QED

Queen's Economics Department Working Paper No. 1188

The Information Content of Treasury Bond Options Concerning Future Volatility and Price Jumps

Thomas Busch
Danske Bank and CREATES

Bent Jesper Christensen
University of Aarhus and CREATES

Morten Ørregaard Nielsen
Queen's University and CREATES

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

2-2006

The Information Content of Treasury Bond Options Concerning Future Volatility and Price Jumps*

Thomas Busch
University of Aarhus

Bent Jesper Christensen
University of Aarhus

Morten Ørregaard Nielsen†
Cornell University

February 16, 2006

Abstract

We study the relation between realized and implied volatility in the bond market. Realized volatility is constructed from high-frequency (5-minute) returns on 30 year Treasury bond futures. Implied volatility is backed out from prices of associated bond options. Recent nonparametric statistical techniques are used to separate realized volatility into its continuous sample path and jump components, thus enhancing forecasting performance. We generalize the heterogeneous autoregressive (HAR) model to include implied volatility as an additional regressor, and to the separate forecasting of the realized components. We also introduce a new vector HAR (VecHAR) model for the resulting simultaneous system, controlling for possible endogeneity of implied volatility in the forecasting equations. We show that implied volatility is a biased and inefficient forecast in the bond market. However, implied volatility does contain incremental information about future volatility relative to both components of realized volatility, and even subsumes the information content of daily and weekly return based measures. Perhaps surprisingly, the jump component of realized bond return volatility is, to some extent, predictable, and bond options appear to be calibrated to incorporate information about future jumps in Treasury bond prices, and hence interest rates.

Keywords: Bipower variation, bond futures options, HAR, Heterogeneous Autoregressive Model, implied volatility, jumps, realized volatility, VecHAR, volatility forecasting.

JEL classification: C22, C32, G1.

*We are grateful to Tim Bollerslev for providing the realized volatility, bipower variation, and tripower quarticity data used in this study, and to the Center for Analytical Finance (CAF), Aarhus, and the Danish Social Sciences Research Council (SSF) for research support.

†Corresponding author. Please address correspondence to: Morten Ørregaard Nielsen, Department of Economics, Cornell University, 482 Uris hall, Ithaca, NY 14853, USA; phone: +1 607 255 6338; fax: +1 607 255 2818; email: mon2@cornell.edu

1 Introduction

In the theoretical as well as empirical finance literatures, volatility is generally viewed as one of the most important determinants of risky asset prices, such as stock and bond prices, and hence of interest rates. Since the distribution and riskiness of future returns are priced, the forecasting of future volatility is particularly important for asset pricing as well as derivative pricing, hedging, and risk management. In the recent literature, statistical techniques have been developed that allow separating the continuous sample path and jump components of the return volatility process and using them individually and in new combinations to build volatility forecasts. It has been conjectured that using these new econometric methods together with high-frequency data will provide hard-to-beat volatility forecasts. However, an alternative route is to include derivative prices, forecasting future volatility using implied volatility estimates.

In the present paper, we investigate whether implied volatility from options on 30 year Treasury bond (T-bond) futures contains incremental information when assessed against volatility forecasts based on high-frequency (5-minute) current and past 30 year T-bond futures returns, using the recently available statistical techniques to generate efficient measurements of realized volatility and its separate continuous and jump components. Furthermore, we investigate the predictability of the separate volatility components, including the role played by implied volatility in forecasting these.

The construction and analysis of realized volatility (essentially, the summation of squared returns over a specified time interval) from high-frequency return data as a consistent estimate of conditional integrated volatility have received much attention in recent literature, see e.g. French, Schwert & Stambaugh (1987), Schwert (1989), Andersen & Bollerslev (1998*a*), Andersen, Bollerslev, Diebold & Ebens (2001), Andersen, Bollerslev, Diebold & Labys (2001), Barndorff-Nielsen & Shephard (2002*a*), Andersen, Bollerslev & Diebold (2004), and Barndorff-Nielsen & Shephard (2007). In particular, Andersen, Bollerslev, Diebold & Labys (2003) and Andersen, Bollerslev & Meddahi (2004) show that simple reduced form time series models for realized volatility constructed from historical returns outperform commonly used GARCH and related stochastic volatility models in forecasting future volatility. In recent theoretical contributions, Barndorff-Nielsen & Shephard (2004, 2006) derive a fully nonparametric separation of the continuous sample path and jump components of realized volatility. Applying this nonparametric separation technique, Andersen, Bollerslev & Diebold (2005) extend results of Andersen, Bollerslev, Diebold & Labys (2003) and Andersen, Bollerslev & Meddahi (2004) by including both the continuous and jump components of past realized volatility as separate regressors in the forecasting of future realized volatility in the foreign exchange and U.S. stock and bond markets. They show that the continuous sample path and jump components play very different roles in volatility forecasting in all markets. Significant gains in forecasting performance are achieved by splitting the explanatory variables into the separate continuous and jump components, compared to using only total past realized volatility. While the continuous component of past realized volatility is strongly serially correlated, the jump component is found to be distinctly less persistent, and almost not forecastable.

While we focus on the forecasting of volatility in the bond market, many recent studies have stressed the importance of separate treatment of the jump and continuous sample path components in other markets, particularly the stock and foreign exchange markets. This work has considered both the estimation of parametric stochastic volatility models (e.g. Andersen, Benzoni & Lund (2002), Chernov, Gallant, Ghysels & Tauchen (2003), Eraker, Johannes & Polson (2003), Eraker (2004), Ait-Sahalia (2004), and Johannes (2004), who considers interest rates), nonparametric realized volatility modeling (e.g. Barndorff-Nielsen & Shephard (2004, 2006), Huang & Tauchen (2005), and Andersen et al. (2005), who also consider the bond market), empirical option pricing (e.g. Bates (1996) and Bakshi, Cao & Chen (1997)), and information arrival processes (e.g. Andersen & Bollerslev (1998*b*) and Andersen, Bollerslev, Diebold & Vega (2003)). Indeed, in the stochastic volatility and realized volatility literatures, the jump component is found to be far less predictable than the continuous sample path component, clearly indicating separate roles for the two components in volatility forecasting.

In view of the improved realized volatility forecasting performance from return based measures achieved by using high-frequency return data and differentiating the continuous and jump components, the next important question that arises is whether implied volatility from option prices contains incremental information about future realized volatility, relative to return based measures. Recently, this question was addressed by Christensen & Nielsen (2005) for the stock market (the S&P 500 index and associated SPX options) and by Busch, Christensen & Nielsen (2005) for the foreign exchange market (the $/DM exchange rate and associated futures options). The results for the stock market show that implied volatility is a nearly unbiased forecast of future realized volatility, containing incremental forecasting power relative to both past realized volatility and the continuous and jump components of this. Nevertheless, past realized volatility and its continuous component remain significant after introducing implied volatility if variables are measured in logarithms (the transformation leaving them closest to Gaussian), so implied volatility does not appear to be a fully efficient forecast in the stock market. In the foreign exchange market, implied volatility is an informationally efficient but biased forecast of future realized exchange rate volatility. Moreover, in both markets the jump component of future realized volatility is to some extent predictable, and option implied volatility is the dominant forecast of the future jump component.

So far, no study has compared the volatility forecasting performance obtained in the bond market by using return based volatility measures computed from high-frequency data with that obtained using implied volatility extracted from associated bond options. Amin & Morton (1994) use the Heath, Jarrow & Morton (1992) approach to calculate time series of implied interest rate volatilities. They consider Eurodollar futures and futures options, in contrast to our T-bond futures and futures options. Their data are daily, from 1987-1992, while we use high-frequency (5-minute) data for the longer and more recent period 1990-2002. They consider one-day ahead forecasts of implied volatility in the sense that they compare observed option prices with model prices based on current futures prices and one-day lagged implied volatilities, and their results show clear mispricing. They do not consider realized volatility, jumps, or longer term volatility forecasting, whereas we use implied return volatility as a forecast of future realized volatility through expiration of the relevant option one month hence, and compare to forecasting performance using return based forecasts that account for jumps. In another study using the 1987-1992 data, along with daily Eurodollar spot interest rates, Amin & Ng (1997) examine the performance of implied interest rate volatility as a forecast of future interest rate volatility by including implied volatility in GARCH-type equations. They find that option implied interest rate volatility explains much of the variation in future interest rate volatility, but they do not consider bond returns, high-frequency data, or jumps. Bliss & Ronn (1998) use interest rate volatility implied from callable T-bonds to reveal empirical anomalies, but do not consider option data. In contrast to the other studies, we extract option implied bond return volatility and do not explicitly consider interest rate volatility, and our realized measures use high-frequency bond futures returns and the new nonparametric separation of the continuous and jump components of realized volatilty.[1]

We study high-frequency (5-minute) returns on 30 year Treasury bond futures and monthly prices of associated options. We compute alternative volatility measures from the two separate data segments. The return based measures are realized volatility and its continuous and jump components from high-frequency 30 year T-bond futures returns, while the option based measure is implied volatility. The latter is widely perceived as a natural forecast of integrated volatility over the remaining life of the option contract. Since options expire at a monthly frequency, we consider the forecasting of one-month volatility measures. The issue is whether implied volatility retains incremental information on future integrated volatility when assessed against return based measures from the previous month. Here, measures covering the entire previous month may not be the only relevant yardstick, since squared returns nearly one month past may not be as informative about future volatility as squared returns that are only one or a few days old. To accommodate this feature in our econometric framework, we apply the heterogeneous autoregressive (HAR) model proposed by Corsi (2004) for realized volatility analysis and extended by Andersen et al. (2005) to include the separate continuous (C) and jump (J) components of realized volatility as regressors. Specifically, we follow Andersen et al. (2005) and include daily, weekly, and monthly explanatory variables in the HAR forecasting specifications. As a novel feature, we generalize the HAR framework to include implied volatility from option prices as an additional regressor. Furthermore, because of the different time series properties of the continuous and jump components, as documented in Andersen et al. (2005), separate forecasting of these is relevant for pricing and risk management purposes, and we extend the HAR methodology to the forecasting of each of the volatility components C and J individually, again using implied volatility as an additional explanatory variable. The analysis of the implied-realized volatility relation within the HAR framework is unique to the present study, and in particular sets our work apart from that of Christensen & Nielsen (2005) and Busch et al. (2005), who study the stock and foreign markets, respectively, without invoking

[1] Several other related studies have considered the pricing of various fixed income assets and associated contingent claims, including Li, Ritchken & Sankarausbramanian (1995), Das (1999), Goldstein (2000), Cakici & Zhu (2001), Guan, Ting & Warachka (2005), and Gupta & Subrahmanyam (2005), among others, while the pricing of fixed income futures in the presence of jumps is considered in Chiarella & Tô (2003), but none of these consider volatility forecasting.

the HAR methodology.

We show that bond option implied volatility contains incremental information relative to both the continuous and jump components of realized volatility when forecasting subsequently realized return volatility on 30 year T-bond futures. In fact, implied volatility subsumes the information content of the daily and weekly return based measures in most of our specifications. However, implied volatility is not a fully efficient estimate and in particular it should be used in conjunction with the monthly measures of the separate continuous and jump components of realized volatility when forecasting future volatility. This is an important difference from the findings for the foreign exchange market, where implied volatility completely subsumes the information content of realized volatility and its separate components in forecasting future volatility. Our results also show that, although there is clearly volatility information in option prices which is not contained in return data, bond market implied volatility is a biased forecast of future realized volatility. This is an important difference from the findings for the stock market, where implied volatility is a nearly unbiased forecast of future volatility.

As an additional novel contribution, we consider separate forecasting of the continuous and jump components of future realized volatility using the HAR framework. Our results show that implied volatility has predictive power for both components, again largely subsuming the information content of the daily and weekly continuous and jump components of realized volatility. The forecasting of the continuous component is very much like the forecasting of realized volatility itself, whereas jumps are forecast quite differently. In particular, our results show that even the jump component of realized volatility is, to some extent, predictable, with both bond option implied volatility and the past monthly jump component significant in the jump forecasting relation.

We also introduce a structural vector heterogeneous autoregressive (labelled VecHAR) model, which is unique to the present study, for the joint modeling of implied volatility and the separate components of realized volatility. This novel system approach allows handling simultaneity issues which may arise from a number of sources. Since implied volatility is the new variable added in our study, compared to the realized volatility literature, and since it may potentially be measured with error stemming from nonsynchronicity between sampled option prices and corresponding futures prices, bid-ask spreads, model error, etc., we take special care in handling this variable. The structural VecHAR analysis controls for possible endogeneity of implied volatility in the forecasting equations. Furthermore, the simultaneous system approach allows testing interesting cross-equation restrictions. The results from full information maximum likelihood (FIML) estimation of the VecHAR model reinforce our earlier conclusions, in particular that bond option implied volatility should be included in the information set, along with monthly, but not weekly or daily, separate continuous and jump components of past realized volatility in forecasting future volatility.

The remainder of the paper is laid out as follows. In the next section we briefly describe realized volatility and the nonparametric identification of its separate continuous sample path and jump components. In Section 3 we discuss the bond option pricing model. Section 4 describes our data and provides summary statistics. In section 5 the empirical results are presented, and section 6 concludes.

2 Testing for Jumps in Bond Prices

Most contemporary continuous time asset pricing theory is cast in terms of a stochastic volatility model, possibly with an additive jump component. Thus, we assume that the logarithm of the Treasury bond price, $p(t)$, follows the general stochastic volatility jump diffusion model

$$dp(t) = \mu(t)\,dt + \sigma(t)\,dw(t) + \kappa(t)\,dq(t), \quad t \geq 0. \tag{1}$$

The mean $\mu(\cdot)$ is assumed continuous and locally bounded and the instantaneous volatility $\sigma(\cdot) > 0$ càdlàg, both independent of the driving standard Brownian motion $w(\cdot)$. The counting process $q(t)$ is normalized such that $dq(t) = 1$ corresponds to a jump at time t and $dq(t) = 0$ otherwise. Hence, $\kappa(t)$ is the random jump size at time t if $dq(t) = 1$. We write $\lambda(t)$ for the possibly time varying intensity of the arrival process for jumps.[2] Stochastic volatility allows returns in the model (1) to have leptokurtic (unconditional) distributions and exhibit volatility clustering, which is empirically relevant.

An important quantity in this model is the integrated volatility (or integrated variance)

$$\sigma^{2*}(t) = \int_0^t \sigma^2(s)\,ds. \tag{2}$$

In option pricing, this is the relevant volatility measure, see Hull & White (1987). Estimation of integrated volatility is studied e.g. in Andersen & Bollerslev (1998*a*). Integrated volatility is closely related to quadratic variation $[p](t)$, defined for any semimartingale (see Protter (2004)) by

$$[p](t) = p\lim \sum_{j=1}^{M} \left(p(s_j) - p(s_{j-1})\right)^2, \tag{3}$$

where $0 = s_0 < s_1 < ... < s_M = t$ and the limit is taken for $\max_j |s_j - s_{j-1}| \to 0$ as $M \to \infty$. In particular, the quadratic variation process for the model (1) is in wide generality given by

$$[p](t) = \sigma^{2*}(t) + \sum_{j=1}^{q(t)} \kappa^2(t_j), \tag{4}$$

where $0 \leq t_1 < t_2 < ...$ are the jump times, $dq(t_j) = 1$. In (4), quadratic variation is decomposed as integrated volatility plus the sum of squared jumps that have occurred through time t (see e.g. Andersen, Bollerslev, Diebold & Labys (2001, 2003)). Recent studies on the stock market (e.g., Andersen et al. (2002), Chernov et al. (2003), Eraker et al. (2003), Eraker (2004), Ait-Sahalia (2004), and Christensen & Nielsen (2005)) and on interest rates (Johannes (2004)) all find that jumps are empirically important. To investigate the importance of jumps in bond market volatility forecasting, we follow Andersen et al. (2005) and include the jump component explicitly in this market, too. Rather than modeling (1) directly at the risk of adopting erroneous parametric assumptions, we use high-frequency bond return data and

[2] Formally, $\Pr(q(t) - q(t-h) = 0) = 1 - \int_{t-h}^{t} \lambda(s)\,ds + o(h)$, $\Pr(q(t) - q(t-h) = 1) = \int_{t-h}^{t} \lambda(s)\,ds + o(h)$, and $\Pr(q(t) - q(t-h) \geq 2) = o(h)$. This rules out infinite activity Lévy processes, e.g. the normal inverse Gaussian process, with infinitely many jumps in finite time.

invoke a powerful nonparametric approach to identification of the two separate components of the quadratic variation process (4), integrated volatility respectively the sum of squared jumps, following Barndorff-Nielsen & Shephard (2004, 2006) and Andersen et al. (2005).

Assume that T months of intra-monthly bond price observations are available and denote the $M+1$ evenly spaced intra-monthly observations for month t on the logarithm of the bond price by $p_{t,j}$. The one month time interval is used in order to match the sequence of consecutive non-overlapping one month option lives available given the monthly option expiration cycle. The M continuously compounded intra-monthly returns for month t are

$$r_{t,j} = p_{t,j} - p_{t,j-1}, \quad j = 1, ..., M, \quad t = 1, ..., T. \tag{5}$$

Realized volatility for month t is given by the sum of squared intra-monthly returns,

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2, \quad t = 1, ..., T. \tag{6}$$

Some authors refer to the quantity (6) as realized variance and reserve the term realized volatility for the square root of (6), e.g. Barndorff-Nielsen & Shephard (2001, 2002*a*, 2002*b*), but we shall use the more conventional term realized volatility. The nonparametric estimation of the separate continuous sample path and jump components of quadratic variation, following Barndorff-Nielsen & Shephard (2004, 2006), also requires the related bipower and tripower variation measures. The (first lag) realized bipower variation is defined as

$$BV_t = \frac{1}{\mu_1^2} \sum_{j=2}^{M} |r_{t,j}| \, |r_{t,j-1}|, \quad t = 1, ..., T, \tag{7}$$

where $\mu_1 = \sqrt{2/\pi}$. In theory, a higher value of M improves the precision of the estimators, but in practice it also makes them more susceptible to market microstructure effects, such as bid-ask bounces, stale prices, measurement errors, etc., see e.g. Hansen & Lunde (2005) and Barndorff-Nielsen & Shephard (2007). These effects potentially introduce artificial (typically negative) serial correlation in returns. Huang & Tauchen (2005) show that the resulting bias in (7) is mitigated by considering the staggered (second lag) realized bipower variation

$$\widetilde{BV}_t = \frac{1}{\mu_1^2(1 - 2M^{-1})} \sum_{j=3}^{M} |r_{t,j}| \, |r_{t,j-2}|, \quad t = 1, ..., T. \tag{8}$$

The staggered version avoids the sharing of the bond price $p_{t,j-1}$ which by (5) enters the definition of both $r_{t,j}$ and $r_{t,j-1}$ in the non-staggered version (7). A further statistic necessary for construction of the relevant tests is the realized tripower quarticity measure

$$TQ_t = \frac{1}{M} \mu_{4/3}^{-3} \sum_{j=3}^{M} |r_{t,j}|^{4/3} \, |r_{t,j-1}|^{4/3} \, |r_{t,j-2}|^{4/3}, \quad t = 1, ..., T, \tag{9}$$

where $\mu_{4/3} = 2^{2/3}\Gamma(7/6)/\Gamma(1/2)$. The associated staggered realized tripower quarticity is

$$\widetilde{TQ}_t = \frac{1}{M\mu_{4/3}^3(1-4M^{-1})}\sum_{j=5}^{M} |r_{t,j}|^{4/3} |r_{t,j-2}|^{4/3} |r_{t,j-4}|^{4/3}, \quad t = 1, ..., T, \tag{10}$$

again avoiding common bond prices in adjacent returns. As the staggered quantities $\widetilde{BV}_t$ and $\widetilde{TQ}_t$ are asymptotically equivalent to their non-staggered counterparts BV_t and TQ_t, staggered versions of test statistics can be applied for robustness against market microstructure effects without sacrificing asymptotic results.

Using (3), RV_t in (6) is by definition a consistent estimator of the monthly increment to the quadratic variation process (4) as $M \to \infty$ (Andersen & Bollerslev (1998*a*), Andersen, Bollerslev, Diebold & Labys (2001) and Barndorff-Nielsen & Shephard (2002*a*, 2002*b*)). At the same time, BV_t is consistent for month t integrated volatility, defined as $\sigma_t^{2*} = \int_{t-1}^{t} \sigma^2(s)\, ds$, the component of the increment to quadratic variation due to continuous sample path movements in the price process (1), i.e.,

$$BV_t \to_p \sigma_t^{2*}, \quad \text{as } M \to \infty, \tag{11}$$

as shown by Barndorff-Nielsen & Shephard (2004). It follows that the difference between realized volatility and realized bipower variation converges to the sum of squared jumps that have occurred during the course of the month. Of course, in finite samples, $RV_t - BV_t$ may be non-zero due to sampling variation even if there is no jump during month t, so a notion of a "significant jump component" is needed. Following Barndorff-Nielsen & Shephard (2004) and Huang & Tauchen (2005) we apply the test statistic

$$Z_t = \sqrt{M}\frac{(RV_t - BV_t)RV_t^{-1}}{\left(\left(\mu_1^{-4} + 2\mu_1^{-2} - 5\right)\max\{1, TQ_t BV_t^{-2}\}\right)^{1/2}}, \tag{12}$$

which, in the absence of jumps, satisfies

$$Z_t \to_d N(0,1), \quad \text{as } M \to \infty.$$

Thus, large positive values of Z_t indicate the presence of jumps during month t. This statistic was shown by Huang & Tauchen (2005) to have better small sample properties than the alternative asymptotically equivalent statistics in Barndorff-Nielsen & Shephard (2004, 2006). Using the staggered versions (8) and (10) of bipower variation and tripower quarticity yields a staggered version $\widetilde{Z}_t$ of the test, and this is recommended by Huang & Tauchen (2005) and Andersen et al. (2005).

With these definitions, the (significant) jump component of realized volatility is given by

$$J_t = I_{\{Z_t > \Phi_{1-\alpha}\}}(RV_t - BV_t), \quad t = 1, ..., T, \tag{13}$$

where $I_{\{A\}}$ is the indicator function of the event A, $\Phi_{1-\alpha}$ is the $100(1-\alpha)\%$ point of the standard normal distribution, and α is the chosen significance level. Accordingly, the estimator of the continuous component of quadratic variation is the remaining portion of realized volatility,

$$C_t = RV_t - J_t, \quad t = 1, ..., T. \tag{14}$$

This way, the month t continuous component equals realized volatility when there is no significant jump in month t, and it equals realized bipower variation when there is a jump, i.e. $C_t = I_{\{Z_t \leq \Phi_{1-\alpha}\}} RV_t + I_{\{Z_t > \Phi_{1-\alpha}\}} BV_t$. Note that for any standard significance level, both J_t and C_t from (13) and (14) are automatically positive, since $\Phi_{1-\alpha} > 0$ for $\alpha < 1/2$. Since Z_t and BV_t enter the definition (13), there are staggered and non-staggered versions of both the continuous and the jump component. Consistency of the separate components of realized volatility as estimators of the corresponding components of quadratic variation, i.e.

$$C_t \to_p \sigma_t^{*2} \text{ and } J_t \to_p \sum_{j=q(t-1)+1}^{q(t)} \kappa^2 (t_j)$$

may be achieved by letting $\alpha \to 0$ and $M \to \infty$ simultaneously, whether staggered or non-staggered versions are used. Hence, this high-frequency data approach allows for month-by-month separate non-parametric consistent estimation of both components of quadratic variation, i.e. the jump component and the continuous sample-path or integrated volatility component, as well as the quadratic variation process for log-bond prices itself.

3 The Bond Option Pricing Formula

We consider options on the 30 year US Treasury bond (T-bond) futures contract. Let c denote the (European call) option price, K the strike price, τ the time to expiration of the option, F the price of the underlying futures contract expiring Δ periods after the option, σ the bond return volatility, and $B(t, t+\tau+\Delta) = e^{-r(\tau+\Delta)}$ the price at time t of a zero-coupon bond paying \$1 at time $t+\tau+\Delta$, where r is the risk free rate. Following Bates (1996), we use the appropriately modified Black (1976) formula, given by

$$\begin{aligned} c(F, K, \tau, \Delta, r, \sigma) &= e^{-r(\tau+\Delta)}[F\Phi(d) - K\Phi(d - \sigma\sqrt{\tau})], \\ d &= \frac{\ln(F/K) + \frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}}, \end{aligned} \tag{15}$$

where $\Phi(\cdot)$ is the standard normal c.d.f.

Given observations on the option price c and the remaining variables F, K, τ, r, and Δ, an implied volatility ($IV^{1/2}$) estimate[3] can be backed out from the option pricing formula in (15) by numerical inversion of the nonlinear equation

$$c = c(F, K, \tau, r, \Delta, IV^{1/2}) \tag{16}$$

with respect to $IV^{1/2}$. We compute IV estimates by iterating the scheme

$$IV_{n+1} = IV_n + \frac{c - c(F, K, \tau, r, \Delta, IV_n^{1/2})}{\mathcal{V}(F, K, \tau, r, \Delta, IV_n^{1/2})}$$

[3] IV is used in the text as a general abbreviation of option implied volatility. When the explicit transformation is relevant, $IV^{1/2}$ and IV denote standard deviation and variance measures, respectively.

until convergence, with $\mathcal{V}(F,K,\tau,r,\Delta,IV_n^{1/2}) = F\sqrt{\tau}\phi(d)e^{-r(\tau+\Delta)}$ the vega of the option formula (see e.g. Hull (2002)) and $\phi(\cdot)$ the standard normal p.d.f. The last factor in vega, $e^{-r(\tau+\Delta)}$, does not enter, e.g., in the vega of the standard Black & Scholes (1973) formula, but enters here since the futures contract can be regarded as an asset paying a continuous dividend yield equal to the risk free rate r. The algorithm is stopped when $\left|c - c(F,K,\tau,r,\Delta,IV_n^{1/2})\right| < 10^{-7}$.

Note that in (15) the term to option expiration, τ, enters d, whereas the term to futures expiration, $\tau+\Delta$, is used for discounting both the futures price and the strike price. In the unmodified Black (1976) formula, τ enters in both places, thus exaggerating the option price by a proportional factor $e^{r\Delta}$. This would lead to a systematic upward bias in implied volatilities when options and underlying futures do not expire simultaneously. In our data, Δ ranges between $\frac{5}{365}$ and $\frac{78}{365}$. Thus, we use the modified formula (15) throughout when calculating implied volatility.

4 Data and Descriptive Statistics

We consider daily open auction closing prices for the 30 year Treasury bond (T-bond) futures options traded at the Chicago Board of Trade (CBOT) over the period October 1, 1990, to December 31, 2002. The data are obtained from the Commodity Research Bureau. The delivery dates of the underlying futures contracts follow the quarterly cycle March, June, September, and December. The particular T-bond serving as underlying asset for the futures contract is not uniquely identified by the contract specifications. It is simply required that the T-bond is not callable for at least 15 years from the first day of the contract month (the delivery month of the underlying futures contract), or, for a noncallable T-bond, has a maturity of at least 15 years from the first day of the contract month. For a detailed description of the 30 year T-bond futures, see e.g. Hull (2002).

In October 1990 serial futures options with monthly expiration cycle were introduced. Thus, some of the options expire in the two months between the quarterly delivery dates of the futures contracts. The options are American and expire on the last Friday followed by at least two business days in the month prior to the contract month. Upon exercise the holder of the option contract is at the end of the trading day provided with a position at the strike in the futures contract. Delivery of the futures contract may be made on any day in the contract month by the party with the short position. The last trading day of the futures contract is the seventh-to-last business day of the month. Since the futures price in our case is an increasing function of time to maturity, it is optimal for the party with the short position to deliver as early as possible (Hull (2002, p. 60)). Consequently, we choose the delivery date for the underlying futures contract as the first business day of the contract month. The delivery lag in calender days upon terminal exercise of the futures contract, Δ, varies between $\frac{5}{365}$ and $\frac{78}{365}$ in our data. The US Eurodollar deposit one-month middle rate (downloaded from Datastream) is used for the risk-free rate.

For the implied volatility (IV) estimates we use at-the-money (ATM) calls with one month to expiration. The prices are recorded on the first business day after the last trading day of the preceding option contract. In total, a sample of 146 annualized monthly IV observations from ATM calls are available. Hence, although the underlying futures contracts expire at a quarterly frequency, the IV estimates are based on option contracts covering non-overlapping monthly time intervals. Furthermore, as suggested by French (1984), the option pricing formula in (15) is extended such that trading days are used for volatilities (τ) and calender days for interest rates ($\tau + \Delta$).

When comparing IV and realized volatility (RV) estimates it is important to ensure that the RV measure corresponds to the underlying asset of the option contract used to compute IV. Rather than calculating RV from returns on T-bonds, which are not uniquely associated with the futures contracts, we therefore consider RV from returns to the futures contract itself, i.e., precisely the underlying asset for the futures options.

The trading hours for the open auction at the CBOT are 7:20 am to 2:00 pm, central time, and our price observations start at 7:25 am. For RV and its separate components we use the same data as Andersen, Bollerslev, Diebold & Vega (2004) and Andersen et al. (2005), which are based on linearly interpolated five-minute observations (following Müller, Dacorogna, Olsen, Pictet, Schwarz & Morgenegg (1990) and Dacorogna, Müller, Nagler, Olsen & Pictet (1993), among others) on the 30 year T-bond futures prices, thus providing us with a total of 79 high-frequency returns per day ($r_{t,j}$ from (5), $M = 79$, $T = 146$). The volatility measures are annualized and constructed on a monthly basis to cover exactly the same periods as the IV estimates. Our time index refers to the calendar month where implied volatility is sampled. Thus, IV_t can be regarded as a forecast of RV_{t+1}, since implied volatility is sampled at the beginning of the time interval covered by realized volatility for period $t + 1$. For example, if t and $t + 1$ correspond to May and June, say, it means that IV_t is sampled on the first business day after the last Friday in May which is followed by at least two business days in this month, and RV_{t+1} is calculated from squared returns covering the period from the sampling of IV_t and until the last Friday followed by at least two business days in June. As suggested by Andersen et al. (2005) a significance level of $\alpha = 0.1\%$ is used to detect jumps, thus providing the series for the jump component J from (13) and continuous component C from (14) of realized volatility RV.

Following the standard in the literature, we consider three different transformations of the volatility measures x (where $x = RV, C, J, IV$): Logarithmically transformed variances, $\ln x$, standard deviations, $x^{1/2}$, and raw variances, x. Note the slight abuse of terminology – there is no correction for sample average. To avoid taking logarithm of zero, the jump component J_t, which equals zero in the case of no significant jump during the month, is for the logarithmic transformation replaced by $J_t + 1$. There are 116 (138) out of 146 months with significant jumps for the non-staggered (staggered) data, corresponding to significant jumps in 79.5% (94.5%) of the months. This indicates a non-negligible difference between staggered and non-staggered data. Thus, results are presented for both versions in the following.

Table 1 about here

Summary statistics for the four different annualized volatility measures are presented in Table 1. Panel A presents statistics for the logarithmic transform, Panel B for the standard deviation form, and

Panel C for the variance form of the volatility measures. For the logarithmic transformation, Panel A, the null of normality cannot be rejected at the 5% significance level by the Jarque & Bera (1980) (JB) test for realized T-bond futures return volatility and its continuous component. This complements the results in Andersen, Bollerslev, Diebold & Labys (2001) for realized exchange rate return volatility and in Andersen, Bollerslev, Diebold & Ebens (2001) for realized stock index return volatility, which were also found to be close to log-normal. The results are also consistent with the skewness and kurtosis measures reported for daily log-realized volatility by Andersen et al. (2005).

Our results, where implied volatility from option prices has been added to the data, reveal that this, too, is nearly log-normal. The JB test for IV in Panel A takes the value 4.56 and is statistically insignificant at the 5% level. Indeed, implied volatility has the lowest excess kurtosis in the panel, although skewness is less for realized volatility and its continuous component. The logarithmic transformation of the jump component is far from Gaussian, so Panel A of Table 1 does provide evidence of the statistical difference between the continuous and jump components of realized volatility. Under the other two transformations (Panels B and C), all volatility measures appear non-Gaussian. There appears to be no clear distributional difference between staggered and non-staggered measures of the continuous and jump components, although the staggered versions have somewhat higher JB statistics. Finally, while implied volatility has higher mean than realized volatility under all three transformations, it has lower standard deviation in Panels A and B and nearly identical standard deviation in Panel C, consistent with the notion of implied volatility as a forecast (conditional expectation) of future realized volatility.

Figures 1-3 about here

Time series plots of the four volatility measures are exhibited in Figures 1-3. Each of the three volatility transformations is shown in a separate figure[4]. In each figure, measures based on non-staggered data appear in Panel A, and their staggered counterparts in Panel B. Consistent with the difference in means of the volatility measures in Table 1, the continuous component of realized volatility is below realized volatility itself in all figures. Interestingly, the two are clearly distinguishable whether using staggered or non-staggered data. Implied volatility is above realized volatility, possibly reflecting that the replication strategy of the option payoff is more expensive than in perfect markets, due to market restrictions such as liquidity constraints, taxes, discrete trading, etc., thereby increasing the price of the option and implicitly also increasing implied volatility relative to realized volatility, consistent with the difference in means in Table 1. Comparing these results for the bond market with those for the stock and foreign exchange markets and associated options in Christensen & Nielsen (2005) and Busch et al. (2005), the difference between RV and C clearly sets the bond market apart from the stock and foreign exchange markets, whereas the difference between IV and the other volatility measures are similar across the three option markets.

There are fewer significant jumps when using staggered data (Panel B) compared to the corresponding

[4]In Figure 1, the depicted jump component series are transformed as $100 \times \ln(J_t + 1) - 6$ in order to fit on the graph. No such transformation is needed in Figures 2 and 3 or in the regressions below.

non-staggered data (Panel A). Hence, although the distributional properties of the two jump components appear similar from Table 1, the two measures still exhibit important differences which could have a significant impact in a forecasting context. There is clear evidence of lower autocorrelation and in general different behavior of the jump series compared to the other series. From Figures 1-3 and the descriptive statistics in Table 1, the importance of analyzing the continuous and jump components separately in the 30 year T-bond market is evident, and clearly neither of the two jump components are negligible.

5 Econometric Models and Empirical Results

In this section we study the relation between realized 30 year T-bond futures return volatility together with its disentangled components and implied volatility from the associated option contract. We apply both standard univariate regression models and heterogeneous autoregressive (HAR) models, and we introduce new multivariate extensions of the latter, which turn out to prove useful in our context with implied volatility as well as separate continuous and jump measures of realized volatility.

Each of the tables in the empirical analysis consists of three panels. Results for the logarithmically transformed measures are presented in Panel A, while Panel B contains results for the standard deviation variables, and Panel C for the variance form. Typeface in *italicss* denotes results where the continuous and jump components are computed using staggered measures of realized bipower variation (8) and realized tripower quarticity (10).

5.1 Forecasting Realized Bond Futures Return Volatility

We consider regression of realized bond futures return volatility, RV, on bond futures option implied volatility, IV, as well as lagged RV or its continuous and jump components. The general form of the one-month ahead forecasting equation is

$$RV_{t+1} = \alpha + \beta IV_t + \gamma x_t + \varepsilon_{t+1}, \quad t = 1, 2, 3, \ldots, \tag{17}$$

where α denotes the intercept, β is the slope parameter for IV, and ε_{t+1} is the forecasting error. One of the variables RV, C, J, or the vector (C, J) is contained in x_t, where γ is the associated coefficient vector. For some of the regressions presented, $\beta = 0$ or $\gamma = 0$ is imposed. In Table 2 we report coefficient estimates (estimated standard errors in parentheses) together with adjusted R^2, and the Breusch (1978) and Godfrey (1978) (henceforth BG) test statistic for residual autocorrelation up to lag 12 (one year), which is used instead of the standard Durbin-Watson statistic due to the presence of lagged endogenous variables in several of the regressions. Under the null hypothesis of absence of serial dependence in the residuals, the BG statistic is asymptotically χ^2 with 12 degrees of freedom. Likelihood ratio (LR) test statistics are reported in the final two columns of the table, where LR_1 denotes the test of the null hypothesis of a unit coefficient on implied volatility when this is included as a regressor. Hence, this is the test of the basic unbiasedness hypothesis, $\beta = 1$. The test in the column denoted LR_2 is for the more restrictive joint hypothesis $\beta = 1$, $\gamma = 0$, i.e. both unbiasedness and efficiency of the implied volatility forecast, against the unrestricted null. The asymptotic distributions of LR_1 and LR_2 are χ^2 on 1 respectively $1+\dim(x_t)$ degrees of freedom under the relevant null hypotheses. Panel A of Table 2 shows results for log-volatilities, using $\ln(J_t + 1)$ for the jump measure in x_t. Panel B depicts results for volatilities in standard deviation form, and Panel C for raw variances.[5]

Table 2 about here

The first order autocorrelation coefficient in the regression of the logarithmic transformation of realized volatility in the first row of Panel A in Table 2 is .59 with an associated t-statistic of 8.8. In the following, alternative volatility predictors are compared to this natural benchmark. We first address the importance of separating realized volatility into its continuous and jump components when forecasting future volatility. This is done for non-staggered and staggered volatility measures in rows two and three, respectively. It is immediately clear that the two components of realized volatility play very different roles in forecasting realized volatility of bond futures returns. The coefficient on the continuous component is .57 and clearly significant with a t-statistic of 9.1, and close to that of the realized volatility regression. In contrast, the coefficient on the jump component is negative and insignificant with a t-statistic of -.12 (-1.03) for the non-staggered (staggered) case, suggesting very little impact from the jump component on future realized volatility. This accords well with the consensus in the empirical finance literature that jumps are very hard to predict. Our results complement this notion by showing that jumps in the bond market are of little importance in predicting future realized volatility. Thus, C and J play very different roles in a forecasting context showing that it is not appropriate to use total lagged realized volatility as in the regression in the first row, or imposing equal coefficients on C and J. Indeed, separating realized volatility into its continuous and jump components increases predictability of future volatility with adjusted R^2 increasing from 34.7% in the first line to 36.1% in line two and a full 39.1% in the third line using staggered data to appropriately separate the continuous and jump components.

As a novel contribution we introduce bond option implied volatility into this forecasting regression framework. This allows assessing the incremental information from bond option prices relative to the nonparametric volatility measures extracted from high-frequency bond futures returns. In particular, the enhanced forecasting performance achieved by separating the continuous and jump components of the returns based measures implies that IV is subjected to a more stringent test in the following.

The fourth equation in Table 2 examines the information content of implied volatility in forecasting future realized volatility. The regression coefficient on implied volatility is .73 and strongly significant with a t-statistic of 9.24, higher than those of past realized volatility or its separate continuous and jump components in rows one and two. Adjusted R^2, at 37%, is higher than that in the regression on past realized volatility in the first row and in between adjusted R^2 coefficients in the staggered and

[5] In the log-volatility regressions, variables in (17) and similar equations are implicitly understood to be in log-form, i.e., for brevity we do not rewrite the equation for the logarithmic (and standard deviation) cases.

non-staggered versions of the regressions on C and J. The LR_1 test of the unbiasedness hypothesis $\beta = 1$ rejects at the 1% level in the implied volatility regression and the BG test shows some evidence of residual correlation.

The incremental information in implied volatility vis-à-vis realized volatility and its separate components is next assessed by including these simultaneously as regressors. The results in the fifth line of Table 2 shows that when both realized and implied volatility are included in the regression, each of them has incremental information and remains significant in the regression. The last two lines of the panel show the results when including both implied volatility and the separate continuous and jump components of realized volatility. Whether using staggered or non-staggered data, the findings are that implied volatility gets a coefficient of about .5 and is strongly significant, the BG test shows no sign of misspecification, and adjusted R^2 is higher than in the previous regressions at 45% or higher. The continuous component remains significant, too, as does the staggered version of the jump component (last line of the panel). The latter result indicates some difference in information content between staggered and non-staggered versions of the jump measure.

Overall, our results so far suggest that implied volatility from option prices is about as important and useful for forecasting future realized volatility as the separate continuous and jump components of realized volatility itself. Implied volatility contains incremental information relative to realized volatility and its separate components. Thus, implied volatility passes a rather stringent test since we also find that forecasting performance is enhanced by separating the C and J components of RV. From univariate tests, implied volatility appears both biased and inefficient as a forecast of future realized volatility, although there is some indication that implied volatility subsumes the information content of the lagged jump component. Our preferred specifications are those including both implied volatility and the separate realized volatility components, where the BG tests show no signs of misspecification and adjusted R^2 is the highest. Focusing, e.g., on the regression using staggered versions of the volatility measures (last line), the t-test of the unbiasedness hypothesis $\beta = 1$ takes the value -4.61, and the likelihood ratio test of the joint unbiasedness and efficiency hypothesis $\beta = 1, \gamma = 0$ takes the value 38.2, strongly significant in the asymptotic χ^2_3 distribution. Thus, both implied volatility and the realized volatility components are important for forecasting future realized volatility, and should be used in conjunction for this purpose based on our results.

For comparison with the results in Andersen et al. (2005), the regression specifications in (17) for the logarithmically transformed variables are repeated for the standard deviation form in Panel B of Table 2 and for the variance form in Panel C. Qualitatively, the results in Panels B and C are similar to those obtained in Panel A. First, forecasting power improves when realized volatility is separated into its continuous and jump components using the new nonparametric methodology, thereby allowing for different coefficients in the forecasting regression, where the continuous component is far more important than the jump component. Hence, the continuous and jump components should not be combined in the form of raw realized volatility when predicting future volatility. Second, implied volatility is an important predictor of future realized volatility throughout containing incremental information relative to realized volatility and its separate components, although our findings show that implied volatility is in fact both biased and inefficient. The conclusion from the results in Table 2 is therefore that both implied volatility and the continuous and jump components should be included in this type of predictive regressions.

Our implied volatility measure is backed out from the modified Black (1976) bond futures option pricing formula (15), as is standard among practitioners and in the empirical literature. The formula is consistent with a time-varying volatility process for a continuous sample path bond futures price process but does not incorporate jumps in prices, and hence the jump component may not be explained by implied volatility. Nevertheless, our results show that implied volatility can in fact predict bond futures return volatility, which does include a jump component. Further results below on the direct forecasting of the jump component of future volatility support this interpretation. This suggests that option prices, at least to some extent, are calibrated to incorporate the effect of jumps, thus reducing the empirical need to invoke a more general option pricing formula allowing explicitly for jumps in bond prices. Such an approach would entail estimating additional parameters, including prices of volatility and jump risk. This would be a considerable complication, but would potentially reveal that even more information is contained in option prices. While leaving the alternative, more complicated analysis for future research, we note that our approach yields a conservative estimate of the information content on future bond return volatility contained in option prices.

5.2 Heterogeneous Autoregressive (HAR) Model

In forecasting future realized volatility, the OLS regressions in Table 2 above use measures of past realized volatility and the components of this, where squared returns are assigned equal weight throughout the month. This way squared returns nearly one month ago are given the same weight as squared returns one or two days ago. This may not be relevant if squared returns observed close to one month ago are nearly obsolete, and recent squared returns much more informative about future volatility. Instead, it may be more relevant to place higher weight on recent squared returns than on squared returns that are more distant in the past, and one way to do so in a parsimonious fashion is to apply the heterogeneous autoregressive (HAR) model proposed by Corsi (2004). When applying this only to realized volatility itself, we follow Corsi (2004) and denote the model HAR-RV. We also follow Andersen et al. (2005) and separate the realized volatility regressors into their continuous (C) and jump (J) components in what they denote the HAR-RV-CJ model.

In our analysis we modify and generalize the HAR-RV-CJ model in four directions. First, we include implied volatility (IV) as an additional regressor and abbreviate the model HAR-RV-CJIV. Secondly, in the following subsection HAR regressions are used to predict each of the separate continuous and jump components rather than total realized volatility, and we denote the corresponding models HAR-C-CJIV respectively HAR-J-CJIV. Thirdly, in HAR type estimations data are measured on different time scales, such as daily, weekly, and monthly. Both Corsi (2004) and Andersen et al. (2005) normalize the time series to the daily frequency. However, in line with our OLS analysis and without loss of generality, we annualize the data instead of normalizing to the daily frequency. Fourth, Andersen et al. (2005) estimate HAR models with the regressand sampled at overlapping time intervals, e.g. monthly RV is sampled at the daily frequency, causing serial correlation in the error term. This does not necessarily invalidate the parameter estimates, although an adjustment must be made to obtain correct corresponding standard errors. However, options expire on a monthly basis and the analysis in Christensen & Prabhala (1998) suggests that use of overlapping data may lead to erroneous inferences in a setting with both implied volatility from option prices and realized volatility. Hence, we sample monthly, weekly, and daily measures of realized volatility and its components at the monthly frequency, with the weekly (daily) measures covering the last five trading days (last day) of the corresponding monthly measures.

For the HAR-RV-CJIV model, which is the HAR equivalent of the least squares regression equation in (17), we denote the h-day realized variation, normalized to the annual frequency by

$$RV_{t,t+h} = 252h^{-1}\left(RV_{t+1} + RV_{t+2} + \ldots + RV_{t+h}\right),$$

where $h = 1, 2, \ldots$. Here, and throughout the rest of the paper, we use t to indicate trading days rather than months for all realized volatility measures. For instance, RV_{t+1} now denotes the daily realized volatility for day $t+1$. Thus, $RV_{t,t+h}$ may denote a daily ($h = 1$), weekly ($h = 5$), or monthly ($h = 22$) realized volatility measure. Similar measures may be computed for the continuous ($C_{t,t+h}$) and jump ($J_{t,t+h}$) components of realized volatility. Note that $RV_{t,t+1} = 252RV_{t+1}$ and under stationarity $E\left[RV_{t,t+h}\right] = 252E\left[RV_{t+1}\right]$ for all h.

The monthly frequency HAR-RV-CJIV model is

$$RV_{t,t+22} = \alpha + \gamma_m x_{t-22,t} + \gamma_w x_{t-5,t} + \gamma_d x_t + \beta IV_t + \varepsilon_{t,t+22}, \quad t = 22, 44, 66, \ldots, \tag{18}$$

where $\varepsilon_{t,t+22}$ is the monthly forecasting error, $x_{t-22,t}$ is either $RV_{t-22,t}$, $C_{t-22,t}$, $J_{t-22,t}$, or the vector $(C_{t-22,t}, J_{t-22,t})$, and similarly for the weekly and daily variables $x_{t-5,t}$ respectively x_t. When a variable is not included in the specific regression, $\beta = 0$ or $\gamma_m = \gamma_w = \gamma_d = 0$ is imposed. Note that $x_{t-22,t}$ corresponds to the one-month lagged realized volatility measures included in the earlier regressions, whereas $x_{t-5,t}$ and x_t allow extracting information about future volatility from the more recent history of past squared returns.

Table 3 about here

Table 3 shows the results for the HAR-RV-CJIV model in (18). The format is the same as in Table 2, so the first regression depicted in Panel A is for the logarithmically transformed realized volatility. The combined impact from realized volatility measures on future realized volatility is .44 + .08 + .09 = .61, strikingly close to .59, i.e. the corresponding estimate from the least squares regression in Table 2. The t-statistics for the monthly, weekly, and daily RV parameter estimates are 4.16, .80, and 1.66, respectively, indicating that the weekly and daily variables contain only minor information concerning future monthly volatility. This is supported by the fact that the adjusted R^2 only increases marginally from 34.7% to 36.5%, comparing the first lines of Tables 2 and 3, a result that generally holds for all the regressions in the two tables.

Turning to the results in rows two and three, where the continuous and jump components of realized volatility enter separately in the regression, the conclusions for the continuous component variable are similar to the those for RV in the first row of the table. As in Table 2 the monthly jump component is insignificant for the non-staggered regression and significant for the staggered measure. The same is true for the daily jump component, whereas both weekly jump components are insignificant.

Next, implied volatility is added to the information set at time t in the HAR regressions.[6] When RV is included together with IV, fifth row, all the realized volatility coefficients turn insignificant and the coefficient on implied volatility only decreases from .44 to .43, providing clear evidence for the relevance of implied volatility in forecasting future volatility. The last two rows of Panel A show the results when including the separate continuous and jump components of realized volatility together with implied volatility, i.e. the new HAR-RV-CJIV model. The most notable difference from the results in the fifth row of the table is that the coefficient estimate for the monthly continuous component is now significant with a t-statistic of 2.05 (2.70) for the non-staggered (staggered) measures. This is consistent with the finding from Table 2 that forecasting performance increases when the continuous and jump components are separated.

Results of the HAR-RV-CJIV analysis for the square-root and raw variance versions of the volatility measures are provided in Panels B and C of Table 3. The main difference compared to the results for the logarithmically transformed measures in Panel A is that the continuous and jump components of realized volatility decline in significance, whereas implied volatility becomes more significant. The finding so far is that implied volatility as a forecast of future volatility contains incremental information relative to return based measures, even when allowing more weight to be placed on more recent squared returns and when separating the continuous and jump components of the realized volatility regressor. We next use the HAR methodology to assess whether the conclusions extend to separate forecasting of the continuous and jump components of future realized volatility.

5.3 Forecasting the Continuous and Jump Components

We now split realized bond return volatility, $RV_{t,t+22}$, into its separate continuous sample path and jump components, $C_{t,t+22}$ and $J_{t,t+22}$, and examine how these are forecast by the variables in the information set at time t. This is particularly interesting since Andersen et al. (2005) have shown that the time series properties of the continuous and jump components are very different, consistent also with our findings in Section 4. This suggests that the two components should be forecast in different ways. In fact, Christensen & Nielsen (2005) and Busch et al. (2005) show that in the stock and foreign exchange markets, the two components are forecast in quite different ways in ordinary regressions of the type (17). In the following we extend the HAR methodology to the forecasting of the separate continuous and jump

[6] The regression in row four of Table 3 duplicates that in row four of Table 2, and is included for clarity.

components of future volatility. Although Andersen et al. (2005) did not consider the forecasting of the separate components, our analysis below shows that the HAR methodology is well suited also for this purpose. In addition, since our generalized specification includes implied volatility as well, we are able to assess the incremental information in option prices on future continuous and jump components of volatility.

Our HAR-C-CJIV model for forecasting the continuous component of future volatility is given by

$$C_{t,t+22} = \alpha + \gamma_m x_{t-22,t} + \gamma_w x_{t-5,t} + \gamma_d x_t + \beta IV_t + \varepsilon_{t,t+22}, \quad t = 22, 44, 66, \ldots, \tag{19}$$

where $C_{t,t+22}$ replaces $RV_{t,t+22}$ on the left-hand side of the regression compared to (18). Table 4 shows the results from estimation of this model. The format is the same as in Table 3, except that C and J are always considered separately rather than combined in the form of RV.

Table 4 about here

The results in Table 4 are similar to the corresponding results in Table 3. A minor difference is the indication that the daily continuous component, C_t, appears to contain some predictability of the future monthly continuous component, $C_{t,t+22}$. For example, the t-statistic for C_t in the third row of Panel A is 2.00. This relationship, however, is lost when implied volatility, IV_t, is included in the information set. The BG tests show no or only very modest signs of misspecification, except when IV_t is the sole regressor. Implied volatility gets higher coefficients and t-statistics than the lagged continuous and jump components of realized volatility, and adjusted R^2 is highest when implied volatility is included along these. However, implied volatility does not completely subsume the information content of the realized volatility measures. Thus, the monthly continuous component remains significant under all transformations (all three panels) in specifications including all explanatory variables, and so does the staggered version of the montly jump component.

We next consider the predictability of the jump component of realized volatility. The relevant HAR-J-CJIV model takes the form

$$J_{t,t+22} = \alpha + \gamma_m x_{t-22,t} + \gamma_w x_{t-5,t} + \gamma_d x_t + \beta IV_t + \varepsilon_{t,t+22}, \quad t = 22, 44, 66, \ldots. \tag{20}$$

Table 5 reports results from regression of the future jump component, $J_{t,t+22}$, on the same variables in the information set at time t as in the previous table.

Table 5 about here

Across all three panels (transformations) the forecasting power as measured by adjusted R^2 is highest when using staggered data, hence verifying the value of this approach. Particularly the lagged monthly jump component is significant in all specifications when using staggered data and in Panels B and C also using non-staggered data. Thus, jumps are to some extent predictable from their own past. When implied volatility is entered, this turns out to have even stronger predictive power for future jumps. It gets higher t-statistics than the lagged jump components and is significant, whether using staggered or non-staggered data. The BG tests show only few signs of misspecification, except again when implied volatility is the only regressor. Finally, our results suggest that, although the monthly and weekly continuous components are insignificant throughtout the table, there is some explanatory power for future jumps in the daily continuous measure, and this gets a negative coefficient whether using staggered or non-staggered data.

Comparing across Tables 3, 4, and 5, it is clear that the results are most similar in Tables 3 and 4 and quite different in Table 5. Clearly, realized volatility and the continuous component of this behave similarly, also in this forecasting context, and our results show that implied volatility from option prices are important in forecasting both. The difference in results when moving to Table 5 again reinforces that the continuous and jump components should be treated separately. Even when doing so, we find both that implied volatility retains incremental information, thus implicating that option prices incorporate jump information, and that jumps are predictable from variables in the information set, both of which are interesting results.

5.4 The Vector Heterogeneous Autoregressive Model

Our results so far show that realized volatility should be separated into its continuous and jump components for forecasting purposes, and that implied volatility from option prices has incremental information for the forecasting of both. We now introduce a simultaneous system approach for the joint analysis of implied volatility and the separate continuous and jump components of realized volatility. The reason a simultaneous system approach is needed is firstly that results up to this point have been obtained in different regression equations which are not independent, so the relevant joint hypotheses actually involve cross-equation restrictions. Secondly, in our setting implied volatility may be contaminated with measurement error, stemming from non-syncronous option and futures prices, misspecification of the option pricing formula, etc. Even a simple errors-in-variables problem in implied volatility of this kind generates correlation between the implied volatility regressor and the error terms in the forecasting equations for the continuous and jump components, and thus a particular case of an endogeneity problem. Our simultaneous system approach provides an efficient method for handling endogeneity more generally.

Thus, we consider the structural vector heterogeneous autoregressive (VecHAR) system

$$\begin{aligned}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ B_{31m} & B_{32m} & 1 \end{pmatrix}\begin{pmatrix} C_{t,t+22} \\ J_{t,t+22} \\ IV_{t+22} \end{pmatrix} &= \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} + \begin{pmatrix} A_{11m} & A_{12m} & \beta_1 \\ A_{21m} & A_{22m} & \beta_2 \\ 0 & 0 & A_{33} \end{pmatrix}\begin{pmatrix} C_{t-22,t} \\ J_{t-22,t} \\ IV_t \end{pmatrix} \\ &+ \begin{pmatrix} A_{11w} & A_{12w} & 0 & 0 \\ A_{21w} & A_{22w} & 0 & 0 \\ 0 & 0 & A_{33w} & A_{34w} \end{pmatrix}\begin{pmatrix} C_{t-5,t} \\ J_{t-5,t} \\ C_{t+17,t+22} \\ J_{t+17,t+22} \end{pmatrix} \\ &+ \begin{pmatrix} A_{11d} & A_{12d} & 0 & 0 \\ A_{21d} & A_{22d} & 0 & 0 \\ 0 & 0 & A_{33d} & A_{34d} \end{pmatrix}\begin{pmatrix} C_t \\ J_t \\ C_{t+22} \\ J_{t+22} \end{pmatrix} + \begin{pmatrix} \varepsilon^1_{t,t+22} \\ \varepsilon^2_{t,t+22} \\ \varepsilon^3_{t,t+22} \end{pmatrix}.\end{aligned} \quad (21)$$

The first two equations in our structural VecHAR system comprise the forecasting equations for the separate components of realized volatility and the third equation endogenizes implied volatility. In the representation (21) of our VecHAR model, the three coefficient matrices on the right hand side are associated with monthly, weekly, and daily volatility measures, respectively. There are two sources of simultaneity in the structural VecHAR system. Firstly, the off-diagonal terms B_{31m} and B_{32m} in the leading coefficient matrix allow that IV_{t+22} depends on both $C_{t,t+22}$ and $J_{t,t+22}$. That is, option prices at the end of the month may reflect return movement over the course of the month, although via the HAR type specification of the third equation, more recent returns may receive higher weight. In addition, our specification accommodates dependence on lagged implied volatility. Secondly, the system errors may be contemporaneously correlated.

Table 6 about here

In Table 6 we present the results of Gaussian full information maximum likelihood (FIML) estimation of the VecHAR system. The results show only small deviations across the three transformations in Panels A-C. Compared to the univariate HAR results in Tables 3-5, the BG tests in most cases show even less sign of misspecification. The general picture for the continuous and jump components is clear. First, weekly and daily measures of the continuous and jump components do not improve forecasting power with respect to their future monthly counterparts. Secondly, only the staggered version of the monthly jump component matters for predicting the continuous component of future realized volatility. Third, the lagged monthly continuous component is insignificant in the jump equation whether using staggered or non-staggered data. Fourth, implied volatility is strongly significant in the forecasting equations for both the continuous and jump components of future volatility, showing that option prices contain incremental information beyond that in the high-frequency volatility measures.

Turning to the implied volatility forecasting equations in the last two rows of each panel, we find that lagged implied volatility is insignificant throughout. On the other hand, various monthly, weekly, and daily realized volatility components are significant depending on the precise transformation and whether staggered or non-staggered data are used. These results are natural and in accord well with normal trading behavior, i.e. market participants incorporate the latest information in setting option prices.

Table 7 about here

Table 7 shows results of likelihood ratio (LR) tests of various hypotheses of interest in the VecHAR model. Overall, earlier conclusions are confirmed. Firstly, implied volatility subsumes the information in the weekly and daily measures of the continuous and jump components. Specifically, the hypotheses $\mathrm{H}_2 : A_{11w} = 0, A_{12w} = 0$ and $\mathrm{H}_3 : A_{11d} = 0, A_{12d} = 0$ in (21) are the informational efficiency hypotheses in the continuous component forecasting equation with respect to the weekly and daily realized volatility components, respectively. From the VecHAR model we get p-values for the weekly measures (H_2) between 41 and 92 percent. For the daily measures (H_3) the p-values range from 19 to 30 percent. Clearly, there is no relevant information about the future continuous component of volatility in the weekly and daily measures of past realized volatility components once implied volatility is included in the specification. However, the evidence on whether implied volatility in addition subsumes the information content of the monthly realized volatility components is slightly mixed, and seems to depend on whether staggered or non-staggered data are used for the test. Thus, the test of $\mathrm{H}_1 : A_{11m} = 0, A_{12m} = 0$ is significant at the 1% level when using staggered data, but only at the 5% level when using non-staggered data. Based on our findings in the previous sections and the recommendation of Andersen et al. (2005), we tend to prefer the results using staggered data. In this case, these suggest that implied volatility is informative about the future continuous component of volatility, and subsumes some, but not all, of the information content of the past realized volatility components, i.e., implied volatility is strictly not fully informationally efficient. It is also somewhat biased, as the test of the unbiasedness hypothesis $\mathrm{H}_4 : \beta_1 = 1$ rejects at conventional levels throughout. In H_5-H_7, the unbiasedness hypothesis H_4 is tested jointly with the efficiency hypotheses H_1-H_3 and is rejected for each of the monthly, weekly, and daily measures. Using the matrix notation

$$\bar{\mathbf{A}}_k = \begin{pmatrix} A_{11k} & A_{12k} \\ A_{21k} & A_{22k} \end{pmatrix}, \quad k = m, w, d,$$

we examine in $\mathrm{H}_8 : \bar{\mathbf{A}}_m = 0, \bar{\mathbf{A}}_w = 0, \bar{\mathbf{A}}_d = 0$ the possibility that all the coefficients in both the continuous and jump equations are jointly insignificant, which is a cross-equation retriction and hence requires the system approach. Here, the informational efficiency hypothesis is tested simultaneously for both the continuous and jump components, and our results from the VecHAR model reject this hypothesis. Clearly, rejection of H_8 is due to the $\bar{\mathbf{A}}_m$ term, i.e., informational inefficiency with respect to the monthly realized volatility components. Finally, in $\mathrm{H}_9 : \beta_2 = 0$, we examine the hypothesis that implied volatility carries no information about the future jump component of realized volatility. This restriction leads to strong rejection, both for staggered and non-staggered versions of the volatility components, thus providing evidence that option prices do contain incremental information about future jumps.

6 Concluding Remarks

This paper is the first to examine the implied-realized volatility relation in the bond market. We consider realized volatility constructed from high-frequency (5-minute) returns on 30 year Treasury bond futures and implied volatility backed out from prices of associated bond futures option contracts. Recent nonparametric statistical techniques are used to separate realized volatility into its continuous sample path and jump components, as Andersen et al. (2005) show that this leads to improved forecasting performance.

On the methodological side, we generalize the heterogeneous autoregressive (HAR) model proposed by Corsi (2004) and applied by Andersen et al. (2005) to include implied volatility from option prices as an additional regressor, and to the forecasting of the separate continuous and jump components of realized volatility. Furthermore, we introduce a new vector HAR (VecHAR) model for the simultaneous modeling of implied volatility and the separate components of realized volatility, controlling for possible endogeneity of implied volatility in the forecasting equations.

On the substantive side, our empirical results show that bond option implied volatility contains incremental information about future bond return volatility relative to both the continuous and jump components of realized volatility. Indeed, implied volatility subsumes the information content of the daily and weekly return based measures in the HAR and VecHAR models. However, implied volatility is not a fully efficient forecast in the bond market, since also monthly return based measures retain incremental information, in contrast to the situation in the foreign exchange market, where implied volatility subsumes the information content of the return based measures, see Busch et al. (2005). Furthermore, the implied volatility forecast is somewhat biased in the bond market, unlike in the stock market, where the bias is negligible and statistically insignificant, see Christensen & Nielsen (2005). Finally, our results show that even the jump component of realized bond return volatility is, to some extent, predictable, and that bond option implied volatility enters significantly in the relevant forecasting equation. This suggests that bond options are calibrated to incorporate information about future jumps in Treasury bond prices, and hence interest rates.

References

Ait-Sahalia, Y. (2004), 'Disentangling diffusion from jumps', *Journal of Financial Economics* **74**, 487–528.

Amin, K. I. & Morton, A. J. (1994), 'Implied volatility functions in arbitrage-free term structure models', *Journal of Financial Economics* **35**, 141–180.

Amin, K. I. & Ng, V. K. (1997), 'Inferring future volatility from the information in implied volatility in Eurodollar options: A new approach', *Review of Financial Studies* **10**, 333–367.

Andersen, T. G., Benzoni, L. & Lund, J. (2002), 'An empirical investigation of continuous-time equity return models', *Journal of Finance* **57**, 1239–1284.

Andersen, T. G. & Bollerslev, T. (1998*a*), 'Answering the skeptics: Yes, standard volatility models do provide accurate forecasts', *International Economic Review* **39**, 885–905.

Andersen, T. G. & Bollerslev, T. (1998*b*), 'Deutchemark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies', *Journal of Finance* **53**, 219–265.

Andersen, T. G., Bollerslev, T. & Diebold, F. X. (2004), Parametric and nonparametric volatility measurement, *in* L. P. Hansen & Y. Ait-Sahalia, eds, 'Handbook of Financial Econometrics (Forthcoming)', North-Holland, Amsterdam.

Andersen, T. G., Bollerslev, T. & Diebold, F. X. (2005), 'Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility', *Working Paper, Universiy of Pennsylvania* .

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2001), 'The distribution of exchange rate volatility', *Journal of the American Statistical Association* **96**, 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2003), 'Modeling and forecasting realized volatility', *Econometrica* **71**, 579–625.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Vega, C. (2003), 'Micro effects of macro announcements: Real-time price discovery in foreign exchange', *American Economic Review* **93**, 38–62.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Vega, C. (2004), 'Real-time price discovery in stock, bond and foreign exchange markets', *Working Paper, University of Pennsylvania* .

Andersen, T. G., Bollerslev, T. & Meddahi, N. (2004), 'Analytical evaluation of volatility forecasts', *International Economic Review* **45**, 1079–1110.

Bakshi, G., Cao, C. & Chen, Z. (1997), 'Empirical performance of alternative option pricing models', *Journal of Finance* **52**, 2003–2049.

Barndorff-Nielsen, O. E. & Shephard, N. (2001), 'Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion)', *Journal of the Royal Statistical Society Series B* **63**, 167–241.

Barndorff-Nielsen, O. E. & Shephard, N. (2002*a*), 'Econometric analysis of realized volatility and its use in estimating stochastic volatility models', *Journal of the Royal Statistical Society Series B* **64**, 253–280.

Barndorff-Nielsen, O. E. & Shephard, N. (2002*b*), 'Estimating quadratic variation using realized variance', *Journal of Applied Econometrics* **17**, 457–477.

Barndorff-Nielsen, O. E. & Shephard, N. (2004), 'Power and bipower variation with stochastic volatility and jumps (with discussion)', *Journal of Financial Econometrics* **2**, 1–57.

Barndorff-Nielsen, O. E. & Shephard, N. (2006), 'Econometrics of testing for jumps in financial economics using bipower variation', *Journal of Financial Econometrics* **4**, 1–30.

Barndorff-Nielsen, O. E. & Shephard, N. (2007), Variation, jumps, market frictions, and high frequency data in financial econometrics, *in* R. Blundell, T. Persson & W. K. Newey, eds, 'Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress', Cambridge University Press, Cambridge, UK.

Bates, D. S. (1996), 'Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options', *Review of Financial Studies* **9**, 69–107.

Black, F. (1976), 'The pricing of commodity contracts', *Journal of Financial Economics* **3**, 167–179.

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**, 637–654.

Bliss, R. R. & Ronn, E. R. (1998), 'Callable U.S. Treasury bonds: Optimal calls, anomalies and implied volatilities', *Journal of Business* **71**, 211–252.

Breusch, T. S. (1978), 'Testing for autocorrelation in dynamic linear models', *Australian Economic Papers* **17**, 334–355.

Busch, T., Christensen, B. J. & Nielsen, M. Ø. (2005), 'Forecasting exchange rate volatility in the presence of jumps', *CAF Working Paper 217, University of Aarhus* .

Cakici, N. & Zhu, J. (2001), 'Pricing Eurodollar futures options with the Heath-Jarrow-Morton model', *Journal of Futures Markets* **21**, 655–680.

Chernov, M., Gallant, A. R., Ghysels, E. & Tauchen, G. (2003), 'Alternative models of stock price dynamics', *Journal of Econometrics* **116**, 225–257.

Chiarella, C. & Tô, T.-D. (2003), 'The jump component of the volatility structure of interest rate futures markets: An international comparison', *Journal of Futures Markets* **23**, 1125–1158.

Christensen, B. J. & Nielsen, M. Ø. (2005), 'The implied-realized volatility relation with jumps in underlying asset prices', *Working paper, Cornell University* .

Christensen, B. J. & Prabhala, N. R. (1998), 'The relation between implied and realized volatility', *Journal of Financial Economics* **50**, 125–150.

Corsi, F. (2004), 'A simple long memory model of realized volatility', *Working Paper, Universiy of Lugano* .

Dacorogna, M. M., Müller, U. A., Nagler, R. J., Olsen, R. B. & Pictet, O. V. (1993), 'A geographical model for the daily and weekly seasonal volatility in the foreign exchange market', *Journal of International Money and Finance* **12**, 413–438.

Das, S. R. (1999), 'A direct discrete-time approach to Poisson-Gaussian bond option pricing in the Heath-Jarrow-Morton model', *Journal of Economic Dynamics and Control* **23**, 333–369.

Eraker, B. (2004), 'Do stock prices and volatility jump? Reconciling evidence from spot and option prices', *Journal of Finance* **59**, 1367–1403.

Eraker, B., Johannes, M. & Polson, N. (2003), 'The impact of jumps in volatility and returns', *Journal of Finance* **58**, 1269–1300.

French, D. W. (1984), 'The weekend effect on the distribution of stock prices: Implications for option pricing', *Journal of Financial Economics* **13**, 547–559.

French, K. R., Schwert, G. W. & Stambaugh, R. F. (1987), 'Expected stock returns and volatility', *Journal of Financial Economics* **19**, 3–30.

Godfrey, L. G. (1978), 'Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables', *Econometrica* **46**, 1293–1302.

Goldstein, R. F. (2000), 'The term structure of interest rates as a random field', *Review of Financial Studies* **13**, 365–384.

Guan, L. K., Ting, C. & Warachka, M. (2005), 'The implied jump risk from LIBOR rates', *Journal of Banking & Finance* **29**, 2503–2522.

Gupta, A. & Subrahmanyam, M. G. (2005), 'Pricing and hedging interest rate options: Evidence from cap-floor markets', *Journal of Banking & Finance* **29**, 701–733.

Hansen, P. R. & Lunde, A. (2005), 'Realized variance and market microstructure noise (with discussion)', *Forthcoming in Journal of Business and Economic Statistics* .

Heath, D., Jarrow, R. & Morton, A. (1992), 'Bond pricing and the term structure of interest rates: A new methodology', *Econometrica* **60**, 77–105.

Huang, X. & Tauchen, G. (2005), 'The relative contribution of jumps to total price variance', *Journal of Financial Econometrics* **3**, 456–499.

Hull, J. C. (2002), *Options, Futures, and Other Derivatives*, 5th edn, Prentice-Hall, Englewood Cliffs, New Jersey.

Hull, J. C. & White, A. (1987), 'The pricing of options on assets with stochastic volatilities', *Journal of Finance* **42**, 281–300.

Jarque, C. M. & Bera, A. K. (1980), 'Efficient tests for normality, homoskedasicity and serial independence of regression residuals', *Economics Letters* **6**, 255–259.

Johannes, M. (2004), 'The statistical and economic role of jumps in interest rates', *Journal of Finance* **59**, 227–260.

Li, A., Ritchken, P. & Sankarausbramanian, L. (1995), 'Lattice models for pricing American interest rate claims', *Journal of Finance* **50**, 719–737.

Müller, U. A., Dacorogna, M. M., Olsen, R. B., Pictet, O. V., Schwarz, M. & Morgenegg, C. (1990), 'Statistical study of foreign exchange rates, empirical evidence of a price scaling law, and intraday analysis', *Journal of Banking and Finance* **14**, 1189–1208.

Protter, P. (2004), *Stochastic Integration and Differential Equations*, 2nd edn, Springer-Verlag, New York.

Schwert, G. W. (1989), 'Why does stock market volatility change over time?', *Journal of Finance* **44**, 1115–1153.

Table 1: Summary statistics

Panel A: Variables in logarithmic form						
Statistic	$\ln\mathrm{RV}_t$	$\ln\mathrm{C}_t$	$\mathit{ln}C_t$	$\ln(\mathrm{J}_t{+}1)$	$\mathit{ln}(J_t{+}1)$	$\ln\mathrm{IV}_t$
Mean	-5.0140	-5.1208	*-5.1962*	0.0007	*0.0011*	-4.8097
Std. dev.	0.3462	0.3728	*0.3800*	0.0005	*0.0007*	0.2886
Skewness	0.1174	0.1120	*0.2454*	1.2373	*1.4987*	0.4243
Kurtosis	3.3061	3.1742	*3.1983*	6.6344	*7.3174*	3.1700
JB	0.9055	0.4899	*1.7043*	117.61**	*168.05***	4.5557
Panel B: Variables in std. dev. form						
Statistic	$\mathrm{RV}_t^{1/2}$	$\mathrm{C}_t^{1/2}$	$C_t^{1/2}$	$\mathrm{J}_t^{1/2}$	$J_t^{1/2}$	$\mathrm{IV}_t^{1/2}$
Mean	0.0827	0.0786	*0.0758*	0.0221	*0.0312*	0.0912
Std. dev.	0.0146	0.0149	*0.0149*	0.0128	*0.0110*	0.0136
Skewness	0.7058	0.7088	*0.8580*	-0.5591	*-0.6017*	0.8462
Kurtosis	4.1130	3.9495	*4.2717*	2.8027	*5.4584*	3.8629
JB	19.659**	17.710**	*27.753***	7.8400*	*45.576***	21.952**
Panel C: Variables in variance form						
Statistic	RV_t	C_t	C_t	J_t	J_t	IV_t
Mean	0.0071	0.0064	*0.0060*	0.0007	*0.0011*	0.0085
Std. dev.	0.0026	0.0025	*0.0025*	0.0005	*0.0007*	0.0027
Skewness	1.3635	1.3621	*1.5539*	1.2404	*1.5026*	1.2828
Kurtosis	6.3468	6.0038	*6.8377*	6.6488	*7.3350*	4.9643
JB	113.38**	100.04**	*148.35***	118.43**	*169.26***	63.515**

Note: The annualized monthly realized volatility, RV_t, and its continuous and jump components, C_t and J_t, are constructed from 5-minute 30 year US Treasury bond futures returns spanning the period October 1990 through December 2002, for a total of 146 monthly observations, each based on about 1600 5-minute returns. Typeface in *italics* indicates that the continuous and jump components are computed using staggered measures of realized bipower variation and realized tripower quarticity. The monthly implied volatility, IV_t, is backed out from the bond option pricing formula (15) applied to the at-the-money call option on the 30 year US Treasury bond futures contract. The option expires on the last Friday preceded by at least two business days in the month prior to the contract month and sampled on the first business day after the last trading day of the preceding option contract. The delivery date of the futures contract is the first trading day of the contract month. Each of the four volatility measures covers the same one-month interval between two consecutive expiration dates. One and two asterisks denote rejection of the null of normality for the Jarque & Bera (1980) test (JB) at the 5% and 1% significance levels, respectively.

Table 2: Realized volatility regression models

Panel A: Dependent variable is $\ln \mathrm{RV}_{t+1}$

Const.	$\ln \mathrm{RV}_t$	$\ln \mathrm{C}_t$	$\ln(\mathrm{J}_t{+}1)$	$\ln \mathrm{IV}_t$	Adj R^2	BG	LR_1	LR_2
−2.0317	0.5944	–	–	–	34.7%	16.21	–	–
(0.3393)	(0.0675)							
−2.1149	–	0.5651	−5.5847	–	36.1%	16.37	–	–
(0.3197)		(0.0619)	(44.9491)					
−1.9632	–	*0.5793*	*−35.1366*	–	*39.1%*	*15.29*	–	–
(0.3160)		*(0.0596)*	*(34.2017)*					
−1.4947	–	–	–	0.7317	37.2%	35.94**	11.20**	–
(0.3815)				(0.0792)				
−1.1099	0.3097	–	–	0.4882	42.5%	26.49**	21.31**	21.33**
(0.3780)	(0.0893)			(0.1080)				
−0.9816	–	0.3017	−61.3425	0.5083	44.6%	19.81	20.61**	27.77**
(0.3799)		(0.0796)	(43.4099)	(0.1060)				
−0.6892	–	*0.3228*	*−89.9567*	*0.5293*	*48.5%*	*17.12*	*20.34***	*38.22***
(0.3807)		*(0.0739)*	*(33.1907)*	*(0.1022)*				

Panel B: Dependent variable is $\mathrm{RV}_{t+1}^{1/2}$

Const.	$\mathrm{RV}_t^{1/2}$	$\mathrm{C}_t^{1/2}$	$\mathrm{J}_t^{1/2}$	$\mathrm{IV}_t^{1/2}$	Adj R^2	BG	LR_1	LR_2
0.0348	0.5807	–	–	–	33.1%	17.96	–	–
(0.0057)	(0.0683)							
0.0369	–	0.5824	0.0067	–	34.5%	19.64	–	–
(0.0058)		(0.0666)	(0.0777)					
0.0411	–	*0.6084*	*−0.1403*	–	*38.6%*	*18.24*	–	–
(0.0057)		*(0.0642)*	*(0.0864)*					
0.0239	–	–	–	0.6453	36.3%	36.17**	23.21**	–
(0.0066)				(0.0712)				
0.0185	0.2887	–	–	0.4442	41.1%	28.00**	29.37**	30.56**
(0.0065)	(0.0911)			(0.0984)				
0.0204	–	0.2988	−0.0712	0.4441	42.7%	20.65	30.93**	35.60**
(0.0065)		(0.0874)	(0.0746)	(0.0960)				
0.0235	–	*0.3114*	*−0.2773*	*0.4877*	*48.5%*	*20.92*	*29.06***	*51.15***
(0.0061)		*(0.0810)*	*(0.0832)*	*(0.0916)*				

Panel C: Dependent variable is RV_{t+1}

Const.	RV_t	C_t	J_t	IV_t	Adj R^2	BG	LR_1	LR_2
0.0031	0.5574	–	–	–	30.5%	19.08	–	–
(0.0005)	(0.0695)							
0.0034	–	0.5865	−0.1337	–	32.0%	20.45	–	–
(0.0005)		(0.0702)	(0.3458)					
0.0037	–	*0.6357*	*−0.3615*	–	*36.0%*	*20.30*	–	–
(0.0005)		*(0.0701)*	*(0.2609)*					
0.0022	–	–	–	0.5683	34.0%	36.43**	38.11**	–
(0.0006)				(0.0659)				
0.0018	0.2709	–	–	0.4005	38.5%	28.97**	38.93**	43.41**
(0.0006)	(0.0921)			(0.0907)				
0.0020	–	0.2920	−0.5515	0.4186	40.8%	24.15*	38.16**	49.95**
(0.0006)		(0.0908)	(0.3348)	(0.0892)				
0.0023	–	*0.3338*	*−0.7802*	*0.4347*	*45.5%*	*24.98**	*39.01***	*62.01***
(0.0006)		*(0.0879)*	*(0.2545)*	*(0.0857)*				

Note: The table shows ordinary least squares results for the regression specification (17) and the corresponding log-volatility and standard deviation regressions. Asymptotic standard errors are in parentheses, Adj R^2 is the adjusted R^2 for the regression, and BG is the Breusch-Godfrey test statistic (with 12 lags) for the null of no serial correlation in the residuals. LR_1 is the test of the unbiasedness null of $\beta = 1$ and LR_2 is the test of the joint null of $\beta = 1$ and all other coefficients except the constant being zero. One and two asterisks denote rejection at the 5% and 1% significance levels, respectively. Typeface in *italics* indicates results where the continuous and jump components are computed using staggered measures of realized bipower variation and realized tripower quarticity.

Table 3: Realized volatility HAR models

Panel A: Dependent variable is $\ln\mathrm{RV}_{t,t+22}$														
Const.	$\ln\mathrm{RV}_{t-22,t}$	$\ln\mathrm{RV}_{t-5,t}$	$\ln\mathrm{RV}_t$	$\ln\mathrm{C}_{t-22,t}$	$\ln\mathrm{C}_{t-5,t}$	$\ln\mathrm{C}_t$	$\ln(\mathrm{J}_{t-22,t}+1)$	$\ln(\mathrm{J}_{t-5,t}+1)$	$\ln(\mathrm{J}_t+1)$	$\ln\mathrm{IV}_t$	Adj R^2	BG	LR_1	LR_2
−1.8995	0.4400	0.0808	0.0940	–	–	–	–	–	–	–	36.5%	16.65	–	–
(0.3391)	(0.1058)	(0.1013)	(0.0567)											
−2.0208	–	–	–	0.4403	0.0684	0.0734	10.0556	−20.3103	27.0867	–	36.4%	18.04	–	–
(0.3239)				(0.1041)	(0.1020)	(0.0574)	(48.2171)	(33.2342)	(39.5732)					
−1.9468	–	–	–	*0.4445*	*0.0952*	*0.0421*	*−12.4903*	*−43.3166*	*41.2666*	–	*40.0%*	*12.68*	–	–
(0.3174)				*(0.1085)*	*(0.1078)*	*(0.0582)*	*(38.8882)*	*(32.6912)*	*(22.7093)*					
−1.4947	–	–	–	–	–	–	–	–	–	0.7317	37.2%	35.94**	11.20**	–
(0.3815)										(0.0792)				
−1.0594	0.1908	0.0951	0.0558	–	–	–	–	–	–	0.4607	43.3%	22.33*	23.40**	25.27**
(0.3770)	(0.1160)	(0.0958)	(0.0544)							(0.1089)				
−0.9758	–	–	–	0.2232	0.0669	0.0365	−47.3636	−12.9848	23.3046	0.4826	44.0%	20.73	22.31**	30.28**
(0.3839)				(0.1091)	(0.0957)	(0.0544)	(47.0318)	(31.2185)	(37.1311)	(0.1084)				
−0.7153	–	–	–	*0.2880*	*0.0396*	*0.0084*	*−90.2080*	*−4.1190*	*30.6914*	*0.5097*	*47.8%*	*14.72*	*19.86***	*40.50***
(0.3967)				*(0.1066)*	*(0.1012)*	*(0.0547)*	*(39.9104)*	*(31.6210)*	*(21.2961)*	*(0.1093)*				
Panel B: Dependent variable is $\mathrm{RV}^{1/2}_{t,t+22}$														
Const.	$\mathrm{RV}^{1/2}_{t-22,t}$	$\mathrm{RV}^{1/2}_{t-5,t}$	$\mathrm{RV}^{1/2}_t$	$\mathrm{C}^{1/2}_{t-22,t}$	$\mathrm{C}^{1/2}_{t-5,t}$	$\mathrm{C}^{1/2}_t$	$\mathrm{J}^{1/2}_{t-22,t}$	$\mathrm{J}^{1/2}_{t-5,t}$	$\mathrm{J}^{1/2}_t$	$\mathrm{IV}^{1/2}_t$	Adj R^2	BG	LR_1	LR_2
0.0328	0.4078	0.1069	0.1033	–	–	–	–	–	–	–	35.2%	17.41	–	–
(0.0057)	(0.1058)	(0.1052)	(0.0648)											
0.0346	–	–	–	0.4223	0.1045	0.0867	0.0291	−0.0414	0.0795	–	35.5%	19.50	–	–
(0.0059)				(0.1096)	(0.1122)	(0.0666)	(0.0799)	(0.0697)	(0.1154)					
0.0401	–	–	–	*0.4396*	*0.1200*	*0.0591*	*−0.1010*	*−0.0698*	*0.1539*	–	*40.0%*	*15.55*	–	–
(0.0058)				*(0.1159)*	*(0.1188)*	*(0.0669)*	*(0.0964)*	*(0.0606)*	*(0.0744)*					
0.0239	–	–	–	–	–	–	–	–	–	0.6453	36.3%	36.17**	23.21**	–
(0.0066)										(0.0712)				
0.0175	0.1360	0.1317	0.0637	–	–	–	–	–	–	0.4261	42.5%	24.04*	31.54**	35.99**
(0.0064)	(0.1178)	(0.0993)	(0.0618)							(0.0984)				
0.0195	–	–	–	0.1717	0.1157	0.0476	−0.0456	−0.0379	0.0573	0.4230	42.9%	22.36*	33.28**	40.22**
(0.0066)				(0.1180)	(0.1055)	(0.0633)	(0.0771)	(0.0656)	(0.1087)	(0.0970)				
0.0234	–	–	–	*0.2378*	*0.0745*	*0.0276*	*−0.2814*	*−0.0006*	*0.1158*	*0.4639*	*48.3%*	*18.95*	*29.48***	*54.67***
(0.0064)				*(0.1155)*	*(0.1107)*	*(0.0624)*	*(0.0970)*	*(0.0580)*	*(0.0695)*	*(0.0965)*				
Panel C: Dependent variable is $\mathrm{RV}_{t,t+22}$														
Const.	$\mathrm{RV}_{t-22,t}$	$\mathrm{RV}_{t-5,t}$	RV_t	$\mathrm{C}_{t-22,t}$	$\mathrm{C}_{t-5,t}$	C_t	$\mathrm{J}_{t-22,t}$	$\mathrm{J}_{t-5,t}$	J_t	IV_t	Adj R^2	BG	LR_1	LR_2
0.0029	0.3535	0.1315	0.1350	–	–	–	–	–	–	–	34.1%	18.21	–	–
(0.0005)	(0.1047)	(0.1073)	(0.0716)											
0.0031	–	–	–	0.3713	0.1387	0.1236	−0.0045	−0.1526	0.1274	–	34.4%	21.47*	–	–
(0.0005)				(0.1129)	(0.1147)	(0.0746)	(0.3652)	(0.2516)	(0.2993)					
0.0035	–	–	–	*0.3988*	*0.1675*	*0.0889*	*−0.1444*	*−0.3440*	*0.3038*	–	*38.4%*	*17.68*	–	–
(0.0005)				*(0.1254)*	*(0.1261)*	*(0.0748)*	*(0.2932)*	*(0.2477)*	*(0.1699)*					
0.0022	–	–	–	–	–	–	–	–	–	0.5683	34.0%	36.43**	38.11**	–
(0.0006)										(0.0659)				
0.0016	0.0741	0.1672	0.0941	–	–	–	–	–	–	0.3868	41.4%	24.86*	42.06**	52.57**
(0.0006)	(0.1179)	(0.1015)	(0.0681)							(0.0893)				
0.0018	–	–	–	0.0999	0.1636	0.0819	−0.4248	−0.0834	0.0997	0.3977	42.3%	26.90**	41.79**	57.91**
(0.0006)				(0.1221)	(0.1077)	(0.0705)	(0.3551)	(0.2364)	(0.2807)	(0.0890)				
0.0022	–	–	–	*0.1917*	*0.1268*	*0.0577*	*−0.6892*	*−0.0651*	*0.2365*	*0.4016*	*45.9%*	*22.91**	*41.05***	*67.23***
(0.0006)				*(0.1262)*	*(0.1185)*	*(0.0705)*	*(0.3002)*	*(0.2403)*	*(0.1599)*	*(0.0894)*				

Note: The table shows HAR-R-CJIV results for the specification (18) and the corresponding log-volatility and standard deviation models. Asymptotic standard errors are in parentheses, Adj R^2 is the adjusted R^2 for the regression, and BG is the Breusch-Godfrey test statistic (with 12 lags) for the null of no serial correlation in the residuals. LR_1 is the test of the unbiasedness null of $\beta = 1$ and LR_2 is the test of the joint null of $\beta = 1$ and all other coefficients except the constant being zero. One and two asterisks denote rejection at the 5% and 1% significance levels, respectively. Typeface in *italics* indicates results where the continuous and jump components are computed using staggered measures of realized bipower variation and realized tripower quarticity.

Table 4: Continuous component HAR models

Panel A: Dependent variable is $\ln C_{t,t+22}$

Const.	$\ln C_{t-22,t}$	$\ln C_{t-5,t}$	$\ln C_t$	$\ln(J_{t-22,t}+1)$	$\ln(J_{t-5,t}+1)$	$\ln(J_t+1)$	$\ln IV_t$	Adj R^2	BG	LR_1	LR_2
−1.8854	0.4644	0.0448	0.1186	–	–	–	–	38.0%	21.03*	–	–
(0.3413)	(0.1095)	(0.1062)	(0.0607)								
−1.7683	*0.4751*	*0.0973*	*0.0848*	–	–	–	–	*42.0%*	*19.12*	–	–
(0.3349)	*(0.1118)*	*(0.1124)*	*(0.0627)*								
−1.8645	0.4821	0.0244	0.1227	−14.7544	−26.1939	39.0388	–	37.4%	21.67*	–	–
(0.3459)	(0.1112)	(0.1090)	(0.0613)	(51.4977)	(35.4954)	(42.2657)					
−1.6303	*0.4917*	*0.0812*	*0.0901*	*−70.3778*	*−59.8763*	*42.9264*	–	*45.4%*	*14.08*	–	–
(0.3323)	*(0.1136)*	*(0.1129)*	*(0.0609)*	*(40.7119)*	*(34.2243)*	*(23.7743)*					
−1.5197	–	–	–	–	–	–	0.7487	33.5%	44.08**	8.09**	–
(0.4227)							(0.0877)				
−1.6436	–	–	–	–	–	–	*0.7386*	*31.5%*	*50.65***	*8.16***	–
(0.4376)							*(0.0908)*				
−0.8980	0.2813	0.0231	0.0886	−67.8591	−19.4188	35.5409	0.4464	42.9%	22.26*	21.66**	33.18**
(0.4173)	(0.1187)	(0.1041)	(0.0592)	(51.1331)	(33.9409)	(40.3691)	(0.1178)				
−0.5774	*0.3579*	*0.0337*	*0.0613*	*−136.8233*	*−26.3640*	*33.8850*	*0.4357*	*50.1%*	*12.13*	*22.57***	*58.24***
(0.4262)	*(0.1145)*	*(0.1087)*	*(0.0588)*	*(42.8712)*	*(33.9669)*	*(22.8760)*	*(0.1175)*				

Panel B: Dependent variable is $C^{1/2}_{t,t+22}$

Const.	$C^{1/2}_{t-22,t}$	$C^{1/2}_{t-5,t}$	$C^{1/2}_t$	$J^{1/2}_{t-22,t}$	$J^{1/2}_{t-5,t}$	$J^{1/2}_t$	$IV^{1/2}_t$	Adj R^2	BG	LR_1	LR_2
0.0296	0.4199	0.0927	0.1186	–	–	–	–	37.3%	22.89*	–	–
(0.0053)	(0.1091)	(0.1084)	(0.0668)								
0.0261	*0.4380*	*0.1398*	*0.0823*	–	–	–	–	*41.8%*	*22.36**	–	–
(0.0050)	*(0.1125)*	*(0.1143)*	*(0.0667)*								
0.0307	0.4381	0.0621	0.1267	−0.0195	−0.0558	0.0971	–	36.5%	22.51*	–	–
(0.0060)	(0.1113)	(0.1139)	(0.0676)	(0.0811)	(0.0708)	(0.1172)					
0.0354	*0.4628*	*0.0775*	*0.1068*	*−0.2213*	*−0.0960*	*0.1288*	–	*46.1%*	*17.15*	–	–
(0.0056)	*(0.1121)*	*(0.1149)*	*(0.0646)*	*(0.0932)*	*(0.0585)*	*(0.0719)*					
0.0217	–	–	–	–	–	–	0.6236	32.3%	44.57**	23.45**	–
(0.0069)							(0.0752)				
0.0208	–	–	–	–	–	–	*0.6024*	*30.5%*	*51.67***	*25.52***	–
(0.0070)							*(0.0758)*				
0.0178	0.2237	0.0717	0.0933	−0.0834	−0.0527	0.0781	0.3620	41.6%	24.09*	37.38**	46.40**
(0.0068)	(0.1222)	(0.1093)	(0.0655)	(0.0798)	(0.0679)	(0.1126)	(0.1005)				
0.0224	*0.3067*	*0.0424*	*0.0824*	*−0.3609*	*−0.0425*	*0.0993*	*0.3589*	*50.7%*	*16.86*	*40.85***	*78.06***
(0.0063)	*(0.1150)*	*(0.1102)*	*(0.0621)*	*(0.0966)*	*(0.0578)*	*(0.0692)*	*(0.0960)*				

Panel C: Dependent variable is $C_{t,t+22}$

Const.	$C_{t-22,t}$	$C_{t-5,t}$	C_t	$J_{t-22,t}$	$J_{t-5,t}$	J_t	IV_t	Adj R^2	BG	LR_1	LR_2
0.0025	0.3548	0.1332	0.1387	–	–	–	–	36.3%	24.53*	–	–
(0.0005)	(0.1081)	(0.1083)	(0.0714)								
0.0021	*0.3777*	*0.1705*	*0.1021*	–	–	–	–	*40.7%*	*25.21**	–	–
(0.0004)	*(0.1130)*	*(0.1144)*	*(0.0697)*								
0.0026	0.3704	0.1105	0.1462	−0.1700	−0.1780	0.1848	–	35.6%	25.60*	–	–
(0.0005)	(0.1095)	(0.1112)	(0.0723)	(0.3540)	(0.2439)	(0.2901)					
0.0028	*0.4044*	*0.1258*	*0.1222*	*−0.4874*	*−0.3777*	*0.2399*	–	*44.0%*	*21.59**	–	–
(0.0005)	*(0.1140)*	*(0.1146)*	*(0.0680)*	*(0.2665)*	*(0.2252)*	*(0.1545)*					
0.0020	–	–	–	–	–	–	0.5205	29.8%	44.61**	45.03**	–
(0.0006)							(0.0665)				
0.0018	–	–	–	–	–	–	*0.4947*	*28.4%*	*51.73***	*50.55***	–
(0.0006)							*(0.0655)*				
0.0016	0.1512	0.1306	0.1125	−0.5095	−0.1220	0.1623	0.3212	40.8%	26.85**	52.09**	70.43**
(0.0006)	(0.1209)	(0.1067)	(0.0699)	(0.3518)	(0.2342)	(0.2781)	(0.0882)				
0.0019	*0.2528*	*0.0960*	*0.0994*	*−0.8863*	*−0.1735*	*0.1906*	*0.2940*	*48.3%*	*21.21**	*61.08***	*99.38***
(0.0005)	*(0.1177)*	*(0.1105)*	*(0.0657)*	*(0.2799)*	*(0.2240)*	*(0.1491)*	*(0.0833)*				

Note: The table shows HAR-C-CJIV results for the specification (19) and the corresponding log-volatility and standard deviation models. Asymptotic standard errors are in parentheses, Adj R^2 is the adjusted R^2 for the regression, and BG is the Breusch-Godfrey test statistic (with 12 lags) for the null of no serial correlation in the residuals. LR_1 is the test of the unbiasedness null of $\beta = 1$ and LR_2 is the test of the joint null of $\beta = 1$ and all other coefficients except the constant being zero. One and two asterisks denote rejection at the 5% and 1% significance levels, respectively. Typeface in *italics* indicates results where the continuous and jump components are computed using staggered measures of realized bipower variation and realized tripower quarticity.

Table 5: Jump component HAR models

Panel A: Dependent variable is $\ln(J_{t,t+22}+1)$											
Const.	$\ln C_{t-22,t}$	$\ln C_{t-5,t}$	$\ln C_t$	$\ln(J_{t-22,t}+1)$	$\ln(J_{t-5,t}+1)$	$\ln(J_t+1)$	$\ln IV_t$	Adj R^2	BG	LR_1	LR_2
0.0005	–	–	–	0.1595	0.0138	−0.0624	–	1.1%	17.90	–	–
(0.0001)				(0.0891)	(0.0610)	(0.0732)					
0.0007	–	–	–	*0.3381*	*0.0229*	*0.0636*	–	*12.8%*	*12.48*	–	–
(0.0001)				*(0.0866)*	*(0.0743)*	*(0.0510)*					
0.0009	0.0000	0.0002	−0.0002	0.1646	0.0277	−0.0578	–	1.7%	23.02*	–	–
(0.0006)	(0.0002)	(0.0002)	(0.0001)	(0.0892)	(0.0615)	(0.0732)					
0.0011	*0.0001*	*0.0001*	*−0.0002*	*0.3237*	*0.0354*	*0.0622*	–	*12.6%*	*12.93*	–	–
(0.0007)	*(0.0002)*	*(0.0002)*	*(0.0001)*	*(0.0894)*	*(0.0752)*	*(0.0522)*					
0.0026	–	–	–	–	–	–	0.0004	5.3%	19.85	1855.75**	–
(0.0007)							(0.0001)				
0.0045	–	–	–	–	–	–	*0.0007*	*9.3%*	*26.16**	*1789.91***	–
(0.0009)							*(0.0002)*				
0.0023	−0.0003	0.0002	−0.0002	0.0902	0.0372	−0.0627	0.0006	7.1%	19.20	1745.29**	1849.92**
(0.0007)	(0.0002)	(0.0002)	(0.0001)	(0.0902)	(0.0599)	(0.0712)	(0.0002)				
0.0036	*−0.0002*	*0.0000*	*−0.0003*	*0.1713*	*0.1123*	*0.0415*	*0.0010*	*20.7%*	*11.51*	*1683.69***	*1801.13***
(0.0009)	*(0.0003)*	*(0.0002)*	*(0.0001)*	*(0.0938)*	*(0.0743)*	*(0.0500)*	*(0.0003)*				
Panel B: Dependent variable is $J^{1/2}_{t,t+22}$											
Const.	$C^{1/2}_{t-22,t}$	$C^{1/2}_{t-5,t}$	$C^{1/2}_t$	$J^{1/2}_{t-22,t}$	$J^{1/2}_{t-5,t}$	$J^{1/2}_t$	$IV^{1/2}_t$	Adj R^2	BG	LR_1	LR_2
0.0172	–	–	–	0.2118	0.0499	−0.1015	–	3.7%	18.94	–	–
(0.0021)				(0.0851)	(0.0720)	(0.1234)					
0.0211	–	–	–	*0.3044*	*0.0219*	*0.0845*	–	*11.0%*	*15.63*	–	–
(0.0026)				*(0.0873)*	*(0.0546)*	*(0.0665)*					
0.0150	0.0251	0.1556	−0.1621	0.2217	0.0726	−0.1101	–	5.4%	21.55*	–	–
(0.0063)	(0.1167)	(0.1195)	(0.0709)	(0.0851)	(0.0743)	(0.1230)					
0.0148	*0.0733*	*0.1109*	*−0.1057*	*0.3000*	*0.0458*	*0.0803*	–	*12.6%*	*21.64**	–	–
(0.0052)	*(0.1057)*	*(0.1083)*	*(0.0609)*	*(0.0879)*	*(0.0552)*	*(0.0678)*					
0.0089	–	–	–	–	–	–	0.1455	2.3%	24.60*	89.24**	–
(0.0072)							(0.0776)				
0.0107	–	–	–	–	–	–	*0.2245*	*7.6%*	*32.31***	*100.91***	–
(0.0060)							*(0.0648)*				
0.0062	−0.1197	0.1620	−0.1846	0.1786	0.0747	−0.1229	0.2445	8.1%	18.53	44.05**	103.33**
(0.0073)	(0.1317)	(0.1178)	(0.0706)	(0.0860)	(0.0732)	(0.1213)	(0.1083)				
0.0049	*−0.0450*	*0.0843*	*−0.1242*	*0.1942*	*0.0863*	*0.0580*	*0.2720*	*17.3%*	*20.51*	*54.39***	*119.43***
(0.0061)	*(0.1104)*	*(0.1058)*	*(0.0596)*	*(0.0927)*	*(0.0555)*	*(0.0664)*	*(0.0922)*				
Panel C: Dependent variable is $J_{t,t+22}$											
Const.	$C_{t-22,t}$	$C_{t-5,t}$	C_t	$J_{t-22,t}$	$J_{t-5,t}$	J_t	IV_t	Adj R^2	BG	LR_1	LR_2
0.0006	–	–	–	0.1594	0.0137	−0.0623	–	1.1%	17.90	–	–
(0.0001)				(0.0891)	(0.0609)	(0.0731)					
0.0007	–	–	–	*0.3381*	*0.0229*	*0.0633*	–	*12.8%*	*12.49*	–	–
(0.0001)				*(0.0866)*	*(0.0744)*	*(0.0509)*					
0.0005	0.0009	0.0282	−0.0226	0.1655	0.0253	−0.0573	–	0.5%	22.55*	–	–
(0.0001)	(0.0278)	(0.0282)	(0.0183)	(0.0898)	(0.0618)	(0.0736)					
0.0007	*−0.0056*	*0.0417*	*−0.0333*	*0.3430*	*0.0337*	*0.0639*	–	*12.5%*	*12.68*	–	–
(0.0002)	*(0.0382)*	*(0.0384)*	*(0.0228)*	*(0.0893)*	*(0.0755)*	*(0.0518)*					
0.0002	–	–	–	–	–	–	0.0478	6.0%	21.04*	479.02**	–
(0.0001)							(0.0157)				
0.0005	–	–	–	–	–	–	*0.0736*	*8.5%*	*28.08***	*406.35***	–
(0.0002)							*(0.0198)*				
0.0003	−0.0513	0.0330	−0.0306	0.0847	0.0387	−0.0627	0.0765	7.5%	22.90*	375.28**	481.53**
(0.0001)	(0.0308)	(0.0272)	(0.0178)	(0.0897)	(0.0597)	(0.0709)	(0.0225)				
0.0003	*−0.0610*	*0.0308*	*−0.0416*	*0.1971*	*0.1084*	*0.0459*	*0.1076*	*20.6%*	*14.81*	*311.83***	*426.66***
(0.0002)	*(0.0391)*	*(0.0367)*	*(0.0218)*	*(0.0930)*	*(0.0744)*	*(0.0495)*	*(0.0277)*				

Note: The table shows HAR-J-CJIV results for the specification (20) and the corresponding log-volatility and standard deviation models. Asymptotic standard errors are in parentheses, Adj R^2 is the adjusted R^2 for the regression, and BG is the Breusch-Godfrey test statistic (with 12 lags) for the null of no serial correlation in the residuals. LR_1 is the test of the unbiasedness null of $\beta = 1$ and LR_2 is the test of the joint null of $\beta = 1$ and all other coefficients except the constant being zero. One and two asterisks denote rejection at the 5% and 1% significance levels, respectively. Typeface in *italics* indicates results where the continuous and jump components are computed using staggered measures of realized bipower variation and realized tripower quarticity.

Table 6: Structural VecHAR models

Panel A: Variables in logarithmic form

Dep. var.	Constant	$\ln C_{t,t+22}$	$\ln C_{t-22,t}$	$\ln C_{t+17,t+22}$	$\ln C_{t-5,t}$	$\ln C_{t+22}$	$\ln C_t$	$\ln(J_{t,t+22}+1)$	$\ln(J_{t-22,t}+1)$	$\ln(J_{t+17,t+22}+1)$	$\ln(J_{t-5,t}+1)$	$\ln(J_{t+22}+1)$	$\ln(J_t+1)$	$\ln IV_t$	BG
$\ln C_{t,t+22}$	−0.7574	–	0.2807	–	0.0102	–	0.0891	–	−87.7410	–	−13.3037	–	39.0812	0.4867	11.73
	(0.4306)		(0.1172)		(0.1035)		(0.0590)		(51.7041)		(34.1138)		(39.7431)	(0.1199)	
	−0.4483	–	*0.3543*	–	*0.0333*	–	*0.0571*	–	*−145.5211*	–	*−25.3529*	–	*32.2383*	*0.4690*	*4.71*
	(0.4380)		*(0.1142)*		*(0.1085)*		*(0.0587)*		*(43.0548)*		*(33.8730)*		*(22.8158)*	*(0.1194)*	
$\ln(J_{t,t+22}+1)$	0.0022	–	−0.0003	–	0.0002	–	−0.0002	–	0.0657	–	0.0188	–	−0.0336	0.0007	21.21*
	(0.0008)		(0.0002)		(0.0002)		(0.0001)		(0.0722)		(0.0506)		(0.0540)	(0.0002)	
	0.0038	–	*−0.0003*	–	*0.0001*	–	*−0.0003*	–	*0.1291*	–	*0.1197*	–	*0.0266*	*0.0011*	*23.92**
	(0.0010)		*(0.0002)*		*(0.0002)*		*(0.0001)*		*(0.0895)*		*(0.0704)*		*(0.0475)*	*(0.0003)*	
$\ln IV_{t+22}$	−1.7962	0.4867	–	0.0224	–	0.0828	–	424.1407	–	−13.6987	–	1.4981	–	0.0510	13.36
	(0.6285)	(0.3171)		(0.0551)		(0.0392)		(307.1435)		(21.1035)		(31.1382)		(0.3665)	
	−1.9122	*0.2716*	–	*0.0713*	–	*0.0751*	–	*218.6496*	–	*−64.5782*	–	*17.9009*	–	*0.1913*	*10.47*
	(0.4882)	*(0.1714)*		*(0.0553)*		*(0.0395)*		*(118.5955)*		*(20.5275)*		*(15.4254)*		*(0.2172)*	

Panel B: Variables in std. dev. form

Dep. var.	Constant	$C^{1/2}_{t,t+22}$	$C^{1/2}_{t-22,t}$	$C^{1/2}_{t+17,t+22}$	$C^{1/2}_{t-5,t}$	$C^{1/2}_{t+22}$	$C^{1/2}_t$	$J^{1/2}_{t,t+22}$	$J^{1/2}_{t-22,t}$	$J^{1/2}_{t+17,t+22}$	$J^{1/2}_{t-5,t}$	$J^{1/2}_{t+22}$	$J^{1/2}_t$	$IV^{1/2}_t$	BG
$C^{1/2}_{t,t+22}$	0.0162	–	0.2167	–	0.0666	–	0.0923	–	−0.1150	–	−0.0345	–	0.0787	0.3973	18.74
	(0.0069)		(0.1216)		(0.1091)		(0.0653)		(0.0816)		(0.0690)		(0.1119)	(0.1029)	
	0.0213	–	*0.2811*	–	*0.0383*	–	*0.0793*	–	*−0.3974*	–	*−0.0418*	–	*0.0832*	*0.4117*	*21.70**
	(0.0064)		*(0.1128)*		*(0.1085)*		*(0.0615)*		*(0.0957)*		*(0.0569)*		*(0.0679)*	*(0.0971)*	
$J^{1/2}_{t,t+22}$	0.0085	–	−0.1619	–	0.1925	–	−0.2079	–	0.1451	–	0.0596	–	−0.0967	0.2578	16.39
	(0.0074)		(0.1127)		(0.1105)		(0.0677)		(0.0763)		(0.0637)		(0.1018)	(0.1078)	
	0.0064	–	*−0.1483*	–	*0.0976*	–	*−0.1412*	–	*0.0980*	–	*0.0906*	–	*0.0035*	*0.3799*	*17.30*
	(0.0061)		*(0.0772)*		*(0.0794)*		*(0.0532)*		*(0.0764)*		*(0.0437)*		*(0.0463)*	*(0.0828)*	
$IV^{1/2}_{t+22}$	0.0103	0.5818	–	−0.0230	–	0.1024	–	0.5935	–	−0.0348	–	0.0314	–	0.1789	23.11*
	(0.0102)	(0.2860)		(0.0670)		(0.0500)		(0.3400)		(0.0512)		(0.0883)		(0.2361)	
	0.0056	*0.6697*	–	*0.0528*	–	*0.0753*	–	*1.2801*	–	*−0.1446*	–	*0.0774*	–	*−0.1367*	*20.83*
	(0.0133)	*(0.3157)*		*(0.0662)*		*(0.0499)*		*(0.6022)*		*(0.0428)*		*(0.0538)*		*(0.3453)*	

Panel C: Variables in variance form

Dep. var.	Constant	$C_{t,t+22}$	$C_{t-22,t}$	$C_{t+17,t+22}$	$C_{t-5,t}$	C_{t+22}	C_t	$J_{t,t+22}$	$J_{t-22,t}$	$J_{t+17,t+22}$	$J_{t-5,t}$	J_{t+22}	J_t	IV_t	BG
$C_{t,t+22}$	0.0014	–	0.1705	–	0.1004	–	0.1096	–	−0.6976	–	−0.0616	–	0.2034	0.3634	19.58
	(0.0006)		(0.1166)		(0.1058)		(0.0690)		(0.3482)		(0.2294)		(0.2655)	(0.0898)	
	0.0018	–	*0.2409*	–	*0.0751*	–	*0.0895*	–	*−1.0009*	–	*−0.2064*	–	*0.0817*	*0.3538*	*21.26**
	(0.0005)		*(0.1041)*		*(0.1027)*		*(0.0637)*		*(0.2553)*		*(0.1985)*		*(0.1286)*	*(0.0809)*	
$J_{t,t+22}$	0.0003	–	−0.0403	–	0.0258	–	−0.0362	–	0.0621	–	0.0251	–	−0.0285	0.0758	17.30
	(0.0001)		(0.0224)		(0.0255)		(0.0167)		(0.0627)		(0.0428)		(0.0439)	(0.0209)	
	0.0003	–	*−0.0677*	–	*0.0242*	–	*−0.0476*	–	*0.1599*	–	*0.0873*	–	*−0.0096*	*0.1268*	*16.35*
	(0.0002)		*(0.0271)*		*(0.0303)*		*(0.0204)*		*(0.0704)*		*(0.0519)*		*(0.0306)*	*(0.0245)*	
IV_{t+22}	−0.0008	1.0604	–	−0.0349	–	0.1012	–	6.1435	–	−0.1601	–	0.0078	–	−0.2233	15.90
	(0.0022)	(0.5945)		(0.0741)		(0.0601)		(4.3529)		(0.1895)		(0.3085)		(0.5325)	
	−0.0019	*1.4317*	–	*0.0779*	–	*0.0728*	–	*6.2075*	–	*−0.5748*	–	*0.1178*	–	*−0.6590*	*16.17*
	(0.0034)	*(0.9599)*		*(0.0755)*		*(0.0610)*		*(4.5353)*		*(0.1868)*		*(0.1421)*		*(0.8865)*	

Note: The table shows FIML results for the simultaneous structural VecHAR system (21) and the corresponding log-volatility and standard deviation models. Asymptotic standard errors are in parentheses and BG is the Breusch-Godfrey test statistic (with 12 lags) for the null of no serial correlation in the residuals. One and two asterisks denote rejection at the 5% and 1% significance levels, respectively. Typeface in *italics* indicates results where the continuous and jump components are computed using staggered measures of realized bipower variation and realized tripower quarticity.

Table 7: LR tests in structural VecHAR models

Panel A: Variables in logarithmic form

Hypothesis	Test statistics		d.f.	p-values	
$\mathrm{H}_1 : A_{11m} = 0, A_{12m} = 0$	9.0489	*19.192*	2	0.0108	*0.0001*
$\mathrm{H}_2 : A_{11w} = 0, A_{12w} = 0$	0.1641	*0.6571*	2	0.9212	*0.7200*
$\mathrm{H}_3 : A_{11d} = 0, A_{12d} = 0$	3.3297	*2.6391*	2	0.1892	*0.2673*
$\mathrm{H}_4 : \beta_1 = 1$	17.833	*18.614*	1	0.0000	*0.0000*
$\mathrm{H}_5 : A_{11m} = 0, A_{12m} = 0, \beta_1 = 1$	27.179	*47.774*	3	0.0000	*0.0000*
$\mathrm{H}_6 : A_{11w} = 0, A_{12w} = 0, \beta_1 = 1$	17.850	*18.704*	3	0.0005	*0.0003*
$\mathrm{H}_7 : A_{11d} = 0, A_{12d} = 0, \beta_1 = 1$	19.505	*19.093*	3	0.0002	*0.0003*
$\mathrm{H}_8 : \bar{\mathbf{A}}_m = 0, \bar{\mathbf{A}}_w = 0, \bar{\mathbf{A}}_d = 0$	36.869	*75.852*	12	0.0002	*0.0000*
$\mathrm{H}_9 : \beta_2 = 0$	1729.8	*1666.6*	1	0.0000	*0.0000*

Panel B: Variables in std. dev. form

Hypothesis	Test statistics		d.f.	p-values	
$\mathrm{H}_1 : A_{11m} = 0, A_{12m} = 0$	6.1179	*22.305*	2	0.0370	*0.0000*
$\mathrm{H}_2 : A_{11w} = 0, A_{12w} = 0$	0.8657	*0.7373*	2	0.5796	*0.6917*
$\mathrm{H}_3 : A_{11d} = 0, A_{12d} = 0$	2.4963	*3.0251*	2	0.1956	*0.2203*
$\mathrm{H}_4 : \beta_1 = 1$	31.077	*32.831*	1	0.0000	*0.0000*
$\mathrm{H}_5 : A_{11m} = 0, A_{12m} = 0, \beta_1 = 1$	41.004	*68.876*	3	0.0000	*0.0000*
$\mathrm{H}_6 : A_{11w} = 0, A_{12w} = 0, \beta_1 = 1$	31.461	*33.182*	3	0.0000	*0.0000*
$\mathrm{H}_7 : A_{11d} = 0, A_{12d} = 0, \beta_1 = 1$	31.667	*33.240*	3	0.0000	*0.0000*
$\mathrm{H}_8 : \bar{\mathbf{A}}_m = 0, \bar{\mathbf{A}}_w = 0, \bar{\mathbf{A}}_d = 0$	40.444	*77.195*	12	0.0000	*0.0000*
$\mathrm{H}_9 : \beta_2 = 0$	42.685	*46.726*	1	0.0000	*0.0000*

Panel C: Variables in variance form

Hypothesis	Test statistics		d.f.	p-values	
$\mathrm{H}_1 : A_{11m} = 0, A_{12m} = 0$	6.5925	*19.959*	2	0.0370	*0.0000*
$\mathrm{H}_2 : A_{11w} = 0, A_{12w} = 0$	1.0907	*1.7866*	2	0.5796	*0.4093*
$\mathrm{H}_3 : A_{11d} = 0, A_{12d} = 0$	3.2632	*2.4076*	2	0.1956	*0.3000*
$\mathrm{H}_4 : \beta_1 = 1$	45.627	*54.723*	1	0.0000	*0.0000*
$\mathrm{H}_5 : A_{11m} = 0, A_{12m} = 0, \beta_1 = 1$	63.966	*92.442*	3	0.0000	*0.0000*
$\mathrm{H}_6 : A_{11w} = 0, A_{12w} = 0, \beta_1 = 1$	46.947	*56.080*	3	0.0000	*0.0000*
$\mathrm{H}_7 : A_{11d} = 0, A_{12d} = 0, \beta_1 = 1$	46.212	*54.743*	3	0.0000	*0.0000*
$\mathrm{H}_8 : \bar{\mathbf{A}}_m = 0, \bar{\mathbf{A}}_w = 0, \bar{\mathbf{A}}_d = 0$	41.806	*81.722*	12	0.0000	*0.0000*
$\mathrm{H}_9 : \beta_2 = 0$	369.00	*304.33*	1	0.0000	*0.0000*

Note: The table shows LR test results for the simultaneous structural VecHAR system (21) and the corresponding log-volatility and standard deviation models. Typeface in *italics* indicates results where the continuous and jump components are computed using staggered measures of realized bipower variation and realized tripower quarticity. The matrix notation $\bar{\mathbf{A}}_k = \begin{pmatrix} A_{11k} & A_{12k} \\ A_{21k} & A_{22k} \end{pmatrix}$, $k = m, w, d$, is used.

Figure 1: Time series plots of volatility measures in logarithmic form

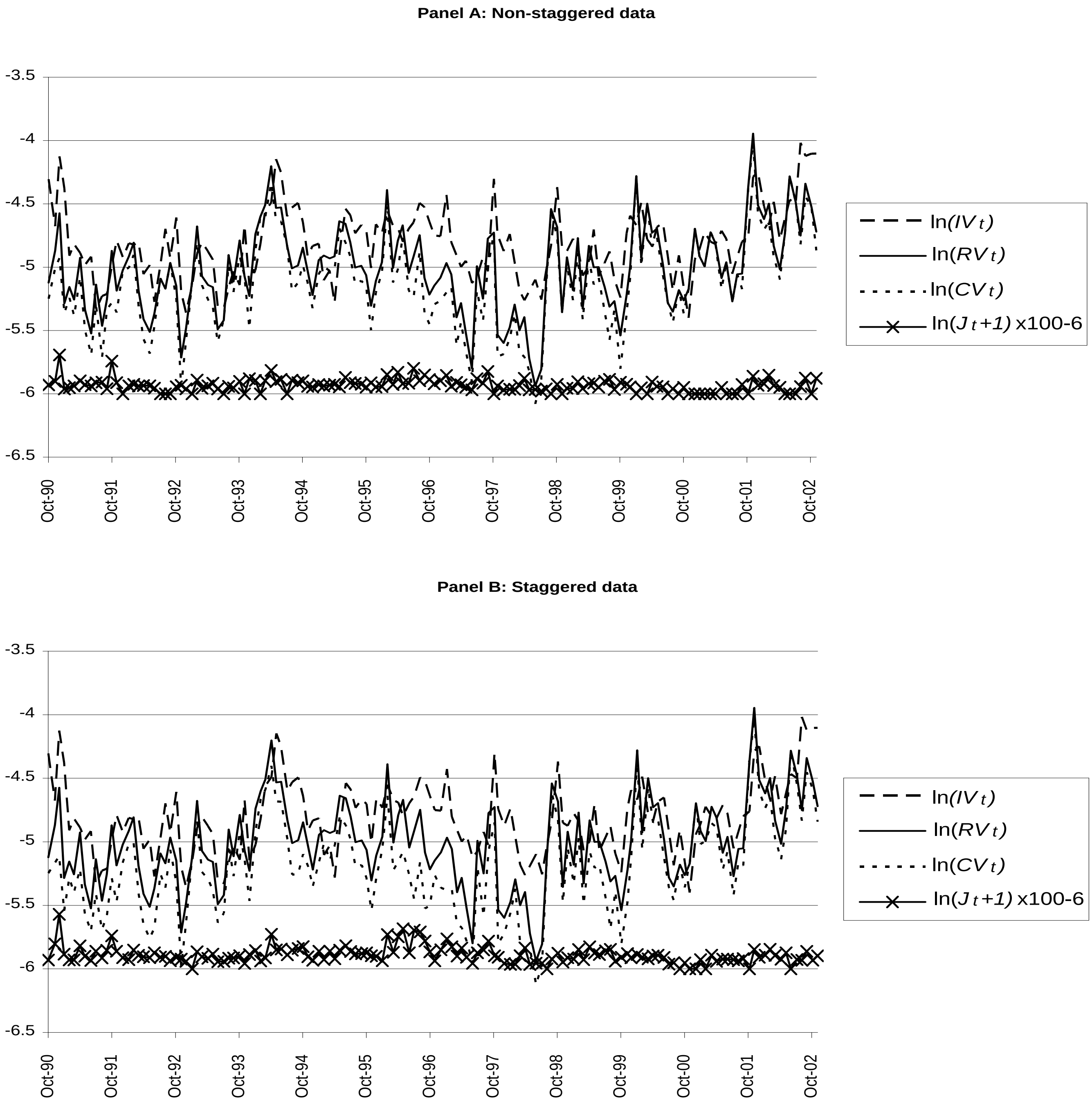
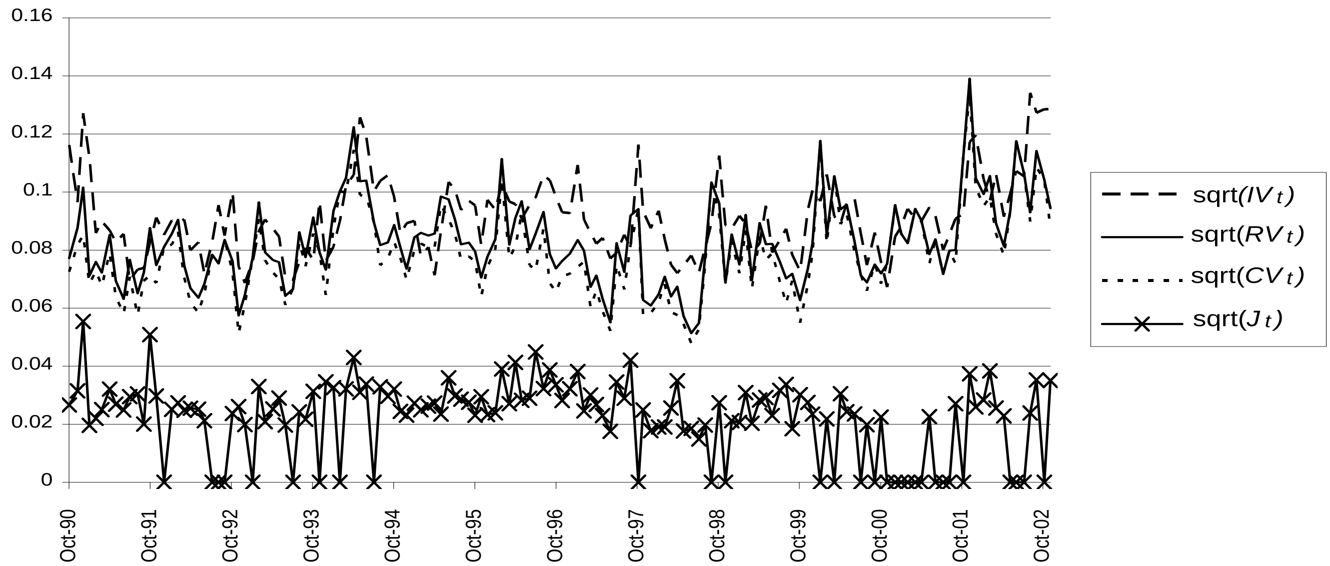


Figure 2: Time series plots of volatility measures in standard deviation form

Panel A: Non-staggered data



Panel B: Staggered data

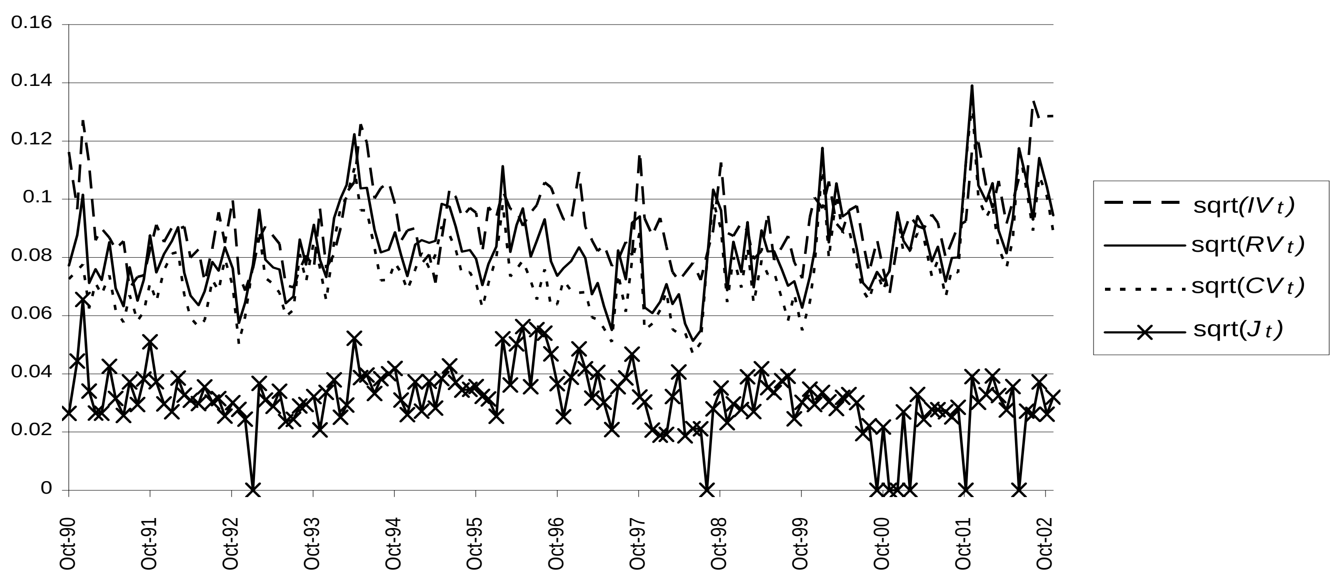


Figure 3: Time series plots of volatility measures in variance form