# Conditional Volatility Persistence*

Jianxin Wang[1] and Minxian Yang[2]

July 2018

Abstract

This study provides evidence on the common determinants for two prominent features of equity market volatility: its persistence over time and its asymmetric dependence on past returns. We show that daily volatility persistence increases with current returns, especially negative returns. It decreases with current volatility. The estimated volatility persistence from the observed variables is termed "conditional volatility persistence". It provides a new economic link from return to future volatility, and a more robust explanation for their asymmetric relationship. By estimating the variations in the latent volatility persistence, our model significantly improves volatility forecasts relative to recent advances in volatility models.

Keywords: realized variance, volatility persistence, asymmetric volatility, volatility forecast

JEL codes: G12, G14, D83, C22

* We appreciate valuable comments from Tony He, Kai Li, Blake LeBaron, Ben Marshall, Andrew Patton, Timo Teräsvirta, Takeshi Yamada, Xiangkang Yin, Yoshi Yoshida, and Qiaoqiao Zhu and from participants of the 2017 Time Series and Forecasting Symposium at the University of Sydney, the 2017 Asian Econometric Society Meeting, the 2018 North American Econometric Society Meeting, and seminars at Australian National University, La Trobe University, Macquarie University, Massey University, Nagasaki University, National ChengChi University, National Taiwan University of Science and Technology, Shiga University, Tianjin University, Tsinghua University, and the University of Adelaide. The usual disclaimer applies.

[1] Corresponding author. The University of Technology Sydney, UTS Business School, Sydney, Australia 2007, e-mail: jianxin.wang@uts.edu.au.
[2] The University of New South Wales, UNSW Business School, Sydney, Australia 2052, email: m.yang@unsw.edu.au.

*"The 'state of the world' is a serially correlated thing; hence, we find ARCH."*

*Diebold and Nerlove (1989)*

# I. **Introduction**

This study examines the determinants of short-run volatility persistence. Our empirical design is motivated by several economic mechanisms linking the information process as well as uninformed trading to volatility persistence. Our evidence supports the above view of Diebold and Nerlove (1989): the overall "state of the world", as measured by daily return and volatility, is an important determinant of volatility persistence. Volatility persistence as a function of return represents a new link from current return to future volatility. It sharply reduces the direct return impact on future volatility and offers a new and more robust explanation for asymmetric volatility. By estimating volatility persistence from the observed market variables, our model significantly improves volatility forecasts.

Since the seminal studies by Engle (1982) and Bollerslev (1986), a vast literature has emerged, extending GARCH-family models to capture a wide range of statistical features. However, the literature on the economic origins of volatility persistence remains relatively small and diverse. The most prominent is the mixture of distribution hypothesis (MDH) where volatility persistence is driven by the persistence in information arrivals.[1] Other explanations include the persistence in wealth distribution (Cabrales and Hoshi, 1996), investor learning about uncertainty (e.g. Brock and LeBaron, 1996; Johnson, 2000; He, Li, and Wang, 2016), information cost (de Fontnouvelle, 2000), time-varying risk aversion (McQueen and Vorkink, 2004), and heterogeneous trading frequencies (Xue and Gençay, 2012), among others. Since the relevant variables are unobservable, empirical tests of these mechanisms are inherently difficult and affected by how well the latent variable is estimated. The MDH as a mechanism for volatility persistence finds mixed empirical support. Several studies, e.g. Laux and Ng (1993), Andersen (1996), and He and Velu (2014), show that the latent information arrivals can partially explain volatility persistence. Other studies, e.g. Lamoureux and Lastrapes (1994) and Liesenfeld (1998), fail to find empirical support for the MDH. Liesenfeld (2001) and Berger, Chaboud, and Hjalmarsson (2009) point to time-varying price sensitivity to new information as an important source for volatility persistence. Lamoureux and Lastrapes (1990) and Hamilton and Susmel (1994) show that persistence is lower after accounting for volatility

[1] The MDH was originally proposed to explain the non-Gaussian distribution of asset returns (Clark, 1973) and the volatility-volume relationship (Tauchen and Pitts, 1983; Andersen, 1996). Diebold (1986) and Gallant, Hsieh, and Tauchen (1991) were among the first to suggest the persistence of exogenous information flow as the source for volatility persistence.

regime shifts. Patton and Sheppard (2015) show that persistence comes mostly from the "bad" volatility associated with negative returns. Ning, Xu, and Wirjanto (2015) report that high volatility is more persistent than low volatility. Bollerslev, Patton, and Quaedvlieg (2016, BPQ hereafter) provide the first evidence on time-varying daily volatility persistence that is inversely related to measurement errors in daily volatility.

Instead of focusing on one particular mechanism, we take a broader view on the origins of volatility persistence and propose several mechanisms linking the overall market state to volatility persistence. To illustrate the idea, let's assume that daily volatility $\sigma_t^2$ follows a simple dynamic process $\sigma_{t+1}^2 = \alpha + \beta\sigma_t^2 + \varepsilon_{t+1}$. In this study, we present evidence that volatility persistence as measured by $\beta$ is a function of market state $s_t$. Using return and volatility as proxies for the market state $s_t = (r_t, \sigma_t^2)$, the dynamic process for $\sigma_t^2$ becomes nonlinear: $\sigma_{t+1}^2 = \alpha + \beta(r_t, \sigma_t^2)\sigma_t^2 + \varepsilon_{t+1}$. In section II, we argue that $\beta(r_t, \sigma_t^2)$ increases with $|r_t|$ but decreases with $\sigma_t^2$. The intuition is that large information shocks cause large $|r_t|$ and high $\sigma_t^2$. They draw greater investor attention, causing further information search and additional information arrivals. They may also trigger portfolio adjustments and herding behaviour which often last several days. Such endogenous information flows and uninformed trading increase the dependence of $\sigma_{t+1}^2$ on $\sigma_t^2$, i.e. $\beta(r_t, \sigma_t^2)$. Negative returns tend to generate greater investor reaction and higher $\beta(r_t, \sigma_t^2)$. On the other hand, volatility is often viewed as an information flow measure (Ross, 1989; Hasbrouck, 1995; Andersen, 1996). Given an information event, a high $\sigma_t^2$ implies more information being priced on a trading day, less unpriced information and less correlation between $\sigma_{t+1}^2$ and $\sigma_t^2$, i.e. $\beta(r_t, \sigma_t^2)$. The economic mechanisms linking return and volatility to volatility persistence are further elaborated in Section II. To the degree that return and volatility reflect the overall "state of the market", our approach echoes the view of Diebold and Nerlove (1989) in the epigraph. Volatility persistence as a function of market state variables, e.g. $\beta(r_t, \sigma_t^2)$, is termed the conditional volatility persistence (CVP), akin to the conditional variance in the GARCH-family models.

Since the proposed mechanisms for CVP are strongest at daily frequency, our empirical analyses focus on daily persistence of realized variance (RV). We modify the heterogeneous autoregressive (HAR) model of Corsi (2009) to allow RV persistence to vary with return and RV. The estimated HAR_CVP model shows that daily RV persistence changes with the size and sign of daily returns and the effect is economically large. For the S&P 500 ETF (ticker SPY), a +1% daily return implies an increase in volatility persistence by 14 to 16% of the average daily RV persistence, and a -1% daily return implies an increase in volatility persistence by 48 to 54%! The percentages are 13% and 29% respectively for large stocks in the S&P 100 index. The asymmetric return impact on volatility persistence remains highly significant after controlling the asymmetric effects of semi-variances in Patton and Sheppard (2015) and the impact of measurement errors in BPQ (2016). We find a small but significant negative impact from RV to its persistence. Overall the evidence strongly supports CVP and the proposed mechanisms for volatility persistence. Return and RV remain significant determinants of volatility persistence in models with non-linear CVP specifications, additional market-state variables, and alternative measures for volatility persistence.

The strong return impact on RV persistence documented here has not been formally studied in the literature. In a HAR model with semi-variances, Patton and Sheppard (2015) estimate the impact of $RV_t I_{(r_t<0)}$ on $RV_{t+1}$ where $I_{(.)}$ is an indicator variable. They report positive and highly significant coefficients of $RV_t I_{(r_t<0)}$, which is consistent with our finding of negative returns associated with high volatility persistence. However, Patton and Sheppard (2015) term $RV_t I_{(r_t<0)}$ as “a simple leverage effect variable” unrelated to volatility persistence. Palandri (2015) examines the persistence of positive and negative semi-variances. He reports sharp differences in the half lives of positive and negative semi-variances when semi-variances are estimated from GARCH-based volatility models. However, the differences largely disappear when semi-variances are calculated from intraday returns.

Evidence linking volatility level with its persistence is limited. Ning, Xu, and Wirjanto (2015) report that daily RV has greater right-tail dependence than left-tail dependence, i.e. high RV levels have greater persistence. Their study does not consider the effect of return on RV persistence. We emphasize the differential impact of return and RV on RV persistence. Our CVP combines a large positive effect from return size with a small negative effect from RV. Daily CVP is indeed positively correlated with RV (Table 3). BPQ (2016) document a negative impact on RV persistence from RV measurement errors captured by realized quarticity (RQ). Since RQ and RV are highly correlated (Table 1), their result is consistent with our finding of a negative impact from RV to RV persistence. However, the underlying mechanisms are very different. They argue that RQ reduces RV’s information content and its impact on future RV. In their model, high RV periods have high RQ and low RV persistence (see the example in their Figure 2). We emphasize the impact of information shocks and price discovery on RV persistence. By conditioning RV persistence on daily returns, daily CVP is high in high RV periods, e.g. during the global financial crisis of 2008-09.

Conditional volatility persistence provides a new link from return to future volatility. Based on Shapley $R^2$, this new link has much higher explanatory power than the direct impact from return to future volatility. It offers a new explanation for asymmetric volatility, in addition to financial leverage (Black, 1976), volatility feedback (Pindyck, 1984), and herding and contrarian trading (Avramov, Chordia, and Goyal, 2006). CVP implies an asymmetric return impact on *volatility persistence*: negative returns increase volatility persistence more than positive returns. Ceteris paribus, higher persistence leads to higher volatility tomorrow. Thus CVP provides an alternative mechanism for the asymmetric return impact on future volatility. In section II, we show that in the well-known Glosten, Jagannathan, and Runkle (1993, GJR hereafter) model, asymmetric volatility comes *entirely* from asymmetric volatility persistence. In section V, we demonstrate that CVP sharply reduces the asymmetry in RV attributed to the direct return impact, by 57 to 67% for SPY and 46 to 58% for stocks. Overall the evidence indicates that at daily frequency, CVP is the dominant link from return to future volatility and the dominant source for volatility asymmetry.

By estimating the dependence of future RV on today's RV, our HAR_CVP model significantly improves daily volatility forecasts relative to the HARSV model of Patton and Sheppard (2015) and the HARQ model of BPQ (2016). Based on four loss functions, the median loss values of HAR_CVP are 8 to 28% lower than HARQ and HARSV for SPY and 10 to 16% lower for individual stocks. The average loss values of HAR_CVP are often more than 40% lower than the competing models. DM tests show the loss reductions are highly significant. HAR_CVP's superior performance is robust to most market conditions. It is even stronger on days with large positive or negative returns, e.g. during the global financial crisis. It is also stronger in periods (and for stocks) with high CVP variations. These findings indicate that the superior forecast accuracy of HAR_CVP comes from its ability to capture large variations in the latent daily volatility persistence. HAR and HARSV have constant volatility persistence for the rolling forecast windows and HARQ adjusts volatility persistence only to realized quarticity (RQ). We find mixed evidence on whether a high CVP value itself is the source for HAR_CVP's superior forecast accuracy.

This paper has the following sections. Section II outlines the economic mechanisms in which daily volatility persistence increases with the size of daily returns and decreases with volatility level. They motivate the empirical specification for conditional volatility persistence. In section III, we review and modify recent models of RV dynamics to allow conditional persistence. Section IV presents empirical evidence on the determinants and the characteristics of daily volatility persistence, together with a range of robustness checks. Section V shows that CVP offers an alternative and robust mechanism for the asymmetric impact of return on future volatility. Section VI compares volatility forecasts of models with constant or conditional volatility persistence. We conclude in Section VII.

## II. Return, Volatility, Volatility Persistence, and Asymmetric Volatility

This section first explores the economic links from daily return and volatility to volatility persistence. The MDH is taken as the baseline case: the clustering of periodic macro news releases and company disclosures, is an important source for volatility persistence. We argue heuristically additional economic mechanisms linking daily return and volatility to volatility persistence. Using the GJR model, we show that the well-known asymmetric volatility can be the results of an asymmetric impact from return to volatility persistence. The GJR model is in fact a model of asymmetric volatility persistence.

### *Return and volatility persistence*

There are several economic mechanisms linking daily return to volatility persistence. The first mechanism is based on the observation that daily information arrivals are partially endogenous. We define correlated information as pieces of information relating to the same information event. A large return $|r_t|$ reflects the net price impact of an information event on day t. It draws greater investor and media attention, triggering further information search and the arrivals of correlated information on day t+1.[2] These correlated information arrivals increase the correlation between $RV_t$ and $RV_{t+1}$. When the information event is ambiguous, one would expect more information searches and more subsequent arrivals of correlated information. Boudoukh, et al. (2015) examine volatility persistence after large daily price changes, e.g. $|r_t|$ above one standard deviation. They show that volatility does not persist when its source can be identified; volatility is persistent only when it is driven by unidentified and complex events. Dimpfl and Jank (2016) report that large price swings drive more Google searches on stocks, leading to higher volatility on the next day. Economic intuition and empirical evidence suggest that following large and complex information shocks, investors increase their effort in searching for information and explanations. The arrivals of correlated

[2] Evidence on investor attention and endogenous information arrivals can be found in Cao, Coval, Hirshleifer (2002), Barber and Odean (2008), Hou, Peng, and Xiong (2009), and Andrei and Hasler (2015) among others. Recent studies link volatility to internet search activities, e.g. Da, Engelberg, and Gao (2011), Drake, Roulstone, and Thornock (2012), and Dimpfl and Jank (2016). Such information search increases short-run information arrivals. It is different from long-run information production, e.g. Veldkamp (2005, 2006) and Brockman, Liebengerg, and Schutte (2010). It also differs from price discovery through trading.

information increase volatility persistence. On the other hand, a small $|r_t|$ is unlikely to motivate investors to seek further information; volatility persistence is driven by the persistence in exogenous information arrivals as in the MDH.

There are non-information channels linking large returns to high volatility persistence. For example, large returns can trigger portfolio adjustments by passive investors. Since investors cannot perfectly separate informed and uninformed trading, such non-information trading may increase volatility and volatility persistence. Large returns may also cause herding trades, e.g. momentum traders trading in one direction and contrarian traders trading in the opposite direction. Kremer and Nautz (2013a and 2013b) find strong evidence of short-term institutional herding from daily data, with lagged daily return as a significant determinant for institutional herding. Persistence in herding is likely to increase volatility persistence. In addition, large returns may trigger long-run investors to react to short-run information shocks. As illustrated by Xue and Gençay (2012), the interaction of short and long-run investors can lead to volatility persistence.

We expect large negative $r_t$ to increase volatility persistence more than large positive $r_t$. Investors tend to have strong loss aversion (Kahneman and Tversky, 1979). They pay more attention to bad news, resulting in greater endogenous information search and correlated information arrivals after large negative $r_t$. Andrei and Hasler (2015) provide theoretical and empirical evidence that greater investor attention is a source for stock volatility especially during bad times. Dzielinski, Rieger, and Talpsepp (2018) show that investor attention is a source for the asymmetric return-volatility relation. Consistent with greater investor attention to negative returns, Patton and Sheppard (2015) show that future volatility depends more on today's negative semi-variance that reflects bad news.

***Volatility level and its persistence***

The economic link between volatility level and persistence is centred on price discovery. Andersen (1996) defines the price discovery process as "private information arrivals induce a dynamic learning process that results in prices fully revealing the content of the private information through the sequence of trades and transaction prices." In his MDH model, "each informational arrival induces a price discovery phrase followed by an equilibrium phrase." While investors acquire information through information search discussed above, price discovery is the trading process incorporating new information into a new equilibrium price. Greater price discovery implies that more information is being priced and less correlated information arrivals on the next day. Therefore, price discovery reduces volatility persistence.

Hasbrouck (1993) develops a widely used measure for price discovery. The observed return ($r_t$) is decomposed into a random-walk component ($m_t$) capturing the price impact of information arrivals, and a serially correlated noise ($n_t$) reflecting the effects of microstructure frictions, transaction costs, behavioural biases, etc. Price discovery is measured by var($m_t$) and pricing error is measured by var($n_t$).[3] Based on a sample of 1361 NYSE stocks, Boehmer and Wu (2013) report the average $\mathrm{var}(n_t)^{1/2}\mathrm{var}(r_t)^{-1/2}$ to be 0.095 with a median of 0.062. Based on 50 stocks in the S&P 500 index, Ozturk, van der Wel, and van Dijk (2017) report that var($m_t$) is about 20 times larger than var($n_t$). These results suggest that var($r_t$) is dominated by, thus, a close proxy for var($m_t$). Since price discovery reduces volatility persistence, we expect a negative relationship between var($r_t$) and its persistence.

The proposed inverse relationship between volatility level and its persistence appears to contradict the common perceptions that "volatility drives volatility", e.g. high-volatility days tend to be followed by high volatilities. Under the proposed mechanisms, persistence is due to large swings in daily return $r_t$, not high $RV_t$ itself. When $RV_t$ is high, the market is experiencing an active price discovery period. If the resulting $r_t$ is small, it indicates that investors' prior expectation has been reaffirmed; there is less information search and less endogenous information arrivals on day t+1. Therefore, high $RV_t$ and small $r_t$ lead to low volatility persistence. Empirically the uncertainty associated with a large information shock, as approximated by $r_t$, is usually not resolved by daily price discovery approximated by $RV_t$. Our estimated daily $CVP_t$ is indeed positively correlated with $RV_t$ (Table 3).

***Conditional Volatility Persistence***

The above discussion implies that (1) daily volatility persistence increases with daily return size; (2) negative returns increase volatility persistence more than positive returns; and (3) daily volatility level reduces daily volatility persistence. In this study, we test the above implications using a simple linear specification:

$$CVP_t = \beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{RV} RV_t$$

Let $r_t^- = r_t I_{(r_t<0)}$ and $r_t^+ = r_t I_{(r_t>0)}$. An equivalent specification is $CVP_t = \beta_0 + (\beta_r - \beta_{|r|})r_t^- + (\beta_r + \beta_{|r|})r_t^+ + \beta_{RV} RV_t$. The above implications become $\beta_{|r|} > 0$, $\beta_r < 0$, $\beta_{RV} < 0$, and $|\beta_r - \beta_{|r|}| > \beta_r + \beta_{|r|} > 0$. The return-based component of CVP is termed the asymmetric volatility persistence (AsyVP). It can be a source for asymmetric volatility as shown below and in section V. Return

[3] Various extensions to the Hasbrouck model have been developed, e.g. De Jong and Schotman (2010), Yan and Zivot (2010), Wang and Yang (2011), and Putnins (2013).

and volatility are the outcomes of the overall trading process, e.g. the information arrivals and uninformed trading. The above CVP improves upon the MDH where volatility persistence is determined only by the number of information arrivals.

We note that CVP is economically and statistically different from models of volatility regime switching, where regimes are based on volatility levels and persistence is constant within each regime. It also differs from regressions with time-varying coefficients where the determinants for time-varying coefficients are unknown. The proposed mechanisms work better for short-run persistence than for long-run persistence. Large daily returns are better proxies for information shocks and have greater effect on investor attention than large weekly or monthly return. Therefore, endogenous information arrivals and non-information trading are likely to have a strong effect on daily volatility persistence. Long-run persistence may be driven by other mechanisms.[4] Recent studies by Patton and Sheppard (2015) and BPQ (2016) also focus on daily persistence while controlling long-run dependence.

***Asymmetric Volatility Persistence and Asymmetric Volatility***

As discussed above, negative returns should increase volatility persistence more than positive returns. Ceteris paribus, higher persistence leads to higher future volatility. Thus the asymmetric volatility persistence may lead to asymmetric volatility, without invoking financial leverage or volatility feedback.

We illustrate this mechanism using the GJR model for volatility dynamics. Assuming $E(r_t) = 0$, the variance equation of the GJR model is

$$\sigma_{t+1}^2 = \omega + \alpha\sigma_t^2 + \left(\beta + \lambda I_{(r_t<0)}\right) r_t^2$$

Since $E(r_t^2|\mathfrak{F}_{t-1}) = \sigma_t^2$, where $\mathfrak{F}_{t-1}$ is the information set at t-1, $\lambda$ can be interpreted as the incremental effect of $\sigma_t^2$ on $\sigma_{t+1}^2$ when $r_t < 0$. The variance equation can be written as

$$\sigma_{t+1}^2 = \omega + \left(\alpha + \beta + \lambda I_{(r_t<0)}\right)\sigma_t^2 + \left(\beta + \lambda I_{(r_t<0)}\right)\varepsilon_t$$

where $\varepsilon_t \equiv r_t^2 - \sigma_t^2$. Here volatility persistence is time-varying and asymmetric. Since return is the only conditioning variable, $CVP_t = AsyVP_t = \alpha + \beta + \lambda I_{(r_t<0)}$. Since $E(\varepsilon_t|\mathfrak{F}_{t-1}) = 0$, return impact on $\sigma_{t+1}^2$ comes from AsyVP: $E(\sigma_{t+1}^2|r_t>0,\mathfrak{F}_{t-1}) < E(\sigma_{t+1}^2|r_t<0,\mathfrak{F}_{t-1})$ is caused by low persistence $\alpha+\beta$ when $r_t > 0$ versus high persistence $\alpha+\beta+\lambda$ when $r_t < 0$.[5] Therefore, *in the GJR*

[4] Liesenfeld (2001) shows that "the short-run volatility dynamics are directed by the information arrival process, whereas the long-run dynamics are associated with the sensitivity to new information."

[5] Metaphorically the flow of water depends not only on the water level in the tank ($\sigma_t^2$) but also on the time-varying size of the pipe ($\alpha+\beta$ or $\alpha+\beta+\lambda$).

*model, asymmetric volatility comes entirely from AsyVP.* Many other GARCH models have the same property, i.e. volatility persistence depends on recent returns.

The significance of the alternative representation of the GJR model is to reveal CVP as a potential new source for asymmetric volatility. Neither the original GJR study nor subsequent studies on asymmetric volatility have explored the modelling and implications of time-varying volatility persistence. However, the empirical success of the GJR model lends support to the alternative representation and interpretation. In section V, we further explore the degree of asymmetry in RV persistence and present evidence that CVP sharply reduces the asymmetry in the news impact curve of Engle and Ng (1993) and Chen and Ghysels (2010).

## III. Modelling Volatility Dynamics

Investors in financial markets have different investment horizons: high-frequency traders often reverse their positions within a few seconds or shorter while pension funds typically hold their positions for several months if not longer. Investors trade on different information (e.g. order flow versus fundamental value) at different frequencies (e.g. intraday versus quarterly), therefore have different impact on future volatility. Müller, et al. (1997) suggests a heterogeneous ARCH model to capture the impact from investors trading at different frequencies. Corsi (2009) argues that investors can be broadly classified as trading at daily, weekly, and monthly frequencies. He proposes a heterogeneous autoregressive (HAR) model that allows differential impact from three volatility components: the daily $RV_t$, the average weekly $RV_{t,W} \equiv \frac{1}{5}\sum_{i=0}^{4} RV_{t-i}$ and the average monthly $RV_{t,M} \equiv \frac{1}{22}\sum_{i=0}^{21} RV_{t-i}$. Future RV follows an autoregressive structure of the three volatility components:

$$(1) \qquad RV_{t+1} = \alpha + \beta_D RV_t + \beta_W RV_{t,W} + \beta_M RV_{t,M} + \varepsilon_{t+1}$$

The coefficients, $\beta_D$, $\beta_W$, and $\beta_M$ measure the step-wise dependence of $RV_{t+1}$ on short and long-run volatility. In spite of its simplicity, the HAR model can generate long memory in volatility and has good out-of-sample forecast performance. Corsi (2009), Craioveanu and Hillebrand (2012), and Audrino and Knaus (2016) show that it has equal or better forecasts than the fractionally integrated ARMA model, models with flexible lag structures, and models selected by the lasso-based procedures. It is closely related to the mixed data sampling (MIDAS) models of Ghysels, Santa-Clara, and Valkanov (2006) and Ghysels, Sinko, and Valkanov (2007), and has better volatility forecasts than a freely parameterized MIDAS model (Clements, Galvao, and Kim, 2008). According to Bollerslev, et al. (2017), it has become "a benchmark in the financial econometrics literature for judging other RV-based forecasting procedures."

The asymmetric relationship between asset return and volatility has been extensively documented. Corsi and Reno (2012) demonstrate that the return impact on future volatility is highly persistent and propose a HAR structure for returns to capture their heterogeneous effects. The lagged weekly ($r_{t,W}$) and monthly ($r_{t,M}$) returns are similarly defined as $RV_{t,W}$ and $RV_{t,M}$. The heterogeneous return impact is estimated in a modified HAR model:

(2) $$RV_{t+1} = \alpha + \beta_D RV_t + \beta_W RV_{t,W} + \beta_M RV_{t,M} + \theta_D r_t + \theta_W r_{t,W} + \theta_M r_{t,M} + \varepsilon_{t+1}$$

We take the model in (2) as our baseline model for RV dynamics and further calibrate daily RV persistence ($\beta_D$) in terms of the observed market state variables below.

Empirical evidence on time-varying RV persistence as captured by $\beta_D$ in (2) remains limited. Forsberg and Ghysels (2007) allow $\beta_D$ to vary with a dummy for volatility jumps but report no change in RV dynamics. Forsberg and Ghysels (2007) and BPQ (2016) show theoretically that the autocorrelation of RVs is inversely related to integrated quarticity, which is the variance of measurement errors of daily $RV_t$. Integrated quarticity can be consistently estimated by realized quarticity defined as $RQ_t \equiv \frac{n}{3}\sum_{i=1}^{n} r_{i,t}^4$. BPQ (2016) use a simple linear function of $RQ_t^{1/2}$ to capture the effect of measurement errors on RV persistence:

(3) $$RV_{t+1} = \alpha + (\beta_0 + \beta_{RQ} RQ_t^{1/2}) RV_t + \beta_W RV_{t,W} + \beta_M RV_{t,M} + \varepsilon_{t+1}$$

The estimated $\beta_{RQ}$ is negative and highly significant. We denote the implied RV persistence as $CVP_t^{RQ} \equiv \beta_0 + \beta_{RQ} RQ_t^{1/2}$. The forecasting performance of the model in (3) surpasses several benchmark models. BPQ (2016, page 9) attribute the superior performance to "the model's ability to place a larger weight on the lagged daily RV on days when RV is measured relatively accurately (RQ is low), and to reduce the weight on days when RV is measured relatively poorly (RQ is high)." Their evidence indicates that the time-varying "weight on the lagged daily RV" can better capture the true RV persistence.

The current study explores economic determinants of RV persistence. Section II suggests that persistence increases with returns, especially negative returns, and decreases with RV. A linear representation of the above implications is $CVP_t = \beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{RV} RV_t$. The CVP coefficients are estimated from the modified HAR model:

(4) $$RV_{t+1} = \alpha + (\beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{RV} RV_t) RV_t + \varphi Z_t + \varepsilon_{t+1}$$

where $Z_t = (RV_{t,W}, RV_{t,W}, r_t, r_{t,W}, r_{t,M})'$ captures the dependence at longer horizons as well as the heterogeneous return impact, and $\varphi = (\beta_W, \beta_M, \theta_D, \theta_W, \theta_M)$. As in BPQ (2016), the lagged variables are non-overlapping, e.g. $RV_{t,W} \equiv \frac{1}{4}\sum_{i=1}^{4} RV_{t-i}$, $RV_{t,M} \equiv \frac{1}{17}\sum_{i=5}^{21} RV_{t-i}$, and similarly for $r_{t,W}$ and $r_{t,M}$. The model in (4) is termed the HAR_CVP model and is estimated in Section IV. The mechanism in Section II is supported if the CVP coefficients are significant with the expected signs. The HAR_CVP model also provides a direct comparison between two channels linking $r_t$ to $RV_{t+1}$: the new CVP effect via $\beta_{|r|}$ and $\beta_r$ and the "leverage" effect via $Z_t$ and its coefficient $\varphi$. The comparison is carried out in Section V.

## IV. Evidence on Conditional Volatility Persistence

This section presents the empirical results of the HAR_CVP model in (4). The impact of CVP on RV dynamics is measured through its coefficients and its explanatory power. We also explore the empirical characteristics of the estimated daily CVP. We conduct a range of robustness checks, including alternative models, an alternative measure for RV persistence, and additional conditioning variables. The evidence indicates that CVP has a strong and robust impact on future RV and explain a large portion of RV variations over time.

### (i) Sample and Summary Statistics

Our analyses are based on the S&P 500 ETF (ticker SPY) and the S&P 100 index constituent stocks. Our sample for SPY is from 2 January 2000 to 30 May 2014. From the S&P 100 constituent stocks, we remove seven stocks with less than five years of intraday data and six stocks with share prices below \$5 during the sample period. We find that in 2000 and 2001, several stock-months have less than 15 days of intraday data. Our sample of 87 stocks starts on or after 2 January 2002 and ends on 31 December 2014.

Intraday 5-minute data are extracted from the Thomson Reuters Tick History (TRTH) database, including the first, the high, the low, and the last prices, as well as the volume and the number of trades for each 5-minute interval. Data outside the NYSE trading hours are removed. We also remove short trading days, e.g. the day before July 4 and Christmas, and days with less than 3 hours of data possibly due to missing data or slow trading. To filter out possible data errors, we apply a filter similar to those of Barndorff-Nielsen, et al. (2009). For each 5-minute return, we calculate the standard deviation of the remaining returns on the same day. A return is removed if it is outside 6 standard deviations from zero. The filter removes 246 intervals for SPY, representing 0.088% of the 5-minute sample. It has no effect on 96.3% of the SPY trading days. Of the remaining 3.7% trading days, 2.9% have unfiltered realized variance larger than the filtered ones by 50% or more. Therefore, the filter removes extremely large returns relative to the rest of the trading day.

Our measure for daily volatility is realized variance (RV) based on 5-minute returns. This is a common practice as surveyed by Hansen and Lunde (2006) and remains popular in recent studies.[6] Let $r_{i,t}$ be the log-return in interval i and n (=78) be the number of intraday intervals on a trading day t. $RV_t$ is calculated as $\sum_{i=1}^{n} r_{i,t}^2$. Zhang, Mykland, and Aït-Sahalia (2005) propose to use sub-grids to improve RV precision. Patton and Sheppard (2015) uses ten grids of 5-minute returns to calculate ten RVs and take the average as the daily RV. We use the average RV from two grids of 5-minute returns from the first and the last price in each 5-minute interval. To compare with models proposed by Patton and Sheppard (2015) and BPQ (2016), we also calculate the negative semi-variance $NSV_t = \sum_{i=1}^{n} r_{i,t}^2 I_{(r_{i,t}<0)}$, the positive semi-variance $PSV_t = \sum_{i=1}^{n} r_{i,t}^2 I_{(r_{i,t}>0)}$, and the realized quarticity $RQ_t = \frac{n}{3}\sum_{i=1}^{n} r_{i,t}^4$. As a robustness check, we also use RV measures based on subsampling and realized kernel that reduce the impact of microstructure noise. For SPY, we obtain daily RV and semi-variances from the Oxford-Man Institute (OMI) realized library: the $RV^{OM}$ from 5-minute returns with 1-minute subsampling and the corresponding semi-variance, and the realized kernel $RK^{OM}$ after removing microstructure noise. RV measures from OMI have superscript OM while RV based on TRTH data has no superscript.

Panel A of Table 1 presents summary statistics on our return and RV measures, and Panel B reports their correlations. For stocks, the summary statistics are calculated for each stock and then are averaged across stocks. Since returns are calculated in percentage, the realized variance is inflated by $10^4$. This also appears to be the case for OMI data. The mean, median, and standard deviation of our RV are broadly similar to those of $RV^{OM}$. Our RV has lower skewness and kurtosis than $RV^{OM}$. Both our RV and $RV^{OM}$ have lower mean, median, and standard deviation than the realized kernel $RK^{OM}$. To facilitate replication, we use the OM daily returns which are publicly available. They are very similar to the daily returns from DataStream with a correlation coefficient of 0.97. Not surprisingly, stocks have higher RV than SPY, with an average of 2.50 and standard deviation 5.60. Daily correlations in Panel B are broadly consistent with the literature. Daily returns have negative correlations with different RV measures except PSV. The realized variances and semi-variances have strong positive correlations ranging from 0.85 to 0.97.

[6] E.g. Patton and Ramadorai (2013), Amaya, Christoffersen, Jacobs, and Vasquez (2016), BPQ (2016), Bollerslev, Li, and Zhao (2017).

**Table 1: Data Summary**

For SPY, return is the daily return from the Oxford-Man Institute (OMI) Realized Library. RV is the Realized Variance based on Thomson-Reuters data. NSV and PSV are the negative and positive semi-variances. RQ is the Realized Quarticity. $RV^{OM}$ is the Realized Variance from the OMI Realized Library. $NSV^{OM}$ and $PSV^{OM}$ are the negative and positive semi-variances. $RK^{OM}$ is the Realized Kernel from the OMI Realized Library. All variances and $RQ^{1/2}$ are scaled by $10^4$. LB5 is the Ljung-Box statistic for 5 lags. For stocks, all variables are based on the Thomson-Reuters data. Summary statistics and correlations are calculated for each stock and then averaged across stocks.

Panel A: Summary statistics

| | Mean | Median | St Dev | Skew | Kurt | LB5 |
|---|---|---|---|---|---|---|
| **SPY** | | | | | | |
| RV | 1.13 | 0.551 | 2.35 | 10.6 | 185 | 7074 |
| NSV | 0.558 | 0.268 | 1.09 | 8.04 | 98 | 8028 |
| PSV | 0.570 | 0.261 | 1.34 | 13.6 | 306 | 4909 |
| $RQ^{1/2}$ | 1.29 | 0.625 | 2.88 | 13.2 | 291 | 4872 |
| Return | 0.007 | 0.064 | 1.25 | -0.152 | 10.2 | 41.2 |
| $RV^{OM}$ | 1.10 | 0.537 | 2.49 | 14.8 | 402 | 5946 |
| $NSV^{OM}$ | 0.551 | 0.256 | 1.25 | 13.4 | 329 | 5671 |
| $PSV^{OM}$ | 0.551 | 0.255 | 1.30 | 15.3 | 415 | 4960 |
| $RK^{OM}$ | 1.25 | 0.615 | 2.75 | 14.2 | 378 | 6032 |
| **Stocks** | | | | | | |
| Return | 0.018 | 0.037 | 2.31 | -5.37 | 247 | 19.8 |
| RV | 2.50 | 1.21 | 5.60 | 9.64 | 172 | 6065 |
| NSV | 1.24 | 0.585 | 2.83 | 9.63 | 174 | 5665 |
| PSV | 1.26 | 0.593 | 2.93 | 10.2 | 185 | 4976 |
| $RQ^{1/2}$ | 3.14 | 1.44 | 7.83 | 10.8 | 220 | 4631 |

Panel B: Correlations

| | Return | RV | NSV | PSV | $RQ^{1/2}$ | $RV^{OM}$ | $NSV^{OM}$ | $PSV^{OM}$ |
|---|---|---|---|---|---|---|---|---|
| **SPY** | | | | | | | | |
| RV | -0.067 | | | | | | | |
| NSV | -0.224 | 0.963 | | | | | | |
| PSV | 0.065 | 0.976 | 0.881 | | | | | |
| $RQ^{1/2}$ | -0.053 | 0.976 | 0.929 | 0.961 | | | | |
| $RV^{OM}$ | -0.066 | 0.957 | 0.906 | 0.946 | 0.940 | | | |
| $NSV^{OM}$ | -0.249 | 0.926 | 0.920 | 0.881 | 0.910 | 0.973 | | |
| $PSV^{OM}$ | 0.114 | 0.937 | 0.846 | 0.961 | 0.921 | 0.975 | 0.898 | |
| $RK^{OM}$ | -0.056 | 0.944 | 0.893 | 0.935 | 0.927 | 0.995 | 0.967 | 0.972 |
| **Stocks** | | | | | | | | |
| RV | -0.058 | | | | | | | |
| NSV | -0.171 | 0.966 | | | | | | |
| PSV | 0.052 | 0.970 | 0.875 | | | | | |
| $RQ^{1/2}$ | -0.058 | 0.975 | 0.947 | 0.940 | | | | |

## (ii) Estimation of the HAR Models

Patton and Sheppard (2015) point out that because the dependent variable in a HAR model is a volatility measure, OLS estimates tend to overweigh periods with high volatility and under-weigh periods with low volatility. Therefore, OLS residuals have heteroskedasticity related to the level of RV. We follow Patton and Sheppard (2015) and use the weighted least squares (WLS) to overcome this problem. For SPY, inference is based on the Newey-West robust covariance with automatic lag selection using Bartlett kernel. For individual stocks, the reported coefficients are the cross-sectional averages. Following Hameed, Kang, and Viswanathan (2010), the standard error of the $k^{th}$ average coefficient $\bar{\hat{\beta}}_k$ is given by

$$\text{StDev}\left(\bar{\hat{\beta}}_k\right) = \text{StDev}\left(\frac{1}{N}\sum_{i=1}^{N}\hat{\beta}_{i,k}\right) = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\hat{\omega}_{i,j}\sqrt{\text{Var}(\hat{\beta}_{i,k})\text{Var}(\hat{\beta}_{j,k})}} \tag{5}$$

where $\text{Var}(\hat{\beta}_{i,k})$ is based on the Newey-West standard error of the regression of stock i and $\hat{\omega}_{i,j}$ is the correlation between the regression residuals for stocks i and j.

## (iii) Empirical Evidence on CVP

### *The HAR and HAR_CVP Models for SPY*

Table 2 reports the estimation results of the baseline HAR model in (2) and the HAR_CVP model in (4). Panels A and B report the results for SPY and individual stocks respectively. We focus first on SPY and discuss the results for individual stocks shortly below. The realized variance measures for SPY are $RV_t$ based on TRTH data, $RV_t^{OM}$ and $RK_t^{OM}$ from the OMI realized library. The baseline HAR model restricts $\beta_{|r|} = \beta_r = \beta_{RV} = 0$. As in earlier studies, the lagged weekly and monthly RVs and all lagged returns are significant at 1% or 5%, consistent with high volatility persistence and heterogeneous effects of returns. Overall the RV dynamics based on $RV_t$, $RV_t^{OM}$, and $RK_t^{OM}$ are qualitatively similar.

The HAR_CVP model results confirm that $\beta_{|r|} > 0$, $\beta_r < 0$, and $\beta_{RV} < 0$, supporting the mechanisms and hypotheses on the determinants of CVP in Section II. The F test resoundingly rejects $\beta_{|r|} = \beta_r = \beta_{RV} = 0$. The return size effect $\beta_{|r|}$ is numerically larger and statistically stronger than the return sign effect $\beta_r$. The signs of the coefficients of positive and negative returns, $\beta_{|r|}+\beta_r > 0$ and $\beta_r - \beta_{|r|} < 0$, indicate that they both increase RV persistence with negative returns having greater effects, i.e. $|\beta_r - \beta_{|r|}| > \beta_r + \beta_{|r|} > 0$. For $RV_t$, $CVP_t = 0.222 + 0.108|r_t| – 0.0572r_t – 0.0043RV_t$: a -1% return increases CVP by $(\beta_r-\beta_{|r|})(-1) = 0.165$ and a +1% return increases CVP by $\beta_{|r|}+\beta_r = 0.051$, holding $RV_t$ fixed. Over the sample period, the average of daily $CVP_t$

is 0.308 (Table 3). Ceteris paribus, a -1% return increases CVP by 54% (=0.165/0.308) relative to its average and a 1% return increases CVP by 16% (=0.051/0.308). The numbers are 50% and 14% for $RV_t^{OM}$, and 48% and 15% for $RK_t^{OM}$. Given the daily return standard deviation of 1.25% (Table 1 Panel A), days with $|r_t| > 1\%$ and large changes in volatility persistence occur quite often. The impact of RV on CVP is relatively small. Ceteris paribus, a one-standard deviation (Table 1) increase in $RV_t$ reduces CVP by -0.0101 which is -3.3% of the average CVP. The corresponding values are -2.2% for $RV_t^{OM}$ and -1.9% for $RK_t^{OM}$.

The model diagnostics show remarkable improvements of fit for HAR_CVP. The $\overline{R}^2$ increases by 11.2% for both $RV_t$ (0.729-0.617) and $RV_t^{OM}$ (0.666-0.554), and 9.6% for $RK_t^{OM}$ (0.654-0.558)! The $\overline{R}^2$s for HAR_CVP are 17 to 20% higher than those for HAR. The Akaike information criteria (AIC) for HAR_CVP are 20 to 46% lower than those for HAR. The Ljung-Box statistics for residuals at 5 lags (LB5) are 64 to 78% lower for HAR_CVP.

***The HAR and HAR_CVP Models for Individual Stocks***

Panel B of Table 2 reports the results for individual stocks. We report in Appendix A that in the baseline HAR model, stock returns are not significant when the S&P 500 index returns are included. In the HAR_CVP model, the S&P 500 returns have larger coefficients and greater significance than stock returns.[7] Therefore, in the analyses of individual stocks, daily returns are the S&P 500 index returns. We report the average coefficients across stocks with t statistics based on the Hameed, Kang, and Viswanathan (2010) standard error in (5). Inference for individual stock regressions is based on Newey-West robust covariance. We summarize how the model performs for individual stocks, with %($t \leq -1.96$) and %($t \geq 1.96$) being the percentages of stocks with significant coefficients at 5%.

The results for individual stocks are qualitatively the same as those for SPY. The cross-stock average coefficients $\beta_{|r|} > 0$, $\beta_r < 0$, and $\beta_{RV} < 0$, supporting the CVP mechanisms in Section II. The F test rejects $\beta_{|r|} = \beta_r = \beta_{RV} = 0$. At 5% level of significance, 77% of the stocks have significant $\beta_{|r|} > 0$, 47% have significant $\beta_r < 0$, and 90% have significant $\beta_{RV} < 0$. The cross-sectional median coefficients are similar to the averages. The return effect on CVP is large: across all stocks, a +1% (-1%) return increases RV persistence by 0.0413 (0.0903), which is 13% (29%) of the average CVP (0.434). The mean and median $\hat{\beta}_{RV}$ are larger than those for SPY. Given the standard deviation of 5.60 for stock RVs, a one-standard deviation increase in stock $RV_t$ reduces CVP by -0.0336 which is -10.9% of the average CVP.

[7] Vlastakis and Markellos (2012) show that firm-level RV increases with demand for market-related information but is largely unrelated to demand for firm-specific information.

**Table 2: The HAR and HAR_CVP Models**

This table reports the estimation of the following regression:

$$RV_{t+1} = \alpha + (\beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{RV} RV_t) RV_t + \beta_W RV_{t,W} + \beta_M RV_{t,M} + \theta_D r_t + \theta_W r_{t,W} + \theta_M r_{t,M} + \varepsilon_{t+1}$$

The variables are the same as in Table 1. AIC is the Akaike information criteria. F is the F statistic for $\beta_{|r|} = \beta_r = \beta_{RV} = 0$. LB5 is the Ljung-Box statistic for residuals at 5 lags. In Panel A, the t statistics are based on the Newey–West robust covariance with Bartlett kernel. In Panel B, the reported t statistics are based on the standard error for the average coefficient proposed by Hameed, Kang, and Viswanathan (2010). The t statistics for individual stock regressions are based on the Newey–West robust covariance with Bartlett kernel. %(t≤-1.96) and %(t≥1.96) are the percentage of stocks with significant coefficients at 5%. The asterisks ***, **, * indicate significance at 1%, 5%, and 10% respectively. In Panel C, Δ% is the percentage change of Shapley $R^2$ relative to the standard HAR model with $\beta_{|r|} = \beta_r = \beta_{RV} = 0$.

Panel A: SPY

| | $\beta_0$ | $\beta_{\lvert r\rvert}$ | $\beta_r$ | $\beta_{RV}$ | $\beta_W$ | $\beta_M$ | $\theta_D$ | $\theta_W$ | $\theta_M$ | $\bar{R}^2$ | F | AIC | LB5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RV | 0.223* | | | | 0.490*** | 0.151*** | -0.396*** | -0.559*** | -0.361** | 0.617 | | 2692 | 259 |
| *t-stat* | *1.90* | | | | *3.16* | *2.89* | *-2.93* | *-2.85* | *-2.47* | | | | |
| | 0.222** | 0.108*** | -0.0572** | -0.0043*** | 0.316*** | 0.093* | -0.096 | -0.385*** | -0.331** | 0.729 | 489 | 1467 | 57 |
| | *2.58* | *3.49* | *-2.13* | *-3.57* | *4.62* | *1.86* | *-1.39* | *-3.34* | *-2.10* | | | | |
| $RV^{OM}$ | 0.207 | | | | 0.455*** | 0.188*** | -0.394*** | -0.654** | -0.408*** | 0.554 | | 3627 | 140 |
| *t-stat* | *1.64* | | | | *3.76* | *3.30* | *-2.68* | *-2.37* | *-2.82* | | | | |
| | 0.234** | 0.099*** | -0.0562* | -0.0028*** | 0.274*** | 0.112** | -0.086 | -0.464*** | -0.329** | 0.666 | 397 | 2604 | 34 |
| | *2.34* | *2.76* | *-1.92* | *-2.59* | *3.73* | *2.32* | *-1.16* | *-2.74* | *-2.46* | | | | |
| $RK^{OM}$ | 0.225** | | | | 0.434*** | 0.190*** | -0.417*** | -0.724** | -0.477*** | 0.558 | | 4308 | 132 |
| *t-stat* | *1.96* | | | | *3.82* | *3.33* | *-2.84* | *-2.37* | *-2.96* | | | | |
| | 0.217** | 0.098*** | -0.0503* | -0.0021** | 0.263*** | 0.129*** | -0.114 | -0.549*** | -0.414*** | 0.654 | 328 | 3444 | 47 |
| | *2.16* | *2.93* | *-1.90* | *-2.12* | *3.59* | *2.70* | *-1.51* | *-2.74* | *-2.83* | | | | |

Panel B: Individual stocks

| | $\beta_0$ | $\beta_{\|r\|}$ | $\beta_r$ | $\beta_{RV}$ | $\beta_W$ | $\beta_M$ | $\theta_D$ | $\theta_W$ | $\theta_M$ | $\bar{R}^2$ | F | AIC | LB5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ave Coeff | 0.275*** | | | | 0.370*** | 0.210*** | -0.602*** | -1.069*** | -0.971*** | 0.585 | | 6327 | 71 |
| Median | 0.285 | | | | 0.375 | 0.202 | -0.458 | -0.899 | -0.666 | 0.631 | | 5347 | 53 |
| *t stat* | *4.70* | | | | *5.54* | *4.51* | *-4.13* | *-4.53* | *-3.40* | | | | |
| *%(t≤-1.96)* | *0%* | | | | *0%* | *0%* | *99%* | *98%* | *72%* | | | | |
| *%(t≥1.96)* | *87%* | | | | *94%* | *87%* | *0%* | *0%* | *0%* | | | | |
| Ave Coeff | 0.390*** | 0.0658*** | -0.0245*** | -0.0060*** | 0.273*** | 0.144*** | -0.267*** | -0.770*** | -0.662*** | 0.638 | 181 | 5845 | 36 |
| Median | 0.381 | 0.0673 | -0.0250 | -0.0053 | 0.270 | 0.148 | -0.142 | -0.666 | -0.493 | 0.687 | 163 | 4810 | 25 |
| *t stat* | *6.42* | *4.40* | *-2.62* | *-6.26* | *5.36* | *3.41* | *-2.85* | *-3.98* | *-2.91* | | | | |
| *%(t≤-1.96)* | *0%* | *0%* | *47%* | *90%* | *0%* | *0%* | *47%* | *92%* | *41%* | | | | |
| *%(t≥1.96)* | *91%* | *77%* | *1%* | *0%* | *93%* | *67%* | *0%* | *0%* | *0%* | | | | |

Panel C: Shapley $R^2$

| | $RV_t$ | $\|r_t\|RV_t$ | $r_tRV_t$ | $RV_t^2$ | $\|r_t\|RV_t+r_tRV_t+RV_t^2$ | $RV_{t,W}+RV_{t,M}$ | $r_t$ | $r_{t,W}+r_{t,M}$ | $R^2$ | $(\|r_t\|RV_t+r_tRV_t+RV_t^2)/R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| **SPY** | | | | | | | | | | |
| RV | | | | | | | | | | |
| HAR | 17.2% | | | | | 33.1% | 4.3% | 7.1% | 61.8% | |
| HAR_CVP | 13.4% | 19.6% | 4.6% | 3.8% | 28.0% | 24.4% | 2.1% | 5.2% | 73.0% | 38.4% |
| Δ% | -22.5% | | | | | -26.5% | -51.2% | -26.8% | | |
| $RV^{OM}$ | | | | | | | | | | |
| HAR | 14.9% | | | | | 30.1% | 3.7% | 6.8% | 55.5% | |
| HAR_CVP | 12.2% | 18.9% | 4.2% | 2.7% | 25.8% | 22.0% | 1.7% | 5.1% | 66.6% | 38.7% |
| Δ% | -18.4% | | | | | -26.9% | -54.1% | -25.0% | | |
| $RK^{OM}$ | | | | | | | | | | |
| HAR | 15.6% | | | | | 29.9% | 3.5% | 6.9% | 55.9% | |
| HAR_CVP | 12.4% | 18.3% | 3.4% | 2.6% | 24.2% | 22.0% | 1.7% | 5.2% | 65.5% | 36.9% |
| Δ% | -20.6% | | | | | -26.5% | -52.3% | -24.2% | | |
| **Stocks** | | | | | | | | | | |
| HAR | 17.9% | | | | | 31.6% | 2.2% | 6.8% | 58.6% | |
| HAR_CVP | 13.5% | 12.4% | 1.9% | 5.4% | 19.7% | 24.3% | 1.2% | 5.1% | 63.9% | 30.1% |
| Δ% | -24.6% | | | | | -23.2% | -45.5% | -25.0% | | |

The model diagnostics are strongly in favour of the HAR_CVP model. The average $\overline{R}^2$ increases by 5.3% and the median $\overline{R}^2$ increases by 5%. The $\overline{R}^2$s for HAR_CVP are 7.8 to 9.1% higher than those for HAR. The AIC and LB5 are substantially lower for HAR_CVP. We note that the median diagnostic statistics are better than the cross-stock averages, indicating outliers with unfavourable HAR_CVP diagnostics. Overall the evidence from individual stocks supports the mechanisms and hypotheses for volatility persistence outlined in Section II.

***Contribution of CVP to RV Dynamics***

A striking feature in Panels A and B is that when CVP is included, the coefficients of all lagged returns as well as the lagged weekly and monthly RVs are sharply reduced. The large and pervasive impact of CVP inclusion on the coefficients of other variables indicates that CVP captures important features in RV dynamics. Without CVP, the direct impact of return on future RV and the long-run RV persistence are significantly overstated.

To further assess the impact of CVP on RV dynamics, we compute the Shapley decomposition of the regression $R^2$. The Shapley $R^2$ of an explanatory variable measures its incremental explanatory power while controlling its correlations with all other explanatory variables.[8] We defined the CVP variables as $|r_t|RV_t$, $r_tRV_t$, and $RV_t^2$ in (4), excluding $RV_t$ whose coefficient is the constant volatility persistence in HAR. Panel C of Table 2 reports the Shapley $R^2$ of individual or group of variables and has several important features.

First, CVP explains large portions of changes in RV. Shapley $R^2$ of the CVP variables is 24~28% for SPY and 20% for individual stocks. The last column of Panel C shows that CVP variables account for 30 to 39% of the explanatory power of all variables. The CVP variable $|r_t|RV_t$ has the highest Shapley $R^2$ across all explanatory variables for SPY, and the third highest Shapley $R^2$ for individual stocks.

Second, for SPY, the combined Shapley $R^2$ for $|r_t|RV_t$ and $r_tRV_t$ is around 22~25%, much higher than the combined Shapley $R^2$ of $r_t$, $r_{t,W}$, and $r_{t,M}$ at around 7%. For stocks, the combined Shapley $R^2$ for $|r_t|RV_t$ and $r_tRV_t$ is 14.3% compared to 6.3% for returns. Almost 90% of individual stocks have higher Shapley $R^2$ for $|r_t|RV_t$ and $r_tRV_t$ than for returns. Therefore, CVP as a link from $r_t$ to $RV_{t+1}$ is two to three times more important than the impact of $r_t$, $r_{t,W}$, and $r_{t,M}$ on $RV_{t+1}$.

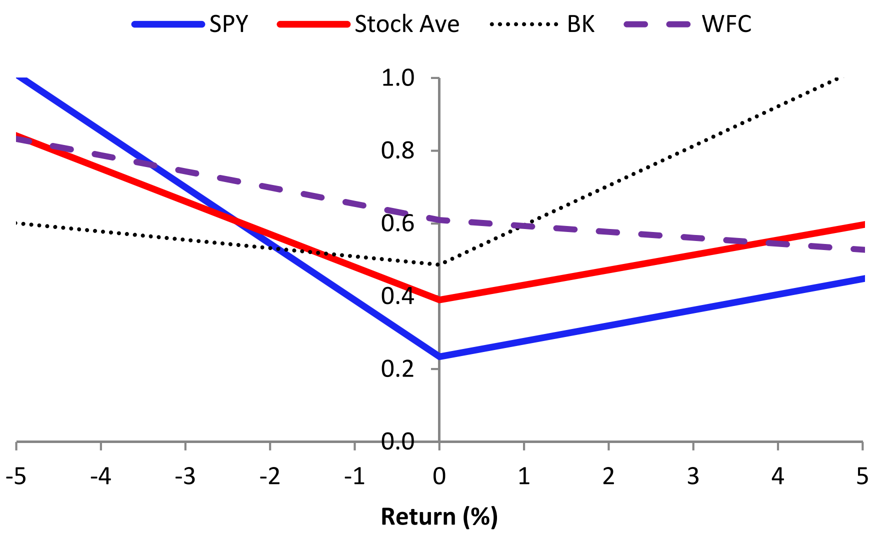
---

[8] Shapley decomposition provides a linear attribution of the regression $R^2$ to each explanatory variable. The Shapley $R^2$ of a group of variables is the sum of the Shapley $R^2$s within the group. The sum of Shapley $R^2$s across all variables is the total $R^2$. Lahaye and Neely (2016) provide a brief literature review, an illustrative example of the calculation, and a finance application. Owen and Prieur (2017) demonstrate the advantage of Shapley decomposition over ANOVA-based decompositions.

Third, although Panel B shows that the estimated coefficient of $RV_t$ ($\beta_0$) increases in size and significance in the HAR_CVP model, the Shapley $R^2$ of $RV_t$ in the HAR_CVP model for stocks is 25% lower than in the HAR model. In other words, around 25% of the explanatory power of $RV_t$ in the HAR model for stocks are crowed out when CVP variables are included. The corresponding number for SPY is 18 to 22%. Overall Panel C of Table 2 shows that CVP variables explain a very large portion of variations in RV.

**Figure 1: Asymmetric Volatility Persistence**

This figure depicts the relationship $AsyVP_t = \beta_0 + \beta^+ r_t^+ + \beta^- r_t^-$ where the coefficients are calculated from Table 2. The RV measure for SPY is $RV^{OM}$. The individual stocks are Bank of New York Mellon Corp (BK) and Wells Fargo (WFC).




***Asymmetric Volatility Persistence***

From the linear CVP specification in section II, the asymmetric volatility persistence is

$$AsyVP_t \equiv \beta_0 + \beta_{|r|}|r_t| + \beta_r r_t = \beta_0 + (\beta_r - \beta_{|r|})r_t^- + (\beta_r + \beta_{|r|})r_t^+.$$

From Panel A of Table 2, the estimated coefficients for SPY with $RV^{OM}$ implies that

$$AsyVP_t = 0.234 - 0.155 r_t^- + 0.0428 r_t^+$$

The average coefficients for stocks in Panel B of Table 2 implies that

$$AsyVP_t = 0.390 - 0.0903 r_t^- + 0.0413 r_t^+$$

These linear relations are depicted in Figure 1 together with those of two stocks. For SPY and stocks on average, the impact of $r_t^-$ on CVP is over twice the impact of $r_t^+$. This is similar to the AsyVP implied by the GJR model and appears to be a property of equity returns. The asymmetry is stronger for SPY than for stocks on average. Both the degree and the direction of the asymmetry vary across stocks. Bank of New York Mellon (BK) is one of the five stocks (out of 87) with the opposite asymmetry. Wells Fargo (WFC) is one of the five stocks where persistence is lower for large positive returns, i.e. $\beta_r+\beta_{|r|} < 0$. How AsyVP relates to the well-known asymmetric volatility is further explored in section V.

***Statistical Characteristics of Daily CVP***

We now summarize the statistical properties of the estimated daily CVP. Figure 2 depicts the time series and the histogram of the daily CVP for SPY based on $RV^{OM}$. CVP is very high during the height of the financial crisis in October 2008 due to large and mostly negative daily returns. Table 3 reports the summary statistics of the estimated daily CVP. They are qualitatively similar across RV, $RV^{OM}$, and $RK^{OM}$ for SPY. The mean values are significant at 1%. Medians are below the means, indicating some extremely high CVP values. The CVP histogram has a long right tail. Depending on the RV measure used, the estimated CVP > 1 on 10~14 days, representing 0.28% to 0.4% of the sample days. Although $\beta_{RV} < 0$, the correlations between CVP and the RV measures are above 0.3. This is consistent with the findings of Ning, Xu, and Wirjanto (2015) and is due to positive correlations between CVP and the absolute return (about 0.8) and positive correlations between the absolute return and RV (about 0.45). Ljung-Box statistics indicate that CVP is highly autocorrelated.

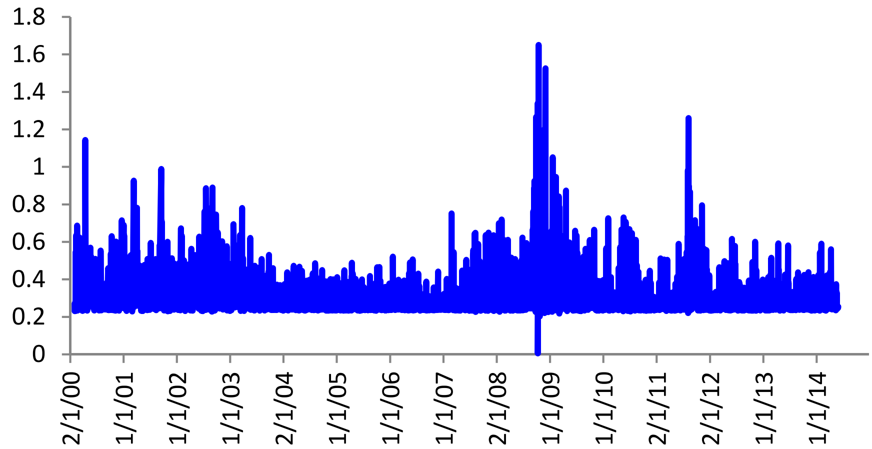
For stocks, we report cross-sectional mean and median summary statistics. Table 3 shows that across all stocks the average CVP is slightly higher and the average CVP standard deviation is slight lower than those of SPY. The median correlations are larger than the mean, indicating the presence of extremely low correlations.

**Table 3: Characteristics of the estimated CVP**

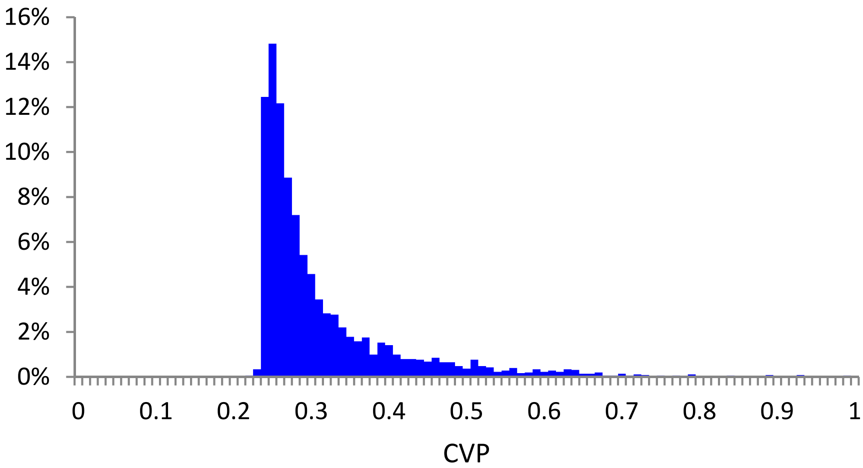
| | Summary Statistics | | | | | | Correlations | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | Median | St Dev | Skew | Kurt | LB5 | \|Return\| | Return | RV |
| **SPY** | | | | | | | | | |
| CVP(RV) | 0.308 | 0.264 | 0.122 | 3.53 | 20.3 | 573 | 0.812 | -0.627 | 0.313 |
| CVP($RV^{OM}$) | 0.314 | 0.272 | 0.116 | 3.58 | 20.6 | 600 | 0.797 | -0.648 | 0.321 |
| CVP($RK^{OM}$) | 0.297 | 0.258 | 0.110 | 3.55 | 20.6 | 715 | 0.823 | -0.614 | 0.349 |
| **Stocks** | | | | | | | | | |
| CVP(RV) | 0.434 | 0.412 | 0.133 | 2.71 | 25.4 | 356 | 0.763 | -0.474 | 0.067 |
| Median | 0.431 | 0.407 | 0.126 | 3.23 | 24.1 | 301 | 0.828 | -0.530 | 0.124 |

**Figure 2: Daily Conditional Volatility Persistence**

(A) Daily CVP($RV^{OM}$)




(B) Histogram of CVP($RV^{OM}$)




### (iv) Robustness Checks

We conduct several robustness checks for the HAR_CVP results in Table 2. We test whether CVP survives the inclusion of (1) RQ which affects RV persistence as shown by BPQ (2016); (2) semi-variances or jump variations as in Patton and Sheppard (2015), and (3) an alternative measure for daily volatility persistence. Results with additional conditioning variables and sub-period estimations are reported in the Appendix.

***CVP and Measurement Errors***

The HAR_CVP model in (4) offers an alternative mechanism for time-varying RV persistence to that of BPQ (2016) in (3). It would be of interest to jointly estimate strength of the two alternatives. However, the correlation of $RV_t$ and $RQ_t^{1/2}$ is extremely high at 0.976 for

SPY and 0.975 on average for the individual stocks, making it difficult to separate the impact of $RV_t$ and $RQ_t^{1/2}$. We test the impact of returns on CVP after taking into account of $RQ_t^{1/2}$:

(6) $$RV_{t+1} = \alpha + (\beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{RQ} RQ_t^{1/2}) RV_t + \varphi Z_t + \varepsilon_{t+1}$$

Table 4 reports the estimation results for (6). When $\beta_{|r|}$ and $\beta_r$ are set to zero, our results are qualitatively similar to those of BPQ (2016, Table 3): $\hat{\beta}_{RQ}$ is negative and highly significant,[9] confirming the negative impact of measurement errors on RV persistence. When $\beta_{|r|}$ and $\beta_r$ are not restricted to zero, they are numerically and statistically similar to the values in the HAR_CVP model in Table 2, and so are the model diagnostics. The combined Shapley $R^2$ is over 20% for $|r_t|RV_t$ and $r_tRV_t$ in (6) and is around 3% for $RQ_t^{1/2}RV_t$. The adjusted $R^2$ when $\beta_{|r|} = \beta_r = 0$ is much lower than the adjusted $R^2$ of unrestricted models. This helps to explain why HAR_CVP has better out-of-sample forecast accuracy than HARQ in section VI.

***CVP and Semi-variances***

Barndorff-Nielsen, Kinnerbrock, and Shephard (2010) show that daily RV can be decomposed into the positive and negative semi-variances: $RV_t = PSV_t + NSV_t$ where $PSV_t$ and $NSV_t$ are defined above in IV(i). The difference between $PSV_t$ and $NSV_t$ is termed the signed jump variations $SJV_t \equiv PSV_t - NSV_t$, with negative and positive jump variations defined as $NJV_t = SJV_t \times I_{(SJV<0)}$ and $PJV_t = SJV_t \times I_{(SJV>0)}$ respectively. The continuous component of $RV_t$ is the bi-power variation defined as $BV_t = \frac{\pi}{2}\sum_{i=2}^{n} |r_{i-1,t}||r_{i,t}|$. Patton and Sheppard (2015) estimate the impact of these components on $RV_{t+1}$:

(7) $$RV_{t+1} = \alpha + \beta_{NSV} NSV_t + \beta_{PSV} PSV_t + \theta_r RV_t I_{(r_t<0)} + \beta_W RV_{t,W} + \beta_M RV_{t,M} + \varepsilon_{t+1}$$

(8) $$RV_{t+1} = \alpha + \beta_{NJV} NJV_t + \beta_{PJV} PJV_t + \beta_{BV} BV_t + \beta_W RV_{t,W} + \beta_M RV_{t,M} + \varepsilon_{t+1}$$

They report that $NSV_t$ has much greater impact on $RV_{t+1}$ than $PSV_t$, i.e. $\beta_{NSV} >> \beta_{PSV} > 0$, suggesting different information content in the two components. Similarly they find $\beta_{NJV} < 0$ and highly significant but $\beta_{PJV}$ has mixed signs. Therefore, NJV increases volatility while PJV has mixed effects. These return-based RV decompositions may affect the working of CVP. For example, it is unclear whether the impact of negative returns on CVP would remain if $NSV_t$ and $NJV_t$ are included. We test the impact of RV decompositions on CVP in the following two regressions:

[9] Our $\hat{\beta}_{RQ}$ is similar to those of Berkierman and Manner (2017) but is much smaller than the one reported in Table 3 of BPQ (2016). Using demeaned $RQ_t^{1/2}$ (as they do) only affects $\hat{\beta}_0$, not $\hat{\beta}_{RQ}$.

**Table 4: The HAR_CVP-RQ Model**

This table reports the daily persistence coefficients of the following models:

$$RV_{t+1} = \alpha + (\beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{RQ} RQ_t^{1/2})RV_t + \varphi Z_t + \varepsilon_{t+1}$$

RV and RQ are the Realized Variance and the Realized Quarticity constructed from the TRTH data. $RV^{OM}$ and $RK^{OM}$ are the Realized Variance and the Realized Kernel taken from the OMI Realized Library. The other variables are the same as in Table 2. $Z_t = (RV_{t,W}, RV_{t,M}, r_t, r_{t,W}, r_{t,M})'$. F is the F statistic for $\beta_{|r|} = \beta_r = 0$. AIC is the Aikaike information criteria. LB5 is the Ljung-Box statistic for residuals at 5 lags. The t-statistics are based on the Newey–West robust covariance with automatic lag selection using Bartlett kernel. $\%(t \leq -1.96)$ and $\%(t \geq 1.96)$ are the percentage of stocks with $t \leq -1.96$ or $t \geq 1.96$ respectively. The asterisks ***, **, * indicate significance at 1%, 5%, and 10% respectively.

| | $\beta_0$ | $\beta_{\|r\|}$ | $\beta_r$ | $\beta_{RQ}$ | $\bar{R}^2$ | F | AIC | LB5 |
|---|---|---|---|---|---|---|---|---|
| **SPY** | | | | | | | | |
| RV | 0.614*** | | | -0.0075*** | 0.651 | - | 2362 | 168 |
| *t-stat* | *9.04* | | | *-8.26* | | | | |
| | 0.254*** | 0.103*** | -0.057** | -0.0037*** | 0.732 | 536 | 1428 | 57 |
| | *3.26* | *3.32* | *-2.17* | *-4.38* | | | | |
| $RV^{OM}$ | 0.647*** | | | -0.0070*** | 0.593 | - | 3307 | 115 |
| *t-stat* | *6.60* | | | *-4.92* | | | | |
| | 0.207* | 0.101*** | -0.054* | -0.0024** | 0.665 | 382 | 2617 | 31 |
| | *1.79* | *2.73* | *-1.80* | *-1.96* | | | | |
| $RK^{OM}$ | 0.579*** | | | -0.0061*** | 0.586 | - | 4079 | 96 |
| *t-stat* | *6.98* | | | *-6.17* | | | | |
| | 0.171 | 0.102*** | -0.050* | -0.0016 | 0.653 | 339 | 3460 | 44 |
| | *1.48* | *2.83* | *-1.73* | *-1.33* | | | | |
| **Stocks** | | | | | | | | |
| Ave Coeff | 0.530*** | | | -0.0038*** | 0.599 | - | 6224 | 66 |
| Median | 0.519 | | | -0.0033 | 0.642 | | 5328 | 48 |
| *t stat* | *8.16* | | | *-3.76* | | | | |
| *%(t≤-1.96)* | *0%* | | | *71%* | | | | |
| *%(t≥1.96)* | *99%* | | | *0%* | | | | |
| Ave Coeff | 0.400*** | 0.0672*** | -0.0269*** | -0.0050*** | 0.640 | 223 | 5823 | 37 |
| Median | 0.390 | 0.0707 | -0.0278 | -0.0048 | 0.689 | 209 | 4822 | 24 |
| *t stat* | *6.44* | *4.52* | *-2.97* | *-6.38* | | | | |
| *%(t≤-1.96)* | *0%* | *0%* | *55%* | *92%* | | | | |
| *%(t≥1.96)* | *91%* | *78%* | *0%* | *0%* | | | | |

(9) $RV_{t+1} = \alpha + \beta_{NSV}NSV_t + \beta_{PSV}PSV_t + \left(\beta_{|r|}|r_t| + \beta_r r_t + \beta_{RV}RV_t\right)RV_t + \varphi Z_t + \varepsilon_{t+1}$

(10) $RV_{t+1} = \alpha + \beta_{NJV}NJV_t + \beta_{PJV}PJV_t + (\beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{BV}BV_t)BV_t + \varphi Z_t + \varepsilon_{t+1}$

The model in (9) splits $\beta_0 RV_t$ in (4) into $\beta_{NSV}NSV_t$ and $\beta_{PSV}PSV_t$. The model in (10) includes the negative and positive jump variations; daily RV persistence is measured by its continuous component. The estimation results are reported in Table 5.

Panel A of Table 5 reports the estimation results for (9). When $\beta_{|r|} = \beta_r = \beta_{RV} = 0$, $\beta_{NSV} > 0$ is significant but $\beta_{PSV}$ is not significant: $NSV_t$ dominates volatility spillover as in Patton and Sheppard (2015). When RV persistence is allowed to vary with $r_t$ and $RV_t$, $\beta_{NSV}$ is insignificant for SPY. The significance of $\beta_{PSV}$ depends on the RV measure used. For stocks, $\beta_{PSV}$ becomes significant and $\beta_{NSV}$ no longer dominates the impact on $RV_{t+1}$. For both SPY and stocks, the values and significance of the CVP coefficients remain largely intact as in Table 2. F test strongly rejects $\beta_{|r|} = \beta_r = \beta_{RV} = 0$. Model diagnostics clearly favour the one with CVP. The results for (10) in Panel B show that the strong impact of $NJV_t$ on $RV_{t+1}$ does not survive the inclusion of CVP. On the other hand, the CVP variables are significant and have the same signs as in Table 2; CVP works for both $RV_t$ and $BV_t$. Overall the evidence indicates that CVP is robust to the inclusion of NSV and PSV or their difference.

***An Alternative Persistence Measure***

As a further robustness check, we test our findings in Table 2 using a persistence measure that does not depend on a specific model of volatility dynamics. We may measure RV persistence as $\rho = \frac{E[(RV_{t+1}-\mu)(RV_t-\mu)]}{E(RV_t-\mu)^2}$ where $\mu = E(RV_t)$. It is the coefficient on $RV_t$ in the regression $RV_{t+1} = \alpha + \rho RV_t + \varepsilon_{t+1}$. Let $\rho_{t,t+1} = \frac{(RV_{t+1}-\mu)(RV_t-\mu)}{E(RV_t-\mu)^2}$ therefore $\rho = E\left(\rho_{t,t+1}\right)$. We use $E\left(\rho_{t,t+1}\middle|r_{t-1}, RV_{t-1}\right)$ as an alternative conditional measure of volatility persistence. It can be estimated from the regression $\rho_{t,t+1} = E\left(\rho_{t,t+1}\middle|r_{t-1}, RV_{t-1}\right) + e_{t+1}$ with a suitable specification for $E\left(\rho_{t,t+1}\middle|r_{t-1}, RV_{t-1}\right)$ when $E(e_t|r_{t-1}, RV_{t-1}) = 0$. Since $\rho_{t,t+1}$ is unobservable, we use its sample counterpart $\tilde{\rho}_{t,t+1} \equiv \frac{(RV_{t+1}-\overline{RV})(RV_t-\overline{RV})}{s^2}$ for the above regression with $\overline{RV} = \frac{1}{T}\sum_{t=1}^{T} RV_t$ and $s^2 = \frac{1}{T-1}\sum_{t=1}^{T}(RV_t - \overline{RV})^2$.

Figure 3 explores the characteristics of $\tilde{\rho}_{t,t+1}$ for SPY. Panel A shows the asymmetric relationships between $\tilde{\rho}$ and returns. Large $r_{t-1}$, especially large negative $r_{t-1}$, are associated with high $\tilde{\rho}_{t,t+1}$. This is consistent with the sign and size of return coefficients of the CVP in Table 2. Since $\tilde{\rho}_{t,t+1}$ is not exactly a correlation, it can be outside [-1,1] with some high values.

**Table 5: Semi-Variances, Jump variations, and Conditional Persistence**

Panel A of this table reports selected coefficients of the following regression:

$$RV_{t+1} = \alpha + \beta_{NSV} NSV_t + \beta_{PSV} PSV_t + (\beta_{|r|}|r_t| + \beta_r r_t + \beta_{RV} RV_t) RV_t + \varphi Z_t + \varepsilon_{t+1}$$

$NSV_t$ and $PSV_t$ are the negative and positive semi-variances respectively. The other variables are the same as in Table 2. Panel B reports selected coefficients of the following regression:

$$RV_{t+1} = \alpha + \beta_{NJV} NJV_t + \beta_{PJV} PJV_t + (\beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{BV} BV_t) BV_t + \varphi Z_t + \varepsilon_{t+1}$$

$NJV_t$ and $PJV_t$ are the negative and positive jump variations respectively. $Z_t = (RV_{t,W}, RV_{t,M}, r_t, r_{t,W}, r_{t,M})'$. For stocks, %(t ≤ -1.96) and %(t ≥ 1.96) are the percentage of stocks with robust t ≤ -1.96 or t ≥ 1.96 respectively. The asterisks ***, **, * indicate significance at 1%, 5%, and 10% respectively.

Panel A: Impact of negative and positive semivariances

| | $\beta_{NSV}$ | $\beta_{PSV}$ | $\beta_{\|r\|}$ | $\beta_r$ | $\beta_{RV}$ | $\bar{R}^2$ | F | AIC | LB5 |
|---|---|---|---|---|---|---|---|---|---|
| **SPY** | | | | | | | | | |
| RV | 1.327** | -0.491 | | | | 0.645 | - | 2426 | 259 |
| *t-stat* | *2.19* | *-1.03* | | | | | | | |
| | 0.365 | 0.096 | 0.105*** | -0.053** | -0.0037* | 0.729 | 370 | 1464 | 61 |
| | *0.94* | *0.25* | *3.31* | *-2.02* | *-1.71* | | | | |
| $RV^{OM}$ | 0.934** | -0.460 | | | | 0.563 | - | 3556 | 119 |
| *t-stat* | *2.28* | *-0.80* | | | | | | | |
| | -0.633 | 1.040** | 0.099*** | -0.085* | -0.0028*** | 0.672 | 390 | 2548 | 35 |
| | *-0.98* | *2.08* | *2.72* | *-1.88* | *-2.63* | | | | |
| **Stocks** | | | | | | | | | |
| Ave Coeff | 0.438*** | 0.146 | | | | 0.595 | - | 6264 | 68 |
| Median | 0.487 | 0.090 | | | | 0.634 | - | 5321 | 53 |
| *t stat* | *2.65* | *0.84* | | | | | | | |
| *%(t≤-1.96)* | *1%* | *0%* | | | | | | | |
| *%(t≥1.96)* | *53%* | *16%* | | | | | | | |
| Ave Coeff | 0.413*** | 0.359*** | 0.0668*** | -0.0229** | -0.0062*** | 0.643 | 171 | 5807 | 34 |
| Median | 0.416 | 0.322 | 0.0684 | -0.0226 | -0.0057 | 0.690 | 148 | 4807 | 24 |
| *t stat* | *2.80* | *2.59* | *4.57* | *-2.55* | *-6.52* | | | | |
| *%(t≤-1.96)* | *1%* | *0%* | *0%* | *47%* | *92%* | | | | |
| *%(t≥1.96)* | *56%* | *38%* | *79%* | *1%* | *0%* | | | | |

Panel B: Impact of negative and positive jump variations

| | $\beta_{NJV}$ | $\beta_{PJV}$ | $\beta_0$ | $\beta_{\|r\|}$ | $\beta_r$ | $\beta_{RV}$ | $\bar{R}^2$ | F | AIC |
|---|---|---|---|---|---|---|---|---|---|
| **SPY** | | | | | | | | | |
| RV | -1.721* | -0.331 | 0.305*** | | | | 0.654 | - | 2336 |
| *t-stat* | *-1.85* | *-1.01* | *4.56* | | | | | | |
| | -0.646 | 0.341* | 0.234*** | 0.105*** | -0.059** | -0.0067*** | 0.740 | 392 | 1323 |
| | *-0.95* | *1.90* | *3.14* | *3.02* | *-2.36* | *-3.86* | | | |
| $RV^{OM}$ | -1.867** | 0.316 | 0.162 | | | | 0.592 | - | 3312 |
| *t-stat* | *-2.11* | *0.82* | *1.59* | | | | | | |
| | 0.956 | 0.525 | 0.246** | 0.110** | -0.083* | -0.0035*** | 0.677 | 309 | 2492 |
| | *1.24* | *1.09* | *2.01* | *2.16* | *-1.77* | *-2.66* | | | |
| **Stocks** | | | | | | | | | |
| Ave Coeff | -0.715** | 0.308 | 0.254*** | | | | 0.604 | - | 6184 |
| Median | -0.664 | 0.262 | 0.270 | | | | 0.638 | - | 5195 |
| *HKV t stat* | *-2.22* | *1.56* | *3.27* | | | | | | |
| *%(t≤-1.96)* | *28%* | *1%* | *0%* | | | | | | |
| *%(t≥1.96)* | *1%* | *14%* | *60%* | | | | | | |
| Ave Coeff | -0.271 | 0.196 | 0.390*** | 0.076*** | -0.031*** | -0.0076*** | 0.649 | 162 | 5749 |
| Median | -0.266 | 0.128 | 0.382 | 0.078 | -0.030 | -0.0065 | 0.690 | 137 | 4766 |
| *HKV t stat* | *-1.01* | *1.18* | *5.14* | *4.98* | *-3.12* | *-6.36* | | | |
| *%(t≤-1.96)* | *11%* | *1%* | *0%* | *0%* | *55%* | *87%* | | | |
| *%(t≥1.96)* | *0%* | *14%* | *84%* | *77%* | *0%* | *0%* | | | |

Panels B and C depict the relationship between $\tilde{\rho}_{t,t+1}$ and $RV_{t-1}$. Panel B shows a nonlinear relationship between $\tilde{\rho}_{t,t+1}$ and $RV_{t-1}$. This is not surprising since $\tilde{\rho}_{t,t+1}$ involves the product of $RV_t$ and $RV_{t+1}$, which are highly correlated with $RV_{t-1}$. From Table 1 we see that the mean and median RV are 1.13 and 0.551 respectively for SPY, indicating daily RVs are heavily concentrated in the low end of the RV range. Panel C zooms into the range $RV_{t-1} < 3$. This is the normal range of daily RV, accounting for 93% of trading days for SPY. Higher $RV_{t-1}$ is associated with lower $\tilde{\rho}_{t,t+1}$, consistent with the negative RV coefficients in CVP in Table 2. On most trading days, higher volatility is associated with lower future volatility persistence.

Figure 3 motivates the following regression specification for $E(\rho_{t,t+1}|r_{t-1}, RV_{t-1})$:[10]

(11) $$E(\rho_{t,t+1}|r_{t-1}, RV_{t-1}) = \alpha + \beta_{|r|}|r_{t-1}| + \beta_r r_{t-1} + \beta_{RV}RV_{t-1} + \beta_{RV2}RV_{t-1}^2$$

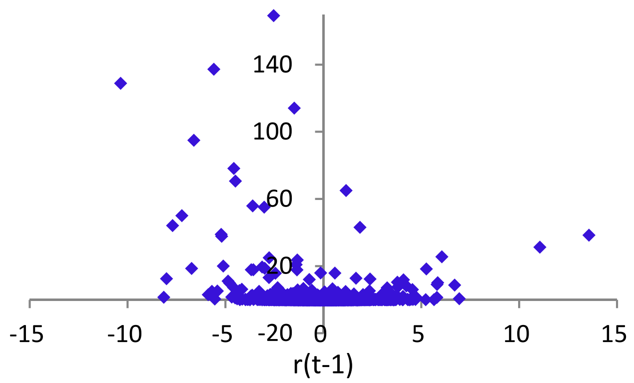
The quadratic term $RV_{t-1}^2$ captures the relationship in Panel B of Figure 3, which shows extremely high $\tilde{\rho}_{t,t+1}$ associated with extremely high $RV_{t-1}$. We winsorize the top 1% RV to

[10] We do not include the lagged dependent variable $\tilde{\rho}_{t-1,t}$ because it is a function of $RV_t$ which is not observed on day t-1. It is also a function of $RV_{t-1}$; so its inclusion is likely to distort the estimated coefficients of $RV_{t-1}$.

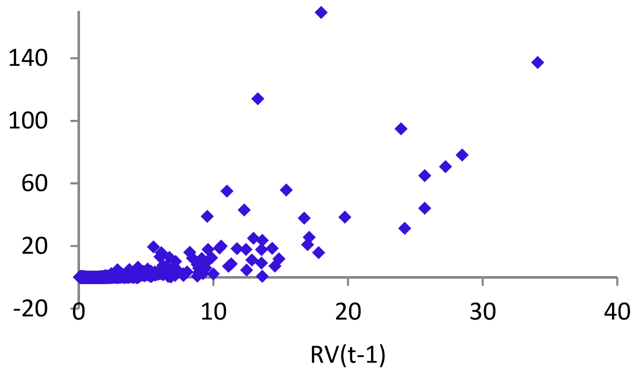
**Figure 3: SPY Return, RV, and the Proxy for RV Persistence $\tilde{\rho}_{t,t+1}$**

This figure plots $\tilde{\rho}_{t,t+1} \equiv \frac{(RV_{t+1}-\overline{RV})(RV_t-\overline{RV})}{s^2}$ against return and RV. $\overline{RV} = \frac{1}{T}\sum_{t=1}^{T} RV_t$ and $s^2 = \frac{1}{T-1}\sum_{t=1}^{T}(RV_t - \overline{RV})^2$.

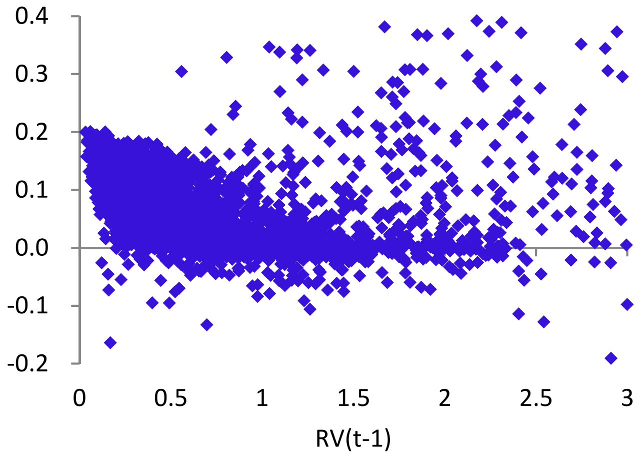
(A) Return and the proxy for RV persistence $\tilde{\rho}_{t,t+1}$ defined in eq. (8)




(B) RV and the proxy for RV persistence $\tilde{\rho}_{t,t+1}$




(C) Small RV and the proxy for RV persistence $\tilde{\rho}_{t,t+1}$



reduce their impact. Table 6 reports the estimated coefficients. It holds the qualitative features of the estimated CVP in Table 2: $\beta_{|r|} > 0$, $\beta_r < 0$, and $\beta_{RV} < 0$, all significant at 5% except $\beta_{RV}$ for stocks. Both positive and negative returns increase $\tilde{\rho}_{t,t+1}$: $\beta_{|r|}+\beta_r > 0$ and $\beta_r - \beta_{|r|} < 0$. Negative returns increase $\tilde{\rho}_{t,t+1}$ more than positive returns: $|\beta_r - \beta_{|r|}| > \beta_r + \beta_{|r|} > 0$. In the case of $\beta_{RV}$ for stocks, 23% of stocks have negative and significant $\beta_{RV}$ with only 3% having positive and significant $\beta_{RV}$.

**Table 6: Alternative Volatility Persistence Measure**

This table reports the results of the following regression for SPY:

$$\tilde{\rho}_{t,t+1} = \beta_0 + \beta_{|r|}|r_{t-1}| + \beta_r r_{t-1} + \beta_{RV} RV_{t-1} + \beta_{RV2} RV_{t-1}^2 + e_{t+1}$$

where $\tilde{\rho}_{t,t+1} \equiv \frac{(RV_{t+1}-\overline{RV})(RV_t-\overline{RV})}{s^2}$ with $\overline{RV} = \frac{1}{T}\sum_{t=1}^{T} RV_t$ and $s^2 = \frac{1}{T-1}\sum_{t=1}^{T}(RV_t - \overline{RV})^2$. The other variables are the same as Table 2. %(t ≤ -1.96) and %(t ≥ 1.96) are the percentage of stocks with t ≤ -1.96 or t ≥ 1.96 respectively. The asterisks ***, **, and * indicate statistical significance at 1%, 5%, and 10% respectively.

| | $\beta_0$ | $\beta_{\|r\|}$ | $\beta_r$ | $\beta_{RV}$ | $\beta_{RV2}$ | $\overline{R}^2$ | AIC | LB5 |
|---|---|---|---|---|---|---|---|---|
| **SPY** | | | | | | | | |
| RV | -0.068 | 0.640*** | -0.189** | -0.487** | 0.256*** | 0.595 | 6166 | 355 |
| *t stat* | *-0.54* | *3.50* | *-2.38* | *-2.45* | *6.44* | | | |
| $RV^{OM}$ | -0.062 | 0.562*** | -0.219*** | -0.444** | 0.251*** | 0.623 | 6616 | 322 |
| | *-0.53* | *3.26* | *-2.94* | *-2.15* | *6.34* | | | |
| $RK^{OM}$ | 0.006 | 0.533*** | -0.276*** | -0.454** | 0.197*** | 0.607 | 6671 | 280 |
| | *0.05* | *3.14* | *-3.52* | *-2.39* | *6.37* | | | |
| **Stocks** | | | | | | | | |
| Ave Coeff | -0.154 | 0.513*** | -0.180*** | -0.127 | 0.094*** | 0.577 | 5288 | 423 |
| Median | -0.156 | 0.534 | -0.175 | -0.101 | 0.056 | 0.611 | 5306 | 354 |
| *t stat* | *-1.35* | *4.19* | *-3.23* | *-1.29* | *7.71* | | | |
| *%(t≤-1.96)* | *8%* | *0%* | *57%* | *23%* | *0%* | | | |
| *%(t≥1.96)* | *0%* | *90%* | *0%* | *3%* | *97%* | | | |

***Other Robustness Checks***

Andersen (1996) reports that volatility persistence is much lower when volume data are used in estimating a stochastic volatility model. We add additional conditioning variables and report the results in Appendix B. The additional variables are volatility jumps, number of trades, Amihud illiquidity, and trade imbalance. Overall the CVP results in Table 2 remain qualitatively the same when additional conditioning variables are included. Some of the new variables are significant for SPY but none is significant for stocks. Appendix C reports the results of the HAR_CVP model in 2-year sub-periods. At least two of the CVP coefficients are significant in each sub-period. Therefore, CVP is present in all sub-periods.

Additional robustness results are reported in the internet Applendix. We estimate CVP as a threshold function and a power function of returns. The results are qualitatively the same. Allowing conditional persistence for the lagged weekly and monthly RVs does not change the results for daily CVP. Results in Table 2 hold qualitatively after daily RV is adjusted for time-trend and seasonality, same as in Chordia, Sarkar, and Subrahmanyam (2005). They hold for the square root of daily RV and when RV is replaced by the high-low based volatility measure of Parkinson (1980). They hold when CVP includes a holiday or weekend dummy that may affect RV autocorrelation.

## V. Conditional Volatility Persistence and Asymmetric Volatility

The asymmetric return impact on future volatility is "one of the most enduring empirical regularities in equity markets" (Hasanhodzic and Lo, 2013) and is becoming stronger in many markets around the world (Talpsepp and Rieger, 2010). There is a sizable literature and an ongoing debate on financial leverage versus volatility feedback as the main reason for asymmetric volatility, e.g. Choi and Richardson (2016) and Engle and Siriwardane (2016). In this section, we present evidence that CVP sharply reduces volatility asymmetry attributed to financial leverage and volatility feedback, thus offers a new and robust explanation for asymmetric volatility.

To further understand the asymmetric return impact on RV level and persistence, we divide daily returns into 10 ranges in percentage, {-∞, -3, -2, -1, -0.5, 0, 0.5, 1, 2, 3, ∞}, indexed by k = -5, …, 5 with k ≠ 0. The size of the range is smaller around zero since daily returns are heavily concentrated around zero. We estimate a modified HAR_CVP model

$$\mathrm{RV}_{t+1} = \alpha + \left(\beta_0 + \textstyle\sum_{k=-5}^{5} \beta_k r_t D_{t,k} + \beta_{RV}\mathrm{RV}_t\right)\mathrm{RV}_t + \textstyle\sum_{k=-5}^{5} \theta_k r_t D_{t,k} + \varphi Z_t + \varepsilon_{t+1} \quad (12)$$

where $Z_t = (\mathrm{RV}_{t,W}, \mathrm{RV}_{t,M}, r_{t,W}, r_{t,M})'$ and $D_{t,k} = 1$ if $r_t$ is in the $k^{th}$ range, 0 otherwise. AsyVP is captured by $\sum_{k=-5}^{5} \beta_k r_t D_{t,k}$ which allows step-wise changes in return impact on CVP. The direct return impact on volatility is captured by $\sum_{k=-5}^{5} \theta_k r_t D_{t,k}$ and is typically attributed to financial leverage (Black, 1976), volatility feedback (Pindyck, 1984), and herding and contrarian trading (Avramov, Chordia, and Goyal, 2006).

The estimated $\beta_k$ and $\theta_k$ in (12) are reported in Table 7. Consistent with AsyVP in Figure 2 and Table 2, Panel A shows that RV persistence increases with return size, especially negative returns. We note that the size of $\beta_k$ is almost strictly inversely related to the size of return, while the size of t statistics generally increases with the size of return. Let $r_k^m$ be the

**Table 7: Return Impact on RV and RV Persistence**

This table reports the estimation results of the following regression:

$$RV_{t+1} = \alpha + \left(\beta_0 + \sum_{k=-5}^{5} \beta_k r_t D_{t,k} + \beta_{RV} RV_t\right) RV_t + \sum_{k=-5}^{5} \theta_k r_t D_{t,k} + \varphi Z_t + \varepsilon_{t+1}$$

where $Z_t = (RV_{t,W}, RV_{t,M}, r_{t,W}, r_{t,M})'$ and $D_{t,k} = 1$ if $r_t$ is in the $k^{th}$ range, and 0 otherwise. There are 10 return ranges in percentage (-∞, -3, -2, -1, -0.5, 0, 0.5, 1, 2, 3, ∞) which are indexed by k = -5,…,5 (k ≠ 0). The asterisks ***, **, and * indicate statistical significance at 1%, 5%, and 10% respectively.

Panel A: Return impact on RV persistence

| | $\beta_{-5}$ | $\beta_{-4}$ | $\beta_{-3}$ | $\beta_{-2}$ | $\beta_{-1}$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ |
|---|---|---|---|---|---|---|---|---|---|---|
| **SPY** | | | | | | | | | | |
| RV | -0.178** | -0.215*** | -0.201* | -0.488** | -1.152*** | 0.791* | 0.580*** | 0.387* | 0.179** | 0.078*** |
| *t stat* | *-2.07* | *-3.44* | *-1.83* | *-2.17* | *-2.78* | *1.64* | *2.36* | *1.88* | *2.18* | *3.12* |
| $RV^{OM}$ | -0.185* | -0.161** | -0.187 | -0.383 | -1.128*** | 0.749 | 0.542** | 0.335 | 0.117 | 0.067*** |
| *t stat* | *-1.77* | *-2.19* | *-1.61* | *-1.50* | *-2.54* | *1.08* | *1.96* | *1.53* | *1.29* | *2.39* |
| **Stocks** | | | | | | | | | | |
| Ave | -0.096*** | -0.114** | -0.169* | -0.180 | -0.369 | 0.449 | 0.151 | 0.145 | 0.048 | 0.054*** |
| Med | -0.096 | -0.124 | -0.199 | -0.173 | -0.447 | 0.502 | 0.208 | 0.148 | 0.053 | 0.059 |
| *t stat* | *-3.11* | *-2.24* | *-1.88* | *-1.00* | *-0.99* | *1.15* | *1.00* | *1.51* | *0.95* | *2.71* |

Panel B: Return impact on RV

| | $\theta_{-5}$ | $\theta_{-4}$ | $\theta_{-3}$ | $\theta_{-2}$ | $\theta_{-1}$ | $\theta_1$ | $\theta_2$ | $\theta_3$ | $\theta_4$ | $\theta_5$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | No AsyVP ($\beta_k = 0$) | | | | | |
| **SPY** | | | | | | | | | | |
| RV | -1.449*** | -0.444*** | -0.388*** | -0.341** | -0.665* | 0.439 | 0.149 | 0.016 | 0.010 | 0.109 |
| *t stat* | *-3.07* | *-5.16* | *-4.06* | *-2.47* | *-1.86* | *1.57* | *1.19* | *0.22* | *0.16* | *0.92* |
| $RV^{OM}$ | -1.406*** | -0.422*** | -0.322*** | -0.293** | -0.746** | 0.351 | 0.121 | -0.028 | -0.044 | 0.082 |
| *t stat* | *-2.77* | *-4.92* | *-4.06* | *-2.37* | *-2.04* | *1.44* | *1.04* | *-0.49* | *-0.84* | *0.79* |
| **Stocks** | | | | | | | | | | |
| Ave | -1.731*** | -0.597*** | -0.350*** | -0.472*** | -0.679** | 0.695** | 0.091 | -0.031 | -0.063 | 0.045 |
| Med | -1.361 | -0.510 | -0.291 | -0.407 | -0.587 | 0.566 | 0.105 | -0.022 | -0.050 | 0.103 |
| *t stat* | *-3.85* | *-4.90* | *-3.35* | *-2.88* | *-2.05* | *2.18* | *0.59* | *-0.28* | *-0.62* | *0.23* |
| | | | | | AsyVP ($\beta_k \neq 0$) | | | | | |
| **SPY** | | | | | | | | | | |
| RV | -0.346 | -0.086 | -0.182 | 0.152 | 0.616** | -0.448 | -0.405*** | -0.400** | -0.164* | 0.019 |
| *t stat* | *-1.10* | *-1.38* | *-1.47* | *1.05* | *2.22* | *-1.53* | *-2.57* | *-2.14* | *-1.97* | *0.23* |
| $RV^{OM}$ | -0.079 | -0.164** | -0.105 | 0.123 | 0.470 | -0.456 | -0.387** | -0.381* | -0.144 | -0.053 |
| *t stat* | *-0.18* | *-2.27* | *-1.09* | *0.71* | *1.56* | *-1.03* | *-2.06* | *-1.87* | *-1.37* | *-0.63* |
| **Stocks** | | | | | | | | | | |
| Ave | -0.573 | -0.209 | 0.077 | -0.063 | 0.065 | -0.215 | -0.168 | -0.345** | -0.160 | -0.283 |
| Med | -0.288 | -0.096 | 0.135 | 0.019 | 0.522 | -0.544 | -0.340 | -0.352 | -0.121 | -0.183 |
| *t stat* | *-1.62* | *-1.53* | *0.40* | *-0.21* | *0.12* | *-0.39* | *-0.71* | *-2.07* | *-1.29* | *-1.33* |

mid-point of the return range k with $r^m_{-5}$ = -3.5 and $r^m_5$ = 3.5. The impact of return in range k on CVP is approximated by $\mathrm{CVP}(r^m_k) = \hat{\beta}_0 + \hat{\beta}_k r^m_k$. We estimate $\mathrm{CVP}(r^m_k)$ using RV and $RV^{OM}$ of SPY and the average and median coefficients of stocks. Figure 4 (A) and (B) plot the resulting $\mathrm{CVP}(r^m_k)$. Although large returns have small $\beta_k$, large negative returns are associated with high CVP. The linear representations in Figure 2 are reasonable proxies for $\mathrm{CVP}(r^m_k)$ when return is negative. When return is positive, $\mathrm{CVP}(r^m_k)$ is non-monotonic. It peaks when daily return is around 1.5% for both SPY and stocks. Consistent with AsyVP in Figure 2, positive returns on average have lower CVP than negative returns.

To explore how the inclusion of CVP affect the direct return impact on RV, we estimate (12) with and without restricting $\beta_k$ = 0 for all k. This allows us to measure volatility asymmetry implied by $\theta_k$ with and without AsyVP. The results reported in Panel B of Table 7 show that, without AsyVP (i.e. $\beta_k$ = 0), $\hat{\theta}_k$ for negative returns are all negative and highly significant: negative returns significantly increase future RV, echoing typical findings in the literature. On the other hand, only one $\hat{\theta}_k$ for positive returns is positive and significant: positive returns have no significant impact on future RV. With AsyVP (i.e. $\beta_k \neq 0$), only one $\hat{\theta}_k$ for negative returns remains negative and significant. One $\hat{\theta}_k$ for small negative returns actually becomes positive and significant. Therefore, *AsyVP almost entirely accounts for the impact of negative $r_t$ on $RV_{t+1}$.* This posts a strong challenge to existing explanations for asymmetric volatility. Interestingly, most $\hat{\theta}_k$ for positive returns are now negative and some are significant at 10% or above. When AsyVP is present, positive returns in the range of 0.5 to 2% often significantly reduces future RV, especially for SPY. Overall Table 7 shows richer RV dynamics from both CVP and the direct return impact on RV than revealed in Table 2.

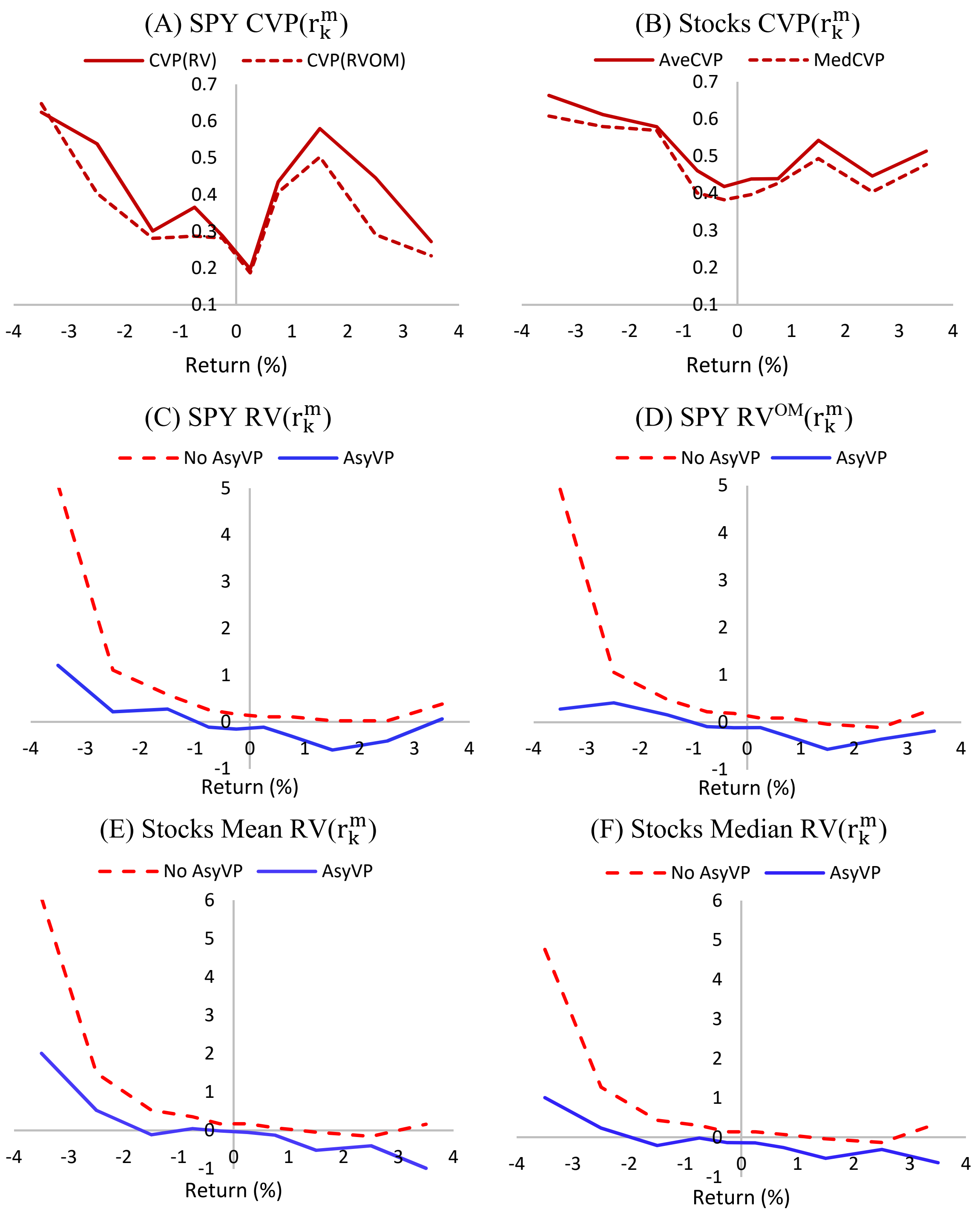
The direct impact of $r_t$ in range k on $RV_{t+1}$ is approximated by $\mathrm{RV}(r^m_k) = \hat{\theta}_k r^m_k$ and varies with $r^m_k$ for k = -5, …, 5 with k ≠ 0. $\mathrm{RV}(r^m_k)$ can be visualized in Figure 4 (C) to (F) which are similar in spirit to the news impact curve of Engle and Ng (1993) and Chen and Ghysels (2010). Again we estimate $\mathrm{RV}(r^m_k)$ with and without restricting $\beta_k$ = 0 for all k. $\mathrm{RV}(r^m_k|\beta_k = 0)$ measures the direct return impact on RV when return does not affect RV persistence. The corresponding lines are denoted as "No AsyVP". $\mathrm{RV}(r^m_k|\beta_k \neq 0)$ measures the direct return impact on RV when there is AsyVP. The corresponding lines are denoted as "AsyVP". Figures (C) and (D) draw $\mathrm{RV}(r^m_k|\beta_k = 0)$ and $\mathrm{RV}(r^m_k|\beta_k \neq 0)$ for RV and $RV^{OM}$ of SPY. Figures (E) and (F) draw $\mathrm{RV}(r^m_k|\beta_k = 0)$ and $\mathrm{RV}(r^m_k|\beta_k \neq 0)$ based on the average and median coefficients across stocks. Two features are present in all four figures: when AsyVP

## Figure 4: Conditional Volatility Persistence and Asymmetric Volatility

The graphs below are based on the following regression:

$$RV_{t+1} = \alpha + \left(\beta_0 + \sum_{k=-5}^{5} \beta_k r_t D_{t,k} + \beta_{RV} RV_t\right) RV_t + \sum_{k=-5}^{5} \theta_k r_t D_{t,k} + \varphi Z_t + \varepsilon_{t+1}$$

where $Z_t = (RV_{t,W}, RV_{t,M}, r_{t,W}, \ r_{t,M})'$ and $D_{t,k} = 1$ if $r_t$ is in the $k^{th}$ range, and 0 otherwise. There are 10 return ranges in percentage by $(-\infty, -3, -2, -1, -0.5, 0, 0.5, 1, 2, 3, \infty)$. Let $r_k^m$ be the mid-point of the return range k with $r_1^m = -3.5$ and $r_{10}^m = 3.5$. Return impact on CVP is $CVP(r_k^m) = \hat{\beta}_0 + \hat{\beta}_k r_k^m$. The direct return impact on RV is $RV(r_k^m) = \hat{\theta}_k r_k^m$.



is allowed ($\beta_k \neq 0$), the RV impact of large negative returns are sharply reduced, and positive returns are associated with lower RV, particularly for returns around 0.5 to 2%.

To quantify the change in the degree of volatility asymmetry when AsyVP is included, we use the following measure which is similar to the one proposed by Daouk and Ng (2011). For a return size $|r_k^m|$, the asymmetric impact of negative and positive returns on RV is

$$RV(r_{-k}^m) - RV(r_k^m) = (\hat{\theta}_{-k} - \hat{\theta}_k)|r_k^m| \text{ for } k = 1,\ldots,5.$$

The asymmetry in RV is defined as AsyRV = $\sum_{k=1}^{5}[RV(r_{-k}^m) - RV(r_k^m)]$. Again we compute AsyRV with and without AsyVP. For SPY RV, AsyRV without AsyVP is 6.533 and AsyRV with AsyVP is 2.789, a reduction of 57%. For SPY $RV^{OM}$, AsyRV without AsyVP is 6.552 and AsyRV with AsyVP is 2.156, a reduction of 67%! For stocks, AsyRV without AsyVP is 8.404 based on average coefficients and 6.506 based on median coefficients; AsyRV with AsyVP is 4.533 based on average coefficients and 2.760 based on median coefficients. AsyVP reduces stocks' AsyRV by 46 to 58%.

Overall the results in Table 7 and Figure 4 show the presence of complex nonlinear AsyVP. As in the GJR model, CVP almost entirely accounts for the impact of negative daily returns on RV and sharply reduces volatility asymmetry associated with direct impact from daily returns. These findings support CVP as a dominant and robust explanation for asymmetric volatility at daily frequency.

## VI. Conditional Volatility Persistence and Volatility Forecast

A key contribution of the recent advances in modelling RV dynamics is the improved volatility forecasts. The HAR model of Corsi (2009) generates more accurate forecasts than a true long-memory model with fractional integration. The semivariance HAR model of Patton and Sheppard (2015) has better RV forecasts than the original HAR and a RV-based GJR model. BPQ (2016) show that the HAR model with realized quarticity outperforms a range of models including the semivariance HAR. This section compares volatility forecasts of the HAR_CVP model in (4) against these three competing models.

To isolate the effect of CVP on forecast performance, we control the direct impact from returns on $RV_{t+1}$ by adding $(r_t, r_{t,W}, r_{t,M})$' to all models. Let $Z_t = (RV_{t,W}, RV_{t,W}, r_t, r_{t,W}, r_{t,M})$'. The HAR in (2) can be written as

(13) $$RV_{t+1} = \alpha + \beta_0 RV_t + \varphi Z_t + \varepsilon_{t+1}$$

The HAR with realized quarticity (HARQ) of BPQ (2016) can be written as

(14) $$\mathrm{RV}_{t+1} = \alpha + (\beta_0 + \beta_{RQ}\mathrm{RQ}_t^{1/2})\mathrm{RV}_t + \varphi Z_t + \varepsilon_{t+1}$$

The semivariance HAR (HARSV) of Patton and Sheppard (2015) is

(15) $$\mathrm{RV}_{t+1} = \alpha + \beta_{NSV}\mathrm{NSV}_t + \beta_{PSV}\mathrm{PSV}_t + \varphi Z_t + \varepsilon_{t+1}$$

The HAR_CVP model in (4) is

(16) $$\mathrm{RV}_{t+1} = \alpha + (\beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{RV}\mathrm{RV}_t)\mathrm{RV}_t + \varphi Z_t + \varepsilon_{t+1}$$

We compare the forecast performance of the HAR_CVP model against HARi = HAR, HARQ, and HARSV. If CVP is a better proxy for the unobserved true RV persistence, HAR_CVP should have better forecasts than the competing models. Furthermore, its performance should be stronger for assets with greater variations in RV persistence.

### (i) Loss Functions

Forecast accuracy can be measured by a wide range of loss functions, e.g. mean squared error (MSE), mean absolute error (MAE), etc. The choice of loss functions may affect the ranking of forecasting models, e.g. Hamilton and Susmel (1994), Bollerslev and Ghysels (1994), Hansen and Lunde (2005), Patton and Sheppard (2009), and Patton (2011). Let $r_t$ be the daily return with $E(r_t|\mathcal{F}_{t-1}) = 0$ and $\sigma_t^2 \equiv \mathrm{Var}(r_t|\mathcal{F}_{t-1})$ where $\mathcal{F}_{t-1}$ is the information set up to day t-1. $\mathrm{RV}_t$ is a noisy ex-post measure for $\sigma_t^2$ and let $\widehat{\mathrm{RV}}_t$ be the forecast of $\sigma_t^2$. Competing models for $\widehat{\mathrm{RV}}_t$ are ranked based on the distance between $\mathrm{RV}_t$ and $\widehat{\mathrm{RV}}_t$. Patton (2011) shows that for some loss functions, the noise in $\mathrm{RV}_t$ may distort the ranking of competing models. He proposes a class of robust loss functions indexed by the parameter b:

(17) $$L(\mathrm{RV}_t, \widehat{\mathrm{RV}}_t; b) = \begin{cases} \frac{1}{(b+1)(b+2)}\left(\mathrm{RV}_t^{b+2} - \widehat{\mathrm{RV}}_t^{b+2}\right) - \frac{1}{b+1}\widehat{\mathrm{RV}}_t^{b+1}\left(\mathrm{RV}_t - \widehat{\mathrm{RV}}_t\right), & \text{for } b \notin \{-1,-2\} \\ \widehat{\mathrm{RV}}_t - \mathrm{RV}_t + \mathrm{RV}_t \ln\frac{\mathrm{RV}_t}{\widehat{\mathrm{RV}}_t}, & \text{for } b = -1 \\ \frac{\mathrm{RV}_t}{\widehat{\mathrm{RV}}_t} - \ln\left(\frac{\mathrm{RV}_t}{\widehat{\mathrm{RV}}_t}\right) - 1, & \text{for } b = -2 \end{cases}$$

The above loss function is robust to the noise in $\mathrm{RV}_t$: the ranking of competing models based on the noisy $\mathrm{RV}_t$ is the same as the ranking based on the true $\sigma_t^2$. Two popular loss functions are part of this family, subject to additive and multiplicative constants. The mean-squared error $\mathrm{MSE} \equiv (\mathrm{RV}_t - \widehat{\mathrm{RV}}_t)^2$ is a special case with b = 0 and the quasi-likelihood function QLIKE is the case with b = -2 given above. As noted in Proposition 2 of Patton (2011), MSE is the only robust loss function that depends solely on $\mathrm{RV}_t - \widehat{\mathrm{RV}}_t$ and QLIKE is the only robust loss function that depends solely on $\mathrm{RV}_t/\widehat{\mathrm{RV}}_t$. However, many studies avoid using MSE because it is often heavily influenced by a few large forecast errors.[11] In addition, Patton and Sheppard (2009) show that the Diebold-Mariano tests using QLIKE have higher power than those using MSE. Patton (2011) advocates the use of QLIKE instead of MSE.

We compare forecast accuracy based on four loss functions. The first two are MSE and QLIKE defined above. The third is the case of b = -1 in (17) which is a combination of the **f**orecast **e**rror $\mathrm{RV_t} - \widehat{\mathrm{RV}}_\mathrm{t}$ and **r**atio $\mathrm{RV_t}/\widehat{\mathrm{RV}}_\mathrm{t}$. We term this the FER loss function. In addition, we compare forecasts based on MSE-ln ≡ $[\ln(\mathrm{RV_t}) - \ln(\widehat{\mathrm{RV}}_\mathrm{t})]^2$, a popular alternative to MSE that mitigates the impact of a few large forecast errors. We report the results for SPY based on RV and $\mathrm{RV^{OM}}$ to conserve space.

### (ii) Forecast Procedure and Comparison

Following Patton and Sheppard (2015) and BPQ (2016), our forecasts are based on 4-year rolling windows, starting in 2004 for SPY and in 2006 for stocks. We have 72 stocks in the S&P 100 index with continuous data from 2002. An "insanity filter" is used to replace a negative RV forecast with the lowest RV in the rolling window.[12] For the HAR_CVP and HARQ models, if the estimated $\mathrm{CVP_t}$ or $\mathrm{CVP_t^{RQ}}$ is below 0 (above 1), it is replaced by the minimum (maximum) value within the rolling window.

Inference on the difference in loss values is based on the Diebold-Mariano (1995) test. While HAR is nested in HAR_CVP, Giacomini and White (2006) show that the DM test remains asymptotically valid when the estimation period is finite. Define the pair-wise loss difference against HAR_CVP as

(18) $$\mathrm{d_t(HARi)} = \mathrm{L(RV_t}, \widehat{\mathrm{RV}}_\mathrm{t}; \mathrm{b}; \mathrm{HAR\_CVP}) - \mathrm{L(RV_t}, \widehat{\mathrm{RV}}_\mathrm{t}; \mathrm{b}; \mathrm{HARi})$$

where HARi = HAR, HARQ, and HARSV. Let $\bar{\mathrm{d}}\mathrm{(HARi)}$ and Var[d(HARi)] be the time-series average and long-run variance of $\mathrm{d_t(HARi)}$ respectively. The number of autocovariances in Var[d(HARi)] is based on Andrews (1991). The DM test statistic for E[d(HARi)] = 0 is $\mathrm{DM}[\bar{\mathrm{d}}\mathrm{(HARi)}] = \frac{\bar{\mathrm{d}}\mathrm{(HARi)}}{\sqrt{\mathrm{Var[d(HARi)]/T}}}$ with T being the number of forecasts.

The above DM test is applied to SPY and individual stocks. For stocks as a whole, we test whether the expected cross-sectional average difference in loss values is zero. Let $\bar{\mathrm{d}}_\mathrm{j}\mathrm{(HARi)}$

[11] The impact of large forecast errors is related to the impact of data scale or data unit discussed in Proposition 3 of Patton (2011) and may affect model ranking based on MSE.
[12] Our insanity filter is the same as in Patton and Sheppard (2015). BPQ (2016) replace RV forecasts outside the high-low range of the rolling window by the average RV in the rolling window. The results based on this alternative filter are reported in the internet Appendix and are strongly in favour of HAR_CVP.

be the average loss difference for stock j and $\bar{\bar{d}}(\text{HARi}) = \frac{1}{N}\sum_{j=1}^{N} \bar{d}_j(\text{HARi})$ be average across N stocks. To takes into account both time-series and cross-sectional variations in loss difference, the standard deviation of $\bar{\bar{d}}(\text{HARi})$ is calculated as

$$(19) \qquad \text{StDev}[\bar{\bar{d}}(\text{HARi})] = \frac{1}{N}\sqrt{\sum_{j=1}^{N}\sum_{k=1}^{N} \widehat{\omega}_{j,k}\sqrt{\text{Var}[\bar{d}_j(\text{HARi})]\text{Var}[\bar{d}_k(\text{HARi})]}}$$

where $\widehat{\omega}_{j,k}$ the sample correlation between $d_{j,t}(\text{HARi})$ and $d_{k,t}(\text{HARi})$ defined in (18), and $\text{Var}[\bar{d}_j(\text{HARi})] = \text{Var}[d_j(\text{HARi})]/T$. The above standard error is analogous to the one in eq. (5). The DM statistic is defined as $\text{DM}[\bar{\bar{d}}(\text{HARi})] = \frac{\bar{\bar{d}}(\text{HARi})}{\text{StDev}[\bar{\bar{d}}(\text{HARi})]}$.

**(iii) Forecast Performance**

Table 8 reports forecast comparisons for SPY. Panel A is a summary of loss values for four models and loss functions. For both RV and $RV^{OM}$, HAR_CVP has the lowest average and median loss values for FER, QLIKE, and MSE-ln. It has the lowest median MSE but not the lowest average MSE. MSE has extremely high mean-to-media ratios and extremely high standard deviations of loss values. Both are consistent with strong impact from a few large errors, which has been documented in many studies, e.g. Pagan and Schwart (1990), Diebold and Lopez (1996), Poon and Granger (2003), and Hansen and Lunde (2005). Model ranking based on average MSE is likely to reflect such impact.

The performance of HAR_CVP is measured by the change in loss values defined as $\Delta(\text{HARi}) \equiv \frac{f(\text{HAR\_CVP})}{f(\text{HARi})} - 1$ with f being the mean or median of the loss values and HARi = HAR, HARQ, and HARSV. It is reported in the right three columns of Panel A. For RV and $RV^{OM}$, the changes in the median FER, QLIKE, and MSE-ln are around -25%. The changes in the average QLIKE and MSE-ln are much larger at -32 to -57%, suggesting the presence of a few days with very large loss reductions. Overall the improvement from HAR_CVP is larger than those reported in similar studies. Panel B of Table 8 reports the DM test for equal average loss values between HAR_CVP and HARi = HAR, HARQ, and HARSV. The DM statistics are highly significant for FER, QLIKE, and MSE-ln. Not surprisingly, the high volatility of MSE values makes the differences in average MSE not statistically significant.

Table 9 reports forecast performance for individual stocks. Loss value comparisons in Panel A are qualitatively the same as for SPY. HAR_CVP has the lowest average and median loss values for three of the four loss functions. The exception is MSE where HAR_CVP has the lowest median loss but not the lowest average loss. Compared to SPY, stock MSE has even

**Table 8: Volatility Forecast Comparison for SPY**

This table reports summary for MSE, FER, QLIKE, and MSE-ln loss functions for SPY.

Panel A: Summary statistics for loss function values. The lowest values across four models are in bold numbers. "Δ(HARi)" reports $\frac{f(HAR\_CVP)}{f(HARi)} - 1$ where f is the mean or median and HARi = HAR, HARQ, and HARSV.

| | HAR_CVP | HAR | HARQ | HARSV | Δ(HAR) | Δ(HARQ) | Δ(HARSV) |
|---|---|---|---|---|---|---|---|
| **RV** | | | | | | | |
| MSE | | | | | | | |
| Mean | 2.973 | 3.167 | 3.920 | **2.836** | -6% | -24% | 5% |
| Median | **0.0297** | 0.0357 | 0.0342 | 0.0323 | -17% | -13% | -8% |
| StDev | 55.4 | 50.8 | 84.7 | 41.0 | | | |
| FER | | | | | | | |
| Mean | **0.211** | 0.296 | 0.275 | 0.276 | -29% | -24% | -24% |
| Median | **0.037** | 0.0487 | 0.0514 | 0.0447 | -24% | -28% | -17% |
| StDev | 1.03 | 1.13 | 1.29 | 1.03 | | | |
| QLIKE | | | | | | | |
| Mean | **0.519** | 1.191 | 1.199 | 1.084 | -56% | -57% | -52% |
| Median | **0.0872** | 0.114 | 0.114 | 0.107 | -24% | -24% | -19% |
| StDev | 3.53 | 8.15 | 9.45 | 8.31 | | | |
| MSE-ln | | | | | | | |
| Mean | **0.527** | 0.892 | 0.896 | 0.814 | -41% | -41% | -35% |
| Median | **0.182** | 0.241 | 0.243 | 0.223 | -24% | -25% | -18% |
| StDev | 1.32 | 2.06 | 1.98 | 1.95 | | | |
| **$RV^{OM}$** | | | | | | | |
| MSE | | | | | | | |
| Mean | 7.129 | 5.193 | 6.206 | **4.839** | 37% | 15% | 47% |
| Median | **0.0312** | 0.0372 | 0.0360 | 0.0359 | -16% | -13% | -13% |
| StDev | 197 | 113 | 157 | 98 | | | |
| FER | | | | | | | |
| Mean | **0.283** | 0.360 | 0.324 | 0.359 | -21% | -13% | -21% |
| Median | **0.0382** | 0.0495 | 0.0508 | 0.0478 | -23% | -25% | -20% |
| StDev | 2.23 | 1.95 | 2.10 | 2.00 | | | |
| QLIKE | | | | | | | |
| Mean | **0.531** | 1.241 | 1.097 | 0.928 | -57% | -52% | -43% |
| Median | **0.0854** | 0.111 | 0.116 | 0.104 | -23% | -26% | -18% |
| StDev | 2.85 | 7.81 | 6.87 | 3.83 | | | |
| MSE-ln | | | | | | | |
| Mean | **0.564** | 0.912 | 0.867 | 0.827 | -38% | -35% | -32% |
| Median | **0.181** | 0.233 | 0.242 | 0.223 | -22% | -25% | -19% |
| StDev | 1.34 | 2.12 | 1.94 | 1.80 | | | |

Panel B: DM test against HAR_CVP. A negative value indicates that HAR_CVP has lower loss values than the competing model. The asterisks ***, **, and * indicate statistical significance at 1%, 5%, and 10% respectively.

| | HAR | HARQ | HARSV | HAR | HARQ | HARSV |
|---|---|---|---|---|---|---|
| | | RV | | | $RV^{OM}$ | |
| MSE | -0.849 | -0.993 | 0.176 | 0.969 | 1.023 | 0.850 |
| FER | $-6.19^{***}$ | $-3.62^{***}$ | $-3.37^{***}$ | $-3.31^{***}$ | $-4.77^{***}$ | $-2.62^{***}$ |
| QLIKE | $-4.00^{***}$ | $-3.08^{***}$ | $-3.19^{***}$ | $-4.00^{***}$ | $-4.50^{***}$ | $-4.54^{***}$ |
| MSE-ln | $-6.97^{***}$ | $-6.46^{***}$ | $-5.93^{***}$ | $-6.36^{***}$ | $-6.97^{***}$ | $-6.74^{***}$ |

higher mean-to-media ratios and higher standard deviations of loss values, indicating longer right tails for loss values. The reductions in loss values from using HAR_CVP are somewhat smaller than for SPY. The changes in median loss values range from -10 to -16%. The changes in average loss values range from -9 to -53%, indicating the presence of a few stocks with very large reductions. HAR_CVP has the lowest volatility of loss values and achieves larger loss reductions for QLIKE and MSE-ln, the same as SPY. The right three columns of Panel A shows that based on FER, QLIKE, and MSE-ln, HAR_CVP has the lower average or median loss values for 83 to 99% of the stocks. It has lower average MSE for 56 to 63% of the stocks. Even though HARQ has lowest cross-sectional average MSE, HAR_CVP has lower average MSE than HARQ for 57% of stocks.

The DM tests for stocks are presented in Panel B of Table 9. Based on FER, QLIKE, and MSE-ln, DM[$\bar{\bar{d}}$(HARi)] statistics indicate that HAR_CVP has significantly lower cross-sectional average loss than HARi = HAR, HARQ, and HARSV. DM tests based on MSE show no significant difference in cross-sectional average loss values. A summary of individual stock DM statistics shows that for 72 to 89% of the stocks, HAR_CVP has lower average FER and MSE-ln at 5% significance. The percentage drops to 28 to 36% for QLIKE, and to 4 to 7% for MSE. On the other hand, virtually no stocks have HAR_CVP significantly underperforming than the competing models. Over the forecast periods of 2004-2014/5 for SPY and 2006-2014/12 for stocks, HAR_CVP significantly outperforms the competing models.

**Table 9: Volatility Forecast Comparison for Stocks**

This table reports summary for MSE, FER, QLIKE, and MSE-ln loss functions for individual stocks. The mean, median, and standard deviation are the average values across stocks.

Panel A: Summary statistics for loss values. The lowest values across four models are in bold numbers. "Δ(HARi)" reports $\frac{\mathrm{f(HAR\_CVP)}}{\mathrm{f(HARi)}} - 1$ where f is the mean or median and HARi = HAR, HARQ, and HARSV. "Δ(HARi)<0" reports the percentage of stocks with $\frac{\mathrm{f(HAR-CVP)}}{\mathrm{f(HARi)}} < 1$.

| | HAR_CVP | HAR | HARQ | HARSV | Δ(HAR) | Δ(HARQ) | Δ(HARSV) | Δ(HAR)<0 | Δ(HARQ)<0 | Δ(HARSV)<0 |
|---|---|---|---|---|---|---|---|---|---|---|
| MSE | | | | | | | | | | |
| Mean | 33.4 | 63.8 | **29.3** | 70.7 | -48% | 14% | -53% | 56% | 57% | 63% |
| Median | **0.288** | 0.344 | 0.318 | 0.343 | -16% | -10% | -16% | 97% | 88% | 97% |
| StDev | 645 | 1981 | 568 | 2138 | | | | | | |
| FER | | | | | | | | | | |
| Mean | **0.663** | 0.729 | 0.728 | 0.737 | -9% | -9% | -10% | 96% | 92% | 97% |
| Median | **0.109** | 0.125 | 0.122 | 0.125 | -13% | -10% | -13% | 96% | 92% | 99% |
| StDev | 4.65 | 4.46 | 5.16 | 4.58 | | | | | | |
| QLIKE | | | | | | | | | | |
| Mean | **0.585** | 0.944 | 1.168 | 0.938 | -38% | -50% | -38% | 83% | 89% | 85% |
| Median | **0.0879** | 0.100 | 0.0987 | 0.0993 | -12% | -11% | -12% | 96% | 94% | 99% |
| StDev | 6.44 | 13.4 | 20.5 | 14.8 | | | | | | |
| MSE-ln | | | | | | | | | | |
| Mean | **0.543** | 0.673 | 0.683 | 0.663 | -19% | -20% | -18% | 96% | 97% | 96% |
| Median | **0.185** | 0.212 | 0.208 | 0.210 | -13% | -11% | -12% | 96% | 96% | 99% |
| StDev | 1.24 | 1.60 | 1.68 | 1.58 | | | | | | |

Panel B: Summary of DM test against HAR_CVP. DM[$\bar{\bar{d}}$(HARi)] is based on the cross-sectional average difference in loss values and the cross-sectional standard deviation in (19). DM < 0 indicates that HAR_CVP has lower loss function values. The asterisks ***, **, and * indicate statistical significance at 1%, 5%, and 10% respectively. DM[$\bar{d}$(HARi)] is the stock-level DM statistic for testing equal average loss values for HAR_CVP against HARi. DM(HARi) < -1.96 and DM(HARi) > 1.96 are the percentages of stocks satisfying the respective conditions.

| | MSE | FER | QLIKE | MSE-ln |
|---|---|---|---|---|
| DM[$\bar{\bar{d}}$(HAR)] | -1.00 | -4.02*** | -5.17*** | -11.6*** |
| DM[$\bar{\bar{d}}$(HARQ)] | 0.83 | -4.10*** | -3.90*** | -12.2*** |
| DM[$\bar{\bar{d}}$(HARSV)] | -1.01 | -4.85*** | -4.28*** | -12.6*** |
| DM[$\bar{d}$(HAR)] < -1.96 | 7% | 79% | 31% | 89% |
| DM[$\bar{d}$(HARQ)] < -1.96 | 4% | 72% | 36% | 89% |
| DM[$\bar{d}$(HARSV)] < -1.96 | 7% | 75% | 28% | 89% |
| DM[$\bar{d}$(HAR)] > 1.96 | 0% | 0% | 0% | 0% |
| DM[$\bar{d}$(HARQ)] > 1.96 | 0% | 0% | 0% | 1% |
| DM[$\bar{d}$(HARSV)] > 1.96 | 0% | 0% | 1% | 0% |

**(iv) Market Conditions and Forecast Performance**

To examine the robustness and the source of HAR_CVP's forecast performance, we investigate how HAR_CVP performance varies with market conditions. Let $r_{25}$ ($r_{75}$) and $RV_{25}$ ($RV_{75}$) be the 25 (75) percentile values for daily return and RV respectively.[13] They form a 3×3 matrix of market conditions, e.g. low return ($r_t < r_{25}$) and high volatility ($RV_t > RV_{75}$). As before, the performance of HAR_CVP is measured by Δ(HARi). Tables 8 and 9 show that Δ(HARi) is more stable when f is the median. We use Δ(HARi) with f being the median to avoid the undue impact of a few extremely high or low loss values.

Table 10 reports Δ(HARi) under different market conditions.[14] We use bold-faced numbers to highlight the cases where Δ(HARi) under specific market conditions is more negative than the overall Δ(HARi) under the same lost function.[15] We can see that most bold

[13] In our sample, the quartiles for SPY are $r_{25}$ = -0.44%, $r_{75}$ = 0.54%, $RV_{25} = 0.23\times10^{-4}$, and $RV_{75} = 0.82\times10^{-4}$.
[14] To conserve space, we only report the results based on $RV^{OM}$ for SPY. The results based on RV and $RK^{OM}$ are reported in the internet appendix.
[15] For example, in Table 8, median Δ(HAR) based on $RV^{OM}$ and FER is -23%. In Table 10 for SPY $RV^{OM}$, the numbers in the 3×3 matrix for FER and Δ(HAR) are compared against -23%. They are in bold-faced if they are more negative than -23%, e.g. -44% when $r_t > r_{75}$ and $RV_t < RV_{25}$.

## Table 10: Market Conditions and Forecast Performance

This table compares forecast performance under different market conditions. Threshold values $r_{25}$ ($r_{75}$) and $RV_{25}$ ($RV_{75}$) are the 25 (75) percentile values for daily return and RV respectively. Forecast comparison is based on $\Delta(\text{HARi}) = \frac{f(\text{HAR\_CVP})}{f(\text{HARi})} - 1$ where f is the median. Bold numbers indicate that HAR_CVP has better relative performance than the full sample values in Tables 8 and 9.

| | Δ(HAR) | | | Δ(HARQ) | | | Δ(HARSV) | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_t<r_{25}$ | $r_{25}<r_t<r_{75}$ | $r_t>r_{75}$ | $r_t<r_{25}$ | $r_{25}<r_t<r_{75}$ | $r_t>r_{75}$ | $r_t<r_{25}$ | $r_{25}<r_t<r_{75}$ | $r_t>r_{75}$ |
| **SPY $RV^{OM}$** | | | | | | | | | |
| MSE | | | | | | | | | |
| $RV_t<RV_{25}$ | **-20%** | -14% | **-35%** | 10% | **-28%** | **-47%** | 16% | **-17%** | **-36%** |
| $RV_{25}<RV_t<RV_{75}$ | -15% | **-23%** | -10% | **-18%** | **-20%** | -12% | **-19%** | **-26%** | 0% |
| $RV_t>RV_{75}$ | **-41%** | **-33%** | **-27%** | **-33%** | **-14%** | -10% | **-33%** | **-34%** | **-30%** |
| FER | | | | | | | | | |
| $RV_t<RV_{25}$ | -1% | -17% | **-44%** | 11% | **-33%** | **-46%** | 15% | -18% | **-39%** |
| $RV_{25}<RV_t<RV_{75}$ | -15% | **-26%** | **-33%** | -13% | -24% | **-38%** | -11% | **-26%** | -5% |
| $RV_t>RV_{75}$ | **-36%** | -22% | **-34%** | -25% | -17% | **-27%** | **-32%** | **-23%** | **-32%** |
| QLIKE | | | | | | | | | |
| $RV_t<RV_{25}$ | -21% | -13% | **-37%** | -5% | **-35%** | **-44%** | **-20%** | -13% | **-30%** |
| $RV_{25}<RV_t<RV_{75}$ | -14% | -18% | **-45%** | -12% | -22% | **-50%** | -13% | **-21%** | -10% |
| $RV_t>RV_{75}$ | **-33%** | -15% | **-25%** | **-28%** | -5% | 0% | **-25%** | -11% | **-25%** |
| MSE-ln | | | | | | | | | |
| $RV_t<RV_{25}$ | -9% | -10% | **-34%** | 10% | **-31%** | **-49%** | 3% | -17% | **-34%** |
| $RV_{25}<RV_t<RV_{75}$ | -23% | -17% | **-34%** | -18% | -24% | **-44%** | **-20%** | **-23%** | -6% |
| $RV_t>RV_{75}$ | **-32%** | **-26%** | -24% | -25% | -16% | 5% | **-24%** | -16% | **-25%** |
| **Stocks** | | | | | | | | | |
| MSE | | | | | | | | | |
| $RV_t<RV_{25}$ | **-32%** | -2% | 14% | **-17%** | 12% | 18% | **-24%** | 1% | 13% |
| $RV_{25}<RV_t<RV_{75}$ | **-26%** | -14% | -9% | **-17%** | -9% | **-11%** | **-22%** | -14% | -8% |
| $RV_t>RV_{75}$ | -8% | -14% | -6% | **-13%** | **-18%** | -4% | -7% | -15% | -7% |
| FER | | | | | | | | | |
| $RV_t<RV_{25}$ | **-28%** | -3% | 0% | **-16%** | 5% | -3% | **-21%** | 0% | 2% |
| $RV_{25}<RV_t<RV_{75}$ | **-23%** | -12% | **-17%** | **-16%** | -9% | **-19%** | **-19%** | -13% | **-14%** |
| $RV_t>RV_{75}$ | -5% | -12% | -12% | -10% | **-15%** | -8% | -3% | **-15%** | -13% |
| QLIKE | | | | | | | | | |
| $RV_t<RV_{25}$ | **-24%** | -3% | -9% | **-15%** | -3% | **-20%** | **-15%** | -1% | -7% |
| $RV_{25}<RV_t<RV_{75}$ | **-18%** | -10% | **-20%** | -11% | -9% | **-25%** | **-16%** | -10% | **-17%** |
| $RV_t>RV_{75}$ | -5% | -10% | **-15%** | -7% | -10% | -8% | -5% | -12% | **-15%** |
| MSE-ln | | | | | | | | | |
| $RV_t<RV_{25}$ | **-26%** | -4% | **-16%** | -6% | 2% | **-22%** | **-24%** | -2% | -11% |
| $RV_{25}<RV_t<RV_{75}$ | **-21%** | -12% | **-24%** | **-14%** | -8% | **-28%** | **-21%** | **-14%** | **-24%** |
| $RV_t>RV_{75}$ | -4% | -11% | -13% | -8% | **-12%** | -7% | -6% | **-14%** | **-14%** |

**Table 11: Forecast Performance during the Global Financial Crisis**

This table reports forecast comparison during the global financial crisis from 1 July 2008 to 30 June 2009. Forecast comparison is based on $\Delta(\text{HARi}) = \frac{\text{f(HAR\_CVP)}}{\text{f(HARi)}} - 1$ where f is the mean or median and HARi = HAR, HARQ, and HARSV. Bold numbers indicate that HAR_CVP has better relative performance than the full sample values in Tables 8 and 9.

| | SPY RV | | | SPY $RV^{OM}$ | | | Stocks | | |
|---|---|---|---|---|---|---|---|---|---|
| | Δ(HAR) | Δ(HARQ) | Δ(HARSV) | Δ(HAR) | Δ(HARQ) | Δ(HARSV) | Δ(HAR) | Δ(HARQ) | Δ(HARSV) |
| MSE | | | | | | | | | |
| Mean | -5% | **-25%** | 6% | 42% | 17% | 54% | **-49%** | 15% | **-55%** |
| Median | **-57%** | **-53%** | **-51%** | **-50%** | **-52%** | **-52%** | -15% | **-14%** | **-17%** |
| FER | | | | | | | | | |
| Mean | -25% | -13% | -21% | -19% | -3% | **-22%** | -6% | -6% | -9% |
| Median | **-50%** | **-34%** | **-35%** | **-44%** | -17% | **-38%** | **-20%** | **-14%** | **-19%** |
| QLIKE | | | | | | | | | |
| Mean | -31% | 24% | -35% | **-63%** | **-53%** | **-55%** | **-45%** | -47% | **-41%** |
| Median | **-39%** | -19% | **-24%** | **-33%** | -12% | **-28%** | **-17%** | **-12%** | **-19%** |
| MSE-ln | | | | | | | | | |
| Mean | -38% | -7% | **-37%** | **-50%** | -22% | **-44%** | **-27%** | -19% | **-27%** |
| Median | **-42%** | -19% | **-28%** | **-34%** | -15% | **-29%** | **-19%** | **-13%** | **-19%** |

numbers are on days with $r_t < r_{25}$ or $r_t > r_{75}$: HAR_CVP performs stronger when daily return is large, either positive or negative. The effect is stronger at the top return quartile for SPY but at the bottom return quartile for stocks. If $CVP_t = \beta_0 + \beta_- r_t^- + \beta_+ r_t^+ + \beta_{RV} RV_t$ is a good proxy for the latent RV persistence, large returns implies large changes in RV persistence. The superior forecasts by HAR_CVP when return is large suggests that CVP is able to capture a significant portion of the daily variations in RV persistence. On the other hand, HAR and HARSV have constant RV persistence over the 4-year estimation window; HARQ captures the changes in RV persistence associated with RQ.[16]

To further test the performance of HAR_CVP under extreme market conditions, we next compare forecast accuracy during the global financial crisis (GFC) period of 1 July 2008 to 30 June 2009.[17] The GFC was a period of extreme economic, financial, and policy uncertainty. The heightened uncertainty was reflected in large daily swings in market indices and prolonged high volatility. If large returns are a key driver of the latent volatility persistence, one would expect HAR_CVP to *perform even better* during the GFC relative to models with constant volatility persistence (e.g. HAR and HARSV) and models where volatility persistence is unrelated to returns (e.g. HARQ).

Table 11 reports $\Delta(\text{HARi}) = \frac{f(\text{HAR\_CVP})}{f(\text{HARi})} - 1$ during the GFC period. For SPY and stocks and f being the mean or the median, Δ(HARi) is mostly negative: HAR_CVP maintains its superior forecast accuracy during the GFC. Again we use bold-faced numbers to indicate that HAR_CVP has greater percentage loss reduction during GFC than for the full sample.[18] For SPY and stocks, 30 out of the 36 median loss values and 14 out of the 36 average loss values are in bold numbers, indicating even stronger relative performance by HAR_CVP during the GFC. HAR_CVP is able to capture the surge in volatility persistence driven by larges swings in daily returns (Figure 1). Models that are slow to capture the large increase in volatility persistence tend to do relatively worse during the GFC. Through RQ, HARQ can partially capture daily variations in RV persistence. We note that during the GFC, the loss reductions from HAR_CVP are often smaller against HARQ than against HAR and HARSV.

[16] There is no strong unconditional effect from RV on HAR_CVP performance, perhaps due to the small $\beta_{RV}$. Conditional on $r_t < r_{25}$, Δ(HARi) becomes more negative for SPY and less negative for stocks as RV increases. Conditional on $r_t > r_{75}$, Δ(HARi) becomes less negative for SPY and has not strong trend as RV increases. These empirical patterns should be further explored in future studies.
[17] A shorter volatility-based GFC period from 2008/9/1 to 2009/4/30 produces the same qualitative results.
[18] For example, the change in the median FER against HAR for SPY $RV^{OM}$ is -44% during GFC vs -23% in the full sample in Table 8. The same comparison is -50% vs -24% for SPY RV and -20% vs -13% for stocks.

### (v) CVP and HAR_CVP Forecast Performance

This subsection explores the direct link between CVP and HAR_CVP forecast accuracy. The high CVP during the GFC (Figure 1) may leave the impression that high CVP values improve HAR_CVP forecast accuracy. This is true only if the true volatility persistence is also high, as during the GFC. In general, CVP improves forecast accuracy when it is a good proxy for the latent RV persistence. As such, one may not expect a positive relation between CVP values and HAR_CVP's forecast accuracy. On the other hand, if the variations of CVP capture the variations of the latent RV persistence, it would give HAR_CVP an advantage over models with constant persistence (e.g. HAR and HARSV) or partially adjusting persistence (e.g. HARQ). While the issue deserves greater attention given here, we provide some preliminary evidence that highlights the importance of time-varying volatility persistence.

We examine the link between HAR_CVP forecast performance and the characteristics of daily CVP. Again HAR_CVP performance is measured by $\Delta(\text{HARi}) \equiv \frac{f(\text{HAR\_CVP})}{f(\text{HARi})} - 1$ with f being the median loss values. The basic characteristics of daily CVP are its average and standard deviation. For SPY, we first compute CVP average, standard deviation, and Δ(HARi) for each quarter from 2004 to 2014/5. We then compute the time-series correlations across 42 quarters between CVP average and Δ(HARi), and between CVP standard deviation and Δ(HARi). For stocks, we first compute each stock's CVP average, standard deviation, and Δ(HARi), then the correlations across 72 stocks between CVP average and Δ(HARi), and between CVP standard deviation and Δ(HARi). These correlations shed light on the direction and strength of the relations between CVP and HAR_CVP's relative forecast performance.

The correlations are reported in Table 12. For SPY, across three competing models and four loss functions, the average correlation between CVP average and Δ(HARi) is 0.107 and significant at 5%:[19] high CVP is associated with less negative Δ(HARi), i.e. poor HAR_CVP forecasts. However, most quarterly correlations are not significant and one of them is actually negative and significant. On the other hand, all of the correlations between CVP standard deviation and Δ(HARi) are negative. While only three of the twelve correlations are statistical significant, the average correlations across three models and four loss functions are -0.185 and significant at 1%. There is strong evidence that high CVP standard deviation is associated with more negative Δ(HARi), i.e. superior HAR_CVP forecasts.

---

[19] Test for zero correlation is based on the estimated correlation $\hat{\rho}$ and its standard error $\sqrt{\frac{1-\hat{\rho}^2}{N-2}}$ where N is the sample size.

**Table 12: CVP and HAR_CVP Forecast Performance**

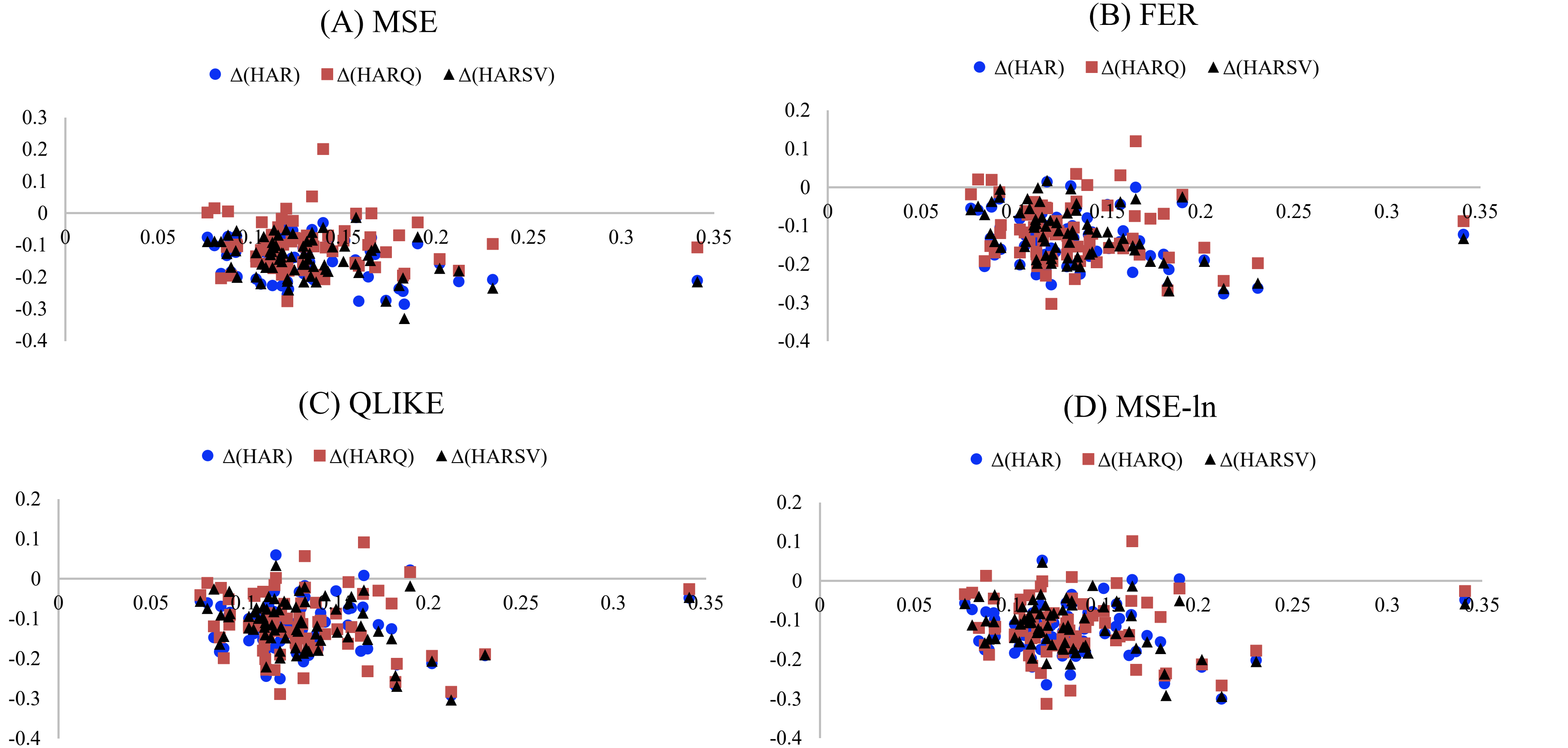
This table reports the correlations between CVP average, standard deviation, and the relative forecast performance of HAR_CVP based on $\Delta(\text{HARi}) = \frac{f(\text{HAR\_CVP})}{f(\text{HARi})} - 1$ with HARi = HAR, HARQ, and HARSV. For SPY, f is the quarterly median loss values. For stocks, f is the median loss values of each stock. The asterisks ***, **, and * indicate significance at 1%, 5%, and 10% respectively.

| | SPY | | Stocks | |
|---|---|---|---|---|
| | Ave(CVP) | StDev(CVP) | Ave(CVP) | StDev(CVP) |
| MSE | | | | |
| Cor[CVP, Δ(HAR)] | 0.119 | -0.367** | -0.339*** | -0.247** |
| Cor[CVP, Δ(HARQ)] | -0.320** | -0.139 | -0.256** | -0.055 |
| Cor[CVP, Δ(HARSV)] | 0.091 | -0.427*** | -0.265** | -0.352*** |
| FER | | | | |
| Cor[CVP, Δ(HAR)] | 0.183 | -0.198 | -0.188 | -0.219* |
| Cor[CVP, Δ(HARQ)] | -0.057 | -0.033 | -0.196* | -0.092 |
| Cor[CVP, Δ(HARSV)] | 0.127 | -0.294* | -0.109 | -0.371*** |
| QLIKE | | | | |
| Cor[CVP, Δ(HAR)] | 0.265* | -0.136 | -0.002 | -0.088 |
| Cor[CVP, Δ(HARQ)] | 0.157 | -0.007 | -0.126 | -0.044 |
| Cor[CVP, Δ(HARSV)] | 0.156 | -0.183 | 0.023 | -0.216* |
| MSE-ln | | | | |
| Cor[CVP, Δ(HAR)] | 0.255* | -0.177 | -0.008 | -0.078 |
| Cor[CVP, Δ(HARQ)] | 0.107 | -0.063 | -0.113 | -0.034 |
| Cor[CVP, Δ(HARSV)] | 0.197 | -0.195 | 0.033 | -0.217* |
| Average | 0.107** | -0.185*** | -0.129*** | -0.168*** |

For stocks, the evidence is stronger for a negative correlation between CVP standard deviation and Δ(HARi). All of the correlations between CVP standard deviation and Δ(HARi) are negative and six of them are significant. The average correlations across three models and four loss functions are -0.168 and significant 1%. Figure 5 plots stocks' Δ(HARi) against their CVP standard deviations for MSE, FER, QLIKE, and MSE-ln. In all four cases, stocks' Δ(HARi) are mostly negative, and become more negative for stocks with high CVP standard deviations. While most stocks have CVP standard deviations ranging from 0.08 to 0.23, there is one stock GILD (Gilead Sciences Inc.) with exceptionally high CVP standard deviation at 0.34. Removing this stock increases the number of significant correlations between CVP standard deviation and Δ(HARi) to eight out of twelve and the average correlation to -0.255. Unlike SPY, stocks have a significantly negative average correlation between CVP average

**Figure 5: CVP Variation and Relative Forecast Performance of HAR_CVP for Stocks**

HAR_CVP performance is measured by $\Delta(\text{HARi}) = \frac{\text{f(HAR\_CVP)}}{\text{f(HARi)}} - 1$ with f being the median loss values of each stock and HARi = HAR, HARQ, and HARSV. The loss functions are MSE, FER, QLIKE, and MSE-ln**.** The horizonal axis is the CVP standard deviation of each stock.



and Δ(HARi) of -0.129: high CVP values are associated with better HAR_CVP performance. Removing GILD increases the average correlation to -0.137.

Overall we find a strong link between CVP variations and HAR_CVP forecast performance. This is consistent with the high Shapley $R^2$ from CVP variables reported in Panel C of Table 2. While HAR_CVP performs better than the competing models, it performs even stronger when there is high variation in RV persistence. We find mixed evidence on the link between CVP level and HAR_CVP forecast performance.

## VII. Conclusion and Future Research

This study shows that the persistence in daily volatility is mainly driven by return. Large returns, positive or negative, are associated with higher dependence of tomorrow's volatility on today's volatility. The dependence is lower when today's return is small but volatility is high. Negative returns increase volatility persistence more than positive returns. This asymmetric volatility persistence has much higher explanatory power than the direct return impact on future volatility and offers a new explanation for asymmetric volatility. By estimating the time-varying dependence between the current and future volatilities, our model improves volatility forecast relative to recent advances in the literature.

The key finding of our study is that volatility persistence varies with the market state. This can be further tested in a more flexible multi-frequency model such as the MIDAS model of Ghysels, Santa-Clara, and Valkanov (2006). It can also be tested in the panel-based HAR models in Bollerslev et al. (2017) for commonality in volatility persistence. Different functional forms for CVP, e.g. a logit function, can be experimented. Extensions to long-run volatility and different asset classes should also be investigated. While we outline the heuristic arguments for CVP in Section II, formal economic modelling is needed to provide theoretical support. Alternative mechanisms for CVP should be explored. Given the on-going debate on leverage effect versus volatility feedback as the main channel for asymmetric volatility, e.g. Choi and Richardson (2016) and Engle and Siriwardane (2016), a valuable empirical analysis would be to assess these explanations while controlling the asymmetric return impact on volatility persistence. At a more fundamental level, the return impact on volatility persistence should be considered in modelling and testing the risk-return relationship and in studies of volatility risk premium.

**Reference**


Amaya, D., Christoffersen, P., Jacobs, K., & Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, *118*(1), 135-167.

Andersen, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. *The Journal of Finance*, *51*(1), 169-204.

Andrei, D., & Hasler, M. (2015). Investor attention and stock market volatility. *The Review of Financial Studies*, *28*(1), 33-72.

Andrews, D. W. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica: Journal of the Econometric Society*, 817-858.

Audrino, F., & Knaus, S. D. (2016). Lassoing the HAR model: A model selection perspective on realized volatility dynamics. *Econometric Reviews*, *35*(8-10), 1485-1521.

Avramov, D., Chordia, T., & Goyal, A. (2006). The impact of trades on daily volatility. *The Review of Financial Studies*, *19*(4), 1241-1277.

Barber, B. M., & Odean, T. (2007). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies*, *21*(2), 785-818.

Barndorff-Nielsen, O. E., P. R.Hansen, A. Lunde, And N. Shephard (2009): "Realized kernels in practice: trades and quotes," *Econometrics Journal*, 12(3), C1–C32.

Berger, D., Chaboud, A., & Hjalmarsson, E. (2009). What drives volatility persistence in the foreign exchange market? *Journal of Financial Economics*, *94*(2), 192-213.

Black, F., (1976). Studies of stock price volatility changes, in Proceedings of the 1976 meetings of the American Statistical Association, Business and Economic Statistics Section.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307-327.

Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, *192*(1), 1-18.

Bollerslev, T., Li, S. Z., & Zhao, B. (2016). Good Volatility, Bad Volatility and the Cross-Section of Stock Returns.

Bollerslev, T., Hood, B., Huss, J., & Pedersen, L. H. (2017). Risk everywhere: Modeling and managing volatility.

Brock, W.A., LeBaron, B., 1996. A dynamic structural model for stock return volatility and trading volume. Review of Economics and Statistics 78, 94–110.

Cabrales, A., Hoshi, T., 1996. Heterogeneous beliefs, wealth accumulation, and asset price dynamics. Journal of Economic Dynamics and Control 20, 1073–1100.

Campbell, J. Y., & Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, *31*(3), 281-318.

Cao, H. H., Coval, J. D., & Hirshleifer, D. (2002). Sidelined investors, trading-generated news, and security returns. *Review of Financial Studies*, *15*(2), 615-648.

Chen, X. and Ghysels, E., 2010. News—good or bad—and its impact on volatility predictions over multiple horizons. *The Review of Financial Studies*, *24*(1), pp.46-81.

Chordia, T., Sarkar, A., & Subrahmanyam, A. (2005). An empirical analysis of stock and bond market liquidity. *Review of Financial Studies*, *18*(1), 85-129.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica: journal of the Econometric Society*, 135-155.

Clements, M. P., Galvão, A. B., & Kim, J. H. (2008). Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. *Journal of Empirical Finance*, *15*(4), 729-750.

Clements, A., & Liao, Y. (2017). Forecasting the variance of stock index returns using jumps and cojumps. *International Journal of Forecasting*, *33*(3), 729-742.

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

Corsi, F., & Renò, R. (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics*, *30*(3), 368-380.

Craioveanu, M., & Hillebrand, E. (2012). Why it is OK to use the HAR-RV (1, 5, 21) model. Working Paper, CREATES, University of Aarhus, Denmark.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461-1499.

Daouk, H., & Ng, D. (2011). Is unlevered firm volatility asymmetric? *Journal of Empirical Finance*, *18*(4), 634-651.

Dieobold, F. X. (1986). Modeling the persistence of conditional variances: A comment. *Econometric Reviews*, *5*(1), 51-56.

Diebold, F. X., & Nerlove, M. (1989). The dynamics of exchange rate volatility: a multivariate latent factor ARCH model. *Journal of Applied econometrics*, *4*(1), 1-21.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & economic statistics*, 13(3), 134-144.

Dimpfl, T., & Jank, S. (2016). Can internet search queries help to predict stock market volatility? *European Financial Management*, *22*(2), 171-192.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of finance*, *48*(5), 1749-1778.

Engle, R. F., & Siriwardane, E. (2016). Structural GARCH: the volatility-leverage connection. HBS working paper 16-009.

de Fontnouvelle, P., 2000. Information dynamics in financial markets. Macroeconomic Dynamics 4, 139–169.

Gallant, A. R., Hsieh, D. A., & Tauchen, G. E. (1991). On fitting a recalcitrant series: the pound/dollar exchange rate, 1974-1983. *Nonparametric and semiparametric methods in econometrics and statistics*, 199-240.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, *131*(1), 59-95.

Ghysels, E., Sinko, A., & Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, *26*(1), 53-90.

Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, *74*(6), 1545-1578.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, *48*(5), 1779-1801.

Goodhart, C. A., & O'Hara, M. (1997). High frequency data in financial markets: Issues and applications. *Journal of Empirical Finance*, *4*(2), 73-114.

Hameed, A., Kang, W., & Viswanathan, S. (2010). Stock market declines and liquidity. *The Journal of Finance*, *65*(1), 257-293.

Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of econometrics*, *64*(1), 307-333.

Hansen, P. R., & Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, *24*(2), 127-161.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The journal of Finance*, *50*(4), 1175-1199.

Hasanhodzic, J., & Lo, A. W. (2013). Black's leverage effect is not due to leverage.

He, X., & Velu, R. (2014). Volume and volatility in a common-factor mixture of distributions model. *Journal of Financial and Quantitative Analysis*, *49*(01), 33-49.

Hou, K., & Moskowitz, T. J. (2005). Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies*, *18*(3), 981-1020.

Hou, K., Xiong, W., & Peng, L. (2009). A tale of two anomalies: The implications of investor attention for price and earnings momentum.

Johnson, T. C. (2001). Return dynamics when persistence is unobservable. *Mathematical Finance*, *11*(4), 415-445.

Kremer, S. and Nautz, D., 2013. Causes and consequences of short-term institutional herding. *Journal of Banking & Finance*, *37*(5), pp.1676-1686.

Kremer, S. and Nautz, D., 2013. Short-term Herding of Institutional Traders: New Evidence from the German Stock Market. *European Financial Management*, *19*(4), pp.730-746.

Lahaye, J., & Neely, C. J. (2016). The role of jumps in volatility spillovers in foreign exchange markets: meteor shower and heat waves revisited. *Federal Reserve Bank of St. Louis Working Paper Series*.

Lamoureux, C. G., & Lastrapes, W. D. (1990). Persistence in variance, structural change, and the GARCH model. *Journal of Business & Economic Statistics*, *8*(2), 225-234.

Lamoureux, C. G., & Lastrapes, W. D. (1994). Endogenous trading volume and momentum in stock-return volatility. *Journal of Business & Economic Statistics*, *12*(2), 253-260.

Laux, P. A., & Ng, L. K. (1993). The sources of GARCH: Empirical evidence from an intraday returns model incorporating systematic and unique risks. *Journal of International Money and Finance*, *12*(5), 543-560.

Liesenfeld, R. (1998). Dynamic BivarSate Mixture Models: Modeling the Behavior of Prices and Trading Volume. *Journal of Business & Economic Statistics*, *16*(1), 101-109.

Liesenfeld, R. (2001). A generalized bivariate mixture model for stock price volatility and trading volume. *Journal of Econometrics*, *104*(1), 141-178.

McQueen, G., & Vorkink, K. (2004). Whence GARCH? A preference-based explanation for conditional volatility. *Review of Financial Studies*, *17*(4), 915-949.

Ning, C., Xu, D., & Wirjanto, T. S. (2015). Is volatility clustering of asset returns asymmetric? *Journal of Banking & Finance*, *52*, 62-76.

Ozturk, S.R., Van der Wel, M. and van Dijk, D., 2017. Intraday price discovery in fragmented markets. *Journal of Financial Markets*, *32*, pp.28-48.

Owen, A., Prieur, C. (2017). On Shapley value for measuring importance of dependent variables. Working paper. Stanford University.

Parkinson, M. (1980), "The Extreme Value Method for Estimating the Variance of the Rate of Return," *Journal of Business*, 53, 61-65

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, *160*(1), 246-256.

Patton, A. J., & Sheppard, K. (2009). Evaluating volatility and correlation forecasts. In *Handbook of financial time series* (pp. 801-838). Springer Berlin Heidelberg.

Patton, A. J., & Ramadorai, T. (2013). On the High‐Frequency Dynamics of Hedge Fund Risk Exposures. *The Journal of Finance*, *68*(2), 597-635.

Patton, A. J., & Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics*, 97(3), 683-697.

Pindyck, R., (1984). Risk, Inflation, and the Stock Market. *American Economic Review 74,* pp.335–351.

Ross, S. A. (1989). Information and volatility: The no‐arbitrage martingale approach to timing and resolution irrelevancy. *The Journal of Finance*, *44*(1), 1-17.

Talpsepp, T., & Rieger, M. O. (2010). Explaining asymmetric volatility around the world. *Journal of Empirical Finance*, *17*(5), 938-956.

Tauchen, G. E., & Pitts, M. (1983). The price variability-volume relationship on speculative markets. *Econometrica: Journal of the Econometric Society*, 485-505.

Wang, J., & Yang, M. (2011). Housewives of Tokyo versus the gnomes of Zurich: Measuring price discovery in sequential markets. *Journal of Financial Markets*, *14*(1), 82-108.

Vlastakis, N. and Markellos, R.N., 2012. Information demand and stock market volatility. *Journal of Banking & Finance*, *36*(6), pp.1808-1821.

Xue, Y., & Gençay, R. (2012). Trading frequency and volatility clustering. *Journal of Banking & Finance*, *36*(3), 760-773.

Zhang, L., Mykland, P. A., & Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, *100*(472), 1394-1411.

## Appendix A: Volatility Impact of Stock and Market Returns

We show that in the HAR and HAR_CVP models for individual stocks, the effects of the S&P index returns dominate those of stock returns. Panel A reports the estimation results of the HAR model with both stock returns and the S&P index returns:

$$\mathrm{RV}_{t+1} = \alpha + \beta_D \mathrm{RV}_t + \beta_W \mathrm{RV}_{t,W} + \beta_M \mathrm{RV}_{t,M} + \varphi_D r_t^S + \varphi_W r_{t,W}^S + \varphi_M r_{t,M}^S + \theta_D r_t + \theta_W r_{t,W} + \theta_M r_{t,M} + \varepsilon_{t+1}$$

where $r_t^S$ is the stock return and $r_t$ is the S&P 500 index return. It shows that coefficients of stock returns are not significant when the S&P index returns are included. Panel B reports the estimation results of the HAR_CVP model in (4) with stock returns, not S&P 500 index returns. Compared to Panel B of Table 2, the CVP return coefficients $\beta_{|r|}$ and $\beta_r$ and the leverage coefficients $\theta_D$, $\theta_W$, and $\theta_M$ are smaller with lower t statistics.

Panel A: HAR

| | $\beta_D$ | $\beta_W$ | $\beta_M$ | $\varphi_D$ | $\varphi_W$ | $\varphi_M$ | $\theta_D$ | $\theta_W$ | $\theta_M$ | $\alpha$ | $\bar{R}^2$ | AIC | LB5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ave Coeff | 0.312$^{***}$ | 0.392$^{***}$ | 0.192$^{***}$ | -0.240$^{***}$ | -0.360$^{***}$ | -0.295$^{**}$ | | | | 0.293 | 0.567 | 6464 | 77 |
| *t stat* | *5.61* | *5.57* | *3.97* | *-3.24* | *-3.37* | *-2.27* | | | | *4.18* | | | |
| *%(t≤-1.96)* | *0%* | *0%* | *0%* | *66%* | *53%* | *23%* | | | | *0%* | | | |
| *%(t≥1.96)* | *93%* | *97%* | *79%* | *0%* | *0%* | *0%* | | | | *91%* | | | |
| Ave Coeff | 0.270$^{***}$ | 0.371$^{***}$ | 0.214$^{***}$ | -0.0943 | -0.127 | -0.134 | -0.453$^{***}$ | -0.899$^{***}$ | -0.820$^{***}$ | 0.420$^{***}$ | 0.587 | 6315 | 70 |
| *t stat* | *4.61* | *5.67* | *4.65* | *-1.52* | *-1.28* | *-0.87* | *-3.36* | *-3.56* | *-2.33* | *5.59* | | | |
| *%(t≤-1.96)* | *0%* | *0%* | *0%* | *10%* | *8%* | *5%* | *87%* | *90%* | *57%* | *0%* | | | |
| *%(t≥1.96)* | *85%* | *95%* | *90%* | *1%* | *1%* | *0%* | *0%* | *0%* | *0%* | *100%* | | | |

Panel B: HAR_CVP

| | $\beta_0$ | $\beta_{\|r\|}$ | $\beta_r$ | $\beta_{RV}$ | $\beta_W$ | $\beta_M$ | $\varphi_D$ | $\varphi_W$ | $\varphi_M$ | $\bar{R}^2$ | F | AIC | LB5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Return = Stock Returns | | | | | | | | | | | | | |
| Ave Coeff | 0.494$^{***}$ | 0.0323$^{***}$ | -0.0160$^{**}$ | -0.0069$^{***}$ | 0.316$^{***}$ | 0.129$^{***}$ | -0.090$^{*}$ | -0.293$^{***}$ | -0.231$^{***}$ | 0.608 | 126 | 6124 | 56 |
| *t stat* | *7.19* | *3.19* | *-2.26* | *-4.86* | *6.12* | *2.72* | *-1.66* | *-3.35* | *-2.06* | | | | |
| *%(t≤-1.96)* | *0%* | *1%* | *28%* | *89%* | *0%* | *0%* | *21%* | *54%* | *16%* | | | | |
| *%(t≥1.96)* | *99%* | *49%* | *0%* | *0%* | *93%* | *47%* | *0%* | *0%* | *0%* | | | | |

## Appendix B: Additional Conditioning Variables

In addition to daily return and RV, we examine whether daily volatility persistence is affected by volatility jumps (VJ), number of trades (NT), illiquidity (IL), and the imbalance of buyer- and seller-initiated trades (TI). Volatility jumps have been shown to help forecast future volatility, e.g. Corsi and Reno (2012). The continuous component of $RV_t$ is the bipower variation defined as $BV_t = \frac{\pi}{2}\sum_{i=2}^{n}|r_{i,t}||r_{i-1,t}|$. It converges to the integrated variance as $n \to \infty$ and is calculated using the skip-4 method of Huang and Tauchen (2005) and Patton and Sheppard (2015) to improve its properties. Volatility jump is $VJ_t = RV_t - BV_t$. Barndorff-Nielsen and Shephard (2006) suggest the following statistic for testing $VJ_t = 0$:

$$Z_t = \frac{n^{1/2}(BV_t/RV_t - 1)}{(\pi^2/4 + \pi - 5)^{1/2} \times \max\{1, QV_t^{1/2}/BV_t\}}$$

where $QV_t \equiv \frac{\pi^2 n}{4}\sum_{i=4}^{n}|r_{i,t}||r_{i-1,t}||r_{i-2,t}||r_{i-3,t}|$. $Z_t$ is asymptotically N(0,1) in the absence of jumps. Let $z_\alpha$ be the left tail of N(0,1) with $P(Z<z_\alpha) = \alpha$. Volatility jump is defined as $VJ_t = I_{(Z_t<z_\alpha)}(RV_t - BV_t)$. With $\alpha = 1\%$ and $z_\alpha = -2.326$, jumps occur on 9% of trading days for SPY and more often for individual stocks. Daily illiquidity is measured by the Amihud (2000) measure defined as $IL_t = |r_t|/Vol_t$ where $Vol_t$ is trading volume in unit of million. We use the bulk volume classification of O'Hara, et al. (2012) to partition the 5-minute trades into buyer- and seller-initiated portions, with the difference being the trade imbalance ($TI_t$). Recently O'Hara, et al. (2015) show that order imbalance based on bulk volume classifications are good proxies of information-based trading.

Let $Y_t$ be $VJ_t$, $NT_t$, $IL_t$, or $TI_t$. To assess the impact of these variables on volatility persistence, we extend the model in (4) to include $Y_t$ and its interaction with $RV_t$:

$$RV_{t+1} = \alpha + \left(\beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{RV}RV_t + \beta_Y Y_t\right)RV_t + \delta_Y Y_t + \varphi Z_t + \varepsilon_{t+1}$$

Table B reports the impact of the additional conditioning variables. While some of the additional variables have a significant impact on future volatility level, their effects on volatility persistence are either insignificant or discrepant. Volatility jumps reduces volatility persistence for SPY but have no effect for individual stocks. Forsberg and Ghysels (2007) find that “the dynamics of the [volatility] process do not change whether we have a jump at time t or not, not even at the one day prediction horizon.” Clements and Liao (2017) show that the jump intensity as estimated from a Hawkes model increases RV persistence. They do not control the impact of return and RV. The lack of impact from trading intensity (NT) is consistent with those of Fleming, Kirby, and Ostdiek (2006) who find trading volume does not explain the ARCH effect in volatility. For SPY, both IL and TI increase future volatility level but decrease volatility persistence. For individual stocks, IL reduces future volatility and VI has no effect. Any information asymmetry embedded in VI appears to be subsumed by return and RV. Table B shows that the CVP coefficients remain significant in almost all cases. For SPY, $\hat{\beta}_{|r|}$ is largely intact as in Table 2; $\hat{\beta}_{r}$ is significant except for TI; $\hat{\beta}_{RV}$ is also smaller but highly significant. For stocks, additional variables are largely insignificant and the CVP coefficients are similar to those in Table 2.

**Table B: Additional Conditioning Variables**

This table reports the estimated coefficients of the following regression:

$$RV_{t+1} = \alpha + \left(\beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{RV} RV_t + \beta_Y Y_t\right) RV_t + \delta_Y Y_t + \varphi Z_t + \varepsilon_{t+1}$$

$Y_t$ is one of the following variables: volatility jump (VJ), the number of trades (NT), illiquidity (IL), and trade imbalance (TI). $Z_t = (RV_{t,W},\ RV_{t,M},\ r_t,\ r_{t,W},\ r_{t,M})'$. The other variables are the same as in Table 3. The asterisks ***, **, * indicate significance at 1%, 5%, and 10% respectively.

| | $\beta_0$ | $\beta_{\|r\|}$ | $\beta_r$ | $\beta_{RV}$ | $\beta_Y$ | $\delta_Y$ | $\bar{R}^2$ |
|---|---|---|---|---|---|---|---|
| **SPY:** | | | | | | | |
| VJ | 0.322*** | 0.106*** | -0.036** | -0.011*** | -0.070** | 0.222 | 0.718 |
| *t-stat* | *3.97* | *3.38* | *-2.12* | *-5.63* | *-2.02* | *0.65* | |
| NT | 0.304*** | 0.100*** | -0.033** | -0.011*** | 0.019 | -0.031 | 0.713 |
| | *3.51* | *3.40* | *-2.05* | *-4.11* | *0.20* | *-0.13* | |
| IL | 0.361*** | 0.102*** | -0.034*** | -0.011*** | -0.088* | 0.381*** | 0.716 |
| | *4.60* | *3.56* | *-2.15* | *-5.59* | *-1.79* | *2.78* | |
| TI | 0.322*** | 0.105*** | -0.013 | -0.009*** | -0.003* | 0.006** | 0.720 |
| | *4.36* | *3.60* | *-0.92* | *-6.52* | *-1.76* | *2.28* | |
| **Stocks** | | | | | | | |
| VJ | 0.404*** | 0.068*** | -0.027*** | -0.006*** | -0.005 | -0.015 | 0.643 |
| *t-stat* | *6.64* | *4.62* | *-2.95* | *-6.18* | *-0.28* | *-0.05* | |
| *%(t≤-1.96)* | *0%* | *0%* | *54%* | *87%* | *22%* | *16%* | |
| *%(t≥1.96)* | *92%* | *82%* | *0%* | *0%* | *21%* | *5%* | |
| NT | 0.353*** | 0.067*** | -0.026*** | -0.006*** | 0.000 | 0.000 | 0.644 |
| *t-stat* | *4.92* | *4.62* | *-2.84* | *-4.89* | *0.08* | *0.01* | |
| *%(t≤-1.96)* | *0%* | *0%* | *51%* | *80%* | *15%* | *8%* | |
| *%(t≥1.96)* | *85%* | *80%* | *0%* | *0%* | *9%* | *7%* | |
| IL | 0.478*** | 0.067*** | -0.029*** | -0.006*** | -0.004 | -0.461* | 0.648 |
| *t-stat* | *6.56* | *4.82* | *-3.19* | *-2.77* | *-0.25* | *-1.73* | |
| *%(t≤-1.96)* | *0%* | *0%* | *57%* | *53%* | *20%* | *18%* | |
| *%(t≥1.96)* | *94%* | *82%* | *0%* | *3%* | *11%* | *2%* | |
| TI | 0.388*** | 0.069*** | -0.020* | -0.006*** | -0.015 | 0.118 | 0.643 |
| *t-stat* | *6.77* | *4.88* | *-1.82* | *-6.01* | *-0.44* | *0.88* | |
| *%(t≤-1.96)* | *0%* | *0%* | *36%* | *92%* | *9%* | *1%* | |
| *%(t≥1.96)* | *93%* | *83%* | *3%* | *0%* | *7%* | *10%* | |

## Appendix C: Sub-period Results

Sub-period analyses allow us to assess how the statistical relationship has evolved over time and when the relationship is strong and weak. We require all stocks have 6 sub-periods, i.e. starting from 2002, which leaves 72 stocks in the sample. Table C reports the results of sub-period estimation. In most sub-periods, at least two of the three conditioning variables are significant with the same signs as in Table 2. There is no significant coefficient of the opposite sign. Negative returns have greater impact than positive returns in almost all sub-periods. Therefore the return-size effect is greater than the return-direction effect in most sub-periods. The HAR_CVP model works particularly well during the crisis period of 2008-09, with all conditioning variables highly significant. The results indicate that the qualitative relationship in Table 2 holds in most sub-periods regardless the level of volatility.

**Table C: Sub-period estimations**

This table reports the daily persistence coefficients of the following models:

$$RV_{t+1} = \alpha + (\beta_0 + \beta_{|r|}|r_t| + \beta_r r_t + \beta_{RV} RV_t) RV_t + \varphi Z_t + \varepsilon_{t+1}$$

$Z_t = (RV_{t,W},\ RV_{t,M},\ r_t,\ r_{t,W},\ r_{t,M})'$. %(t ≤ -1.96) and %(t ≥ 1.96) are the percentage of stocks with t ≤ -1.96 or t ≥ 1.96 respectively. The asterisks ***, **, * indicate significance at 1%, 5%, and 10% respectively.

Panel A: Sub-periods for SPY

| | $\beta_0$ | $\beta_{\|r\|}$ | $\beta_r$ | $\beta_{RV}$ | $\bar{R}^2$ |
|---|---|---|---|---|---|
| 2000 – 2001 | 0.455** | 0.002 | -0.022 | -0.014* | 0.287 |
| *t stat* | *2.55* | *0.08* | *-1.07* | *-1.67* | |
| 2002 – 2003 | 0.311*** | 0.177*** | -0.066*** | -0.021** | 0.798 |
| | *2.88* | *5.04* | *-6.34* | *-2.27* | |
| 2004 – 2005 | 0.260** | 0.196*** | 0.022 | -0.178 | 0.437 |
| | *2.14* | *4.55* | *0.30* | *-1.59* | |
| 2006 – 2007 | 0.292*** | 0.160*** | -0.020 | -0.037 | 0.583 |
| | *2.95* | *2.67* | *-0.33* | *-1.56* | |
| 2008 – 2009 | 0.388*** | 0.098*** | -0.035* | -0.012*** | 0.726 |
| | *2.78* | *2.71* | *-1.78* | *-3.31* | |
| 2010 – 2011 | 0.413** | 0.079*** | -0.046** | 0.002 | 0.697 |
| | *2.46* | *2.94* | *-2.19* | *0.12* | |
| 2012 – 2014/5 | 0.582*** | 0.037 | -0.04 | -0.387** | 0.336 |
| | *4.13* | *0.98* | *-0.96* | *-2.25* | |

Panel A: Sub-periods for stocks

| | $\beta_0$ | $\beta_{\lvert r \rvert}$ | $\beta_r$ | $\beta_{RV}$ | $\bar{R}^2$ |
|---|---|---|---|---|---|
| 2002 – 2003 | | | | | |
| Average | 0.344*** | 0.106*** | -0.040*** | -0.015*** | 0.550 |
| *t stat* | *5.71* | *5.18* | *-3.35* | *-3.35* | |
| *%(t≤-1.96)* | *0%* | *0%* | *47%* | *42%* | |
| *%(t≥1.96)* | *76%* | *78%* | *1%* | *3%* | |
| 2004 – 2005 | | | | | |
| Average | 0.367*** | 0.079*** | -0.018 | -0.042*** | 0.259 |
| *t stat* | *7.01* | *2.89* | *-0.59* | *-3.42* | |
| *%(t≤-1.96)* | *0%* | *0%* | *6%* | *46%* | |
| *%(t≥1.96)* | *79%* | *35%* | *6%* | *3%* | |
| 2006 – 2007 | | | | | |
| Average | 0.363*** | 0.070** | -0.031 | -0.025*** | 0.379 |
| *t stat* | *5.18* | *2.52* | *-1.08* | *-2.82* | |
| *%(t≤-1.96)* | *0%* | *0%* | *14%* | *42%* | |
| *%(t≥1.96)* | *86%* | *42%* | *0%* | *1%* | |
| 2008 – 2009 | | | | | |
| Average | 0.389*** | 0.066*** | -0.022* | -0.0069*** | 0.632 |
| *t stat* | *3.41* | *3.48* | *-1.72* | *-3.68* | |
| *%(t≤-1.96)* | *0%* | *0%* | *39%* | *71%* | |
| *%(t≥1.96)* | *63%* | *65%* | *1%* | *0%* | |
| 2010 – 2011 | | | | | |
| Average | 0.454*** | 0.070*** | -0.013 | -0.013 | 0.411 |
| *t stat* | *4.27* | *3.14* | *-0.94* | *-1.30* | |
| *%(t≤-1.96)* | *0%* | *0%* | *24%* | *49%* | |
| *%(t≥1.96)* | *83%* | *67%* | *7%* | *1%* | |
| 2012 – 2014 | | | | | |
| Average | 0.428*** | 0.033 | -0.066** | -0.039*** | 0.318 |
| *t stat* | *7.99* | *1.40* | *-2.49* | *-2.67* | |
| *%(t≤-1.96)* | *0%* | *0%* | *36%* | *50%* | |
| *%(t≥1.96)* | *90%* | *10%* | *3%* | *1%* | |