapid succession after the announcement, followed in turn by a cascade of RRVs larger than 2 converging to the mean as the effects of the announcement wore off. The cascade bears a remarkable resemblance, on a larger scale, to what happens at the beginning of most trading days. Figures 15–18 in Appendix I contain RRV plots for four other good pools with similar cascades triggered by FOMC announcements.

Figure 14 for bad pool 169 contains the notorious Flash Crash starting at 2:20 PM on the second day. The vertical axis is truncated at 11.[26] There were 29 bad blocks: 17 of category 1, 11 of category 2 and 1 of category 3. The bad blocks marking the beginning of the Flash Crash were followed by a cascade of RRVs above 2, which continued through the following day. The day before the crash there were 5 category 2 and 1 category 1 bad blocks, suggestive of foreshocks preceding an earthquake (the category 1 shock can be seen in the figure).

Cascades greatly simplify the task of dealing with jumps because, when they occur, they totally dominate the process. In the cases we examined, we found that re-estimating the pool with the cascade removed almost always turned bad pools into good pools with reasonable parameter estimates.[27] Dealing with jumps in a more comprehensive way will require a considerable effort, but it clearly seems worthwhile.

5 Conclusion

Our approach to volatility estimation is data driven, intended to take full advantage of the richness of high-frequency stock price data. Rather than impose a single structural model on the eight years of data for SPY, we estimate the parameters of a sequence of

[26] See Figure 19 in Appendix I for the graph without truncation.

[27] See Yan (2017), Chapter 3.

daily structural models. Remarkably, the estimation performs well for nearly half of the trading days and three quarters of 5-day pools.

The Heston model provides a natural interpretation of what we are seeing in the data. We can interpret the asymptotic mean and volatility-of-volatility parameters as characteristics of a local equilibrium, reflecting the beliefs of traders regarding the uncertainty associated with stock returns: the instantaneous volatility represents the *risk* associated with the log return at a given instant; the asymptotic mean and volatility-of-volatility parameters characterize *uncertainty* regarding that risk.

When we pool, the estimates of the speed of mean reversion compress into a relatively narrow band, suggesting that the speed of mean reversion may be time invariant. The fact that mean reversion is quite rapid within the trading day supports the interpretation of the remaining parameters as equilibrium parameters. Their time series are smooth but vary over a large range. Our local estimation allows us to estimate these equilibrium parameters over time, an approach that can be applied to a vast number of stocks and exchange-traded funds representing many different sectors of the economy.[28]

For SPY we know that the local Heston model performs well for 71% of pools but fails for the rest. Jumps are clearly part of the story behind the failures. So are structural breaks that wreak havoc with our assumption that structural parameters are invariant across the five days of every five-day pool. Our model clearly can be improved. However, in doing so it is important to give the data a chance to tell their story.

[28]In addition to SPY, we also estimated the local Heston model for two other exchange-traded funds (ticker symbols IWM and EMM) and five Dow-Jones stocks (ticker symbols BAC, CVX, IBM, INTC and MSFT). IWM and EMM track the Russell 2000 index of small-cap stocks and the MSCI emerging markets index respectively. The stocks are Bank of America, Chevon, IBM, Intel and Microsoft. See Yan (2017) for discussion of the results in comparable detail to our estimation for SPY. The estimation performed well for these assets, with similar results and some interesting differences.

REFERENCES

Aït-Sahalia, Yacine, and Jean Jacod (2014), *High-Frequency Financial Econometrics*, Princeton University Press.

Aït-Sahalia, Yacine, Per A. Mykland and Lan Zhang (2005), "How often to sample a continuous-time process in the presence of market microstructure noise," *The Review of Financial Studies*, 18, 351–416.

Bollerslev, Tim and Hao Zhou (2002), "Estimating stochastic volatility diffusion using conditional moments of integrated volatility," *Journal of Econometrics* 109, 33–65.

Bregantini, Daniele (2013), "Moment-based estimation of stochastic volatility," *Journal of Banking & Finance* 37, 4755–4764.

Chaussé, Pierre (2010), "Computing Generalized Method of Moments and Generalized Empirical Likelihood with R," *Journal of Statistical Software* 34(11), 1–35.

Corradi, Valentina and Walter Distaso (2006), "Semi-parametric comparison of stochastic volatlity models using realized measures," *Review of Economic Studies* 73, 635–667.

Cox, John, Jonathan Ingersoll, and Stephen Ross (1993), "A theory of the term structure of interest rates," *Econometrica* 53(2), 385–407.

Feller, William (1951), "Two singular diffusion problems," *Annals of Mathematics* 54(2), 173–182.

Garcia, René, Marc-André Lewis, Sergio Pastorello and Éric Renault (2011), "Estimation of objective and risk-neutral distributions based on moments of integrated volatility," *Journal of Econometrics* 160, 22–32.

Hall, Alastair R. (2005), *Generalized Method of Moments*, Oxford University Press.

Hansen, Lars Peter (1982), "Large sample properties of generalized method of moments estimators," *Econometrica* 50, 1029-1054.

Heston, Steven (1993), "A closed-form solution for options with stochastic volatility with applications to bonds and currency options," *Review of Financial Studies* 6, 327–343.

Newey, Whitney and Kenneth West (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55 (3), 703-708.

Protter, Philip (2004), *Stochastic Integration and Differential Equations: Second Edition*, Springer-Verlag.

Shreve, Steven (2004), *Stochastic Calculus for Finance: II*, Springer-Verlag.

Sun, Miao (2016), *Modeling Volatility Using High Frequency Data*, Ph.D. thesis, Department of Economics, UCLA.

Whang, Duke (2012), *Volatility at High Frequency*, Ph.D. thesis, Department of Economics, UCLA.

Yan, Sibo (2017), *Estimating Stochastic Volatility within a Trading Day*, Ph.D. thesis, Department of Economics, UCLA.

APPENDIX

A Proof of Proposition 1

Proof. First we prove that for any block $[t, t+h]$

$$c_{t+h} = \alpha c_t + \beta \bar{c} + e^{-\kappa(t+h)} \gamma \int_t^{t+h} e^{\kappa s} \sqrt{c_s}\, dB_s \tag{31}$$

This proof borrows directly from the analysis of the CIR model in Shreve (2004), page 152.[29] Define $f(t, c) = e^{\kappa t} c$. The function f has partial derivatives $f_t = \kappa e^{\kappa t} c$, $f_c = e^{\kappa t}$ and $f_{cc} = 0$. The Itô-Doeblin formula implies

$$\begin{aligned} d(e^{\kappa t} c_t) &= f_t(t, c_t)\, dt + f_c(t, c_t)\, dc_t + \frac{1}{2} f_{cc}(t, c_t)\, dC_t \\ &= \kappa e^{\kappa t} c_t\, dt + e^{\kappa t} \left[\kappa(\bar{c} - c_t)\, dt + \gamma \sqrt{c_t}\, dB_t\right] \\ &= \kappa e^{\kappa t} \bar{c}\, dt + \gamma e^{\kappa t} \sqrt{c_t}\, dB_t \end{aligned}$$

Integrating both sides over the interval $[t, t+h]$ yields

$$e^{\kappa(t+h)} c_{t+h} - e^{\kappa t} c_t = \bar{c}(e^{\kappa(t+h)} - e^{\kappa t}) + \gamma \int_t^{t+h} e^{\kappa s} \sqrt{c_s}\, dB_s$$

Using the definitions of α and β, the result follows.

Because the Itô integral $\int_0^{t+h} e^{\kappa s} \sqrt{c_s}\, dB_s$ is a martingale with zero expectation, equation (31) implies that

$$\mathbb{E}[c_{t+h} \mid \mathcal{F}_t] = \alpha c_t + \beta \bar{c} \tag{32}$$

Using the definition of $C_{t,t+h}$, interchanging expectation and integration with respect to time and substituting for $\mathbb{E}[c_s \mid \mathcal{F}_t]$,

$$\begin{aligned} \mathbb{E}\left[C_{t,t+h} \mid \mathcal{F}_t\right] &= \mathbb{E}\left[\frac{1}{h} \int_t^{t+h} c_s ds \mid \mathcal{F}_t\right] = \frac{1}{h} \int_t^{t+h} \mathbb{E}\left[c_s \mid \mathcal{F}_t\right] ds \\ &= \frac{1}{h} \int_t^{t+h} \left[e^{-\kappa(s-t)} c_t + \left(1 - e^{-\kappa(s-t)}\right) \bar{c}\right] ds \end{aligned}$$

Evaluating the integrals,

$$\mathbb{E}\left[C_{t,t+h} \mid \mathcal{F}_t\right] = a_1 c_t + b_1 \bar{c} \tag{33}$$

[29] Equation (2) has the same form as the stochastic differential equation for the Cox-Ingersoll-Ross (1985) model of the interest rate process, $dR_t = \kappa(\bar{R} - R_t)\, dt + \gamma \sqrt{R_t}\, dB_t$.

where $a_1 = (1 - e^{-\kappa h})/\kappa h$ and $b_1 = 1 - a_1$. By the law of iterated expectations,

$$\begin{aligned}\mathbb{E}\left[C_{t+h,t+2h} \mid \mathcal{F}_t\right] &= \mathbb{E}\left[\mathbb{E}\left[C_{t+h,t+2h}|\mathcal{F}_{t+h}\right]|\mathcal{F}_t\right] = \mathbb{E}\left[a_1 c_{t+h} + b_1\bar{c} \mid \mathcal{F}_t\right] \\ &= a_1\, \mathbb{E}\left[c_{t+h} \mid \mathcal{F}_t\right] + b_1\, \bar{c} = a_1\left(\alpha c_t + \beta\bar{c}\right) + b_1\bar{c} \\ &= a_1\alpha c_t + (a_1\beta + b_1)\bar{c}\end{aligned}$$

where equation (33) is used to substitute for $\mathbb{E}\left[C_{t+h,t+2h}|\mathcal{F}_{t+h}\right]$ and equation (32) to substitute for $\mathbb{E}\left[c_{t+h} \mid \mathcal{F}_t\right]$. Using equation (33) to replace c_t

$$\begin{aligned}\mathbb{E}\left[C_{t+h,t+2h} \mid \mathcal{F}_t\right] &= a_1\alpha\frac{\mathbb{E}\left[C_{t,t+h} \mid \mathcal{F}_t\right] - b_1\bar{c}}{a_1} + (a_1\beta + b_1)\bar{c} \\ &= \alpha\mathbb{E}\left[C_{t,t+h} \mid \mathcal{F}_t\right] + (a_1\beta + (1-\alpha)b_1)\bar{c}\end{aligned}$$

Because

$$a_1\beta + (1-\alpha)b_1 = \left(\frac{\beta}{kh}\right)\beta + \beta\left(1 - \frac{\beta}{kh}\right) = \beta$$

the result follows. □

B Proof of Lemma 1

Proof. Letting $t + h = T$ and $h = T - t$ in equation (33),

$$\kappa\left(T-t\right)\mathbb{E}\left[C_{t,T}|\mathcal{F}_t\right] = \left(1 - e^{-\kappa(T-t)}\right)c_t + \left(\kappa(T-t) - \left(1 - e^{-\kappa(T-t)}\right)\right)\bar{c}$$

Fix T and let t vary. Define

$$f(t,c) = \left(1 - e^{-\kappa(T-t)}\right)c + \left(\kappa(T-t) - \left(1 - e^{-\kappa(T-t)}\right)\right)\bar{c}$$

The partial derivatives of f are

$$f_t = -\kappa e^{-\kappa(T-t)}c - \kappa(1 - e^{-\kappa(T-t)})\bar{c} \qquad f_c = 1 - e^{-\kappa(T-t)} \qquad f_{cc} = 0$$

Applying the Itô-Doeblin formula gives

$$\begin{aligned}d\left(\kappa\left(T-t\right)\mathbb{E}\left[C_{t,T}|\mathcal{F}_t\right]\right) &= \left[-\kappa e^{-\kappa(T-t)}c_t - \kappa(1 - e^{-\kappa(T-t)})\bar{c}\right]dt \\ &\quad + (1 - e^{\kappa(T-t)})[\kappa(\bar{c} - c_t)\,dt + \gamma\sqrt{c_t}\,dB_t] \\ &= -\kappa c_t dt + \left(1 - e^{-\kappa(T-t)}\right)\gamma\sqrt{c_t}\,dB_t\end{aligned}$$

Integrating from t to T yields

$$0 - \kappa\left(T-t\right)\mathbb{E}\left[C_{t,T}|\mathcal{F}_t\right] = -\kappa\int_t^T c_s ds + \int_t^T \left(1 - e^{-\kappa(T-s)}\right)\gamma\sqrt{c_s}\,dB_s$$

Dividing both sides by $\kappa(T-t)$, writing T once again as $t+h$ and using the definition of $C_{t,t+h}$,

$$C_{t,t+h} - \mathbb{E}\left[C_{t,t+h}|\mathcal{F}_t\right] = \int_t^{t+h} \left(\frac{1-e^{-\kappa(t+h-s)}}{\kappa h}\right) \gamma\sqrt{c_s}\, dB_s$$

Squaring both sides, taking the conditional expectation of each side with respect to $\mathcal{F}_t$ and appealing to Itô's isometry,

$$\begin{aligned}
\mathrm{Var}\left(C_{t,t+h} \mid \mathcal{F}_t\right) &:= \mathbb{E}\left[\left(C_{t,t+h} - \mathbb{E}[C_{t,t+h} \mid \mathcal{F}_t]\right)^2 \mid \mathcal{F}_t\right] \\
&= \mathbb{E}\left[\left(\int_t^{t+h} \left(\frac{1-e^{-\kappa(t+h-s)}}{kh}\right) \gamma\sqrt{c_s}\, dB_s\right)^2 \mid \mathcal{F}_t\right] \\
&= \mathbb{E}\left[\int_t^{t+h} \left[\left(\frac{1-e^{-\kappa(t+h-s)}}{kh}\right) \gamma\sqrt{c_s}\right]^2 ds \mid \mathcal{F}_t\right] \\
&= \left(\frac{\gamma}{\kappa h}\right)^2 \int_t^{t+h} \left(1-e^{-\kappa(t+h-s)}\right)^2 \mathbb{E}[c_s \mid \mathcal{F}_t]\, ds
\end{aligned}$$

Using equation (32) to substitute for $\mathbb{E}[c_s \mid \mathcal{F}_t]$ and evaluating the integrals,

$$\mathrm{Var}\left(C_{t,t+h} \mid \mathcal{F}_t\right) = \left(\frac{\gamma}{\kappa h}\right)^2 \left(a_2\, c_t + b_2\, \bar{c}\right)$$

where

$$\begin{aligned}
a_2 &= \int_t^{t+h} (1-e^{-\kappa(t+h-s)})^2 e^{-\kappa(s-t)}\, ds = \frac{1}{\kappa} - 2he^{-\kappa h} - \frac{1}{\kappa}e^{-2\kappa h} \\
a_2 + b_2 &= \int_t^{t+h} (1-e^{-\kappa(t+h-s)})^2\, ds = h - \frac{3}{2\kappa} + \frac{2}{k}e^{-\kappa h} - \frac{1}{2\kappa}e^{-2\kappa h} \\
b_2 &= (a_2+b_2) - a_2 = h - \frac{5}{2h} + \left(\frac{2}{\kappa} + 2h\right) e^{-\kappa h} + \frac{1}{2\kappa}e^{-2\kappa h}
\end{aligned}$$

□

C Proof of Proposition 3

Proof. By the law of total variance

$$\mathrm{Var}(C_{t,t+h}) = \mathrm{Var}(\mathbb{E}(C_{t,t+h} \mid \mathcal{F}_t]) + \mathbb{E}[\mathrm{Var}(C_{t,t+h} \mid \mathcal{F}_t)]$$

Using equation (33), $\mathrm{Var}(\mathbb{E}[C_{t,t+h} \mid \mathcal{F}_t]) = a_1^2\, \mathrm{Var}(c_t)$ and so

$$\lim_{t\to\infty} \mathrm{Var}(\mathbb{E}[C_{t,t+h} \mid \mathcal{F}_t]) = a_1^2 \lim_{t\to\infty} \mathrm{Var}(c_t) = \left(\frac{\gamma}{\kappa h}\right)^2 \frac{(1-e^{-\kappa h})^2}{2\kappa}\bar{c}$$

Using equation (22),

$$\lim_{t\to\infty} \mathbb{E}[\mathrm{Var}(C_{t,t+h} \mid \mathcal{F}_t)] = \left(\frac{\gamma}{\kappa h}\right)^2 (a_2 \lim_{t\to\infty} \mathbb{E}c_t + b_2\bar{c})$$
$$= \left(\frac{\gamma}{\kappa h}\right)^2 \left(\frac{2\kappa h - 3 + 4e^{-\kappa h} - e^{-2\kappa h}}{2\kappa}\right) \bar{c}$$

Consequently the asymptotic variance is

$$\lim_{t\to\infty} \mathrm{Var}(C_{t,t+h}) = \lim_{t\to\infty} \mathrm{Var}(\mathbb{E}[C_{t,t+h} \mid \mathcal{F}_t]) + \lim_{t\to\infty} \mathbb{E}[\mathrm{Var}(C_{t,t+h} \mid \mathcal{F}_t)]$$
$$= \left(\frac{\gamma}{\kappa h}\right)^2 \left(h - \frac{1}{\kappa} + \frac{1}{\kappa}e^{-\kappa h}\right) \bar{c} = \left(\frac{\gamma}{\kappa h}\right)^2 \left(h - \frac{\beta}{\kappa}\right) \bar{c}$$

□

D Proof of Lemma 2

Proof. Using equations (32) and (33), the error

$$\eta_{t,t+2h} = C_{t+h,t+2h} - \alpha\, C_{t,t+h} - \beta\, \bar{c}$$

can be expressed as the Itô integral

$$\eta_{t,t+2h} = \int_t^{t+2h} \psi_s \frac{\gamma}{\kappa h} \sqrt{c_s}\, dB_s \tag{34}$$

where[30]

$$\psi_s = \begin{cases} e^{-\kappa(t+h-s)} - e^{-\kappa h} & \text{if } s \in [t, t+h) \\ 1 - e^{-\kappa(t+2h-s)} & \text{if } s \in [t+h, t+2h] \end{cases} \tag{35}$$

Squaring both sides of equation (34), taking the conditional expectation with respect to $\mathcal{F}_t$ and appealing to Itô's isometry,

$$\mathrm{Var}(\eta_{t,t+2h} \mid \mathcal{F}_t) = \mathbb{E}\left[\eta_{t,t+2h}^2 \mid \mathcal{F}_t\right] = \mathbb{E}\left[\left(\int_t^{t+2h} \psi_s \frac{\gamma}{\kappa h}\sqrt{c_s}\, dB_s\right)^2 \mid \mathcal{F}_t\right]$$
$$= \mathbb{E}\left[\int_t^{t+2h} \left(\psi_s \frac{\gamma}{\kappa h}\sqrt{c_s}\right)^2 ds \mid \mathcal{F}_t\right]$$
$$= \left(\frac{\gamma}{\kappa h}\right)^2 \int_t^{t+2h} \psi_s^2\, \mathbb{E}[c_s \mid \mathcal{F}_t]\, ds$$

[30] Note that if $s = t+h$, then the two lines in the definition of the function ψ agree: $e^{-\kappa(t+h-s)} - e^{-\kappa h} = 1 - e^{-\kappa h}$ and $1 - e^{-\kappa(t+2h-s)} = 1 - e^{-\kappa h}$.

Using equation (32) to substitute for $\mathbb{E}[c_s \mid \mathcal{F}_t]$ and evaluating the integrals gives the result with

$$a_3 = \frac{2}{k}e^{-\kappa h} - 4he^{-2\kappa h} - \frac{2}{\kappa}e^{-3\kappa h}$$
$$b_3 = h - \frac{1}{\kappa} + \left(5h + \frac{1}{\kappa}\right)e^{-2\kappa h} + \frac{2}{\kappa}e^{-3\kappa h}$$

□

E Proof of Proposition 4

Proof. Because $\mathbb{E}[\eta_{t,t+2h} \mid \mathcal{F}_t] = 0$ the law of total variance implies that

$$\mathrm{Var}(\eta_{t,t+2h}) = \mathbb{E}[\mathrm{Var}(\eta_{t,t+2h} \mid \mathcal{F}_t)] + \mathrm{Var}(\mathbb{E}[\eta_{t,t+2h} \mid \mathcal{F}_t]) = \mathbb{E}[\mathrm{Var}(\eta_{t,t+2h} \mid \mathcal{F}_t)]$$

Using Lemma 2

$$\lim_{t\to\infty} \mathrm{Var}\left(\eta_{t,t+2h}\right) = \lim_{t\to\infty} \mathbb{E}[\mathrm{Var}(\eta_{t,t+2h} \mid \mathcal{F}_t)] = \left(\frac{\gamma}{\kappa h}\right)^2 \left[a_3 \lim_{t\to\infty} \mathbb{E}c_t + b_3\,\bar{c}\right]$$
$$= \left(\frac{\gamma}{\kappa h}\right)^2 (a_3 + b_3)\bar{c}$$

where $a_3 + b_3$ is the sum of the formulas for a_3 and b_3 in Lemma 2. □

F Proof of Proposition 5

Proof. V_0 is the asymptotic variance derived in Proposition 3. For $j = 1$

$$\mathbb{E}\left[(C_{t+h,t+2h} - \alpha C_{t,t+h} - \beta\bar{c})^2\right] = \mathbb{E}\left[((C_{t+h,t+2h} - \bar{c}) - \alpha\left(C_{t,t+h} - \bar{c}\right))^2\right]$$

which implies that

$$\lim_{t\to\infty} \mathbb{E}\left[(C_{t+h,t+2h} - \alpha C_{t,t+h} - \beta\bar{c})^2\right] = \left(1 + \alpha^2\right)V_0 - 2\alpha V_1$$

and hence

$$V_1 = \frac{1}{2\alpha}\left(\left(1 + \alpha^2\right)V_0 - \lim_{t\to\infty} \mathbb{E}\left[(C_{t+h,t+2h} - \alpha C_{t,t+h} - \beta\bar{c})^2\right]\right) \tag{36}$$

As proved in Proposition 4

$$\lim_{t\to\infty} \mathbb{E}\left[(C_{t+h,t+2h} - \alpha C_{t,t+h} - \beta\bar{c})^2\right] = \left(\frac{\gamma}{\kappa h}\right)^2 (a_3 + b_3)\,\bar{c} \tag{37}$$

where

$$a_3 + b_3 = h - \frac{1}{\kappa} + \left(h + \frac{1}{\kappa}\right) e^{-2\kappa h}$$

Plugging equation (37) and the equation for V_0 into equation (36) and simplifying yields the formula for V_1.

To compute the autocovariances for lags $j \geq 2$, recall that

$$\mathbb{E}\left[C_{t+h,t+2h} - \alpha C_{t,t+h} - \beta\bar{c} \mid \mathcal{F}_t\right] = 0$$

Because $C_{t-(j-1)h,t-(j-2)h} - \bar{c}$ is measurable with respect to $\mathcal{F}_t$ when $j \geq 2$,

$$\mathbb{E}\left[\left(C_{t+h,t+2h} - \alpha C_{t,t+h} - \beta\bar{c}\right)\left(C_{t-(j-1)h,t-(j-2)h} - \bar{c}\right)\right] = 0$$

which implies

$$\mathbb{E}\left[\left(\left(C_{t+h,t+2h} - \bar{c}\right) - \alpha\left(C_{t,t+h} - \bar{c}\right)\right)\left(C_{t-(j-1)h,t-(j-2)h} - \bar{c}\right)\right] = 0$$

and hence $V_j - \alpha V_{j-1} = 0$ for $j \geq 2$. Replacing α by $1 - \beta$ gives the result. □

G Proof of Corollary 1

Proof. The formula for R_0 is trivial. Using the definition of R_j, the formulas for V_0 and V_1 in Proposition 5 and $\beta = 1 - e^{-\kappa h}$ gives

$$R_1 := \frac{V_1}{V_0} = \frac{\beta^2/2\kappa}{h - \beta/\kappa} = \frac{\beta^2}{2\left(-\log\left(1-\beta\right) - \beta\right)}$$

The formula for R_j for $j \geq 2$ follows directly from the definition of R_j and the formula for V_2. □

H Proof of Proposition 6

Proof. Using Assumption 2 and the definitions of $\nu_{t,t+h}$ and $\widehat{V}_j$,

$$\begin{aligned}
\widehat{V}_0 &= \lim_{t\to\infty} \mathbb{E}\left[\left(\widehat{C}_{t,t+h} - \bar{c}\right)^2\right] = \lim_{t\to\infty} \mathbb{E}\left[\left(C_{t,t+h} + \nu_{t,t+h} - \bar{c}\right)^2\right] \\
&= \lim_{t\to\infty} \mathbb{E}\left[\left(C_{t,t+h} - \bar{c}\right)^2 + \nu_{t,t+h}^2 + 2\left(C_{t,t+h} - \bar{c}\right)\nu_{t,t+h}\right] \\
&= V_0 + \nu^2 + 0
\end{aligned}$$

$$
\begin{aligned}
\widehat{V}_1 &= \lim_{t\to\infty} \mathbb{E}\left[\left(\widehat{C}_{t,t+h} - \bar{c}\right)\left(\widehat{C}_{t+h,t+2h} - \bar{c}\right)\right] \\
&= \lim_{t\to\infty} \mathbb{E}\left[\left(C_{t,t+h} + \nu_{t,t+h} - \bar{c}\right)\left(C_{t+h,t+2h} + \nu_{t+h,t+2h} - \bar{c}\right)\right] \\
&= \lim_{t\to\infty} \mathbb{E}\left[\left(C_{t,t+h} - \bar{c}\right)\left(C_{t+h,t+2h} - \bar{c}\right) + \nu_{t,t+h}\ \nu_{t+h,t+2h}\right. \\
&\qquad \left. + \left(C_{t,t+h} - \bar{c}\right)\nu_{t+h,t+2h} + \left(C_{t+h,t+2h} - \bar{c}\right)\nu_{t,t+h}\right] \\
&= V_1 + 0 + 0 + 0
\end{aligned}
$$

Similarly, $\widehat{V}_j = V_j$ for $j \geq 2$. $\square$

I Relative RV plots

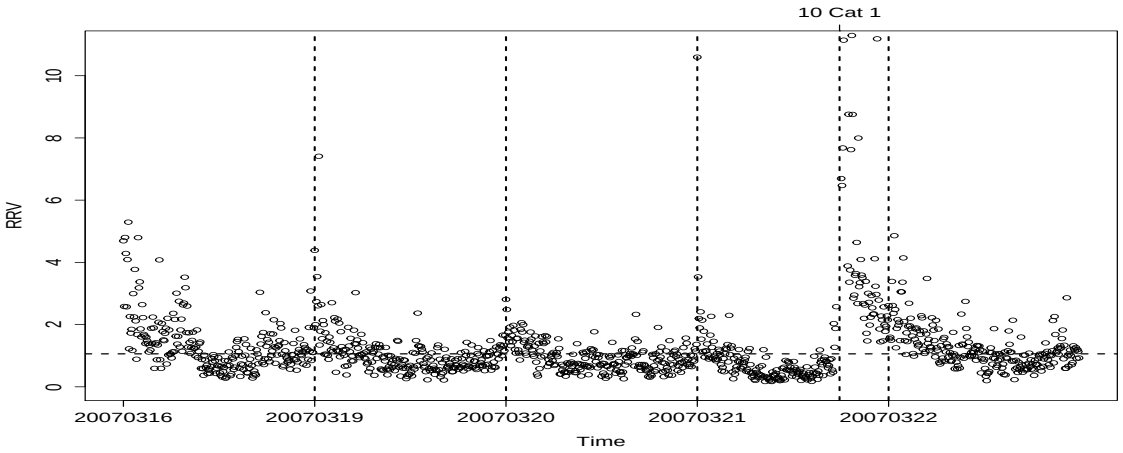

Figure 15: RRV for pool 11 (FOMC announcement on 3/21/2007)

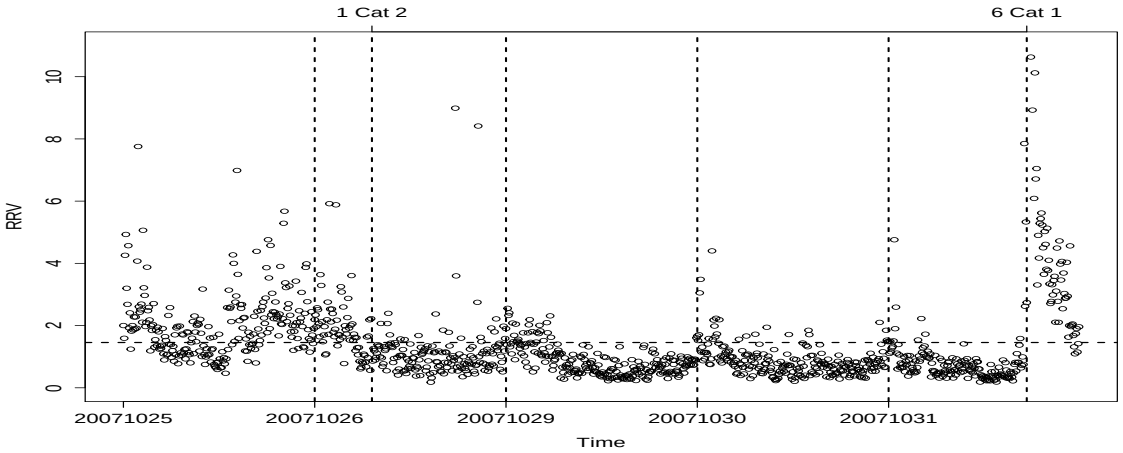

Figure 16: RRV for pool 42 (FOMC announcement on 10/31/2007)

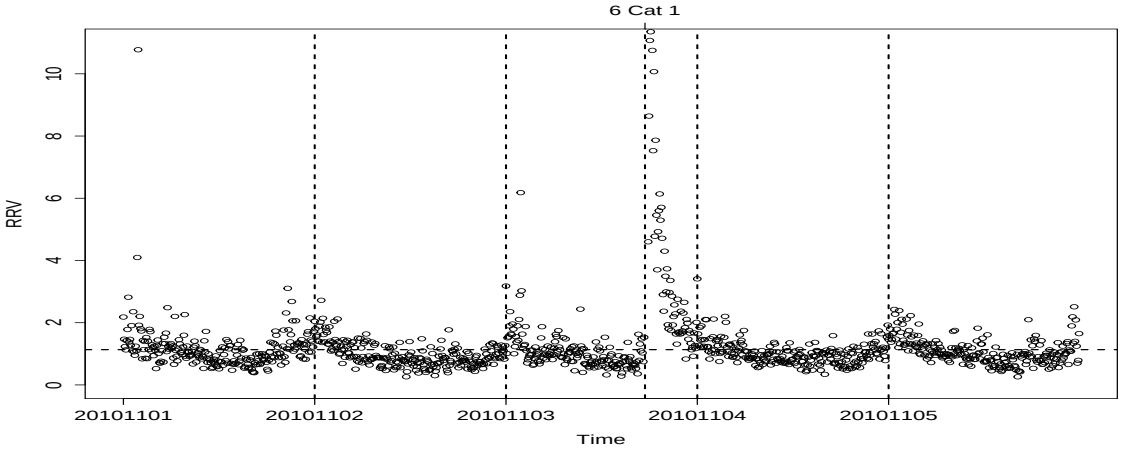

Figure 17: RRV for pool 194 (FOMC announcement on 11/3/2010)

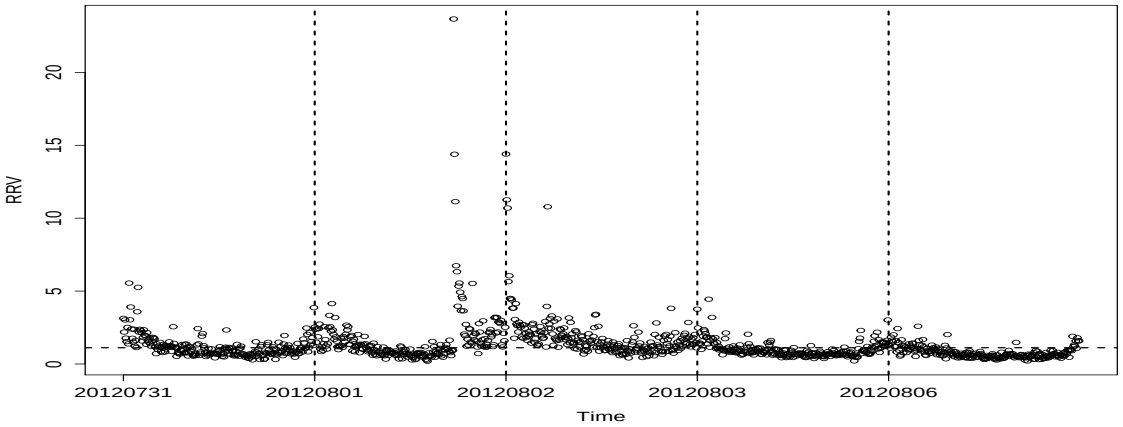

Figure 18: RRV for pool 282 (FOMC announcement on 8/1/2012)

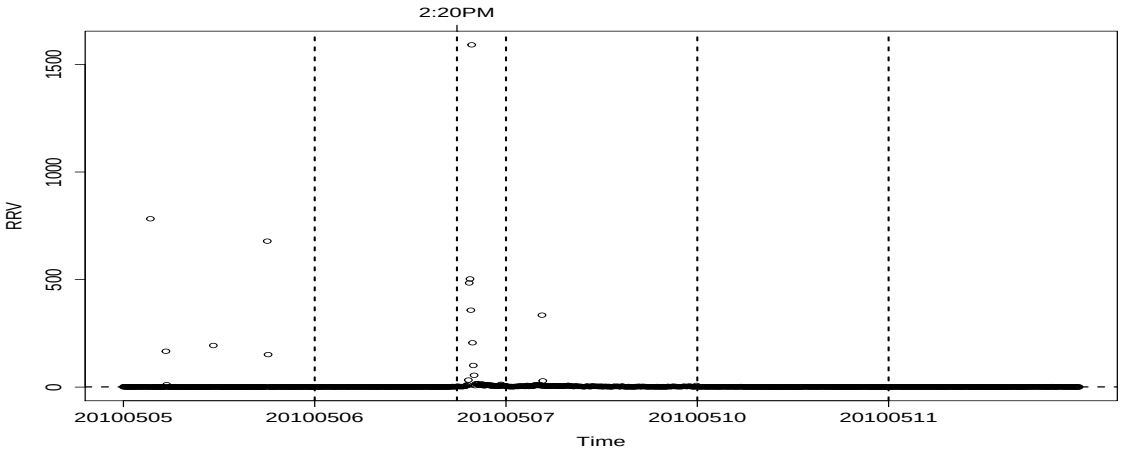

Figure 19: RRV for pool 169 (Flash Crash, no truncation)