Queen's Economics Department Working Paper No. 1187

Forecasting Exchange Rate Volatility in the Presence of 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

12-2005

Forecasting Exchange Rate Volatility in the Presence of Jumps*

Thomas Busch
University of Aarhus

Bent Jesper Christensen
University of Aarhus

Morten Ørregaard Nielsen†
Cornell University

December 10, 2005

Abstract

We study measures of foreign exchange rate volatility based on high-frequency (5-minute) $/DM exchange rate returns using recent nonparametric statistical techniques to compute realized return volatility and its separate continuous sample path and jump components, and measures based on prices of exchange rate futures options, allowing calculation of option implied volatility. We find that implied volatility is an informationally efficient but biased forecast of future realized exchange rate volatility. Furthermore, we show that log-normality is an even better distributional approximation for implied volatility than for realized volatility in this market. Finally, we show that the jump component of future realized exchange rate volatility is to some extent predictable, and that option implied volatility is the dominant forecast of the future jump component.

Keywords: bipower variation, currency options, exchange rates, implied volatility, jumps, realized volatility

JEL classification: C1, F31, G1

*We are grateful to seminar participants at the Duke Financial Econometrics Lunch Group for comments, 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

The analysis and forecasting of asset return volatility is of great importance in the pricing and hedging of financial assets and derivatives. Both current and past return records and contemporaneous derivative price observations may be used in constructing forecasts of unknown future volatility. In an early study using daily data, Jorion (1995) documents the incremental information on future volatility in derivative prices relative to that in past realized return volatility on the foreign exchange market. This market is particularly important because of its sheer size and liquidity, and it is furthermore interesting due to the round-the-clock trading feature of the spot exchange market. Recently, Andersen, Bollerslev, Diebold & Labys (2001) study the properties of the volatility process in the foreign exchange market, showing in particular that realized exchange rate return volatility is close to log-normally distributed.

Besides adding derivative prices to the return data set as in Jorion (1995), another route to improvement of volatility forecasts involves using high-frequency return data and recent statistical techniques 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. Andersen, Bollerslev & Diebold (2005) present results from such an analysis for the foreign exchange market, as well as for the U.S. stock and Treasury bond markets. They show that for all markets, improved volatility forecasts may be obtained by splitting realized return volatility into its continuous and jump components and combining these optimally.

In the present paper, we investigate whether implied volatility from options on foreign currency futures retains the incremental information discovered by Jorion (1995) in the daily data even when assessed against improved volatility forecasts based on high-frequency (5-minute) current and past spot exchange rate 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 these 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 volatility has received much attention in recent literature on the stock, bond and foreign exchange markets, see e.g. French, Schwert & Stambaugh (1987), Schwert (1989), Andersen & Bollerslev (1998), Andersen, Bollerslev, Diebold & Ebens (2001), Andersen, Bollerslev, Diebold & Labys (2001), Barndorff-Nielsen & Shephard (2002*a*), and Andersen, Bollerslev & Diebold (2004). 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 (2003*a*, 2003*b*, 2004*a*, 2004*b*) derive a fully nonparametric separation of the continuous sample path and jump components of realized volatility. They show that realized volatility is a consistent estimate of conditional volatility as the frequency of return observations is increased even in the case of asset price processes that include both stochastic volatility and jump components. Furthermore, the nonparametric estimates of the separate components of realized volatility are consistent for the corresponding continuous and jump components of true conditional volatility. Applying this nonparametric separation technique, Andersen et al. (2005) extend results of Andersen et al. (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 stock, bond and foreign exchange markets. They show that the continuous sample path and jump components of total volatility 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.

Many recent studies have stressed the importance of separate treatment of the jump and continuous sample path components in other markets, particularly the stock market. 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), and Ait-Sahalia (2004)), nonparametric realized volatility modeling (e.g. Barndorff-Nielsen & Shephard (2003*a*, 2004*b*) and Andersen et al. (2005), who also consider the foreign exchange market, and Huang & Tauchen (2005)), and empirical option pricing (e.g. Bates (1996) for the foreign exchange market, and Bates (1991) and Bakshi, Cao & Chen (1997)). 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.

Practitioners in the foreign exchange market typically consider implied volatility a much more precise forecast of future volatility than anything based on past returns, as current option prices avoid obsolete information and are assumed to incorporate all relevant information efficiently. Complete reliance on return data, even of high frequency (say, 5 minutes), may not provide an efficient volatility forecast, given that option prices are clearly in investors' information set. Jorion (1995) considers more than seven years of daily data on $/DM currency futures and associated options and finds that implied volatility outperforms return based alternatives as a forecast of future realized volatility, although it remains a biased forecast. Similar results have been found recently by Covrig & Low (2003). The improved realized volatility forecasting performance from return based measures achieved by using high-frequency return data and differentiating the continuous and jump components begs the question of whether implied volatility continues to be an even better forecast of future realized return volatility, once option prices are added to the data set. This question was addressed recently for the stock market in Christensen & Nielsen (2005), but has never been investigated for the foreign exchange market. In the stock market (the S&P 500 index and the associated SPX options), implied volatility is a nearly unbiased forecast of high-frequency return based realized volatility, and contains 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 retain incremental information relative to implied volatility when 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.

There are reasons to believe that the results may be different in the foreign exchange market. First, volume is tremendous in the currency options market, and combined with the round-the-clock trading feature it is natural to expect an absence of frictions and a high degree of efficiency in this market. Secondly, the relevant foreign exchange options are written on a currency futures contract readily available for hedging purposes, whereas the SPX options are written on the index, leaving hedging using SPX futures slightly imperfect and hedging using the individual stocks comprising the index exceedingly costly. Lack of frictions, market efficiency and inexpensive hedging suggest that arbitrage pricing should work particularly well in the foreign exchange options market. Thirdly, exchange rate returns are generally less skewed than stock index returns. Fourth, no dividends are paid to the exchange rate, whereas the stocks comprising the index pay dividends. Lesser skewness and no dividends imply that standard option pricing formulas should work better for foreign exchange options than for stock index options. In sum, implied volatility may well be a better estimate of unknown future volatility in the foreign exchange market than in the stock market. In particular, this raises the question of whether the incremental forecasting power of past realized volatility and its continuous sample path component relative to implied volatility from option prices in the stock market is retained or disappears when moving to the foreign exchange market.

In this paper, we include implied volatility from option prices in the analysis, thus expanding the set of variables from the information set used for forecasting purposes. Given that Andersen et al. (2005) show that splitting past realized volatility into its separate components yields an improved forecast, adding implied volatility allows examining whether the continuous and jump components of past realized volatility span the relevant part of the information set. Similarly, as Jorion (1995) and Covrig & Low (2003) show that implied volatility outperforms past realized volatility as a forecast, it is of interest to test whether this conclusion holds up after allowing the two components of past realized volatility to act separately. In addition, the earlier literature on the relation between implied and realized volatility has considered realized volatility constructed from daily return observations, due to data limitations, and this could be one reason for imprecise measurement of realized volatility and might have biased the results on forecasting performance in favor of implied volatility from option prices, c.f. Poteshman (2000). In sum, by providing a joint analysis of the forecasting power of both implied volatility and the separate continuous and jump components of realized volatility, based on high-frequency returns, we are able to address a host of issues from the literature in the present paper.

We study high-frequency (5-minute) returns to the \$/DM exchange rate and \$/DM futures options. We compute alternative volatility measures from the two separate data segments: The return based measures, i.e., realized volatility and its continuous and jump components from high-frequency \$/DM exchange rate returns, and the measure based on option prices, i.e., implied volatility. We first show that the logarithm of implied volatility is very close to Gaussian, closer than implied volatility and implied variance, and closer than realized volatility or any of its continuous or jump components under any of the three transformations. This adds to the results of Andersen, Bollerslev, Diebold & Labys (2001), who showed that the logarithm of realized volatility is quite close to Gaussian, closer than realized volatility and realized variance. We then show that implied volatility contains incremental information relative to both the continuous and jump components of realized volatility when forecasting subsequently realized index return volatility. Indeed, we show that in the foreign exchange market implied volatility subsumes the information content of both components of realized volatility. This is an important difference from the findings for the stock market, where specifications using log-volatilities indicate that past realized volatility and its continuous component retain incremental information relative to implied volatility. Confirming the results of Jorion (1995), we find that some degree of bias remains in the implied volatility forecast. However, this bias is not explained by the components of realized volatility. This shows that there is volatility information in option prices which is not contained in return data, and that the continuous and jump components of realized volatility do not span investors' information set, whereas option prices fully reflect all relevant information in both components of realized volatility. Furthermore, implied volatility from option prices retains its dominant role in a forecasting context even when compared to realized volatility split into its separate components and even when using high-frequency (as opposed to daily) returns in constructing these.

As an additional novel contribution, we consider separate forecasting of the continuous and jump components of future realized volatility. 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. Our results show that implied volatility has predictive power for both components, and in particular that even the jump component of realized volatility is, to some extent, predictable.

To examine the robustness of our conclusions, we conduct an number of additional analyses. 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 non-synchronicity between sampled option prices and corresponding futures prices, bid-ask spreads, model error, etc., we take special care in handling this variable. In particular, we consider an instrumental variables approach, using lagged values of implied volatility along with the separate components of past realized volatility as instruments. In addition, we provide a structural vector autoregressive (VAR) analysis of the system consisting of implied volatility in conjunction with the two separate components of realized volatility. Both the instrumental variables analysis and the structural VAR analysis control for possible endogeneity of implied volatility in the forecasting regression. Furthermore, the simultaneous system approach allows testing interesting cross-equation restrictions. The results from these additional analyses reinforce our earlier conclusions, in particular that implied volatility is the dominant forecasting variable in investors' information set, subsuming the information content of both the continuous and jump components of past realized volatility, and that even the jump component of realized volatility is, to some extent, predictable.

The results are interesting and complement both of the above mentioned strands of literature. Firstly, although implied volatility had earlier been found to forecast better than past realized volatility, it might have been speculated that it would be possible to construct an even better forecast of future volatility than that contained in option prices, either by simply measuring past realized volatility more precisely, using high-frequency return data (Poteshman (2000) and Blair, Poon & Taylor (2001) suggest this in the context of the implied-realized volatility relation), or by using the high-frequency data to extract and combine the separate continuous and jump components of realized volatility optimally, e.g. with unequal coefficients. We find that this is not so. Secondly, since recent high-frequency data analysis shows that forecasts are improved by splitting realized volatility into its separate components, it might have been anticipated that these together summarize the relevant information set. Again, we reject the conjecture, showing that incremental information is contained in option prices.

The remainder of the paper is laid out as follows. In the next section we consider realized volatility and the nonparametric identification of its separate continuous sample path and jump components. In Section 3, we discuss the exchange rate derivative pricing model. Section 4 presents our data and Section 5 the empirical results. Finally, Section 6 offers some concluding remarks.

2 The Econometrics of Jumps

A typical assumption in asset pricing is that the log-price $p(t)$ is governed by a continuous time stochastic volatility model (see e.g. Ghysels, Harvey & Renault (1996), Barndorff-Nielsen & Shephard (2001) and the references therein) with an additive jump component. Thus, in our foreign exchange case, we assume that the logarithm of the exchange rate, $p(t)$, follows the general stochastic volatility jump diffusion model

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

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

An important feature of the model (1) is that, in the absence of jumps, the conditional distribution of the log-exchange rate given integrated drift and volatility is normal,

$$p(t)|\int_0^t \mu(s)\,ds, \sigma^{2*}(t) \sim N\left(\int_0^t \mu(s)\,ds, \sigma^{2*}(t)\right). \tag{2}$$

Here, the integrated volatility (or integrated variance)

$$\sigma^{2*}(t) = \int_0^t \sigma^2(s)\,ds \tag{3}$$

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

$$[p](t) = p\lim \sum_{j=1}^{M} \left(p(s_j) - p(s_{j-1})\right)^2, \tag{4}$$

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

$$[p](t) = \sigma^{2*}(t) + \sum_{j=1}^{q(t)} \kappa^2(t_j), \tag{5}$$

where $0 \leq t_1 < t_2 < ...$ are the jump times, $dq(t_j) = 1$. From (5), jumps show up very clearly in quadratic variation, which is written 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 in other markets including Andersen et al. (2002), Chernov et al. (2003), Eraker et al. (2003), Eraker (2004), Ait-Sahalia (2004), and Johannes (2004) all find that jumps are an empirically important part of the price process. To investigate the importance of jumps in the foreign exchange market, 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 exchange rate return data and invoke a powerful nonparametric approach to identification of the two separate components of the quadratic variation process (5), integrated volatility respectively

[1] 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.

the sum of squared jumps, following Barndorff-Nielsen & Shephard (2003*a*, 2003*b*, 2004*a*, 2004*b*), and Andersen et al. (2005).

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

$$r_{t,j} = p_{t,j} - p_{t,j-1}, \quad j = 1, ..., M, \quad t = 1, ..., T. \tag{6}$$

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

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2, \quad t = 1, ..., T. \tag{7}$$

Some authors refer to the quantity (7) as realized variance and reserve the term realized volatility for the square root of (7), 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 (2003*a*, 2003*b*, 2004*a*, 2004*b*), requires also the related bipower and tripower variation measures. The (first lag) realized bipower variation is defined as

$$BV_t = \frac{1}{\mu_1^2} \sum_{j=2}^{M} |r_{t,j}| \, |r_{t,j-1}|, \quad t = 1, ..., T, \tag{8}$$

where $\mu_1 = \sqrt{2/\pi}$. Both realized volatility and realized bipower variation are estimated with a coarseness depending on the number of intra-monthly observations M. Theoretically, a higher value of M improves the precision of the estimator, but in practice it also makes it more susceptible to market microstructure effects, such as bid-ask bounces, stale prices, measurement errors, etc., see Campbell, Lo & MacKinlay (1997). These effects potentially introduce artificial (typically negative) serial correlation in returns. Huang & Tauchen (2005) show that the resulting bias in (8) is mitigated by considering the staggered (second lag) realized bipower variation

$$\widetilde{BV}_t = \frac{1}{\mu_1^2 (1 - 2M^{-1})} \sum_{j=3}^{M} |r_{t,j}| \, |r_{t,j-2}|, \quad t = 1, ..., T. \tag{9}$$

By inserting an additional time interval between the two intervals covered by a pair of returns multiplied together in the definition of the volatility measure, the staggered version avoids the sharing of the price data $p_{t,j-1}$ which by (6) enters the definition of both $r_{t,j}$ and $r_{t,j-1}$ in the non-staggered version (8). A further statistic necessary for construction of the relevant tests is the realized tripower quarticity measure

$$TQ_t = \frac{1}{M} \mu_{4/3}^{-3} \sum_{j=3}^{M} |r_{t,j}|^{4/3} \, |r_{t,j-1}|^{4/3} \, |r_{t,j-2}|^{4/3}, \quad t = 1, ..., T, \tag{10}$$

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

$$\widetilde{TQ}_t = \frac{1}{M\mu_{4/3}^3(1-4M^{-1})}\sum_{j=5}^{M}|r_{t,j}|^{4/3}|r_{t,j-2}|^{4/3}|r_{t,j-4}|^{4/3}, \quad t = 1, ..., T, \tag{11}$$

which again avoids common 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 constructed for robustness against market microstructure effects without sacrificing asymptotic results.

As noted by Andersen & Bollerslev (1998), Andersen, Bollerslev, Diebold & Labys (2001) and Barndorff-Nielsen & Shephard (2002*a*, 2002*b*), RV_t in (7) is by definition a consistent estimator of the monthly increment to the quadratic variation process (5) as $M \to \infty$, using (4), but not of month t integrated volatility, defined as $\sigma_t^{2*} = \int_{t-1}^{t}\sigma^2(s)\,ds$. The latter is the component of the increment to quadratic variation due to continuous sample path movements in the price process (1). Therefore, realized volatility is a consistent estimator of the key integrated volatility measure, σ_t^{2*}, only in the absence of jumps. As shown by Barndorff-Nielsen & Shephard (2004*b*), an estimator that is consistent even in the presence of jumps is given by realized bipower variation from (8), i.e.,

$$BV_t \to_p \sigma_t^{2*}, \quad \text{as } M \to \infty. \tag{12}$$

It follows that the jump component of the increment to quadratic variation is estimated consistently as

$$RV_t - BV_t \to_p \sum_{j=q(t-1)+1}^{q(t)} \kappa^2(t_j). \tag{13}$$

That is, 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. In applications, non-negativity of the estimate of the jump component must be ensured, and this can be done simply by imposing a non-negativity truncation on $RV_t - BV_t$. Of course, in finite samples, $RV_t - BV_t$ may be positive due to sampling variation even if there is no jump during month t, so a notion of a "significant jump component" is needed. To this end, we employ the test statistic Z_t with the following definition and convergence property in the absence of jumps:

$$Z_t = \sqrt{M}\frac{(RV_t - BV_t)RV_t^{-1}}{\left(\left(\mu_1^{-4} + 2\mu_1^{-2} - 5\right)\max\{1, TQ_t BV_t^{-2}\}\right)^{1/2}} \to_d N(0,1), \quad \text{as } M \to \infty. \tag{14}$$

Thus, Z_t measures whether realized volatility exceeds realized bipower variation by more than what can be ascribed to chance, so large positive values of Z_t indicate the presence of jumps during month t in the underlying price process. This statistic was introduced by Barndorff-Nielsen & Shephard (2004*b*) and studied by Huang & Tauchen (2005), who showed that it has better small sample properties than the alternative asymptotically equivalent statistics in

Barndorff-Nielsen & Shephard (2003a, 2004b). Note that Z_t depends on all of RV_t, BV_t and TQ_t. By choosing the staggered versions (9) and (11) of the latter two, a staggered version $\widetilde{Z}_t$ of the test is available, 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

$$J_t = I_{\{Z_t > \Phi_{1-\alpha}\}} \left(RV_t - BV_t \right), \quad t = 1, ..., T, \tag{15}$$

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. Thus, J_t is exactly the portion of realized volatility not explained by realized bipower variation, and hence attributable to jumps in the sample path. Accordingly, the estimator of the continuous component of quadratic variation is

$$C_t = RV_t - J_t, \quad t = 1, ..., T, \tag{16}$$

ensuring that the estimators of the jump and continuous sample path components add up to total realized volatility (otherwise we could have just used the realized bipower variation defined in (8)). 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. $C_t = I_{\{Z_t \leq \Phi_{1-\alpha}\}} RV_t + I_{\{Z_t > \Phi_{1-\alpha}\}} BV_t$. Since Z_t and BV_t enter the definition (15), 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. $C_t \to_p \sigma_t^{*2}$ and $J_t \to_p \sum_{j=q(t-1)+1}^{q(t)} \kappa^2(t_j)$ as $M \to \infty$ may be achieved if also $\alpha \to 0$ (possibly as a function of M). This should hold whether staggered or non-staggered versions are used. Finally, for any standard significance level $\alpha < 1/2$, both J_t and C_t from (15) and (16) are automatically positive, since $\Phi_{1-\alpha} > 0$ for $\alpha < 1/2$. Hence, this high-frequency data approach allows for month-by-month separate nonparametric 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 itself.

3 The Exchange Rate Derivative Pricing Model

Besides computing volatility measures from observed returns, it is possible to get a volatility estimate by comparing the current level of the exchange rate with a contemporaneous price of an exchange rate derivative security and backing out the volatility that would justify the derivative price for the given exchange rate. This is the implied volatility approach, and it involves a choice of derivative pricing formula. None of the existing work on the continuous and jump components of realized volatility from the previous section (e.g. Andersen et al. (2005)) has compared with such implied volatilities from option prices when assessing the volatility forecasting performance of realized volatility and its components. This is perhaps surprising, since if option market participants are rational and markets are efficient, the exchange rate derivative price should reflect all publicly available information about expected future exchange rate volatility over the life of the option. The empirical findings of Jorion (1995) support this notion.

Jorion (1995) uses the Black (1976) and Garman & Kohlhagen (1983) version of the Black & Scholes (1973) and Merton (1973) (BSM) option pricing formula. This formula applies to a European call option with τ periods to expiration and strike price K, written on a currency futures contract with futures price F, and involves replacing the asset price in the BSM formula with the discounted futures price $e^{-r\tau}F$, where r is the riskless U.S. interest rate. However, in currency markets, the underlying futures contract typically expires Δ time periods later than the option contract, where Δ is several weeks or even months. Consequently, as shown by Bates (1996), the option formula should be further modified to

$$\begin{aligned} c(F,K,\tau,\Delta,r,\sigma) &= e^{-r(\tau+\Delta)}[F\Phi(d) - K\Phi(d-\sigma\sqrt{\tau})], \\ d &= \frac{\ln(F/K)+\frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}}, \end{aligned} \tag{17}$$

where Φ is the standard normal c.d.f. and σ is the exchange rate volatility coefficient. Based on an observed option price c, the associated implied volatility ($IV^{1/2}$) estimate[2] is backed out from the option pricing formula in (17) by numerical inversion of the nonlinear equation with respect to $IV^{1/2}$,

$$c = c(F,K,\tau,r,\Delta,IV^{1/2}). \tag{18}$$

Newton's method may be applied to compute the IV estimates by iterating on the equation

$$IV_{n+1}^{1/2} = IV_n^{1/2} + \frac{c - c(F,K,\tau,\Delta,r,IV_n^{1/2})}{\mathcal{V}(F,K,\tau,r,\Delta,IV_n^{1/2})} \tag{19}$$

until convergence, where $\mathcal{V}(F,K,\tau,r,\Delta,IV_n^{1/2}) = F\sqrt{\tau}\phi(d)e^{-r(\tau+\Delta)}$ is the vega of the option formula (see e.g. Hull (2002)) and ϕ is the p.d.f. of the standard normal distribution. The last (extra) term in vega enters since the futures price can be regarded as an asset paying a continuous dividend yield equal to the risk free rate r. In our empirical work, the algorithm is stopped when $\left|c - c(F,K,\tau,r,\Delta,IV_n^{1/2})\right| < 10^{-7}$.

Note that in (17) the term to option expiration, τ, enters d, whereas term to futures expiration, $\tau+\Delta$, is used for discounting both the futures price and the strike price. The upshot is that although d is correct for this application in the Black (1976) and Garman & Kohlhagen (1983) formula, the option price is exaggerated, by a proportional factor $e^{r\Delta}$. This leads to a systematic upward bias in implied volatilities. Consider for example the markets for \$/DM futures and associated options as in Jorion (1995) and our empirical work below.

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

In our data Δ ranges between $\frac{12}{365}$ and $\frac{76}{365}$, which is not negligible. Using the upward biased implied volatilities would generate a downward bias in the coefficient on implied volatility in the forecasting relations, potentially explaining the finding of a bias in Jorion (1995). Thus, we use the corrected formula (17) throughout when calculating implied volatility.

4 Data and Descriptive Statistics

Options on Deutsche Mark (DM) futures traded on the Chicago Mercantile Exchange (CME) over the period January 1987 to May 1999 are used in the analysis. The delivery dates of the underlying futures contract follows the quarterly cycle March, June, September, and December. In 1987 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 futures options are American with expiration dates two Fridays prior to the third Wednesday of each month. The delivery dates of the underlying futures contracts are on the third Wednesday of each of the months March, June, September, and December. Upon exercise the holder of the option contract is provided with a position at the strike price in the underlying futures contract on the following trading day. The delivery lag Δ upon terminal exercise varies between $\frac{12}{365}$ and $\frac{76}{365}$ in our data.

The data consist of daily closing prices obtained from the Commodity Research Bureau. The US Eurodollar deposit 1 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 two business days after the last trading day of the preceding option contract. In total, a sample of 148 annualized monthly IV observations of ATM calls are available. Hence, although the underlying futures contract expires at a quarterly frequency, the IV estimates are based on option contracts covering non-overlapping time intervals. Furthermore, as suggested by French (1984) and Hull (2002), the option pricing formula in (17) is extended such that trading days are used for volatilities (τ) and calender days for interest rates ($\tau + \Delta$).

For estimation of realized volatility (RV from (7)) and its separate components we follow Müller, Dacorogna, Olsen, Pictet, Schwarz & Morgenegg (1990), Dacorogna, Müller, Nagler, Olsen & Pictet (1993) and Barucci & Reno (2002), among others, and use linearly interpolated five-minute spot rates from the \$/DM foreign exchange market, providing us with a total of 288 high-frequency returns per day ($r_{t,j}$ from (6), $M = 288$, $T = 148$). The different measures are annualized and constructed on a monthly basis to cover exactly the same period as the IV estimates. Our time index refers to the month where implied volatility is sampled. Furthermore, we use the timing convention that IV_t is sampled two business days after the recording of the last return entering the computation of RV_t and its components C_t and J_t. Thus, IV_t can be regarded as a forecast of RV_{t+1}, since implied volatility is sampled at the beginning of the month covered by realized volatility for time $t+1$. 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 (15) and continuous component C from (16) of realized volatility RV.

The \$/DM spot exchange rate differs from the futures rate, which is the price of the underlying asset for the option contract. However, through the interest rate parity $\ln F = p + (r_{\$} - r_{DM})\tau$, well known from international finance, it is clear that the futures and spot \$/DM exchange rates only differ by the discounted interest rate differential. Using the spot rate instead of the futures price for realized quantities implies that our estimates of the forecasting power of IV (calculated from futures options) are on the conservative side.

The implied-realized volatility relation is examined for the following three different transformations of the volatility measures x (where $x = RV$, C, J, IV): 1) logarithmically transformed variances, $\log x$; 2) standard deviations, $x^{1/2}$; and 3) raw variances, x. Note the slight abuse of terminology – there is no correction for sample average, and 3) is simply RV from (7). To avoid taking the 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 1) replaced by J_t^*, obtained by substituting the smallest non-zero value from the time series for each zero observation. The smallest non-zero observation is in standard deviation form 0.014594 for the non-staggered data and 0.014773 for the staggered counterpart. There are 51 out of 148 months (34.5%) without significant jumps for the non-staggered data. Perhaps surprisingly, in the case of staggered data there is only one month without significant jumps. Our results therefore indicate that there may be non-negligible differences between the statistical properties of staggered and non-staggered data. Consequently, when relevant, results are reported for both measures.

Table 1 about here

Table 1 presents summary statistics for the four different annualized volatility measures, using all three functional specifications. Furthermore, for the continuous component and the jump component, statistics are shown for both staggered and non-staggered versions. Panel A shows results for the logarithmic transformation of the variance measures, and Panels B and C for the standard deviation and variance measures, respectively.

Confirming the results of Andersen, Bollerslev, Diebold & Labys (2001), the logarithmic transform produces volatility measures closest to Gaussianity. In Panel A, the Jarque & Bera (1980) test only rejects the null hypothesis of normality at the 5% level for the non-staggered version of the jump component.

As a new result from our analysis, Table 1 reveals that option based IV is much closer to Gaussianity than the other (realized) volatility measures. For the logarithmic transform, the Jarque & Bera (1980) statistic is 3.3 for realized volatility, but as low as 0.4 for the corresponding transformation of IV. Even for the standard deviation measure, Panel B, IV

does not depart significantly from Gaussianity, whereas RV does.

Figures 1-3 about here

Figures 1-3 exhibit time series plots of the four volatility measures, with non-staggered data in Panel A, and the staggered counterparts in Panel B of each figure. Each of the three volatility transformations is provided in a separate figure. From the figures, the continuous component of realized volatility is close to realized volatility itself. The new variable in our analysis, implied volatility, is also close to realized volatility, but not as close as the continuous component. The jump component computed using staggered data (Panel B) clearly behaves differently from that using non-staggered data (Panel A), as expected from Table 1. None of the two measures of the jump component is negligible, and the jump series clearly exhibit less serial dependence and behave differently compared to the other series. Hence, Figures 1-3 provide clear indication of the importance of analyzing the continuous and jump components separately in foreign exchange markets.

5 Empirical Results

In this section empirical results on the relation between realized exchange rate volatility, its disentangled components and implied volatility for the \$/DM currency and futures options markets are provided. All tables are divided into three panels. Panel A contains results for the logarithmically transformed volatility measures, Panel B for the square-root variables (standard deviation form), and Panel C for the volatility measures in raw variance form. Typeface in *italic* denotes results where the continuous and jump components are computed using staggered measures of realized bipower variation (9) and realized tripower quarticity (11).

5.1 Forecasting Realized Exchange Rate Volatility

Table 2 shows results of univariate and multivariate regressions of future realized exchange rate volatility on variables in the information set at the beginning of the period. The general form of the regressions is

$$RV_{t+1} = \alpha + \beta IV_t + \gamma x_t + \varepsilon_{t+1}, \tag{20}$$

where α is the intercept, β is the coefficient on implied volatility, hence measuring the degree of bias in this forecast, x_t is one of the lagged realized volatility measures RV_t, C_t, J_t, or the vector (C_t, J_t), ε_{t+1} is the forecast error, and $\beta = 0$ or $\gamma = 0$ is imposed if the corresponding variable is not included in the particular regression specification. Panel A of Table 2 shows the results for the log-volatilities (recall that J_t^* replaces any J_t term in x_t in the log-regressions), Panel

B for volatilities in standard deviation form, and Panel C for the raw variances.[3] Numbers reported are coefficient estimates (estimated standard errors in parentheses), adjusted R^2, and the Breusch (1978)-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 some of our specifications. The BG statistic is asymptotically χ^2 on 12 degrees of freedom under the null of no residual autocorrelation. The final two columns of the table show likelihood ratio (LR) test statistics. Here, LR_1 denotes the test of the hypothesis of a coefficient of unity on implied volatility when this is included as a regressor, i.e., this is the basic unbiasedness hypothesis that $\beta = 1$. LR_2 is the test of the stronger hypothesis of unbiasedness and efficiency of the implied volatility forecast against the unrestricted null, i.e., the joint hypothesis $\beta = 1$, $\gamma = 0$. The asymptotic distributions of LR_1 and LR_2 are χ^2 on 1 resp. $1+\dim(x_t)$ degrees of freedom under the relevant null hypotheses.

Table 2 about here

The results from the first regression in Panel A (log-volatility) show that as expected lagged realized volatility, RV_t, does have significant explanatory power for the future realization, RV_{t+1}. The first-order autocorrelation coefficient is .52, with an associated t-statistic of 7.5. This serves as a useful benchmark for assessing the new nonparametric tools, as well as the incremental information in option prices. Starting with the separation of the realized volatility forecast RV_t into its continuous and jump components C_t and J_t, the second regression in the table allows investigating whether these play different roles in forecasting future volatility RV_{t+1}. The results show that they clearly do. The coefficient on the jump component is significantly lower than that on the continuous component, showing the relevance of allowing the two components to act separately in a forecasting context. Furthermore, the regression has strong implications regarding the relative predictive powers of the continuous and jump components. In particular, quite strikingly, the results in the second line of the table shows that in fact all the information in RV_t about future exchange rate volatility stems from the continuous sample path component. Thus, C_t enters significantly in the regression, with coefficient and standard error almost identical to those for RV_t from the first regression, whereas the jump component is entirely insignificant (t-statistic of .36). This shows that jumps, which, by their very nature, are hard to predict (see Andersen et al. (2005)), also are of little use in forecasting. The results are confirmed by the regression on the staggered versions of the separate volatility components, shown in the third line of the table. This suggests that market microstructure issues, though apparently present (as seen from the differences in Table 1 and Figures 1-3 between measures using staggered and non-staggered data), are of limited consequence for forecasting purposes.

[3] In the log-volatility regressions, variables in (20) and similar equations are implicitly understood to be in log-form, i.e., we do not rewrite the equation for the logarithmic and standard deviation cases, for space considerations.

The main contributions of this paper are adding option prices to the data set and investigating the incremental forecasting power of implied volatility relative to measurements of realized volatility that are based on high-frequency returns and that separate the continuous sample path and jump components, as well as examining the role of implied volatility in forecasting the separate components of future realized volatility. We turn to the first of these investigations in the next regressions. The regression in the fourth line of Table 2 shows that implied volatility contains considerable forecasting power. The t-statistic exceeds 10, higher than for any of the forecasts considered so far. Furthermore, from the adjusted R^2, implied volatility explains 41% of the variation in future exchange rate volatility, whereas none of the regressions without implied volatility explain more than 28%. This would seem a major gain in information by adding option price data.

To test whether the information obtained by including option prices is really incremental relative to that contained in realized volatility and its components, we next add these as explanatory variables. The results in the fifth line of the table show that when regressing on both realized and implied volatility, the former, RV_t, is completely insignificant (coefficient of .02, t-statistic of .22). The coefficient on implied volatility, IV_t, is hardly different from that in the previous univariate regression, at .75, and remains strongly significant. The same is true when splitting the realized volatility forecast into its separate continuous and jump components, which is done in the next specification (sixth line). Both components of realized volatility are insignificant in the regression when implied volatility is included, and the coefficient on the latter is nearly unchanged and strongly significant. The last line of the table shows that the results are confirmed when using the staggered volatility measures. Based on the BG statistics, which are insignificant throughout the table, the findings do not appear to be hampered by misspecification.

Our results show that not only does implied volatility contain incremental forecasting power relative to high-frequency realized volatility and its separate components, it even subsumes the information content of the latter. All relevant information about future exchange rate volatility is reflected in the option prices. This shows that the conclusions of Jorion (1995) hold up even when adding high-frequency return data and using the new nonparametric techniques to disentangle and optimally combine the separate continuous and jump components of the realized volatility forecast.

One further issue regards the presence of bias in the implied volatility forecast, given that this has emerged from our analysis as the dominant forecasting variable. Jorion (1995) found that implied volatility backed out from a basic Black (1976) and Garman & Kohlhagen (1983) style option pricing formula is a biased forecast. As discussed in Section 3, our option pricing formula has been corrected following Bates (1996), thus avoiding an upward bias in implied volatility due to the delivery lag of the underlying futures contract (and hence a downward bias in the associated coefficient) present in Jorion's (1995) analysis. However, despite a non-negligible delivery lag fluctuating between $\frac{12}{365}$ and $\frac{76}{365}$ in our data and hence suggesting the importance of correcting the bias in the measurement of implied volatility, our results in fact confirm that implied volatility is a biased forecast of future realized volatility. All LR_1-tests in the first panel are significant at the 5% level or better, showing that the unbiasedness hypothesis $\beta = 1$ is rejected. The LR_2-tests examine the joint hypothesis of the IV forecast being unbiased ($\beta = 1$) and simultaneously subsuming all relevant information in other variables ($\gamma = 0$). This unbiasedness and efficiency hypothesis is rejected, too.

Following Andersen et al. (2005), we also consider the corresponding results for the cases where each volatility measure is in standard deviation form (Panel B of Table 2) or in variance form (Panel C). The regression specifications are the same as (20) above, keeping in mind the new definitions of RV_t, IV_t, and x_t (standard deviations respectively variances replace the logarithmic measures, and J_t is used instead of J_t^*). For all three transformations, realized volatility is significant in the univariate regression, and its forecasting power stems from its continuous sample path component. Particularly in the variance regressions (Panel C), where the identity $RV_t = C_t + J_t$ is strictly valid, we thus reject the implicit constraint from the first regression in the panel that the continuous and jump components should be combined in the form of raw realized volatility for the purpose of volatility forecasting. The results show that the two components should indeed be entered separately, using the new nonparametric methodology, and have different coefficients in the forecasting regression. Next, when implied volatility is included in the regression, adjusted R^2 increases dramatically, and all other regressors become insignificant, showing the informational efficiency of the implied volatility forecast, even in the presence of high-frequency realized volatility appropriately separated into its continuous and jump components. The coefficient on implied volatility is higher in the standard deviation and variance regressions than in the log-volatility regressions. Indeed, evidence against either the unbiasedness hypothesis or the joint unbiasedness and informational efficiency hypothesis is weak in Panels B and C.

Recall that our measure of implied volatility is backed out from the modified BSM-type option pricing formula (17), as is standard among practitioners and in the empirical literature on currency options. Since the formula does not account for jumps in asset prices, although it is consistent with a time-varying volatility process for a continuous sample path asset price process, it would perhaps be natural to expect that exactly the jump component would not be fully captured by implied volatility. However, our results show that implied volatility is in fact a precise forecast of future exchange rate volatility, subsuming the information content of past high-frequency return based volatility measures. This suggests that option prices may somehow be calibrated to incorporate the effect of jumps, at least to some extent. Further results below on the direct forecasting of the jump component of future volatility support this interpretation. This reduces the empirical need to invoke a more general option pricing formula allowing explicitly for jumps in exchange rates. 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. Thus, our approach yields a conservative estimate of the information content on future exchange rate volatility contained in option prices. Our results are strong, showing that simple implied volatility plays the dominant role in a forecasting context in the presence of jumps in exchange rates, more important than past realized volatility and its separate continuous and jump components based on high-frequency data, and we leave the alternative, more complicated analysis for future research.

5.2 Forecasting the Components of Exchange Rate Volatility

We now split realized exchange rate volatility, RV_{t+1}, on the left hand side of the regression into its separate continuous sample path and jump components, C_{t+1} and J_{t+1}, and examine which variables in the information set at t forecast each component. Issues regarding which variables carry incremental information in forecasting the separate components of future exchange rate volatility have never been addressed before in settings including implied volatility from the currency option markets. If implied volatility is more closely related to the continuous component of realized volatility than to the jump component, then IV_t should show up in the regressions as more important in forecasting C_{t+1} than in forecasting J_{t+1}. In particular, if jumps are essentially unpredictable, both IV_t and other variables should be insignificant in the J_{t+1} regressions.

Table 3 about here

Table 3 shows the results for forecasting the continuous component, C_{t+1}, of realized volatility. The format is the same as in Table 2. The general regression specification is of the form

$$C_{t+1} = \alpha + \beta IV_t + \gamma x_t + \varepsilon_{t+1}, \tag{21}$$

i.e., C_{t+1} replaces RV_{t+1} on the left hand side of the regression. Results are very similar to the corresponding results in Table 2. This suggests that realized volatility and its continuous component share important features, which seems natural. The BG tests show no sign of misspecification. Again, implied volatility gets higher coefficients and t-statistics than the other variables (lagged realized volatility and its continuous and jump components), and adjusted R^2 is highest when implied volatility is included in the regression. Indeed, implied volatility subsumes the information content of the other variables under all three transformations, showing that implied volatility is an informationally efficient (although slightly biased, from the LR_1 statistics) forecast also of the continuous component of future realized exchange rate volatility.

Table 4 about here

To further investigate whether implied volatility in addition reflects information about future jumps, we turn to Table 4, which reports results from regression of the future jump component, J_{t+1}, on the same explanatory variables as in the two previous tables. The general regression specification is therefore

$$J_{t+1} = \alpha + \beta IV_t + \gamma x_t + \varepsilon_{t+1}. \tag{22}$$

Considering first the initial regressions of the jump component on its own lagged value, the results in all three panels show that this has significant explanatory power if and only if staggered data are used, hence verifying the value of this approach. Furthermore, the continuous component is insignificant when this is added to the regression. Thus, jumps are to some extent predictable, and this is from their own past, not from past continuous sample path movements. Next, 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 component and is significant in univariate and multivariate regressions, whether using staggered or non-staggered data. The BG tests show no sign of misspecification. Indeed, the results show that implied volatility subsumes the information content of both components of realized volatility in forecasting the future jump component.

In general, coefficient estimates are clearly different from the previous two tables, showing that the jump component is quite different from the continuous component, and that the latter is most similar to realized volatility. This reinforces once again that the two components should be treated separately. When doing so, we find both that implied volatility forecasts something more than the continuous component of realized volatility, consistent with the notion that option prices are calibrated to incorporate jump information, and that jumps are predictable from variables in the information set, which is a result of interest in its own right.

5.3 Forecasting Implied Exchange Rate Volatility and Instrumental Variable Analysis

The results so far show that the volatility implied in prices of exchange rate futures options is an extremely powerful forecast. It subsumes the information content of both the continuous sample path and jump components of realized volatility, not only in forecasting future realized volatility, but also in forecasting the future continuous and even the future jump component. It is thus of interest to examine the properties of implied exchange rate volatility in some more detail.

Since implied volatility is derived from the prices option market participants are willing to trade at, and since option traders presumably base their decisions on the information available to them, it is natural to examine which variables in the information set implied volatility itself depends on. Thus, we first investigate the forecasting of implied volatility. In addition, since implied volatility is the new variable in our study, we subject this variable to special scrutiny, considering in particular the possibility that it is measured with error. A classical errors-in-variable (EIV) problem in implied volatility, stemming e.g. from misspecification of the option pricing formula, bid-ask spreads, or nonsynchronicities between sampled futures option prices and corresponding underlying exchange rate futures prices, would induce a downward bias in the coefficient β on implied volatility in (20)-(22), thus potentially explaining the finding so far that implied volatility is a somewhat biased (although informationally efficient) forecast of future realized volatility and its components. We use the results from the forecasting analysis (of implied volatility) to implement a standard instrumental variables two-stages least squares (2SLS) correction of the potential EIV problem and to test for the presence of the latter.

Table 5 about here

The results on forecasting implied volatility are presented in Table 5, which is laid out as the previous tables. The generic forecasting regression takes the form

$$IV_{t+1} = \alpha + \delta z_{t+1} + \varepsilon_{t+1}, \tag{23}$$

where z_{t+1} contains the relevant variables in the information set that implied volatility may depend on. Since IV_{t+1} is measured at the end of month $t+1$, it may depend on the realized volatility measures recorded over the course of month $t+1$, i.e., C_{t+1} and J_{t+1}, as well as its own lagged value, implied volatility at the beginning of the month, IV_t. The results in Table 5 are similar across the three transformations considered, and the BG tests show no sign of misspecification. The findings are that implied volatility depends on its own lag, which gets a coefficient slightly above .5 in the univariate regression and explains about 30% of the variation in implied volatility. However, this increases to more than 50% when the realized volatility measures C_{t+1} and J_{t+1} enter the regression, whether or not lagged implied volatility is retained. The continuous component gets a coefficient of about .5 which is strongly significant, whereas the coefficient on lagged implied volatility (when included) drops to about .1 and is only borderline significant at conventional levels. The coefficient on the jump component is significant when using non-staggered versions of the realized variables and insignificant when using staggered versions, and this holds regardless of whether lagged implied volatility is included in the regression. We place most faith in the results using staggered data, and conclude that implied volatility indeed reacts to variables in the information set, with the continuous sample path movements over the preceding month being the most important, and with lagged implied volatility possibly containing relevant information, too. Thus, option market participants use available information when setting prices, consistent with the interpretation of implied volatility as a conditional expectation.

Table 6 about here

As the forecasting regression explains more than half the variation in implied volatility, the forecasting variables on the right hand side are natural instruments for implied volatility in an EIV context. It was found in Table 2 that implied volatility is an informationally efficient but biased forecast, so the interesting possibility is that the bias is driven by measurement error. Thus, results of 2SLS estimation of the regression of realized on implied volatility are presented in Table 6, using the instrumentation corresponding to the last line of each panel of Table 5 for the implied volatility regressor. The column labeled RSS shows the results of the residual sum-of-squares test of the unbiasedness hypothesis.

The results are similar across the three transformations (panels) and across the use of staggered and non-staggered data in the instrumentation, and the BG statistics show no sign of misspecification. The indication is that implied volatility is in fact slightly biased, judged both by the asymptotic standard error and the RSS test, i.e. the results from Table 2 are confirmed, and a classical EIV problem is not the sole source of the phenomenon. In fact, the Hausman (1978) test in the column labelled EIV is insignificant, thus implying that there is no appreciable measurement error in our implied volatility variable. Neither the delivery lag nor an EIV problem explains the bias in the implied volatility forecast.

5.4 Structural Vector Autoregressive Analysis

We now introduce a simultaneous system approach for the joint analysis of the forecasting equations for implied volatility and the separate continuous and jump components of realized volatility. There are several advantages to the system approach. Firstly, the instrumental variable treatment of the forecasting equation for realized volatility in the previous seubsection could equally be applied to the forecasting equations for both the continuous and jump components, and joint analysis of the resulting equations is natural. Secondly, the main substantive conclusions so far, that implied volatility is an informationally efficient but slightly biased forecast of future realized volatility and its continuous sample path component, and furthermore has predictive power for the future jump component, are based on tests conducted in different regression equations which are not independent, so the relevant joint hypothesis actually involves cross-equation restrictions and should be tested in a system framework. Thirdly, while even the simple EIV problem would generate correlation between the implied volatility regressor and the error term, and thus a particular case of an endogeneity problem, the simultaneous system approach provides an efficient method for handling endogeneity more generally. Thus, we consider the structural vector autoregressive (VAR) system

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ B_{31} & 0 & 1 \end{pmatrix} \begin{pmatrix} C_{t+1} \\ J_{t+1} \\ IV_{t+1} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} + \begin{pmatrix} A_{11} & A_{12} & \beta_1 \\ A_{21} & A_{22} & \beta_2 \\ 0 & 0 & A_{33} \end{pmatrix} \begin{pmatrix} C_t \\ J_t \\ IV_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t+1} \\ \varepsilon_{2,t+1} \\ \varepsilon_{3,t+1} \end{pmatrix}, \tag{24}$$

comprising the forecasting equations for the separate components of realized volatility as well as implied volatility. There are two sources of simultaneity in the structural VAR system. Firstly, the off-diagonal term B_{31} in the leading coefficient matrix allows that IV_{t+1} depends on C_{t+1}. Based on the findings from Table 5, it is imposed that IV_{t+1} does not depend on J_{t+1}, i.e. $B_{32} = 0$ is imposed, but a non-zero value of B_{31} does imply simultaneity. Secondly, the system errors may be contemporaneously correlated.

Table 7 about here

Table 7 shows the results of Gaussian full information maximum likelihood (FIML) estimation of the structural VAR system. The BG test shows no sign of misspecification, and results are similar across the three transformations (panels). It turns out that C_{t+1} gets a coefficient slightly above unity in the IV_{t+1} equation, with an off-setting negative coefficient on lagged implied volatility IV_t. The system results confirm those from Tables 3 and 4, in particular that the coefficient on implied volatility is strongly significant but slightly below unity in the C_{t+1} equation, whereas the lagged continuous and jump components are insignificant, and that implied volatility is also significant in the J_{t+1} equation. When using staggered versions of the variables, both C_t and J_t are also significant in the jump equation under the logarithmic and square-root transformations, and so is C_t in the raw variance representation, and in all cases C_t gets a negative coefficient and J_t a positive coefficient. This suggests that jumps are by no means unpredictable. With non-staggered data, both components of realized volatility are insignificant under the first two transformations, but for the reasons given above we believe that staggering is most appropriate. Comparing to Table 4, the coefficient on implied volatility in the jump equation is higher in the system estimation, reinforcing the impression that option prices reflect future jump information.

Table 8 about here

Table 8 shows results of likelihood ratio (LR) tests of various hypotheses of interest. Overall, the earlier conclusions are confirmed by the structural VAR analysis. Thus, implied volatility is an informationally efficient but possibly slightly biased forecast of future realized volatility. Specifically, $\mathrm{H}_1 : A_{11} = 0, A_{12} = 0$ in (24) is the informational efficiency hypothesis, and it is not rejected even at the 10% level in any of the panels (staggered versions). The unbiasedness hypothesis $\mathrm{H}_2 : \beta_1 = 1$ gets p-values between 1% and 5% in all three panels (again staggered versions). The same is true for H_3, the joint unbiasedness and informational efficiency hypothesis, in Panels B and C, but this drops to a p-value of .6% in Panel A, log-volatilities. For raw variances, Panel C, we cannot reject at the 1% level even the stronger hypothesis H_4 of unbiasedness, informational efficiency, and zero intercept. All these hypotheses are tested in the first equation, but simultaneity implies that testing in the system framework is most appropriate.

Adding the restriction that implied volatility carries no information about the future jump component of realized volatility, $\beta_2 = 0$, to each of the hypotheses H_1-H_4 yields joint hypotheses H_6-H_9 across the two first equations in (24) which strictly require the system approach for their proper testing. Our results show that this restriction, whether tested by itself ($\mathrm{H}_5 : \beta_2 = 0$) or as part of a joint hypothesis, leads to strong rejection, both for staggered or non-staggered data. The implication is that option prices do contain incremental information on future jumps.

6 Concluding Remarks

In this paper, we use measures of foreign exchange rate volatility from two separate data sources, namely, high-frequency (5-minute) $/DM spot exchange rate returns, allowing the computation of realized return volatility and its separate continuous sample path and jump components using recent nonparametric statistical techniques, on the one hand, and prices of exchange rate futures options, allowing calculation of option implied volatility, on the other. We confirm earlier conclusions from Jorion (1995) who used daily data to compute realized volatility and found that implied volatility was an informationally efficient but somewhat biased forecast of future realized exchange rate volatility. Thus, Jorion's conclusions hold up to the introduction of high-frequency returns and the disentangling and optimal combination of the separate continuous and jump components of realized volatility, which was shown by Andersen et al. (2005) to lead to improved forecasting of future realized volatility. We investigate two possible sources of the bias in implied volatility, in particular the delivery lag entailed in the futures options, generating a bias in the formula used by Jorion (1995) that would potentially explain his findings, and the possibility of measurement error in implied volatility, the new variable in our study, compared to the recent realized volatility literature, in particular Andersen et al. (2005). We find that the bias remains even after correcting the option pricing formula and controlling for the potential errors-in-variable problem using an instrumental variables approach. Our work also complements that of Andersen, Bollerslev, Diebold & Labys (2001), who found that realized exchange rate return volatility is approximately log-normally distributed. We show that the log-normal approximation is even better for implied volatility in this market. Finally, we show that the jump component of future realized exchange rate volatility is to some extent predictable, and especially that option implied volatility is the dominant forecast of the future jump component. This suggests that option market participants in part base their trading strategies on information about future jumps.

References

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

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), 'Answering the skeptics: Yes, standard volatility models do provide accurate forecasts', *International Economic Review* **39**, 885–905.

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), 'Modelling and forecasting realized volatility', *Econometrica* **71**, 579–625.

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. (2003*a*), 'Econometrics of testing for jumps in financial economics using bipower variation', *Forthcoming in Journal of Financial Econometrics* .

Barndorff-Nielsen, O. E. & Shephard, N. (2003*b*), 'Realised power variation and stochastic volatility', *Bernoulli* **9**, 243–265.

Barndorff-Nielsen, O. E. & Shephard, N. (2004*a*), 'Multipower variation and stochastic volatility', *Working paper, Oxford University* .

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

Barucci, E. & Reno, R. (2002), 'On measuring volatility and the GARCH forecasting performance', *Journal of International Financial Markets, Institutions and Money* **12**, 183–200.

Bates, D. S. (1991), 'The crash of '87: Was it expected? The evidence from options markets', *Journal of Finance* **46**, 1009–1044.

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.

Blair, B. J., Poon, S. & Taylor, S. J. (2001), 'Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high-frequency index returns', *Journal of Econometrics* **105**, 5–26.

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

Campbell, J. Y., Lo, A. W. & MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton.

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

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

Covrig, V. & Low, B. S. (2003), 'The quality of volatility traded on the over-the-counter market: A multiple horizons study', *Journal of Futures Markets* **23**, 261–285.

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.

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.

Garman, M. B. & Kohlhagen, S. (1983), 'Foreign currency option values', *Journal of International Money and Finance* **2**, 231–237.

Ghysels, E., Harvey, A. C. & Renault, E. (1996), Stochastic volatility, *in* C. R. Rao & G. S. Maddala, eds, 'Statistical Methods in Finance', North-Holland, Amsterdam, pp. 119–191.

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

Hausman, J. A. (1978), 'Specification tests in econometrics', *Econometrica* **46**, 1251–1271.

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.

Jorion, P. (1995), 'Predicting volatility in the foreign exchange market', *Journal of Finance* **50**, 507–528.

Merton, R. C. (1973), 'Theory of rational option pricing', *Bell Journal of Economics and Management Science* **4**, 141–183.

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.

Poteshman, A. M. (2000), 'Forecasting future volatility from option prices', *Working Paper, University of Illinois at Urbana-Champaign* .

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	$\ln RV_t$	$\ln C_t$	*lnC_t*	$\ln J_t^*$	*lnJ_t^**	$\ln IV_t$
Mean	-4.5114	-4.5656	*-4.6352*	-7.6101	*-6.7663*	-4.5528
Std. dev.	0.4808	0.4875	*0.4899*	0.7993	*0.6248*	0.4160
Skewness	0.3417	0.2885	*0.2803*	0.4895	*0.1174*	-0.1288
Kurtosis	3.2690	3.2367	*3.1691*	2.3626	*3.5916*	2.9344
JB	3.3267	2.3989	*2.1144*	8.4162*	*2.4988*	0.4358
Panel B: Variables in std. dev. form						
Statistic	$RV_t^{1/2}$	$C_t^{1/2}$	*$C_t^{1/2}$*	$J_t^{1/2}$	*$J_t^{1/2}$*	$IV_t^{1/2}$
Mean	0,1079	0,1051	*0,1016*	0,0192	*0,0356*	0,1049
Std. dev.	0,0275	0,0270	*0,0262*	0,0161	*0,0120*	0,0217
Skewness	1,2176	1,1561	*1,1397*	0,2687	*1,1429*	0,4441
Kurtosis	5,8727	5,5489	*5,5304*	2,5916	*6,1598*	3,2225
JB	87,459*	73,032*	*71,523**	2,8098	*93,787**	5,1697
Panel C: Variables in variance form						
Statistic	RV_t	C_t	*C_t*	J_t	*J_t*	IV_t
Mean	0,0124	0,0118	*0,0110*	0,0006	*0,0014*	0,0115
Std. dev.	0,0070	0,0066	*0,0062*	0,0008	*0,0011*	0,0048
Skewness	2,4225	2,3002	*2,3058*	2,3664	*2,6638*	1,0313
Kurtosis	12,723	11,644	*11,840*	11,376	*12,509*	4,4281
JB	727,68*	591,25*	*613,06**	570,81*	*732,67**	38,811*

Note: The annualized monthly realized volatility RV_t and its continuous component C_t and jump component J_t are constructed from 5-minute $/DM spot exchange rate returns spanning the period from January 1987 through May 1999, for a total of 148 monthly observations, each based on about 5,750 5-minute returns. Typeface in *italic* denotes that the continuous and jump components are computed using the staggered measures of realized bipower variation (9) and realized tripower quarticity (11). The monthly implied volatility IV_t is backed out from the option pricing formula (17) applied to the at-the-money call option on $/DM futures expiring two Fridays prior to the third Wednesday of the contract month and sampled two business days following the expiration date of the option contract of the previous month. Each of the four volatility measures covers the same one-month interval between two consecutive expiration dates. One asterisk denotes rejection of the null of normality by the Jarque & Bera (1980) test at the 5% significance level.

Table 2: Realized volatility regressions

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

Const.	$\ln \mathrm{RV}_t$	$\ln \mathrm{C}_t$	$\ln \mathrm{J}^*_t$	$\ln \mathrm{IV}_t$	Adj R^2	BG	LR_1	LR_2
−2.1539 (0.3167)	0.5237 (0.0698)	–	–	–	27.5%	7.31	–	–
−2.0663 (0.3916)	–	0.5101 (0.0729)	0.0159 (0.0446)	–	27.3%	7.26	–	–
−2.0345 *(0.3808)*	–	*0.4955* *(0.0902)*	*0.0274* *(0.0709)*	–	*27.5%*	*7.68*	–	–
−1.0855 (0.3319)	–	–	–	0.7525 (0.0726)	41.2%	6.66	11.34^{**}	–
−1.1020 (0.3316)	0.0226 (0.1031)	–	–	0.7275 (0.1192)	42.0%	11.12	5.24^{*}	11.97^{**}
−1.3077 (0.3695)	–	0.0442 (0.0992)	−0.0540 (0.0413)	0.7507 (0.1209)	42.3%	10.50	4.31^{*}	13.87^{**}
−1.1248 *(0.3741)*	–	*0.0443* *(0.1107)*	*−0.0110* *(0.0639)*	*0.7161* *(0.1198)*	*41.6%*	*11.44*	$\mathit{5.67}^{*}$	$\mathit{12.09}^{**}$

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

Const.	$\mathrm{RV}^{1/2}_t$	$\mathrm{C}^{1/2}_t$	$\mathrm{J}^{1/2}_t$	$\mathrm{IV}^{1/2}_t$	Adj R^2	BG	LR_1	LR_2
0.0538 (0.0079)	0.4982 (0.0709)	–	–	–	24.9%	7.90	–	–
0.0540 (0.0079)	–	0.5036 (0.0752)	0.0356 (0.1262)	–	24.6%	7.27	–	–
0.0537 *(0.0079)*	–	*0.5156* *(0.0999)*	*0.0423* *(0.2182)*	–	*24.9%*	*8.15*	–	–
0.0218 (0.0085)	–	–	–	0.8218 (0.0797)	40.7%	6.41	4.98^{*}	–
0.0219 (0.0085)	−0.0050 (0.0986)	–	–	0.8230 (0.1249)	41.9%	13.24	2.04	5.32
0.0195 (0.0085)	–	0.0149 (0.0973)	−0.1951 (0.1152)	0.8606 (0.1265)	42.7%	11.61	1.24	8.29^{*}
0.0219 *(0.0085)*	–	*0.0363* *(0.1150)*	*−0.0869* *(0.1935)*	*0.8114* *(0.1250)*	*41.6%*	*13.68*	*2.32*	*5.54*

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
0.0069 (0.0010)	0.4402 (0.0735)	–	–	–	19.3%	11.19	–	–
0.0068 (0.0010)	–	0.4609 (0.0852)	0.0787 (0.7502)	–	18.8%	10.48	–	–
0.0068 *(0.0010)*	–	*0.5260* *(0.1152)*	*−0.2123* *(0.6793)*	–	*19.2%*	*11.27*	–	–
0.0020 (0.0012)	–	–	–	0.9065 (0.0939)	37.6%	9.55	1.00	–
0.0020 (0.0012)	−0.0463 (0.0953)	–	–	0.9537 (0.1383)	38.9%	18.91	0.11	1.33
0.0017 (0.0012)	–	−0.0014 (0.0977)	−1.2563 (0.6722)	0.9984 (0.1394)	39.8%	15.49	0.00	4.69
0.0019 *(0.0012)*	–	*0.0352* *(0.1230)*	*−0.6613* *(0.5943)*	*0.9523* *(0.1383)*	*38.9%*	*20.63*	*0.12*	*2.46*

Note: The table shows ordinary least squares estimation results for the regression specification (20) and the corresponding standard deviation and log-volatility regressions. Standard errors are in parentheses, Adj R^2 is the adjusted R^2 for the regression, and BG is the Breusch-Godfrey statistic (with 12 lags) of the null of no serial correlation in the residuals. LR_1 is testing, where applicable, the unbiasedness null hypothesis of $\beta = 1$, while LR_2 examines the joint unbiasedness and informational efficiency hypothesis of $\beta = 1$ and $\gamma = 0$. One and two asterisks denote rejection at the 5% and 1% significance levels, respectively. Typeface in *italic* denotes results where the continuous and jump components of realized volatility are computed using the staggered measures of realized bipower variation (9) and realized tripower quarticity (11).

Table 3: Continuous component regressions

Panel A: Dependent variable is $\ln C_{t+1}$

Const.	$\ln C_t$	$\ln J_t^*$	$\ln IV_t$	Adj R^2	BG	LR_1	LR_2
−2.1496 (0.3191)	0.5303 (0.0695)	–	–	28.2%	7.60	–	–
−2.1547 *(0.3222)*	*0.5363* *(0.0691)*	–	–	*28.9%*	*7.57*	–	–
−2.1516 (0.3962)	0.5305 (0.0738)	−0.0004 (0.0451)	–	27.7%	7.64	–	–
−2.1883 *(0.3857)*	*0.5458* *(0.0914)*	*−0.0115* *(0.0718)*	–	*28.4%*	*7.87*	–	–
−1.1729 (0.3422)	–	–	0.7452 (0.0749)	39.3%	10.83	11.30^{**}	–
−1.2368 *(0.3446)*	–	–	*0.7465* *(0.0754)*	*38.9%*	*10.04*	*11.05***	–
−1.4302 (0.3798)	0.0874 (0.1020)	−0.0669 (0.0424)	0.7139 (0.1242)	40.8%	15.25	5.35^{*}	15.12^{**}
−1.3496 *(0.3870)*	*0.1298* *(0.1145)*	*−0.0469* *(0.0662)*	*0.6603* *(0.1239)*	*39.8%*	*15.94*	*7.53***	*12.99***

Panel B: Dependent variable is $C_{t+1}^{1/2}$

Const.	$C_t^{1/2}$	$J_t^{1/2}$	$IV_t^{1/2}$	Adj R^2	BG	LR_1	LR_2
0.0516 (0.0077)	0.5062 (0.0706)	–	–	25.7%	8.81	–	–
0.0490 *(0.0074)*	*0.5141* *(0.0701)*	–	–	*26.6%*	*8.45*	–	–
0.0516 (0.0077)	0.5043 (0.0736)	0.0117 (0.1236)	–	25.2%	9.05	–	–
0.0493 *(0.0074)*	*0.5306* *(0.0943)*	*−0.0540* *(0.2060)*	–	*26.1%*	*8.83*	–	–
0.0219 (0.0085)	–	–	0.7941 (0.0792)	39.4%	10.53	6.70^{**}	–
0.0211 *(0.0082)*	–	–	*0.7673* *(0.0769)*	*39.1%*	*9.78*	*8.99***	–
0.0193 (0.0085)	0.0471 (0.0965)	−0.2042 (0.1143)	0.8052 (0.1255)	41.5%	16.42	2.45	10.65^{*}
0.0212 *(0.0082)*	*0.1077* *(0.1106)*	*−0.1680* *(0.1860)*	*0.7157* *(0.1202)*	*40.3%*	*18.48*	*5.64**	*10.76**

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

Const.	C_t	J_t	IV_t	Adj R^2	BG	LR_1	LR_2
0.0064 (0.0010)	0.4533 (0.0730)	–	–	20.5%	11.58	–	–
0.0058 *(0.0009)*	*0.4630* *(0.0725)*	–	–	*21.4%*	*10.42*	–	–
0.0064 (0.0010)	0.4565 (0.0805)	−0.0690 (0.7088)	–	19.9%	11.67	–	–
0.0059 *(0.0009)*	*0.5081* *(0.1012)*	*−0.3818* *(0.5967)*	–	*21.1%*	*10.59*	–	–
0.0020 (0.0011)	–	–	0.8547 (0.0897)	37.0%	12.51	2.63	–
0.0018 *(0.0010)*	–	–	*0.7987* *(0.0840)*	*36.8%*	*11.94*	*5.72**	–
0.0016 (0.0011)	0.0327 (0.0932)	−1.2932 (0.6417)	0.9156 (0.1331)	39.4%	18.98	0.41	6.92
0.0018 *(0.0010)*	*0.0967* *(0.1096)*	*−0.7581* *(0.5297)*	*0.7983* *(0.1233)*	*38.6%*	*24.12**	*2.73*	*8.14**

Note: The table shows ordinary least squares estimation results for the regression specification (21) and the corresponding standard deviation and log-volatility regressions. Standard errors are in parentheses, Adj R^2 is the adjusted R^2 for the regression, and BG is the Breusch-Godfrey statistic (with 12 lags) of the null of no serial correlation in the residuals. LR_1 is testing, where applicable, the unbiasedness null hypothesis of $\beta = 1$, while LR_2 examines the joint unbiasedness and informational efficiency hypothesis of $\beta = 1$ and $\gamma = 0$. One and two asterisks denote rejection at the 5% and 1% significance levels, respectively. Typeface in *italic* denotes results where the continuous and jump components of realized volatility are computed using the staggered measures of realized bipower variation (9) and realized tripower quarticity (11).

Table 4: Jump component regressions

Const.	Regressor 1	Regressor 2	Regressor 3	Adj R^2	BG	LR_1	LR_2
Panel A: Dependent variable is $\ln J^*_{t+1}$							
Const.	$\ln C_t$	$\ln J^*_t$	$\ln IV_t$	Adj R^2	BG	LR_1	LR_2
−6.6161	–	0.1315	–	1.1%	17.88	–	–
(0.6275)		(0.0820)					
−4.4732	–	*0.3395*	–	*11.0%*	*17.22*	–	–
(0.5292)		*(0.0778)*					
−5.9787	0.2092	0.0898	–	1.9%	15.67	–	–
(0.7588)	(0.1412)	(0.0863)					
−4.2167	*0.2189*	*0.2275*	–	*12.1%*	*16.60*	–	–
(0.5473)	*(0.1296)*	*(0.1019)*					
−5.0870	–	–	0.5542	7.1%	13.36	8.44^{**}	–
(0.6960)			(0.1523)				
−3.2276	–	–	*0.7773*	*25.9%*	*10.01*	*4.38**	–
(0.4862)			*(0.1064)*				
−5.1686	−0.2884	0.0152	0.8017	7.6%	16.44	0.62	10.68^{*}
(0.7804)	(0.2095)	(0.0871)	(0.2553)				
−2.9078	*−0.4302*	*0.1723*	*1.0304*	*29.2%*	*9.79*	*0.03*	*12.60***
(0.5376)	*(0.1590)*	*(0.0919)*	*(0.1721)*				
Panel B: Dependent variable is $J^{1/2}_{t+1}$							
Const.	$C^{1/2}_t$	$J^{1/2}_t$	$IV^{1/2}_t$	Adj R^2	BG	LR_1	LR_2
0.0178	–	0.0655	–	−0.3%	17.84	–	–
(0.0021)		(0.0826)					
0.0242	–	*0.3174*	–	*9.5%*	*12.80*	–	–
(0.0029)		*(0.0784)*					
0.0121	0.0586	0.0386	–	−0.0%	15.73	–	–
(0.0053)	(0.0511)	(0.0858)					
0.0212	*0.0606*	*0.2292*	–	*9.9%*	*12.44*	–	–
(0.0038)	*(0.0481)*	*(0.1050)*					
0.0021	–	–	0.1627	3.8%	14.15	125.61^{**}	–
(0.0064)			(0.0599)				
0.0060	–	–	*0.2814*	*25.1%*	*8.55*	*175.63***	–
(0.0042)			*(0.0394)*				
0.0023	−0.0807	−0.0272	0.2454	3.6%	16.31	52.13^{**}	127.16^{**}
(0.0065)	(0.0744)	(0.0881)	(0.0967)				
0.0063	*−0.1641*	*0.1686*	*0.3804*	*28.8%*	*8.57*	*80.57***	*184.86***
(0.0041)	*(0.0558)*	*(0.0938)*	*(0.0606)*				
Panel C: Dependent variable is J_{t+1}							
Const.	C_t	J_t	IV_t	Adj R^2	BG	LR_1	LR_2
0.0005	–	0.1636	–	2.0%	15.58	–	–
(0.0001)		(0.0816)					
0.0011	–	*0.2429*	–	*5.3%*	*12.09*	–	–
(0.0001)		*(0.0803)*					
0.0005	0.0044	0.1477	–	1.5%	14.48	–	–
(0.0001)	(0.0102)	(0.0898)					
0.0010	*0.0179*	*0.1694*	–	*5.2%*	*12.16*	–	–
(0.0002)	*(0.0190)*	*(0.1119)*					
0.0000	–	–	0.0517	10.0%	13.83	553.29^{**}	–
(0.0002)			(0.0123)				
0.0002	–	–	*0.1078*	*23.5%*	*9.21*	*462.97***	–
(0.0002)			*(0.0158)*				
0.0000	−0.0340	0.0369	0.0829	13.4%	16.70	431.15^{**}	558.21^{**}
(0.0002)	(0.0127)	(0.0877)	(0.0182)				
0.0002	*−0.0614*	*0.0968*	*0.1540*	*27.5%*	*8.76*	*346.02***	*470.66***
(0.0002)	*(0.0204)*	*(0.0985)*	*(0.0229)*				

Note: The table shows ordinary least squares estimation results for the regression specification (22) and the corresponding standard deviation and log-volatility regressions. Standard errors are in parentheses, Adj R^2 is the adjusted R^2 for the regression, and BG is the Breusch-Godfrey statistic (with 12 lags) of the null of no serial correlation in the residuals. LR_1 is testing, where applicable, the unbiasedness null hypothesis of $\beta = 1$, while LR_2 examines the joint unbiasedness and informational efficiency hypothesis of $\beta = 1$ and $\gamma = 0$. One and two asterisks denote rejection at the 5% and 1% significance levels, respectively. Typeface in *italic* denotes results where the continuous and jump components of realized volatility are computed using the staggered measures of realized bipower variation (9) and realized tripower quarticity (11).

Table 5: Implied volatility regressions

Panel A: Dependent variable is $\ln IV_{t+1}$					
Const.	$\ln C_{t+1}$	$\ln J^*_{t+1}$	$\ln IV_t$	Adj R^2	BG
−1.7810 (0.3025)	–	–	0.6094 (0.0662)	36.4%	11.54
−1.0105 (0.2404)	0.6207 (0.0447)	0.0931 (0.0274)	–	64.1%	13.14
−1.2703 *(0.2378)*	*0.6300* *(0.0563)*	*0.0536* *(0.0443)*	–	*63.0%*	*16.33*
−0.7357 (0.2634)	0.5417 (0.0552)	0.0854 (0.0271)	0.1526 (0.0642)	65.2%	13.34
−1.0022 *(0.2583)*	*0.5605* *(0.0622)*	*0.0328* *(0.0443)*	*0.1607* *(0.0657)*	*64.2%*	*16.87*
Panel B: Dependent variable is $IV^{1/2}_{t+1}$					
Const.	$C^{1/2}_{t+1}$	$J^{1/2}_{t+1}$	$IV^{1/2}_t$	Adj R^2	BG
0.0442 (0.0073)	–	–	0.5770 (0.0682)	32.6%	13.21
0.0400 (0.0045)	0.5678 (0.0432)	0.2681 (0.0725)	–	61.0%	11.16
0.0392 *(0.0046)*	*0.5908* *(0.0587)*	*0.1593* *(0.1282)*	–	*59.3%*	*14.48*
0.0333 (0.0056)	0.5016 (0.0545)	0.2585 (0.0720)	0.1317 (0.0672)	61.7%	12.36
0.0327 *(0.0057)*	*0.5343* *(0.0657)*	*0.1213* *(0.1288)*	*0.1289* *(0.0692)*	*60.0%*	*15.46*
Panel C: Dependent variable is IV_{t+1}					
Const.	C_{t+1}	J_{t+1}	IV_t	Adj R^2	BG
0.0053 (0.0009)	–	–	0.5356 (0.0705)	28.0%	14.96
0.0052 (0.0005)	0.4630 (0.0438)	1.3371 (0.3860)	–	55.7%	9.88
0.0051 *(0.0005)*	*0.5153* *(0.0604)*	*0.4714* *(0.3560)*	–	*54.3%*	*12.23*
0.0045 (0.0007)	0.4165 (0.0528)	1.2715 (0.3864)	0.1104 (0.0706)	56.1%	12.08
0.0044 *(0.0007)*	*0.4680* *(0.0669)*	*0.4049* *(0.3564)*	*0.1152* *(0.0717)*	*54.8%*	*13.80*

Note: The table shows ordinary least squares estimation results for the regression specification (23) and the corresponding standard deviation and log-volatility regressions. Standard errors are in parentheses, Adj R^2 is the adjusted R^2 for the regression, and BG is the Breusch-Godfrey statistic (with 12 lags) of 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 *italic* denotes results where the continuous and jump components of realized volatility are computed using the staggered measures of realized bipower variation (9) and realized tripower quarticity (11).

Table 6: Instrumental variables volatility regressions

Panel A: Logarithmically transformed variances					
Dep. var.	Const.	$\ln IV_t$	BG	EIV	RSS
$\ln RV_{t+1}$	−1.1744	0.7340	10.78	0.08	9.14**
	(0.4044)	(0.0886)			
	−1.0952	*0.7514*	*8.52*	*0.00*	*7.87***
	(0.4074)	*(0.0892)*			
$\ln C_{t+1}$	−1.2267	0.7344	12.54	0.02	8.56**
	(0.4174)	(0.0914)			
	−1.1614	*0.7641*	*9.26*	*0.16*	*6.57**
	(0.4231)	*(0.0927)*			
$\ln J^*_{t+1}$	−5.3897	0.4893	15.23	0.30	7.55**
	(0.8544)	(0.1871)			
	−3.8065	*0.6496*	*20.87*	*2.70*	*7.54***
	(0.5866)	*(0.1285)*			
Panel B: Standard deviations					
Dep. var.	Const.	$IV_t^{1/2}$	BG	EIV	RSS
$RV_{t+1}^{1/2}$	0.0262	0.7769	12.12	0.46	5.08*
	(0.0106)	(0.0997)			
	0.0229	*0.8076*	*8.81*	*0.03*	*3.68*
	(0.0107)	*(0.1009)*			
$C_{t+1}^{1/2}$	0.0252	0.7597	14.43	0.26	5.96*
	(0.0105)	(0.0991)			
	0.0199	*0.7761*	*9.67*	*0.04*	*5.36**
	(0.0103)	*(0.0974)*			
$J_{t+1}^{1/2}$	0.0045	0.1387	15.52	0.24	131.11**
	(0.0080)	(0.0757)			
	0.0122	*0.2219*	*13.92*	*3.55*	*237.71***
	(0.0054)	*(0.0508)*			
Panel C: Variances					
Dep. var.	Const.	IV_t	BG	EIV	RSS
RV_{t+1}	0.0029	0.8179	14.19	1.17	2.22
	(0.0015)	(0.1230)			
	0.0026	*0.8499*	*11.91*	*0.43*	*1.48*
	(0.0015)	*(0.1243)*			
C_{t+1}	0.0027	0.7849	16.02	0.78	3.39
	(0.0014)	(0.1175)			
	0.0020	*0.7765*	*11.62*	*0.07*	*4.11**
	(0.0013)	*(0.1110)*			
J_{t+1}	0.0002	0.0330	15.00	3.16	3557.56**
	(0.0002)	(0.0163)			
	0.0006	*0.0733*	*12.14*	*6.24**	*1885.76***
	(0.0003)	*(0.0215)*			

Note: The table shows results from the second stage of 2SLS estimation of the regression specifications (20) with $\gamma = 0$ imposed, and the corresponding standard deviation and log-volatility regressions. The first stage regression is (23) with results in the last line of each panel of Table 5. Standard errors are in parentheses, and BG is the Breusch-Godfrey statistic (with 12 lags) for the residuals. EIV denotes the Hausman (1978) test of measurement error in implied volatility. RSS is the residual sum-of-squares test of the unbiasedness null hypothesis of $\beta = 1$. One and two asterisks denote rejection at the 5% and 1% significance levels, respectively. Typeface in *italic* denotes results where the continuous and jump components of realized volatility are computed using the staggered measures of realized bipower variation (9) and realized tripower quarticity (11).

Table 7: Structural VAR models

Panel A: Variables in logarithmic form

Dep. var.	Constant	$\ln C_{t+1}$	$\ln C_t$	$\ln J^*_t$	$\ln IV_t$	BG
$\ln C_{t+1}$	−1.3684 (0.3697)	–	0.1039 (0.0863)	−0.0613 (0.0392)	0.7023 (0.1075)	14.66
	−1.1748 *(0.3639)*	–	*0.0663* *(0.1003)*	*0.0101* *(0.0401)*	*0.6792* *(0.1239)*	*13.53*
$\ln J^*_{t+1}$	−5.1128 (0.7773)	–	−0.3033 (0.2080)	0.0124 (0.0865)	0.8349 (0.2542)	16.03
	−2.7164 *(0.5260)*	–	*−0.4881* *(0.1532)*	*0.2186* *(0.0843)*	*1.0640* *(0.1704)*	*8.53*
$\ln IV_{t+1}$	−0.5002 (0.6513)	1.1216 (0.4979)	–	–	−0.2349 (0.3787)	10.42
	0.0315 *(1.9770)*	*1.4938* *(1.5776)*	–	–	*−0.5149* *(1.1863)*	*10.46*

Panel B: Variables in standard deviation form

Dep. var.	Constant	$C^{1/2}_{t+1}$	$C^{1/2}_t$	$J^{1/2}_t$	$IV^{1/2}_t$	BG
$C^{1/2}_{t+1}$	0.0184 (0.0085)	–	0.0852 (0.0738)	−0.1811 (0.1057)	0.7699 (0.1035)	16.19
	0.0202 *(0.0082)*	–	*0.0939* *(0.1100)*	*−0.0522* *(0.1262)*	*0.6989* *(0.1160)*	*16.86*
$J^{1/2}_{t+1}$	0.0016 (0.0065)	–	−0.0862 (0.0739)	−0.0334 (0.0877)	0.2579 (0.0965)	16.07
	0.0055 *(0.0041)*	–	*−0.1712* *(0.0552)*	*0.2023* *(0.0873)*	*0.3823* *(0.0597)*	*7.65*
$IV^{1/2}_{t+1}$	0.0185 (0.0139)	1.2244 (0.5331)	–	–	−0.4031 (0.4321)	12.29
	0.0160 *(0.0273)*	*1.3913* *(1.2686)*	–	–	*−0.4980* *(0.9806)*	*12.51*

Panel C: Variables in variance form

Dep. var.	Constant	C_{t+1}	C_t	J_t	IV_t	BG
C_{t+1}	0.0016 (0.0011)	–	0.0889 (0.0708)	−1.1182 (0.6119)	0.8521 (0.1102)	19.31
	0.0017 *(0.0010)*	–	*0.1246* *(0.1039)*	*−0.5593* *(0.4858)*	*0.7517* *(0.1082)*	*23.68**
J_{t+1}	0.0000 (0.0002)	–	−0.0340 (0.0127)	0.0365 (0.0874)	0.0838 (0.0182)	16.52
	0.0001 *(0.0002)*	–	*−0.0597* *(0.0200)*	*0.1183* *(0.0961)*	*0.1513* *(0.0222)*	*8.37*
IV_{t+1}	0.0031 (0.0014)	1.1451 (0.5009)	–	–	−0.4493 (0.4380)	15.05
	0.0032 *(0.0016)*	*1.1773* *(0.7567)*	–	–	*−0.4101* *(0.6115)*	*14.60*

Note: The table shows FIML estimation results for the simultaneous system (24) and the corresponding standard deviation and log-volatility systems. Standard errors are in parentheses and BG is the Breusch-Godfrey statistic (with 12 lags) for the residuals. One and two asterisks denote rejection at the 5% and 1% significance levels, respectively. Typeface in *italic* denotes results where the continuous and jump components of realized volatility are computed using the staggered measures of realized bipower variation (9) and realized tripower quarticity (11).

Table 8: LR tests in structural VAR models

Panel A: Variables in logarithmic form					
Hypothesis	Test statistics		d.f.	p-values	
$\mathrm{H}_1 : A_{11} = 0, A_{12} = 0$	5.4178	*1.5599*	2	0.0666	*0.4584*
$\mathrm{H}_2 : \beta_1 = 1$	6.6489	*6.0220*	1	0.0099	*0.0141*
$\mathrm{H}_3 : A_{11} = 0, A_{12} = 0, \beta_1 = 1$	16.504	*12.437*	3	0.0009	*0.0060*
$\mathrm{H}_4 : A_{11} = 0, A_{12} = 0, \beta_1 = 1, \alpha_1 = 0$	16.946	*20.251*	4	0.0020	*0.0004*
$\mathrm{H}_5 : \beta_2 = 0$	10.555	*34.682*	1	0.0012	*0.0000*
$\mathrm{H}_6 : A_{11} = 0, A_{12} = 0, \beta_2 = 0$	17.767	*47.871*	3	0.0005	*0.0000*
$\mathrm{H}_7 : \beta_1 = 1, \beta_2 = 0$	21.850	*66.944*	2	0.0000	*0.0000*
$\mathrm{H}_8 : A_{11} = 0, A_{12} = 0, \beta_1 = 1, \beta_2 = 0$	31.315	*73.957*	4	0.0000	*0.0000*
$\mathrm{H}_9 : A_{11} = 0, A_{12} = 0, \beta_1 = 1, \beta_2 = 0, \alpha_1 = 0$	31.757	*81.770*	5	0.0000	*0.0000*
Panel B: Variables in std. dev. form					
Hypothesis	Test statistics		d.f.	p-values	
$\mathrm{H}_1 : A_{11} = 0, A_{12} = 0$	7.3064	*1.8779*	2	0.0259	*0.3910*
$\mathrm{H}_2 : \beta_1 = 1$	4.2538	*5.0283*	1	0.0392	*0.0249*
$\mathrm{H}_3 : A_{11} = 0, A_{12} = 0, \beta_1 = 1$	13.882	*10.806*	3	0.0031	*0.0128*
$\mathrm{H}_4 : A_{11} = 0, A_{12} = 0, \beta_1 = 1, \alpha_1 = 0$	13.896	*15.666*	4	0.0076	*0.0035*
$\mathrm{H}_5 : \beta_2 = 0$	7.0634	*36.810*	1	0.0079	*0.0000*
$\mathrm{H}_6 : A_{11} = 0, A_{12} = 0, \beta_2 = 0$	16.504	*50.883*	3	0.0009	*0.0000*
$\mathrm{H}_7 : \beta_1 = 1, \beta_2 = 0$	14.716	*70.790*	2	0.0006	*0.0000*
$\mathrm{H}_8 : A_{11} = 0, A_{12} = 0, \beta_1 = 1, \beta_2 = 0$	23.652	*77.727*	4	0.0001	*0.0000*
$\mathrm{H}_9 : A_{11} = 0, A_{12} = 0, \beta_1 = 1, \beta_2 = 0, \alpha_1 = 0$	23.666	*82.587*	5	0.0003	*0.0000*
Panel C: Variables in variance form					
Hypothesis	Test statistics		d.f.	p-values	
$\mathrm{H}_1 : A_{11} = 0, A_{12} = 0$	8.2285	*3.3805*	2	0.0163	*0.1845*
$\mathrm{H}_2 : \beta_1 = 1$	1.5853	*3.8969*	1	0.2080	*0.0484*
$\mathrm{H}_3 : A_{11} = 0, A_{12} = 0, \beta_1 = 1$	10.786	*9.0517*	3	0.0129	*0.0286*
$\mathrm{H}_4 : A_{11} = 0, A_{12} = 0, \beta_1 = 1, \alpha_1 = 0$	11.014	*11.048*	4	0.0264	*0.0260*
$\mathrm{H}_5 : \beta_2 = 0$	19.996	*40.490*	1	0.0000	*0.0000*
$\mathrm{H}_6 : A_{11} = 0, A_{12} = 0, \beta_2 = 0$	30.659	*56.052*	3	0.0000	*0.0000*
$\mathrm{H}_7 : \beta_1 = 1, \beta_2 = 0$	28.764	*78.028*	2	0.0000	*0.0000*
$\mathrm{H}_8 : A_{11} = 0, A_{12} = 0, \beta_1 = 1, \beta_2 = 0$	37.119	*84.795*	4	0.0000	*0.0000*
$\mathrm{H}_9 : A_{11} = 0, A_{12} = 0, \beta_1 = 1, \beta_2 = 0, \alpha_1 = 0$	37.347	*86.792*	5	0.0000	*0.0000*

Note: The table shows likelihood ratio test results for the simultaneous system (24) and the corresponding standard deviation and log-volatility systems. Typeface in *italic* denotes results where the continuous and jump components of realized volatility are computed using the staggered measures of realized bipower variation (9) and realized tripower quarticity (11).

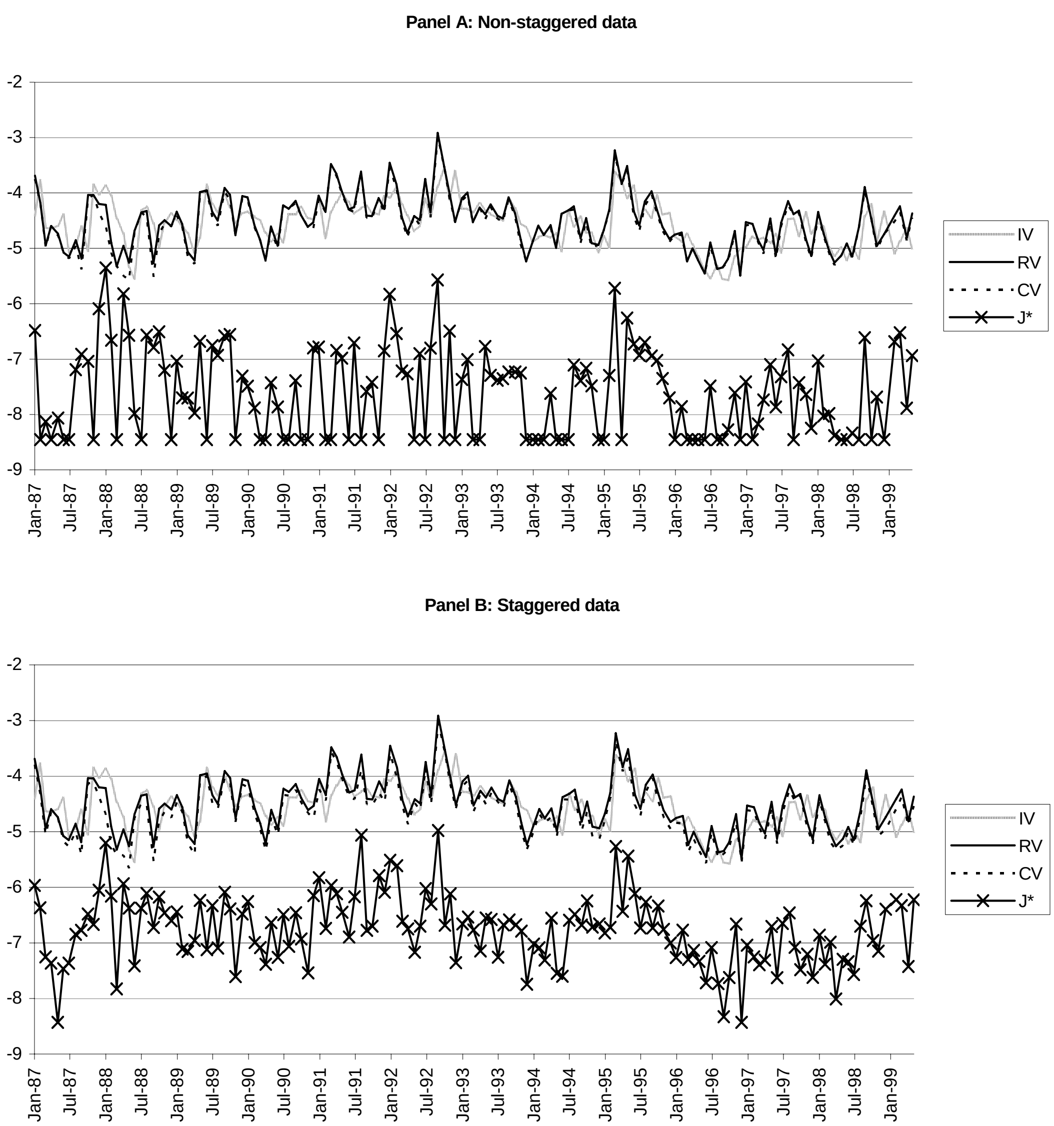


Figure 1: Time series plots of volatility measures in logarithmic form

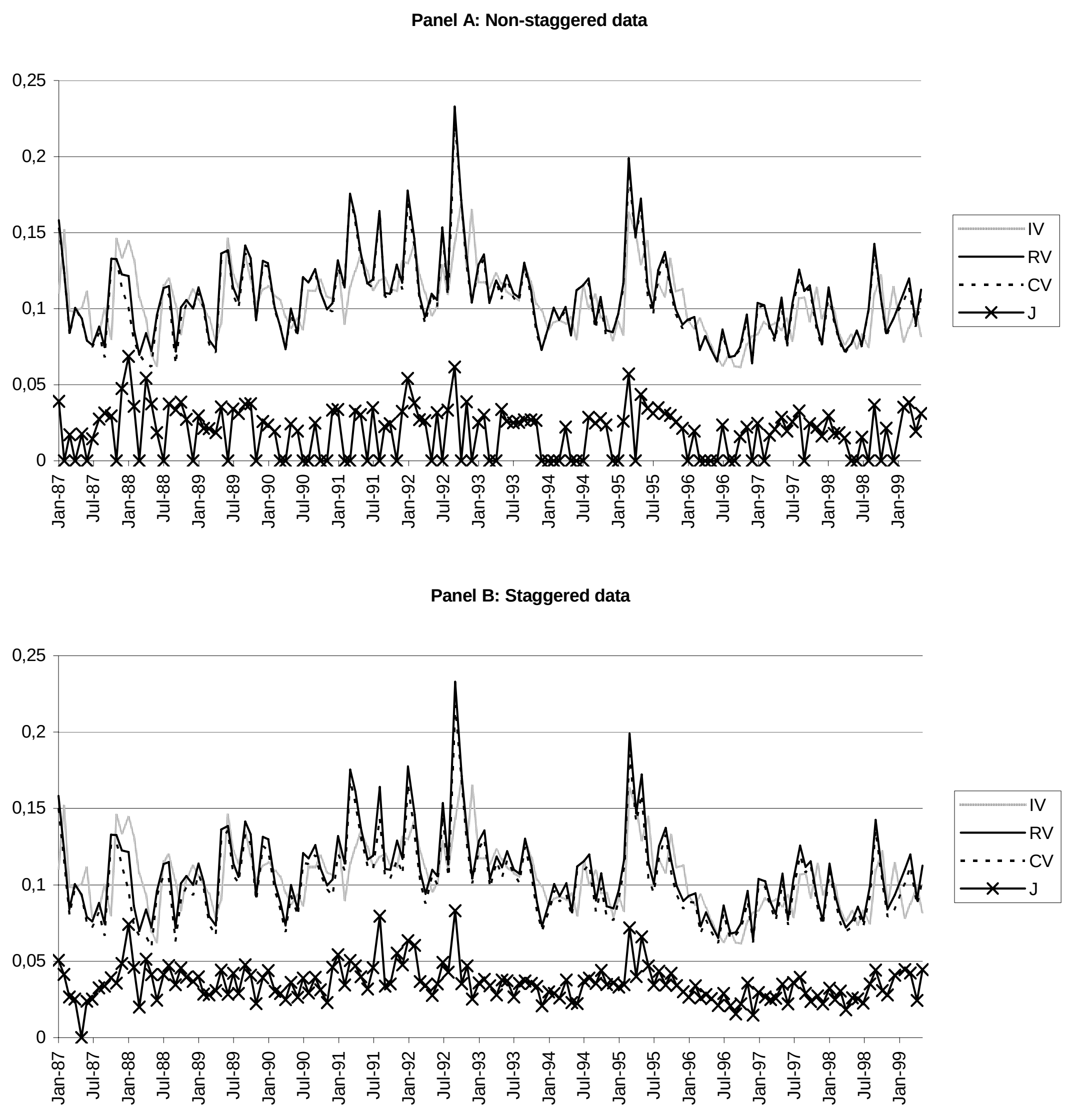
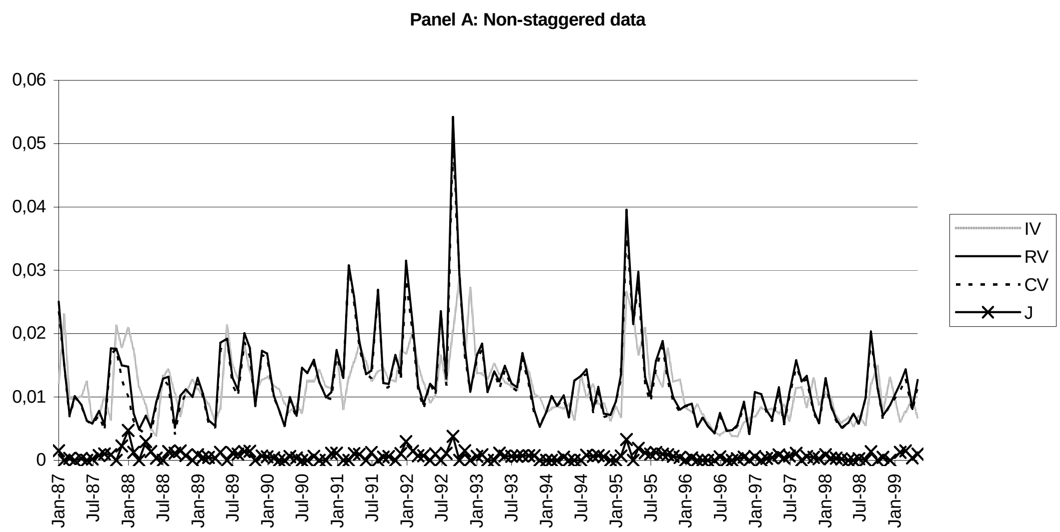


Figure 2: Time series plots of volatility measures in standard deviation form



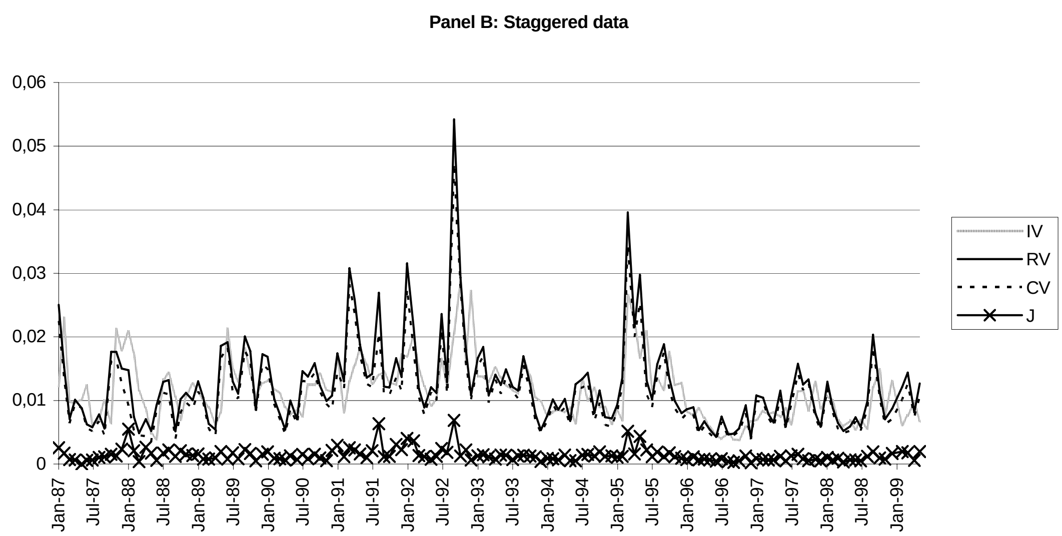


Figure 3: Time series plots of volatility measures in variance form