News - good or bad - and its impact on volatility predictions over multiple horizons*

Xilong Chen† **Eric Ghysels**‡

First Draft: April 2006
This Draft: August 9, 2009

Abstract

We examine whether the sign and magnitude of intra-daily returns have impact on expected volatility the next day or over longer future horizons. We first let the 'data speak', namely with minimal interference we capture the mapping between intra-daily returns and future volatility. We revisit the concept of news impact curves introduced by Engle and Ng (1993). Overall, we find that moderately good (intra-daily) news reduces volatility (the next day), while both very good news (unusual high intra-daily positive returns) and bad news (negative returns) increase volatility, with the latter having a more severe impact. The asymmetries disappear over longer horizons. We also introduce a new class of parametric models which feature asymmetries and with close ties to ARCH-type models, albeit applicable to a mixture of high and low frequency data. Models featuring asymmetries dominate, especially during the 2007-2008 financial crisis.

---

*We like to thank Oliver Linton for comments and sharing with us software. In addition, we like to thank the Referees and the Editor for many helpful suggestions and comments on a previous version of the paper. We also like to thank Robert Engle, Ron Gallant, Benoît Perron, Eric Renault and Neil Shephard for providing helpful comments as well as seminar participants at the CIREQ-CIRANO conference, Columbia University, the New York Fed Research Department, MIT Sloan School of Management, the NYU Volatility Institute Inaugural conference, Princeton University, Stanford University, the University of Chicago Stevanovich Center for Financial Mathematics conference on volatility and UNC. Part of the work was done while the second author was resident scholar at the Federal Reserve Bank of New York, whose support is gratefully acknowledged. Last but not least we would like to thank Michael Sockin for outstanding research assistance.

†Department of Economics, University of North Carolina at Chapel Hill, Gardner Hall CB 3305, Chapel Hill NC 27599-3305, e-mail: `Xilong.Chen@sas.com`

‡Department of Finance, Kenan-Flagler Business School, and Department of Economics, University of North Carolina at Chapel Hill, email: `eghysels@unc.edu`

1 Introduction

After more than two decades of research on volatility forecasting it is still the case that, while most researchers agree that volatility is predictable, there is considerable disagreement on how volatility predictability should be modeled. Persistence of volatility is a well established stylized fact and remains the salient feature of many models. Other empirical regularities, however, are not so commonly embraced in volatility forecasting. One prominent example is the observation that equity returns and volatility are negatively correlated. The phenomenon can be attributed either to a leverage effect, or a volatility feedback effect. Both effects entail different causality patterns: the leverage effect has return shocks affecting conditional volatility, whereas the feedback effect emphasizes the impact of volatility on time-varying equity risk premia. Various papers have documented the empirical evidence and/or disentangled the causal effects, see e.g. Black (1976), Christie (1982), French, Schwert, and Stambaugh (1987), Nelson (1991), Schwert (1990), Campbell and Hentschel (1992), Engle and Ng (1993), Bekaert and Wu (2000), and Bollerslev, Litvinova, and Tauchen (2006) among others. The volatility models that exploit this phenomenon are typically called asymmetric, since 'good' and 'bad' news have a different impact, and to capture the asymmetry Engle and Ng (1993) introduced the notion of *news impact curve*, both as an object of economic interest and a diagnostic tool for volatility modeling.

In recent years, the literature on volatility modeling has put great emphasis on the wide availability of intra-daily financial returns data. Contrary to a decade ago, data of every stock market transaction is now readily available and it enables one to measure so called realized volatility - a post mortem sample realization of the increments in quadratic variation of an underlying price process. Therefore, *measurement* rather than *modeling*, has been the prime focus in the past decade. Measurement, as it turns out, is not easy, as transactions may be affected by microstructure noise and quadratic variation increments may contain a jump component which one might want to extract. The simple measures of realized volatilities involve the intra-daily sum of high frequency *squared* returns. More sophisticated measurements that separate jumps or account for microstructure noise, one way or another, are also based on squared returns.

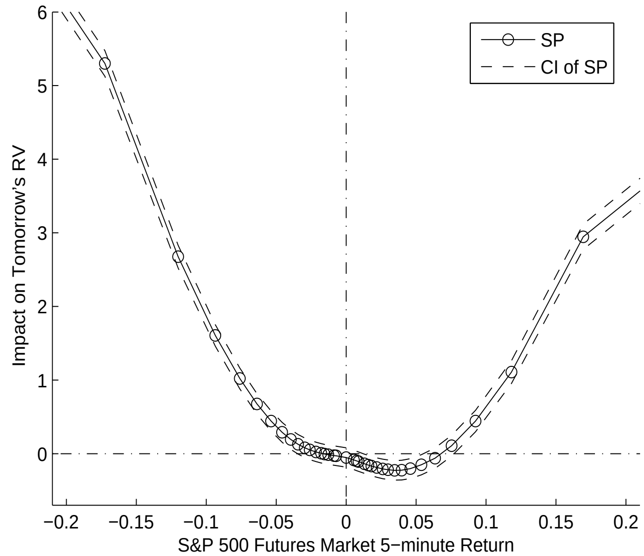
The volatility prediction models that use the new more sophisticated measures of volatility are simple - i.e. usually linear regressions are used involving past realized volatility possibly corrected for jumps or microstructure noise (see Andersen, Bollerslev, and Diebold (2007) and the many references therein). With the focus shifted towards measurement, it is indeed the case that leverage has no impact on the in-sample asymptotic analysis that was developed against the backdrop of increasingly available high frequency financial data (see Jacod (1994), Jacod (1996) and Barndorff-Nielsen and Shephard (2007)). The observation that leverage does not affect measurement appears to have given credence to the fact that asymmetries do not matter for forecasting. It is worth recalling that originally, news impact curves were formulated within the context of daily ARCH type models. One may therefore wonder whether it is because daily data was used, that leverage mattered, and that the use of high frequency now has nullified the issue. The paper shows that asymmetries still matter a lot. To show this, we make various contributions to the existing literature.

It is not obvious how we would go about answering the question whether the sign and magnitude of discretely sampled high frequency returns have any impact on future volatility predictions. First, the raw input is a return over a short interval and the prediction period is not the next short interval, but rather some arbitrary future period - say the next day, week, etc. The mismatch of observation frequency and prediction horizon brings about issues that cannot be handled easily by simple linear models or ARCH-type models. Then there is also the pervasive intra-daily seasonality that prevents one from putting each high frequency interval on equal footing.

We first let the 'data speak', namely with minimal interference we capture the mapping between returns over short horizons and future volatility over longer horizons. To cut straight to the main point, consider an illustrative example of our findings. We take five minute returns on the S&P 500 futures market as the primitive input, and the next day's realized volatility as the future outcome of interest - hence we are thinking along the lines of Engle and Ng (1993) but without a daily volatility model. The typical picture for one day ahead (ignoring the intra-day effects) that emerges from our analysis appears below. The X-axis measures 5-minute returns in the S&P 500 futures market - scaled to annual units. The Y-axis is their expected impact on tomorrow's volatility (with confidence bands) - scaled by average daily realized volatility. The pattern that emerges is interesting. Good news *reduces* tomorrow's expected volatility, i.e. the expected impact dips below zero.[1] In contrast, *very good news* tends to increase volatility, as does *bad* news. This asymmetric pattern has been

[1] For the purpose of clarification we should note that this negative impact pertains to one single five minute interval and the total impact over one or several days is a weighted sum of every five minute's impact. This brings about non-negativity issues regarding the prediction of future volatility - which will be discussed in the paper.

recognized in the past, notably by Engle and Ng (1993). However, here we can carry this further across different horizons using high frequency intra-daily data.



The above plot also reveals that we are essentially dealing with two issues: mis-specification and aggregation. Mis-specification, because measures of quadratic variation are based on squared returns, while the above plot tells us that response functions are not symmetric. Aggregation, because we will build models using high frequency data directly, while all existing models square intra-daily returns and add them up to daily realized variance.

Technically speaking, we start with a semi-parametric MIDAS regressions, inspired by recent work on MIDAS regressions, in particular in the context of volatility as in Ghysels, Santa-Clara, and Valkanov (2006) and Forsberg and Ghysels (2006). The approach adopted in the paper generalizes the original MIDAS regression setup in two important ways. First, it allows for a nonparametric specification of the news impact curve, as depicted in the illustrative plot, and second, a multiplicative scheme is introduced that handles high frequency data well and extends the aforementioned existing MIDAS regression papers by incorporating intra-daily trading patterns. More importantly, we introduce various new parametric models applicable to intra-daily returns, that are inspired by the asymmetric (daily) GARCH models. These simple parametric models do out-perform standard models based on daily realized measures. For example, when we study two specific out-of-sample forecast excercises, one during the relatively calm 2005-2006 low volatility period and the other during the 2007-2008 crisis, we come to the following conclusions. For the calm low volatility period, the standard symmetric models and the new models featuring asymmetries fare equally well. During the financial crisis, however, the out-of-sample forecasting gains of asymmetric models in comparison to the standard ones range from roughly 5 % to as high as 20 % in terms of mean squared error - depending on the series and forecasting horizon. Formal statistical tests show that these gains are statistically significant.

The paper is organized as follows. In sections 2 and 3 we introduce various new models, some involving daily realized measures, some based directly on intra-daily returns in the context of news impact curves and volatility prediction. Empirical results are reported in section 4. Section 5 concludes the paper.

2 Volatility Measurement and Model Specification

In a first subsection we briefly review recent advances in volatility measurement, whereas the second subsection covers the models of recent vintage that use high frequency data volatility measurements. The prime purpose of this section is to provide a succinct overview of existing models. At the end of the section, however, we do introduce a new model using daily realized volatilities that capture asymmetries.

2.1 Measurement and Leverage

The recent vintage of volatility models are based on so called Realized Measures of which Realized Variance (RV) - measured by the sum of squared intra-daily returns - is the most prominent example. More specifically, we think of returns over some short time intervals, say $i = 1, \ldots, 1/\Delta$, on day t. To fix notation, let $r_{t,i\Delta}$, denote the high-frequency return over interval i of length Δ on day t, where $r_{t,i\Delta}$ is the log asset price difference over the interval. Such high frequency intra-daily returns are used to compute Realized Variance, namely:

$$RV_{t,t+k}(\Delta) \equiv \sum_{j=t}^{t+k} \sum_{i=1}^{1/\Delta} r_{j,i\Delta}^2 \tag{2.1}$$

for $k = 1, \ldots$ We will typically drop Δ as an argument of RV to simplify notation.[2] Realized Variance may include jumps and the separation between the continuous path part of integrated volatility (the population counterpart of RV) and the jump component is sometimes taken into account when formulating a prediction model for volatility. The non-jump component is estimated via another Realized Measure, the so called bi-power variation, or:

$$BPV_{t,t+k}(\Delta) \equiv \sum_{j=t}^{t+k} \sum_{i=2}^{1/\Delta} |r_{j,i\Delta}||r_{j,(i-1)\Delta}| \tag{2.2}$$

Without elaborating on the specifics of tests for jumps it suffices to say that we follow Andersen, Bollerslev, and Diebold (2007) who use Hausman-type tests involving RV and BPV. It is important to note, however, that such tests are purely symmetric, i.e. the test statistics involve even powers of intra-daily returns.

The term *leverage* has different meanings, depending on the literature - whether it is corporate finance, empirical asset pricing or mathematical finance. The foundation of realized volatility modeling is the continuous time semi-martingale stochastic process theory, more specifically stochastic volatility continuous time jump-diffusions. In this context, leverage has a specific meaning, as it is defined by the correlation between the Brownian motion that drive returns and the one that drives volatility dynamics (assuming a single factor structure). If we think of asymptotics in terms of sampling at ever finer intra-daily intervals (i.e. $\Delta \to 0$), it has been shown that such leverage does not affect the measurement of realized volatility (see inter alia Jacod and Protter (1998) and Barndorff-Nielsen and Shephard (2007)).[3]

[2] Microstructure noise may mask the true price variation. Various corrected measures of RV have been suggested. See for example Aït-Sahalia, Mykland, and Zhang (2005), Bandi and Russell (2006), Bandi and Russell (2005), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009) and Hansen and Lunde (2006). In our empirical work we consider series for which it is reasonable to assume that they are sampled without microstructure noise - more on this in Section 4. We therefore do not elaborate on the subject in the development of our models.

[3] The standard no-arbitrage semimartingale setting subject only to weak technical regularity conditions suffices to establish this result. Recent work by Barndorff-Nielsen, Kinnebrock, and Shephard (2008) - discussed later - provides a discussion of the technical conditions.

2.2 Forecasting Volatility and Leverage

We are interested in predicting tomorrow's volatility, and consider the regression model:

$$RV_{t,t+1} = \psi_0 + \sum_{j=0}^{\tau} \psi_j(\theta) RV_{t-j-1,t-j} + \varepsilon_{t,t+1} \tag{2.3}$$

Within the class of linear models the specification in (2.3) is quite generic, as it covers models considered by Andersen, Bollerslev, and Diebold (2007), Ghysels, Santa-Clara, and Valkanov (2006), Forsberg and Ghysels (2006), Corsi (2009), among many others. The various approaches differ with respect to the parametrization of $\psi_j(\theta)$. For example, the so called HAR-RV of Andersen, Bollerslev, and Diebold (2007), is:

$$RV_{t,t+k} = \psi_0 + \psi_D RV_{t-1,t} + \psi_W RV_{t-5,t} + \psi_M RV_{t-22,t} + \varepsilon_{t,t+k} \tag{2.4}$$

which predicts future volatility via a linear regression on past daily, weekly and monthly RV. The fact that RV may contain jumps has led to an alternative specification, called HAR-RV-J, as it separates the jump component:

$$RV_{t,t+k} = \psi_0 + \psi_D BPV_{t-1,t} + \psi_W BPV_{t-5,t} + \psi_M BPV_{t-22,t} + \psi_J J_t + \varepsilon_{t,t+k} \tag{2.5}$$

where J_t equals one when day t has a jump and is zero otherwise.

The HAR-RV and HAR-RV-J models share one common feature, namely symmetry. The sign of intra-daily returns does not matter, since both RV is symmetric and tests for jumps are based on even power transformations of intra-daily returns as well. There are some notable exceptions in the recent literature that have tried to accommodate asymmetries, including Barndorff-Nielsen, Kinnebrock, and Shephard (2008) and Engle and Gallo (2006). Both consider some form of 'signed' daily variances, i.e. variance measures multiplied by a sign indicator function. In particular Barndorff-Nielsen, Kinnebrock, and Shephard (2008) introduce the notion of semi-variance:

$$SemiV^{+}_{t+k,t} \equiv \sum_{j=1}^{k} \sum_{i=1}^{1/\Delta} r^2_{j,i\Delta} \mathbf{1}_{\mathbf{r_{j,i\Delta}>0}} \tag{2.6}$$

and $SemiV^{-}_{t+k,t} \equiv RV_{t+k,t}$ - $SemiV^{+}_{t+k,t}$.[4]

2.3 Benchmark Models

The empirical analysis will involve formal statistical tests to assess whether the new models we are about to introduce significantly outperform the state-of-the-art models (2.4) and (2.5) which will be chosen as benchmarks. To raise the bar in terms of benchmark even higher, we also introduce a new model that includes asymmetries in terms of semi-variances. We call this new model HAR-S-RV-J, which is specified as:

$$\begin{aligned} RV_{t,t+k} &= \psi_0 + \psi_D^{+} SemiV^{+}_{t-1,t} + \psi_D^{-} SemiV^{-}_{t-1,t} \\ &\quad + \psi_W^{+} SemiV^{+}_{t-5,t} + \psi_W^{-} SemiV^{-}_{t-5,t} \\ &\quad + \psi_M^{+} SemiV^{+}_{t-22,t} + \psi_M^{-} SemiV^{-}_{t-22,t} + \psi_J J_t + \varepsilon_t \end{aligned} \tag{2.7}$$

so that past daily, weekly and monthly RV are split into semi-variances, and hence are allowed to have an asymmetric response. As noted before, this model is new to the literature, yet it is a natural extension of existing specifications and it is a natural benchmark to compare with the models discussed next. It has one drawback, namely ideally we would like to split BPV instead of RV into semi-BPV measures so that the HAR-S-RV-J model nests the HAR-RV-J one. Unfortunately, defining BPV as a semi-bi-power realized measure is not straightforward. We keep the jump component, nonetheless, as the prior literature has emphasized the importance of jumps.[5] In fact, this shows the limitations of dealing with jumps and asymmetries through aggregate measures. Our approach will circumvent these limitations.

[4] To facilitate the presentation, we use the following indicator process: $\mathbf{1_A}$ which is one when A is true, and equals zero otherwise.

[5] Barndorff-Nielsen, Kinnebrock, and Shephard (2008) discuss how to separate the jump component in positive and negative jumps - something we did not pursue here.

3 A new class of high frequency data volatility prediction models

Consider the prediction of tomorrow's volatility. The models considered in the previous section used past realized volatility, bi-power, jumps and semi-variances as information. The models in this section will formulate a prediction for tomorrow's volatility in terms of all the high frequency returns of today and all previous days as information. Hence, we will *not* aggregate high frequency returns to daily RV measures. Instead, we will use them directly as regressors, for the purpose of forecasting future daily, weekly or monthly volatility. Note two important issues, namely (1) we gain information since we do not aggregate intra-daily returns and (2) we do not impose the quadratic variation transformation - that is squared intra-daily returns - but instead let the regression fit decide what impact news - five minute returns - has on future volatility.

The new class of models relate to several recent contributions pertaining to MIDAS-type regressions, i.e. involving so called mixed data sampling. Examples of MIDAS regressions for the purpose of volatility forecasting include Ghysels, Santa-Clara, and Valkanov (2006) and Forsberg and Ghysels (2006), Becker, Clements, and White (2007) and León, Nave, and Rubio (2007), among others. The approach adopted in the paper generalizes the original MIDAS regression setup in two important ways. First, we introduce semi-parametric MIDAS regressions that will allow us to estimate non-parametrically a news impact curve introduced in the context of ARCH-type models by Engle and Ng (1993). The regressions are semi-parametric because they contain a parametric temporal dependence structure and a non-parametric news impact. Second, the parametric specification of the temporal dependence allows explicitly for intra-daily seasonal patterns, something the prior literature on MIDAS regressions so far ignored.

To predict future volatility with past high-frequency returns, we propose the following regression model (again for simplicity restricting ourselves to a single day forecast):

$$RV_{t,t+1} = \psi_0 + \sum_{j=1}^{\tau} \psi_j(\theta)[\sum_{i=1}^{1/\Delta} \psi_i(\theta) NIC(r_{t-j,i\Delta})] + \varepsilon_{t,t+1} \tag{3.1}$$

where $\psi_j(\theta)$ and $\psi_i(\theta)$ are polynomial lag structures parameterized by θ and $NIC(.)$ is the news impact curve. The regression in (3.1) tells us that each intra-daily return $r_{t-j,i\Delta}$ has an impact on future volatility measured by $NIC(r_{t-j,i\Delta})$ and fading away through time with weights characterized by $\psi_j(\theta)$ (daily) and $\psi_{ij}(\theta)$ (intra-daily). We will refer the above regression as the SP model, since it is a semi-parametric setup.

Note that when the NIC is a quadratic function and $\psi_i(\theta)$ is constant throughout the day, we recover the regression model appearing in (2.3).

We elaborate first on the news impact curve and then discuss the polynomial lag specification. A subsection is devoted to each topic. Before we do, we need to make an observation about comparisons with the models of the previous section. Indeed, some models were based on the separation of increments in quadratic variation into a jump and continuous path component. Such decomposition does not feature in the above regressions as they are rooted in a discrete time modeling approach similar to ARCH. Obviously, extreme high frequency returns - perhaps due to jumps - and their impact will be measured by the behavior of the NIC. Positive and negative jumps may therefore have an asymmetric impact, due to the asymmetry of the curve. This might be viewed as an advantage considering the fact that the discussion in the previous section revealed it is not straightforward to decompose RV into semi-variances and at the same time account for jumps.

3.1 News Impact Curves - Parametric and Non-Parametric

Which shape does an intra-daily news impact curve have? While there is considerable amount of evidence using daily data, we are - as far as intra-daily returns are concerned - in unchartered territory. It is therefore natural to start with a minimal amount of a priori restrictions - the reason why we start with a non-parametric approach. As is typically the case, there is a serious drawback when we simply let the 'data speak'. The estimation procedure is cumbersome, and convergence rates are slow - which means lots of data is required to estimate a NIC. In the present context, it is however, warranted as a starting point. The non-parametric estimates will provide useful guidance for the second approach, which is purely parametric and relies heavily on the existing ARCH literature. It is the purpose of this subsection to discuss both the non-parametric approach and parametric specifications of intra-daily news impact curves.

The non-parametric approach is much inspired by the recent work of Linton and Mammen (2005) who propose the semi-parametric ARCH(∞) model. The difference between the semi-parametric ARCH(∞) and the regression appearing in (3.1) is the mixed data sampling scheme.[6] Their estimation approach, involving non-parametric estimation of NIC and parametric formulation of $\psi_i(\theta)$ and $\psi_j(\theta)$ will be adapted to a mixed frequency sampling setting.

The asymptotic analysis of semi-parametric MIDAS regressions is summarized in Appendix A. The asymptotic distribution of the estimation procedure has a parametric and non-parametric part. The latter is kernel-based and involves solving a so called inverse problem. Chen, Ghysels, and Renault (2009) show that the mixed data sampling scheme in semi-parametric MIDAS regressions only adds an extra term to the asymptotic variance, i.e. when $\Delta = 1$, the asymptotic distribution collapses to the case covered in Linton and Mammen (2005). The main drawback of estimating NIC non-parametrically is computational complexity. The parametric models discussed in the remainder of this section, in contrast, can be estimated in fraction of time compared to the semi-parametric specifications.

There are several other reasons for looking at parametric models. First, the models we will consider relate to ARCH-type models and hence bridge a new and old literature. Second, formal testing in the context of semi-parametric models is quite challenging while this is not in the case for parametric models. Chen, Ghysels, and Renault (2009) derive asymptotic confidence bands for estimated news impact curves, i.e. $\widehat{NIC}$ - appearing also in Appendix A - but simple hypothesis pertaining to asymmetry are not so straightforward to impose on non-parametric estimates.

The parametric models we introduce are both old and new. Not surprisingly, the parametric specifications are inspired by news impact curves adopted in the ARCH literature. Yet, the models we introduce are not autoregressive and hence not ARCH-type models. Instead, they are within the context of MIDAS regressions and based on various parametric functional forms for the NIC, namely:

- $NIC(r) = br^2$, to which we attach the acronym SYMM. The SYMM model can be regarded as a MIDAS extension of ARCH to the case of high-frequency data. Obviously, the SYMM model cannot capture any asymmetries that appear in the data. Note that equation (3.1) with a quadratic NIC is in one very important way different from the RV-based regression (2.3). The latter aggregates on a daily basis all intra-daily squared

[6] We assume that τ is finite - yet we could easily assume it to be infinite. We also assume that Δ is finite. This is a less innocent assumption particularly with respect to the current literature on measurement of realized volatility.

returns whereas SYMM predicts future volatility without the aggregation of intra-daily squared returns - potentially attributing different weights throughout the trading day (the specification of the polynomials will be discussed in the next subsection).

- Inspired by the GJR model proposed by Glosten, Jagannathan, and Runkle (1993), we consider the ASYMGJR model with $NIC(r) = (br^2 + cr^2)\mathbf{1_{r<0}}$.

- Another possible way to allow for asymmetric effects is via a location shift, as in the Asymmetric GARCH model in Engle (1990), yielding the ASYMLS model with $NIC(r) = (b(r-c)^2)$. We also considered two models which combine the GRJ model and Asymmetric GARCH model: ASYMC1 model, $NIC(r) = (b1_{r-d<0}(r-d)^2 + c1_{r-d\geq 0}(r-d)^2)$; and ASYMC2 model, $NIC(r) = (b1_{r<0}(r-d-e)^2 + c1_{r\geq 0}(r-d)^2)$. Due to space limitations we do not report the results as they are roughly similar to the ASYMGJR specification.

3.2 News Impact Throughout the Day

It was noted the SYMM model, i.e. equation (3.1) with a quadratic NIC differs from the RV-based regression (2.3) in an important way, namely that the former does not aggregate throughout the day the intra-daily returns. Inspired by this temporal aggregation argument we design the parametric specification of the polynomial lag structure as:

$$\psi_j(\theta)\psi_i(\theta) = Beta(j, \tau, \theta_1, \theta_2) \times Beta(i, 1/\Delta, \theta_3, \theta_4) \tag{3.2}$$

where the (single) Beta polynomial specification has been used in prior work, notably in Ghysels, Santa-Clara, and Valkanov (2002).[7] Here we accommodate intra-daily patterns according to $Beta(i, 1/\Delta, \theta_3, \theta_4)$ while the daily memory decay is patterned according to $Beta(j, \tau, \theta_1, \theta_2)$. It should first be noted that we impose the restriction that the intra-daily patterns wash out across the entire day, i.e. $\sum_i Beta(i, M, \theta_3, \theta_4) = 1$. We also impose, without loss of generality, a similar restriction on the daily polynomial, which allows us to identify a slope coefficient in the regressions.

The multiplicative specification has several advantages. First, a so called flat aggregation scheme, i.e. all intra-daily weights are equal, yields de facto a daily model with RV when

[7] More specifically: $Beta(k, K, \alpha, \beta) = (k/(K+1))^{\alpha-1}(1-k/(K+1))^{\beta-1}\Gamma(\alpha+\beta)/\Gamma(\alpha)/\Gamma(\beta)$ and $\Gamma(\alpha) = \int_0^{+\infty} e^{-t}t^{\alpha-1}dt$. See also Ghysels, Sinko, and Valkanov (2006) for further discussion.

the news impact curve is quadratic. Or more formally, when $\theta_3 = \theta_4 = 1$, and $NIC(r) = r^2$ we recover RV-based regression appearing in equation (2.3). Second, by estimating $Beta(i, 1/\Delta, \theta_3, \theta_4)$ we let the data decide on the proper aggregation scheme.[8] Obviously, the intra-daily part of the polynomial will pick up how news fades away throughout the day and this - in part - depends on the well known intra-daily seasonal pattern.

Wood, McInish, and Ord (1985), one of the earliest studies employing intra-daily data, documents the well known U-shaped pattern. Much has been written on the topic of seasonality in economic time series (see e.g. Ghysels and Osborn (2001)). Broadly speaking there are two approaches: (1) seasonally adjust series and construct non-seasonal models subsequently, or (2) build seasonal features of the data into the model specification. Intra-daily seasonality has been tackled similarly.[9] The intra-daily seasonal pattern may not be fully captured by the $Beta(i, M, \theta_3, \theta_4)$ polynomial. It appears to work very well empirically, however, and its virtue is that it requires only the estimation of two parameters. Other more complex specifications could be considered - a topic we leave for future research.[10]

To conclude, it should be noted that Ghysels, Santa-Clara, and Valkanov (2006) applied a lag structure that ignored intra-daily seasonality in financial markets, despite it being pervasive. Using a simple monotone MIDAS polynomial, like a single Beta one, is not well suited to exploit the intra-daily data, which is why the results in Ghysels, Santa-Clara, and Valkanov (2006) left the impression that the information in the intra-daily information was not useful. One of the innovations in the current paper is in fact to propose a convenient polynomial specification that enables us to better exploit the intra-daily memory decay pattern.

4 Empirical Results

We analyze four data sets which consist of five-minute intra-daily returns of respectively Dow Jones and S&P 500 cash and futures markets. The data are described in the top

[8] This is a generic issue pertaining to MIDAS regressions and temporal aggregation, see Andreou, Ghysels, and Kourtellos (2009) and Ghysels (2009) for further discussion.

[9] See e.g. Andersen and Bollerslev (1997), Andersen and Bollerslev (1998), Bollen and Inder (2002), Dacorogna, Gençay, Müller, Olsen, and Pictet (2001), Martens, Chang, and Taylor (2002), among others.

[10] We also experimented with 'seasonally adjusted' high frequency returns by demeaning and standardizing as follows $r_{i,t}^{sa} = (r_{i,t} - \overline{r}_i)/s_i$, $i = 1, \ldots, 1/\Delta$, $t = 1, \ldots, T$, where $\overline{r}_i = 1/T \sum_{t=1}^{T} r_{i,t}$ and $s_i = ((1/(T-1) \sum_{t=1}^{T} (r_{i,t} - \overline{r}_i)^2)$. This was only in rare cases an empirically more appealing approach and we therefore do not report the results.

panel of Table 1. The samples of Dow Jones cash and futures markets start in 1993 and 1997 respectively and hence do not include the 1987 crash, while the samples of S&P 500 cash and futures markets start in 1985 and 1982 and therefore include the 1987 crash (for models fitted with S&P 500 data). All samples end December 31, 2008 and therefore include the most recent financial crisis. The 1987 crash, however, will appear in the estimation samples - whereas the 2007-2008 financial crisis will be exploited for the purpose of out-of-sample forecasting. All the models are based on five-minute data, which brings us to the question of microstructure noise. In the case of S&P 500 futures our data sample includes that of Bollerslev, Litvinova, and Tauchen (2006), who document that transactions in the futures market occur on average roughly every 9 seconds. This means that, at least for the futures data, one may safely assume that microstructure effects are negligible, an assumption also underlying the analysis in Bollerslev, Litvinova, and Tauchen (2006). More generally, we also computed signature plots, suggested by Andersen, Bollerslev, Diebold, and Labys (2000), which indicate that 5 minutes appear to be a reasonable sampling frequency for all of our series.[11]

We look at different prediction horizons for future volatility. This will allow us to appraise how asymmetries play out at different horizons. So far we wrote equations predicting RV only one day ahead, and we noted that longer horizons are straightforward extensions. In the empirical work we consider three horizons (1) one day, (2) one week and (3) one month. We start with full-sample estimates, and then we consider various sub-samples. A major concern is that of in-sample over-fitting. We guard against this in two says. First we examine out-of-sample prediction performance for some fixed samples. Second, we also examine rolling sample estimates with repeated out-of-sample predictions. The latter will allow us to test whether forecasting gains from models featuring asymmetries are statistically significant.

In a first subsection we examine estimated news impact curves. The second subsection is devoted to the out-of-sample prediction performance for some fixed samples, notably pertaining to the financial crisis of 2007-2008. The final subsection appraises the statistical significance of the out-of-sample forecasting performance.

[11]Signature plots are omitted but available upon request.

4.1 Is News impact Asymmetric?

We start with the simple question: is there evidence that news has an asymmetric impact when we examine intra-daily data? To answer this question we start with examining the full-sample estimation results of the various models featuring asymmetries introduced in the previous section. This means we look both at the fully parametric specifications, either involving aggregated semi-variances or intra-daily data, and the semi-parametric ones also involving high frequency returns.

The estimation of the parametric models is straightforward, as it involves either OLS (for HAR-RV, HAR-RV-J and HAR-S-RV-J) or nonlinear least squares (for models that use intra-daily returns, namely SYMM, ASYMGJR, and ASYMLS). All the regressions are based on non-overlapping samples of RV, hence the sample sizes for the daily, weekly and monthly horizons are not the same. The advantage, however, is that we do not have to deal with autocorrelation of the residuals. In contrast, heteroskedasticity may still be a concern. A simple test which consists of fitting an AR(1) model to the squared residuals of the regressions shows that heteroskedasticity is present in most regressions with daily and weekly horizons and that heteroskedasticity is not present in the monthly forecast horizon. For the daily and weekly horizons it appears that roughly all models are affected equally. Hence, there is room for improvement roughly equally across all the models. For simplicity, we will proceed with the estimation results, without correcting for heteroskedasticity.

The estimation of the semi-parametric MIDAS regression model also warrants some discussions. It was noted earlier, that the estimation procedure has a parametric and non-parametric part. The latter is kernel-based and involves solving a so called inverse problem. The latter requires - as discussed in Appendix A - a discretization to solve an integral equation. All things considered, many factors affect the estimation results, such as the initial parameters, the lag truncations, the number of grid points for the discretization and the weights of each grid point in numerically solving the integral equation, etc. We learned from experimentation that there appear to be two critical choices that affect the estimation outcome. They are (1) the bandwidth selection for the kernels, and (2) the number of grid points. Regarding the bandwidth, we followed the asymptotically optimal bandwidth selection procedure described in Section 6.1 of Linton and Mammen (2005). In theory, one should choose as many as possible grid choices to appraise robustness; however, in practice, the estimation time is roughly equal to $\Delta T \times \tau/\Delta \times n_g^2$ where ΔT is the sample size, τ/Δ

is the number of lags and n_g is the number of grid points and therefore prohibitively costly with dense grids. In the software used by Linton and Mammen (2005) and kindly supplied by us, the grid number choice is 100 or 200; but they are dealing with low-frequency, daily that is, data. We choose a grid equal to 41, which means the quantile interval is 2.5% and we also tried a grid of 80, which does not improve (or make worse) the results we obtianed significantly. As a result, the estimation time is roughly about 20 hours for PC with P4 2.4G CPU and 1GB of memory.

Since we have four volatility series, we consider four news impact curves in Figures 2 and 3 corresponding to two horizons: one-day and one-week ahead. To facilitate the interpretation and display of results, the news impact is scaled by the mean of RV for each series/horizon - so that the news impact is in terms of fraction of average horizon volatility. Unlike the plot appearing in the Introduction, we consider now four series instead of a single one. For the moment it suffices to look at the dotted lines in each of the figures, which represent the news impact curves obtained via the semi-parametric estimation. It is remarkable to note how similar the shapes are for the SP model across the four different series at the daily forecasting horizon and to a lesser extend at the weekly horizon. The asymmetry of the news impact is obvious from the plots. Negative and positive returns have a different impact. The finding that so called *no news is good news* extensively documented in the literature using daily returns implies that the minimum of the news impact curve is at zero. Instead, with intra-daily data we find in Figures 2 and 3 that the intra-daily news impact curves attain their minimum at some mildly positive return, meaning that such returns result in decreased volatility the next day/week (since the impact is negative).[12] As noted in the Introduction, we also recognize the fact that extremely 'good news', the positive returns (it turns out those larger than the 90% quantiles), cause *increased* future volatility. Finally, as noted earlier, 'bad news' has a more acute impact than positive news.

The units of the plots deserve some attention too. The X-axis measures annualized intra-daily returns.[13] The Y-axis are in terms of fractions of average volatility - namely we scaled

[12] The shape of the news impact curve should bring us back to the issue of positivity constraints. It is important to note that each and every day has many five-minute intervals, and for some $NIC(.)$ is positive, whereas for others the functional yields a negative value. As far as positivity is concerned, what matters is the final model prediction which compounds all the high frequency intervals - so the fact that the function dips below zero over a single interval is not of major concern, as long as the sum of all weighted functionals of five minute returns remains positive. In none of our empirical examples did it ever happen that predictions yielded negative volatilities. The non-negativity constraint that also appears in the Linton and Mammen code supplied to us was never binding in-sample.

[13] The annualized returns calculations are obtained through a simple re-scaling.

the news impact by the mean of the relevant average RV for each series/horizon. Hence, we can actually compare the impact across series and horizons. Here too the similarity is quite striking for the daily horizon - but less so for the weekly horizon. Annualized intra-daily returns less or equal to 10 % yield volatility impacts that appear statistically insignificant except for the S&P 500 futures series, where the impact is statistically significant at the daily horizon. Namely, the 95 % asymptotic confidence intervals around the news impact curves tell us that the dip below zero is statistically significant in the case of S&P 500 futures at the daily horizon. At the weekly horizon, the statistically significant negative impact is more marked, except for the S&P 500 cash series. Returns larger than the 10 % annualized return tend to move volatility up by a fraction equal to between half and four to five-fold average volatility both at the daily and weekly horizon. Are these reasonable numbers? One has to keep in mind that the curves measures the impact of a single five minute interval and every day features 78 such intervals, each with their own impact that fades away across time. Hence the units, while comparable across series are not easy to fully comprehend. What matters most about the exercise is the fact that the figures show clearly the presence of asymmetries. Whether the actual news impact is reasonable can only be assessed by looking at forecasting performance - the topic to which we turn our attention in the next subsection.

While the plots obtained for the semi-parametric estimation are insightful, it is also worth looking at the parameter estimates of models that feature asymmetries and test whether those are significant. Indeed, the parametric models have the advantage that we can test the significance of asymmetries via restrictions on the parameters instead of eyeballing curves. We start with the regression model appearing in (2.7) involving semi-variances. This first model, called HAR-S-RV-J and new to the literature, involves past daily, weekly and monthly RV in the original HAR-RV-J model split into semi-variances. Hence, this first model tells us whether at the *aggregate* level the much neglected asymmetries matter. The results appear in Table 2. The top panel reports tests of the single restrictions that either the daily, weekly, or monthly semi-variances have equal coefficients. The final column contains a joint test of the null with all three restrictions. The asymptotic variances to compute the tests used HAC standard errors with lag lengths that match the inference recommendations of Andersen, Bollerslev, and Diebold (2007). The daily lag coefficients do not seem to feature asymmetries for the four series for predicting future daily volatility, yet it does matter for weekly and monthly forecasts. The weekly lag coefficient only appears to feature asymmetries for predicting future weekly volatility and monthly semi-variances also matter at all but the daily forecast horizon. The final column reveals that asymmetries are present at all forecast horizons, testing all regressors jointly, except for the DJ cash series.

In subsequent tests we will use both the HAR-RV-J and HAR-S-RV-J models as benchmarks. The statistical evidence does imply that the HAR-S-RV-J model is indeed a more stringent benchmark as it captures some of the much neglected asymmetries through daily measures. The tests will therefore reveal whether asymmetries at the intra-daily level matter even more.

Last but not least we turn our attention to Panel B of Table 2, which pertains to parametric models using intra-daily returns and featuring asymmetries. The panel displays the parameter estimates of the lag polynomials involving parameters θ_1 through θ_4, appearing in equation (3.2). The parameters a, b, and c, determine the news impact, where c relates to asymmetries, since for the ASYMGJR model $NIC(r) = (br^2 + c1_{r<0}r^2)$, and for the ASYMLS model $NIC(r) = (b(r-c)^2)$. To save space, the parameter estimates are reported for one representative series, the S&P 500 Futures contract, at one day and one week forecast horizons (more detailed results are available upon request). The evidence of intra-daily asymmetries is overwhelming, judged by the significance of the parameter estimates of c in all the models reported.

We report one final intriguing result in Figure 4. It displays the news impact - parametric and non-parametric - for the one week and one month horizon. For the weekly horizon the two approaches match up reasonably well, except perhaps for the ASYMGJR model. For the monthly horizon they do too, but what is most remarkable is that the curves are symmetric. Hence, over longer horizons asymmetries fade away. We will note later that this indeed matters for forecasting.

The results, so far, clearly indicate that asymmetries matter - whether we look at plots, models using semi-variances, or parametric models that apply directly to intra-daily data. Yet, there appears to be some disagreement with regards to the horizon effect. The evidence from the HAR-S-RV-J model seems to suggest that asymmetries are more important at longer horizons. All other models suggest the opposite - namely at longer horizons asymmetries fade away.

Besides the news impact curves, we need to discuss the parametric part of the semi-parametric MIDAS regression, or more specifically the Beta polynomials appearing in equation (3.2). We plot only one of the four examples, namely the S&P 500 Futures example. There are three plots that appear in Figure 5. The first plot displays the product of the daily and intra-daily lags, hence it contains the profile of the coefficients ψ_{ij}. The second plot displays only the daily coefficients ψ_j and finally the intradaily coefficients ψ_i appears in the third plot. The patterns are not surprising, given the abundant evidence documented in the empirical volatility literature. The daily coefficients decrease monotonically and are close to zero after 6 to 8 days. The intra-daily weights display a somewhat asymmetric U-shaped pattern, perhaps best characterized as a smirk. It means that late afternoon returns, carry relatively more weight that morning returns. The product of the two provides a spiky decay pattern compounding the intra-daily and daily response.

4.2 Out-of-sample Forecasting Performance

We start with a relatively informal statistical analysis. In the next subsection we turn to more formal out-of-sample forecasting evaluations. More specifically, in the next subsection, we will use rolling sample estimates, hence constantly updating the parameter estimates, and perform repeated forecasts out of sample to construct formal statistical tests. Such an exercise is computationally demanding and will therefore exclude the semi-parametric specification of the news impact curve - since the estimation of the latter is prohibitively time-consuming. In this subsection, we therefore focus on two specific out-of-sample forecast comparisons. The advantage is that we can include the computationally intensive semi-parametric approach.

We take a look at two interesting samples, the first ends in 2004, the second in 2006. Why are they interesting? All the models are estimated with at the end of the estimation sample two years of data set apart to perform out of sample forecasting (Table 1 lists the sample configurations). Hence, for the sample ending in 2004 we predict out of sample for 2005-2006 - a relatively calm period as far as volatility goes. The second sample ends right before the financial crisis and parameter estimates are then used to predict during the financial crisis.

Since the focus of the paper is on asymmetries, it is worth first of all taking a look at the two out-of-sample periods, namely the 2005-2006 and 2007-2008 samples. In Figure 1 we plot the histograms - throughout the day in different time zones of the trading day - of the 5 minute returns. The histograms reveal that both samples features roughly symmetrically distributed high frequency returns, with the financial crisis featuring much larger returns, both positive and negative. Hence, the 'news' in both out-of-sample periods is roughly equally positive and negative. During the financial crisis, however, the returns are much larger. The purpose of this comparison is to see how our models cope with the impact of this shift in volatility.

To formalize the discussion, let us denote by $\widehat{RV}^m_{t,t+k}$ the forecast of volatility for horizon k with information up to time t, using model m. To recap, the models using intra-daily data directly are, $m = SP$ for semi-parametric MIDAS, and parametric models SYMM, ASYMGJR, and ASYMLS and models that use daily volatility measures HAR-RV, HAR-RV-J, and HAR-S-RV-J. Suppose we denote the size of the estimation sample T_1 and adding the two-year out-of-sample yields size T_2. For each model we compute out-of-sample the error:

$$u^m_{t,t+k} = RV_{t,t+k} - \widehat{RV}^m_{t,t+k}$$

then the two-year out-of-sample MSFE at the k-horizon is:

$$MSFE^m_k = \frac{1}{N_k}\sum_{i=1}^{N_k}(u^m_{t+(i-1)k,t+ik})^2. \tag{4.1}$$

where N_k is the number of non-overlapping intervals of length k in the out-of-sample period T_2 - T_1. The sample MSFE is computed for each forecasting horizon k and for each forecasting method m. The MSFE is used for the rolling sample estimates (reported in the next subsection) as well as the full-sample ones.[14]

In this subsection, We focus on the one-day and one-week prediction horizons, since the forecast evaluation is based on two years we do not include the monthly horizon due to the small sample size. In the next subsection we will examine monthly horizons with a statistically more elaborate procedure. Table 3 contains one day ahead forecasts for both parametric and semi-parametric model specifications. Note first of all that the differences in out-of-sample forecast performances for the 2005-2006 are small. This is a low volatility era and all competing models produce roughly similar forecasts. Recall that for the moment we do not address whether the differences are statistically different. For the 2007-2008 sample there are two striking observations (a) the general level of MSFE is about 500 times larger (or roughly 20 times in terms of root MSFE) compared to the previous forecasting sample, and (b) the differences between models also are much more pronounced. Overall, all models do poorly, but some do much better nevertheless.

The first striking observation is that the best models - regardless of series or out-of-sample

[14]Patton (2007) showed that when one uses the MSFE criterion, the measurement error in using a proxy (namely the sample RV instead of the true population integrated volatility) will not change the ranking of forecasting methods. This robustness property is not shared by some other popular forecasting evaluation methods, such as the mean absolute forecasting error (MAFE).

period or horizon - always feature asymmetries. Unfortunately, it is not always the same model with asymmetries.

The SP model has the best forecasting performance for daily predictions of the Dow Jones and S&P 500 futures volatility series, in both cases during the financial crisis. It is also the best model for the one-day horizon cash market volatility of the S&P 500 during the financial crisis. This is quite surprising as one would have perhaps expected that in-sample data mining would penalize forecasting performance during unusual events like the crisis. Of course, the differences in performance are sometimes small - and perhaps not statistically significant. This issue will be covered later. Nevertheless, it means that a regression model involving *non-parametric* estimation of a response function applied to high-frequency data, outperforms a fully parametric model involving daily aggregate measures such as RV and even the separation of jumps and continuous path volatility. Typically, well specified parametric models outperform semi-parametric ones. Here, however, the semi-parametric models de facto use more data and are not subject to the pre-specified quadratic transformation of returns.

The most remarkable performance, however, appear to be the ASYMGJR and ASYMLS specifications which are best or nearly best for all one-day horizon predictions during the 2005-2006 period. During the financial crisis it appears to do very well for the one-week horizon - again for various series. However, the HAR-S-RV-J model is quite often the best model too at the weekly horizon, particularly during the financial crisis. Occasionally the SYMM model does well for one-week predictions.

It is also worth noting in Table 3 that models involving intra-daily data typically outperform or perform equally well compared to model using daily aggregates. This clearly shows there are advantages of using the intra-daily information directly while keeping parsimony. It is also worth noting that typically the largest gain comes from the intra-daily asymmetry - not the mere fact that RV is replaced by the symmetric SYMM specification.

The overall picture that emerges can be summarized as follows. We essentially made three types of forecasting performance comparisons. They are (1) semi-parametric with parametric models - both using high frequency data, (2) parametric models that use aggregate measures HAR-RV, and HAR-RV-J - both being implicitly *symmetric* - with a parametric model SYMM using high frequency data directly that are also symmetric, and (3) parametric models that use high frequency data and are *asymmetric* with parametric models also using high frequency data directly but are symmetric. The comparison in (3) deals with parametric mis-specification, i.e. the role played by asymmetry. The comparison (2) tells us something about loss of information due to aggregation and finally (1) reveals mis-specification of the parametric models. Our empirical findings suggest at this point that the forecasting gains are mostly due to asymmetries. The simple asymmetric high frequency data parametric models do out-perform standard models based on daily realized measures. During the financial crisis, the out-of-sample forecasting gains of asymmetric models in comparison to the standard ones range from roughly 5 % to as high as 20 % in terms of mean squared error - depending on the series and forecasting horizon . Formal statistical tests in the next subsection will show that these gains are statistically significant.

4.3 Are the Forecasting Gains Statistically Significant?

In the previous subsection we estimated models and appraised them over a subsequent prediction sample, and we did this for two interesting sample configurations. The out-of-sample comparisons did not involve formal tests - that is we did not provide some type of confidence intervals. To do so, we use a test proposed by Giacomini and White (2006), henceforth denoted GW, which can be viewed as a generalization, or a conditional version of the Diebold and Mariano (1995) and West (1996) tests. Another appeal of using the test is that it can handle non-nested models, which is the case in our application. In fact, Giacomini and White (2006) stress the difference between what they call forecasting methods and as opposed to forecasting models. Loosely speaking forecasting methods are the combination of estimation sample, model specification and prediction sample. This in our application most relevant, as we will not only compare models involving high frequency data directly, or daily aggregate measures, but we will also include ARCH-type models involving daily returns - i.e. the original models that were used in the literature on asymmetries in volatility. Note that original ARCH-type models are estimated via maximum likelihood - in contrast to the models discussed in the previous sections, which are regression-based. Clearly, we are dealing with a plethora of forecasting methods involving different data, models and estimation procedures. For this comparison the GW is most suited.

The GW test is implemented with a 5 year estimation window and 2 year out-of-sample forecasts. The rolling forward is implemented by truncating the earliest month of the data set and adding an additional month to the end, so that the 5 year window rolls one month at a time starting in Jan 1997-Dec 2001. We examine a rolling scheme that ends in 2006 (excluding the financial crisis) and one that ends 2008. The tests involve the sample MSFE appearing in equation (4.1), computed for each forecasting horizon k, and each forecasting model/method m. One picks a benchmark model, for which we will choose two daily realized measure models: HAR-RV-J and HAR-S-RV-J. We will call the benchmark model b. For each horizon k, the empirical efficiency of the forecasts is assessed by comparing the respective MSFEs and testing the following null hypothesis for a benchmark model against any other model m :

$$H_0 : E[\big((u^b_{t,t+k})^2 - (u^m_{t,t+k})^2\big)\,|I_{t-1}] = 0 \tag{4.2}$$

where I_{t-1} is the information set available at time $t-1$. In our setting we use as instruments a constant and $u^b_{t-k,t})^2$, i.e. the most recently available MSFE. Consequently the test statistic will be $\chi^2(1)$.

Table 4 reports the tests of improved predictability with the HAR-RV-J as benchmark model and Table 5 reports the results for the HAR-S-RV-J benchmark. Whenever an entry is boldfaced, the benchmark model is better than the alternative model. More specifically, boldfaced values exceeding the $\chi^2(1)$ critical values of 2.706 (10 %), 3.841 (5 %), 6.635 (1 %) or 10.828 (.1 %) imply that the benchmark model significantly outperforms the alternative model. All other entries exceeding those same critical values mean that the alternative model significantly outperforms the benchmark.

Looking at Table 4 we note that the HAR-RV-J model significantly outperforms the SYMM models at the daily forecast horizon for the 2001-2006 sample, but only for the SPF series when we end in 2008. All other models, including the HAR-S-RV-J outperform the benchmark for the ending 2008, and for most series ending 2006. At the weekly horizon all models typically outperform the benchmark but at the monthly horizon none do. This underscores what we discovered via the estimation of the NIC reported in the lower part of Figure 4 featuring symmetric news impacts at longer horizons.

When we raise the bar to the HAR-S-RV-J, reported in Table 5, we note again that the SYMM models is inferior, whereas in particular the ASYMGJR model featuring asymmetries and using intra-daily data outperforms the benchmark model for the evaluation sample ending 2006. Adding the crisis the evidence weakens the evidence against the ASYMLS specification and the ASYMGJR model remains significantly better at the 5 %. At the weekly horizon this also holds except for the DJF series (ending 2006) and for the the DJC and SPF series when the crisis is added.

The picture that emerges is that asymmetric intra-daily ASYMGJR model outperforms symmetric and asymmetric daily realized measure models - except at the monthly horizon.

Recall that the GW procedure can handle non-nested models, or more general, a diverse set of forecasting methods that are different combinations of estimation sample, model specification and prediction sample. This generality allows us also to include standard daily GARCH models, and compare them with a benchmark model involving high frequency data. Therefore, we conclude this section with a comparison of a symmetric model based on high frequency data, namely the HAR-RV-J model featured as benchmark in Table 4 and the original GJR model of Glosten, Jagannathan, and Runkle (1993) using daily return series for the daily prediction horizon, weekly returns for the same horizon and similar for the monthly one.[15] This implies that, say, daily volatility forecasts are compared with high frequency realized volatility measures - i.e. the same realized series to appraise the other models. The lower panel in the table reports the test statistic results. They clearly show the advantages of using high frequency data, even though the GJR features asymmetries, the daily data is 'too noisy' to provide predictions that outperform the high frequency symmetric benchmark model. The gains achieved by using high frequency data have been abundantly emphasized in the literature. The original daily return GARCH literature emphasized the importance of asymmetries - using daily returns only. The contribution of our paper is to show that asymmetry remains important, but it requires models suitable to handle high frequency data.

5 Conclusions

The paper has two contributions: (1) we show asymmetry matters for volatility forecasting and (2) we show that relatively simple parametric models, sharing features with the original work of Glosten, Jagannathan, and Runkle (1993) but using instead intra-daily returns, outperform models with daily realized variance and semi-variance measures in terms of out-of-sample forecasting - particularly during turbulent volatility periods such as the most recent

[15] This means we deal with direct forecasts, as opposed to iterated forecasts - iterating forward the daily ones - for the weekly and monthly horizon. For further discussion of iterated versus direct forecasts of volatility, see e.g. Ghysels, Rubia, and Valkanov (2009).

financial crisis.

The writing of Engle and Ng (1993) was in part motivated by the recognition that volatility models, including the at time very popular daily GARCH(1,1) model, imposed a symmetric response function of shocks to volatility and that most often such response was probably misspecified. The most preferred model of Engle and Ng (1993), based on their empirical analysis, was that of Glosten, Jagannathan, and Runkle (1993). The main findings of this literature still remains very much part of our core beliefs today regarding the key stylized facts of volatility dynamics. Namely, it is widely believed that *"good"* news and *"bad"* news do not have the same impact on future volatility. This is a theme that resonates in many empirical asset pricing papers, including Campbell and Hentschel (1992), Glosten, Jagannathan, and Runkle (1993), among many others.

Merton (1980) showed that volatility may be measured arbitrarily well through the use of sufficiently finely sampled high-frequency returns. Over the past decade these ideas have been made operational thanks to the wide availability of high frequency financial data. It was shown that better measurement, through high frequency data-based (daily) realized measures, would lead to better predictions of future volatility - and indeed it does as many papers have shown. By construction, these data-driven measures were symmetric, i.e. the sign of (intra-daily) returns did not matter.

Our empirical results clearly show that the findings from the daily volatility models remain important for high frequency data. Moderately good (intra-daily) news reduces volatility (the next day), while both very good news (unusual high positive returns) and bad news (negative returns) increase volatility, with the latter having a more severe impact.

The asymmetry evaporates at longer horizons. Parametric specifications, which bridge the new and old literature, confirm these findings of asymmetry - particularly at shorter horizons - via simple hypotheses imposed on the parameters. Relatively simple asymmetric high frequency data parametric models outperform models with daily realized variance and semi-variance measures in terms of out-of-sample forecasting - particularly during turbulent volatility periods such as the most recent financial crisis, where the out-of-sample forecasting gains of asymmetric models in comparison to the standard ones range from roughly 5 % to as high as 20 % in terms of mean squared error - depending on the series and forecasting horizon . Formal statistical tests show that these gains are statistically significant.

Technical Appendix

A Asymptotic properties of semi-parametric MIDAS regression models

It will be useful to consider a generic setting. In such a setting we consider a regression model involving a low frequency regressand y and regressors x, sampled more frequently with a parametric lag structure for temporal dependence and a nonparametric function $m(.)$. Note that we use a simplified notation, slightly different from the main body in the paper, as it makes the technical results easier to summarize. The regression is written:

$$y_t = \sum_{j=1}^{\tau} B_j(\theta) m(x_{t-1-(j-1)/M}) + \varepsilon_t \tag{A.3}$$

where the residuals ε_t are a martingale difference sequence. The lag coefficients $B_j(.)$, $j = 1, ..., \tau$, are described by a finite dimensional parameter $\theta \in \Theta \subset \mathbb{R}^p$ with $\sum_{j=1}^{\tau} B_j(\theta) = 1$ for identification. Moreover, without loss of generality, we assume $\tau = nM$, $n \in \mathbb{N}$.

Note that this setting easily translate into equation (3.1), namely $y = RV_{t,t+1}$, the returns $r_{t-j,i\Delta}$ are the regressors x, sampled $M = 1/\Delta$ times, the polynomial lag $B_j(\theta)$ is $\psi_{ij}(\theta)$ and finally NIC is the function m. We impose the following key assumption:

Assumption A.1 *The process $\{x_{s/M}\}_{s=-\infty}^{+\infty}$ is stationary; and the processes $\{y_t, x_{t-j/M}\}_{t=-\infty}^{+\infty}$ for $j = 1$, $\ldots$, M are jointly stationary and geometrically α-mixing and $\alpha(k) \leq a\bar{s}^k$ for some constant a and $0 \leq \bar{s} < 1$ when k is big enough.*

The true parameters θ_0 and the true function $m_0(.)$ are defined as the minimizers of the population least squares criterion function

$$S(\theta, m) = E\left[\left\{y_t - \sum_{j=1}^{\tau} B_j(\theta) m(x_{t-1-(j-1)/M})\right\}^2\right] \tag{A.4}$$

Define m_θ as the minimizer of the criterion function for any given $\theta \in \Theta$. A necessary condition for m_θ to be the minimizer of (A.4) is that it satisfies the first order condition

$$E\left[\left\{y_t - \sum_{j=1}^{\tau} B_j(\theta) m(x_{t-1-(j-1)/M})\right\} \sum_{k=1}^{\tau} B_k(\theta) g(x_{t-1-(k-1)/M})\right] = 0 \tag{A.5}$$

for any measurable (and smooth) function g yielding a well-defined expectation. Moreover, the second order condition is $-E\left[\left\{\sum_{k=1}^{\tau} B_k(\theta) g(x_{t-1-(k-1)/M})\right\}^2\right]$. The fact that the latter is negative implies that the solution of the first order condition does indeed (locally) minimize the criterion. The first order condition (A.5) can be rewritten as

$$\sum_{k=1}^{\tau} B_k(\theta)E[y_t g(x_{t-1-(k-1)/M})] \quad - \quad \sum_{k=1}^{\tau}\sum_{j=1,j\neq k}^{\tau} B_k(\theta)B_j(\theta)E[m_\theta(x_{t-1-(j-1)/M})g(x_{t-1-(k-1)/M})]$$
$$= \quad \sum_{k=1}^{\tau} B_k(\theta)^2 E[m_\theta(x_{t-1-(k-1)/M})g(x_{t-1-(k-1)/M})]$$

Taking $g(.)$ to be the Dirac delta function, we have that

$$\sum_{k=1}^{\tau} B_k(\theta)E[y_t|x_{t-1-(k-1)/M}=x] \quad - \quad \sum_{k=1}^{\tau}\sum_{j=1,j\neq k}^{\tau} B_k(\theta)B_j(\theta)E[m_\theta(x_{t-1-(j-1)/M})|x_{t-1-(k-1)/M}=x]$$
$$= \quad \sum_{k=1}^{\tau} B_k(\theta)^2 m_\theta(x)$$

for each x. This is an implicit equation for $m_\theta(.)$ which can be re-expressed as a linear type two integral equation in $L_2(f_0)$, where f_0 is the marginal density of $x_{s/M}$. Define $B_k^*(\theta) = B_k(\theta)/\sum_{j=1}^{\tau} B_j(\theta)^2, k = 1,...,\tau$, and $B_i^+(\theta) = \sum_{k=1}^{\tau-|i|} B_k(\theta)B_{k+|i|}(\theta)/\sum_{j=1}^{\tau} B_j(\theta)^2, i = \pm 1,...,\pm(\tau-1)$. Finally,let $f_{0,j}$ be the joint density of $(x_{s/M}, x_{(s-j)/M})$, then:

$$m_\theta(x) \quad = \quad m_\theta^*(x) + \int H_\theta(x,y)m_\theta(y)f_0(y)dy, \ \textit{or} \ m_\theta = m_\theta^* + H_\theta m_\theta, \tag{A.6}$$

$$m_\theta^*(x) \quad = \quad \sum_{k=1}^{\tau} B_k^*(\theta)E[y_t|x_{t-1-(k-1)/M}=x] \tag{A.7}$$

$$H_\theta(x,y) \quad = \quad -\sum_{i=\pm 1}^{\pm(\tau-1)} B_i^+(\theta)\frac{f_{0,i}(y,x)}{f_0(y)f_0(x)}. \tag{A.8}$$

where the sum in (A.8) runs from i = 1 -τ, ..., τ- 1, excluding 0 (using the same notation as in Linton and Mammen (2005)). Note also that m_0= m_{θ_0}. The general estimation strategy for a given sample $\{\{y_t\}_{t=1}^T, \{x_{s/M}\}_{s=1}^{MT}\}$ is (a) for each θ compute estimators $\widehat{m}_\theta^*, \hat{H}_\theta$ of m_θ^*, H_θ, (b) solve an empirical version of (A.6) to obtain an estimator $\widehat{m}_\theta$ of m_θ and (c) choose $\hat{\theta}$ to minimize the profiled least squares criterion with respect to θ and let $\widehat{m}(x) = \widehat{m}_{\hat{\theta}}(x)$. The practical implementation of the above estimator is basically the same as in Linton and Mammen (2005), and therefore omitted.

Define the functions $\beta_\theta^j(x), j = 1,2$, as solutions to the integral equations $\beta_\theta^j = \beta_\theta^{*,j} + H_\theta\beta_\theta^j$, in which:

$$\beta_\theta^{*,1}(x) \quad = \quad m_\theta^{*\prime\prime}(x),$$
$$\beta_\theta^{*,2}(x) \quad = \quad \sum_{i=\pm 1}^{\pm(\tau-1)} B_i^*(\theta)\left\{E(m_\theta(x_{(s-i)/M}|x_{s/M}=x)\frac{f_0^{\prime\prime}(x)}{f_0(x)}) - \int[\nabla_2 f_{0,s}(x,y)]\frac{m_\theta(y)}{f_0(x)}dy\right\}$$

where the operator ∇_2 is defined as $\nabla_2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$. Then define

$$\omega_\theta(x) = \frac{||K||_2^2}{f_0(x)} var[\eta_{\theta,s} + \zeta_{\theta,s}] \tag{A.9}$$

$$b_\theta(x) = \frac{1}{2}\mu_2(K)\left[\beta_\theta^1(x) + \beta_\theta^2(x)\right] \tag{A.10}$$

Define:

$$\begin{aligned}\omega(x) \quad = \quad & \frac{||K||_2^2 \sum_{j=1}^{\tau} B_j^2(\theta_0) E\left[\varepsilon_t^2 | x_{t-1-(j-1)/M} = x\right]}{f_0(x)\left[\sum_{j=1}^{\tau} B_j^2(\theta_0)\right]^2} \\ & + \frac{M-1}{M} \frac{||K||_2^2 \sum_{j=1}^{\tau}\sum_{k=1,k\neq j}^{\tau} B_j^2(\theta_0) B_k^2(\theta_0)\,\mathrm{var}(m(x_{t+(j-k)/M})|x_t = x)}{f_0(x)\left(\sum_{j=1}^{\tau} B_j^2(\theta_0)\right)^2}\end{aligned} \tag{A.11}$$

$$b(x) = \mu_2(K)\left\{\frac{1}{2}m''(x) + (I - H_\theta)^{-1}\left[\frac{f_0'}{f_0}\frac{\partial}{\partial x}(H_\theta m)\right](x)\right\} \tag{A.12}$$

Let $\varepsilon_t(\theta) = y_t - \sum_{j=1}^{\tau} B_j(\theta) m_\theta(x_{t-1-(j-1)/M})$, and let

$$\Sigma = \left\{E\left[\frac{\partial^2 \varepsilon_t}{\partial\theta\partial\theta^\intercal}(\theta_0)\right]\right\}^{-1} E\left[\frac{\partial\varepsilon_t}{\partial\theta}\frac{\partial\varepsilon_t}{\partial\theta^\intercal}\varepsilon_t^2(\theta_0)\right]\left\{E\left[\frac{\partial^2 \varepsilon_t}{\partial\theta\partial\theta^\intercal}(\theta_0)\right]\right\}^{-1} \tag{A.13}$$

The following theorem establishes the asymptotic properties of the semi-parametric MIDAS regression model:

Theorem A.1 *Suppose that Assumption A.1 and the regularity conditions appearing in Chen, Ghysels, and Renault (2009) hold. Then for each* $\theta \in \Theta$ *and* $x \in (\underline{x}, \overline{x})$

$$\sqrt{Th}\left[\widehat{m}_\theta(x) - m_\theta(x) - h^2 b_\theta(x)\right] \to N(0, \omega_\theta(x))$$

Moreover,

$$\sqrt{T}(\hat{\theta} - \theta_0) \to N(0, \Sigma)$$

Furthermore, for $x \in (\underline{x}, \overline{x})$

$$\sqrt{Th}(\widehat{m}(x) - m(x) - h^2 b(x)) \to N(0, \omega(x))$$

where h *denotes the bandwidth,* Σ *(eq. (A.13)) is variance matrix,* b *(eq. (A.12)) and* b_θ *(eq. (A.10)) are bias functions,* ω *appears below and* ω_θ *(eq. (A.9)) are variance functions.*

In particular,

$$\omega(x) = \frac{||K||_2^2 \sum_{j=1}^{\tau} B_j^2 E\left[\varepsilon_t^2 | x_{t-1-(j-1)/M} = x\right]}{f_0(x)\left(\sum_{j=1}^{\tau} B_j^2\right)^2} + \frac{M-1}{M} \frac{||K||_2^2 \sum_{j=1}^{\tau} \sum_{k=1,k\neq j}^{\tau} B_j^2 B_k^2 \operatorname{var}(m(x_{t+(j-k)/M})|x_t = x)}{f_0(x)\left(\sum_{j=1}^{\tau} B_j^2\right)^2} \tag{A.14}$$

The above expression shows that the mixed data sampling scheme in semi-parametric MIDAS regressions adds an extra term, i.e. the last appearing in the above expression. When $M = 1$, the asymptotic distribution collapses to the case covered in Linton and Mammen (2006). When $M > 1$, the dependent variable is sampled less frequent than the regressor which - compared to the case where all processes are sampled at the frequency $1/M$, implies that $(M-1)/M$ regression equations are missing.

Finally, to calculate the confidence interval, we assume: (1) the sample size is large enough, so the variance of $\widehat{m}(x)$ is the same as the asymptotic variance; (2) $\operatorname{var}(m(x_{s/M})|x_{k/M} = x) = \operatorname{var}(m(x_{s/M})), \forall s \neq k$; (3) $\operatorname{var}(\varepsilon_t | x_{t-j/M} = x) = \operatorname{var}(\varepsilon_t), j = 1, ..., \tau$. The confidence interval is then calculated for any x as,

$$CI(x) = [\widehat{m}(x) + Z_\alpha \hat{s}(x), \widehat{m}(x) + Z_{1-\alpha}\sqrt{\hat{s}(x)}] \tag{A.15}$$

$$\hat{s}(x) = \frac{||K||_2^2}{Th\hat{f}_0(x)\sum_{j=1}^{\tau} B_j^2(\hat{\theta})} \left(\operatorname{var}(\hat{\varepsilon}_t) + \frac{M-1}{M} \frac{\sum_{j=1}^{\tau}\sum_{k=1,k\neq j}^{\tau} B_j^2(\hat{\theta}) B_k^2(\hat{\theta})}{\sum_{j=1}^{\tau} B_j^2(\hat{\theta})} \operatorname{var}(\widehat{m}(x_t)) \right)$$

where Z_α is the α-quantile of the standard normal distribution. We set $\alpha = 0.05$, so that $Z_{0.05} = -1.645$ and $Z_{0.95} = 1.645$.[16]

[16]Note that we are computing the confidence interval of $\widehat{m}(x)$ corresponding to $E(\widehat{m}(x))$ instead of $m_0(x)$, so we omit the discussion related to the bias part. See Wasserman (2006, p.89) for more details related to the confidence interval of nonparametric estimation.

References

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang, 2005, How often to sample a continuous-time process in the presence of market microstructure noise, *Review of Financial Studies* 18, 351–416.

Andersen, T. G., and T. Bollerslev, 1997, Intraday periodicity and volatiliy persistence in financial markets, *Journal of Empirical Finance* 4, 115–158.

——— , 1998, Deutsche Mark–Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies, *Journal of Finance* 53, 219–265.

——— , F.X. Diebold, and P. Labys, 2000, Great realizations, *Risk Magazine* 13, 105 – 108.

Andersen, T. G., T. Bollerslev, and F. X. Diebold, 2007, Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility, *The Review of Economics and Statistics* 89, 701–720.

Andreou, E., E. Ghysels, and E. Kourtellos, 2009, Regression models with mixed sampling frequencies, *Journal of Econometrics* (forthcoming).

Bandi, F. M., and J. R. Russell, 2005, Microstructure noise, realized variance, and optimal sampling, Working paper, University of Chicago.

——— , 2006, Separating microstructure noise from volatility, *Journal of Financial Economics* 79, 655–692.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard, 2009, Regular and modified kernel-based estimators of integrated variance: The case with independent noise, *Econometrica* 76, 1481–1536.

Barndorff-Nielsen, O. E., S. Kinnebrock, and N. Shephard, 2008, Measuring downside risk realised semivariance, Discussion Paper, Oxford.

Barndorff-Nielsen, O. E., and N. Shephard, 2007, Variation, jumps, market frictions and high frequency data in financial econometrics, in Richard Blundell, Persson Torsten, and Whitney K Newey, ed.: *Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress.* Econometric Society Monographs, Cambridge University Press.

Becker, R., A.E. Clements, and S.I. White, 2007, Does implied volatility provide any information beyond that captured in model-based volatility forecasts?, *Journal of Banking and Finance* 31, 2535–2549.

Bekaert, G., and G. Wu, 2000, Asymmetric volatility and risk in equity markets, *Review of Financial Studies* 13, 1–42.

Black, F., 1976, Studies of stock price volatility changes, in *Proceedings of the 1976 Meetings of the Business and Economic Statistics Section, American Statistical Association* vol. 177 p. 81.

Bollen, B., and B. Inder, 2002, Estimating daily volatility in financial markets utilizing intraday data, *Journal of Empirical Finance* 9, 551–562.

Bollerslev, T., J. Litvinova, and G. Tauchen, 2006, Leverage and volatility feedback effects in high-frequency data, *Journal of Financial Econometrics* 4, 353–384.

Campbell, J. Y., and L. Hentschel, 1992, No news is good news: An asymmetric model of changing volatility in stock returns, *Journal of Financial Economies* 31, 281–318.

Chen, X., E. Ghysels, and E. Renault, 2009, Semi-parametric midas regression and index models, Working paper, UNC.

Christie, A.A., 1982, The stochastic behavior of common stock variances: Value, leverage and interest rate effects, *Journal of financial Economics* 10, 407–432.

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

Dacorogna, M., R. Gençay, U. A. Müller, R. B. Olsen, and O. V. Pictet, 2001, *An Introduction of High-Frequency Finance* (Academic Press: San Diego).

Diebold, F., and R. Mariano, 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics* 13, 253–263.

Engle, R., and G. Gallo, 2006, A multiple indicators model for volatility using intra-daily data, *Journal of Econometrics* 131, 3 – 27.

Engle, R. F., 1990, Discussion: Stock market volatility and the crash of '87, *Review of Financial Studies* 3, 103–106.

——— , and V. Ng, 1993, Measuring and testing the impact of news on volatility, *Journal of Finance* 48, 1749–1778.

Forsberg, L., and E. Ghysels, 2006, Why do absolute returns predict volatility so well?, *Journal of Financial Econometrics* 6, 31–67.

French, Kenneth R., W. Schwert, and R. F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economies* 19, 3–29.

Ghysels, E., 2009, Mixed data sampling and temporal aggregation, *Encyclopedia of Quantitative Finance*, Ole Barndorff-Nielsen and Eric Renault (eds) (forthcoming).

——— , and D. Osborn, 2001, *The Econometric Analysis of Seasonal Time Series* (Cambridge University Press, Cambridge).

Ghysels, E., A. Rubia, and R. Valkanov, 2009, Multi-Period Forecasts of Volatility: Direct, Iterated, and Mixed-Data Approaches, Discussion Paper, UNC.

Ghysels, E., P. Santa-Clara, and R. Valkanov, 2002, The MIDAS touch: Mixed data sampling regression models, Working paper, UNC and UCLA.

——— , 2006, Predicting volatility: getting the most out of return data sampled at different frequencies, *Journal of Econometrics* 131, 59–95.

Ghysels, E., A. Sinko, and R. Valkanov, 2006, MIDAS Regressions: Further Results and New Directions, *Econometric Reviews* 26, 53 – 90.

Giacomini, R., and H. White, 2006, Tests of conditional predictive ability, *Econometrica* 74, 1545–1578.

Glosten, L. R., R. Jagannathan, and David E. Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779–1801.

Hansen, P. R., and A. Lunde, 2006, Realized variance and market microstructure noise, *Journal of Business and Economic Statistics* 24, 127–161.

Jacod, J., 1994, Limit of random measures associated with the increments of a brownian semimartingale, Preprint number 120, Laboratoire de Probabilités, Université Pierre et Marie Curie, Paris.

——— , 1996, La variation quadratique du brownian en presence d'erreurs d'arrondi, *Asterisque* 236, 155–162.

——— , and P. Protter, 1998, Asymptotic error distributions for the euler method for stochastic differential equations, *Annals of Probability* 26, 267–307.

León, Á., J.M. Nave, and G. Rubio, 2007, The relationship between risk and expected return in Europe, *Journal of Banking and Finance* 31, 495–512.

Linton, O., and E. Mammen, 2005, Estimating semiparametric ARCH(∞) models by kernel smoothing methods, *Econometrica* 73, 711–836.

——— , 2006, Nonparametric transformation to white noise, Working paper, London School of Economics.

Martens, M., Y. C. Chang, and S. J. Taylor, 2002, A comparison of seasonal adjustment methods when forecasting intraday volatility, *Journal of Financial Research* 25, 283–299.

Merton, R. C., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economies* 8, 323–361.

Nelson, D. B., 1991, Conditional heteroscedasticity in asset returns: A new approach, *Econometrica* 64, 347–370.

Patton, A., 2007, Volatility forecast comparison using imperfect volatility proxies, Working Paper, London School of Economics.

Schwert, William G., 1990, Stock volatility and the crash of 87, *Review of Financial Studies* 3, 77–102.

Wasserman, L., 2006, *All of Nonparametric Statistics* (Springer: New York).

West, K., 1996, Asymptotic inference about predictive ability, *Econometrica* 64, 1067–1084.

Wood, R. A., T. H. McInish, and J. K. Ord, 1985, An investigation of transaction data for nyse stocks, *Journal of Finance* 40, 723–739.

Table 1: Details of the Data Series and Model Acronyms

The top part of the table provides the details of the data used in our study. We analyze four series which consist of intra-day returns of respectively Dow Jones and S&P 500 cash and futures markets. To lower part summarizes all models, showing the equation numbers, the models' acronyms and some details. The generic specification appears in equation (3.1), namely: $RV_{t,t+1} = \psi_0 + \sum_{j=1}^{\tau}[\sum_{i=1}^{1/\Delta} \psi_{ij}(\theta)NIC(r_{t-j,i\Delta})] + \varepsilon_{t,t+1}$, where $\sum_{j=1}^{\tau}\sum_{i=1}^{1/\Delta} \psi_{ij} = 1$ and $NIC(r)$ stands for news impact curve.

			Period	Days	Trading Hours (EST)	$1/\Delta$
			Full Sample			
Dow Jones	Cash	*DJC*	04/01/1993 – 12/31/2008	3969	$9:30-16:00$	78
	Futures	*DJF*	10/06/1997 – 12/31/2008	2834	$9:30-16:00$	78
S&P 500	Cash	*SPC*	09/30/1985 – 12/31/2008	5668	$9:30-16:00$	78
	Futures	*SPF*	04/21/1982 – 12/31/2008	6764	$9:30-16:00$	78

News Impact	Acronym	Explanation
	Intra-daily returns - Parametric	
br^2	SYMM	Symmetric NIC
$(br^2 + c1_{r<0}r^2)$	ASYMGJR	Asymmetric GJR
$((r-c)^2)$	ASYMLS	Asymmetric Location Shift
	Intra-daily returns - Semi-parametric	
Non-parametric NIC NIC	SP	Semi-parametric MIDAS appearing in Eq. (3.1)
	Models with daily volatility	
	HAR-RV	Eq. (2.4) using RV regressors
	HAR-RV-J	Eq. (2.5) using BPV and Jump regressors
	HAR-S-RV-J	Eq. (2.7) using semi-variances and Jump regressors

Table 2: Asymmetry and Parameter Estimates

Panel A pertains to the regression appearing in (2.7) involving semi-variances. Entries pertain to tests of the single restrictions that either the daily, weekly, or monthly semi-variances have equal coefficients. The final column contains a joint test of the null with all three restrictions. The asymptotic variances to compute the tests used HAC standard errors with lag lengths that match the inference recommendations of Andersen, Bollerslev, and Diebold (2007). Panel B pertains to parametric models using intra-daily returns and featuring asymmetries. The panel displays the parameter estimates of the lag polynomials involving parameters θ_1 through θ_4, see equation (3.2). The parameters a, b, and c, determine the news impact, where c relates to asymmetries, since for the ASYMGJR model $NIC(r) = (br^2 + c1_{r<0}r^2)$, and for the ASYMLS model $NIC(r) = (b(r-c)^2)$. The parameter estimates are reported for one representative series, the S&P 500 Futures contract, at one day and one week forecast horizons.

Panel A

Regressors	Daily	Weekly	Monthly	Joint
		Dow Jones - Cash		
Daily	0.4811	0.8780	0.0098	3.1181
Weekly	3.5477	1.1160	6.8195	31.7494
Monthly	58.5899	0.6384	8.6274	794.0652
		Dow Jones - Futures		
Daily	2.2559	1.9559	5.1592	16.5648
Weekly	2.8865	7.4206	0.1964	191.8205
Monthly	5.1708	0.0437	21.7229	53.1562
		S&P Cash		
Daily	3.9291	2.8481	0.0595	13.6250
Weekly	1.9306	0.6735	4.0822	17.4422
Monthly	286.051	2.4145	0.1205	2091.19
		S&P Futures		
Daily	2.6254	1.3337	4.6696	26.0813
Weekly	15.4360	5.8466	2.3991	181.7672
Monthly	2.6423	0.2328	14.0606	35.7201

Table continued on next page ...

Table 2 continued

Panel B

	θ_1	θ_2	θ_3	θ_4	a	b	c
			S&P 500 Futures Market - One-day horizon				
ASYMGJR	2.337e-014	3.051	1.360	1.955	-0.008254	-125.9	428.5
(std. dev.)	(0.005456)	(0.03413)	(0.001195)	(0.004495)	(0.01655)	(4.728)	(10.08)
ASYMLS	2.337e-014	1.974	1.239	2.825	-39.18	67.61	0.7621
(std. dev.)	(0.004674)	(0.0215)	(0.002814)	(0.01108)	(0.4254)	(0.7221)	(0.001995)
			S&P 500 Futures Market - One-week horizon				
ASYMGJR	0.6527	2.672	1.740	1.347	-0.01653	-272.3	727.2
(std. dev.)	(0.02066)	(0.1054)	(0.01897)	(0.01655)	(0.02881)	(16.51)	(34.28)
ASYMLS	1.614e-013	0.6246	78.65	148.1	-0.08167	80.52	0.05015
(std. dev.)	(0.01403)	(0.01892)	(1.300)	(2.554)	(0.02946)	(1.580)	(0.004611)

Table 3: Out-of-sample performance of Semi-Parametric and Parametric MIDAS regression models

The table shows the out-of-sample *MSFE* of the SP, i.e. the semi-parametric MIDAS model, and parametric models SYMM, ASYMGJR, and ASYMLS and models that use daily volatility measures HAR-RV, HAR-RV-J, and HAR-S-RV-J. We consider two sample configurations, one ending in 2004 and out-of-sample appraisal 2005-2006, the other ending in 2006 and out-of-sample appraisal 2007-2008. *DJC* stands for Dow Jones Cash Market, *DJF* for Dow Jones Futures Market, *SPC* for S&P 500 Cash Market, and *SPF* for S&P 500 Futures Market.

		One-day ahead forecasts				One-week ahead forecasts			
		DJC	*DJF*	*SPC*	*SPF*	*DJC*	*DJF*	*SPC*	*SPF*
		Semi-Parametric Intra-daily returns							
SP	2005-2006	0.04	0.05	0.04	0.07	0.02	0.05	0.03	0.08
	2007-2008	19.99	9.63	16.32	11.28	16.94	4.60	20.52	6.96
		Parametric Intra-daily returns							
SYMM	2005-2006	0.04	0.05	0.04	0.04	0.03	0.07	0.02	0.02
	2007-2008	19.07	10.64	29.28	32.49	13.43	5.47	12.36	25.56
ASYMGJR	2005-2006	0.03	0.03	0.03	0.05	0.02	0.07	0.02	0.02
	2007-2008	18.95	13.57	18.19	14.51	11.37	5.09	8.64	46.75
ASYMLS	2005-2006	0.05	0.06	0.04	0.09	0.03	0.07	0.02	0.03
	2007-2008	18.28	9.97	16.46	12.65	12.32	5.30	9.83	25.55
		Parametric Daily Volatility Measures							
HAR-RV	2005-2006	0.05	0.05	0.07	0.04	0.06	0.08	0.09	0.05
	2007-2008	20.76	10.86	17.55	13.20	17.97	5.36	12.16	6.96
HAR-RV-J	2005-2006	0.04	0.04	0.06	0.041	0.06	0.04	0.07	0.05
	2007-2008	20.48	10.53	18.27	13.72	16.69	3.91	12.22	6.93
HAR-S-RV-J	2005-2006	0.04	0.03	0.04	0.03	0.03	0.02	0.03	0.04
	2007-2008	20.15	9.14	16.60	11.71	17.48	3.84	10.86	6.66

Table 4: Tests of Improved Predictability: HAR-RV-J Benchmark

The table reports the conditional forecasting ability test of Giacomini and White for each model against the benchmark HAR-RV-J model appearing in equation (2.5). The model acronyms appear in lower panel of Table 1, except for the GJR entry which pertains to the original specification of Glosten, Jagannathan, and Runkle (1993) with daily data. The tests involve the sample MSFE appearing in equation (4.1), computed for each forecasting horizon k, and each forecasting model/method m. The null hypothesis appears in equation (4.2) with instruments a constant and $u^b_{t-k,t})^2$., The entries are test statistic will be $\chi^2(1)$. *DJC* stands for Dow Jones Cash Market, *DJF* for Dow Jones Futures Market, *SPC* for S&P 500 Cash Market, and *SPF* for S&P 500 Futures Market. Whenever an entry is boldfaced, the benchmark model is better than the alternative model. Boldfaced values exceeding the $\chi^2(1)$ critical values of 2.706 (10 %), 3.841 (5 %), 6.635 (1 %) or 10.828 (.1 %) imply that the benchmark model significantly outperforms the alternative model. Normal-faced entries exceeding those values mean that the alternative model significantly outperforms the benchmark. The tests involve rolling 5-year samples for estimation and 2-year out-of-sample forecast performances rolling over one month at the time - starting with the 1997-2001 estimation sample.

	One-day ahead forecasts				One-week ahead forecasts				One-month ahead forecasts			
	DJC	*DJF*	*SPC*	*SPF*	*DJC*	*DJF*	*SPC*	*SPF*	*DJC*	*DJF*	*SPC*	*SPF*
						2001-2006						
HAR-S-RV-J	36.49	195.87	40.92	**175.19**	8.40	7.60	5.33	7.04	**1.12**	2.28	1.68	1.14
SYMM	**64.02**	**150.77**	**628.02**	**282.50**	7.21	3.11	5.60	6.52	1.74	1.96	2.14	2.15
ASYMGJR	11.64	163.97	11.92	277.58	7.43	4.87	8.01	8.50	1.88	2.11	2.13	2.30
ASYMLS	80.87	**308.13**	145.51	**144.81**	4.86	2.93	5.51	5.97	1.79	1.51	2.02	2.33
GJR	**132.64**	**251.22**	**312.13**	**190.43**	**75.20**	**53.42**	**114.38**	**67.77**	**20.01**	**45.39**	**19.11**	**16.88**
						2001-2008						
HAR-S-RV-J	6.46	9.37	5.70	4.36	2.84	8.00	1.10	5.90	**2.70**	**0.79**	**1.12**	**2.50**
SYMM	7.23	4.72	4.40	**5.15**	6.19	5.78	6.75	4.23	2.84	1.79	2.70	**1.71**
ASYMGJR	5.28	6.12	5.21	4.24	6.42	7.42	5.84	4.10	**1.27**	1.44	2.71	**2.08**
ASYMLS	5.73	5.41	5.13	3.79	6.07	5.75	5.76	4.11	2.51	2.91	2.57	2.15
GJR	**22.18**	**31.57**	**42.41**	**91.10**	**24.05**	**37.82**	**26.99**	**34.56**	**19.20**	**17.50**	**19.56**	**21.45**

Table 5: Tests of Improved Predictability: HAR-S-RV-J Benchmark

The table reports the conditional forecasting ability test of Giacomini and White for each model against the benchmark HAR-S-RV-J model appearing in equation (2.7). The model acronyms appear in lower panel of Table 1. The tests involve the sample MSFE appearing in equation (4.1), computed for each forecasting horizon k, and each forecasting model/method m. The null hypothesis appears in equation (4.2) with instruments a constant and $u^b_{t-k,t})^2$., The entries are test statistic will be $\chi^2(1)$. *DJC* stands for Dow Jones Cash Market, *DJF* for Dow Jones Futures Market, *SPC* for S&P 500 Cash Market, and *SPF* for S&P 500 Futures Market. Whenever an entry is boldfaced, the benchmark model is better than the alternative model. Boldfaced values exceeding the $\chi^2(1)$ critical values of 2.706 (10 %), 3.841 (5 %), 6.635 (1 %) or 10.828 (.1 %) imply that the benchmark model significantly outperforms the alternative model. Normal-faced entries exceeding those values mean that the alternative model significantly outperforms the benchmark. The tests involve rolling 5-year samples for estimation and 2-year out-of-sample forecast performances rolling over one month at the time - starting with the 1997-2001 estimation sample.

	One-day ahead forecasts				One-week ahead forecasts				One-month ahead forecasts			
	DJC	*DJF*	*SPC*	*SPF*	*DJC*	*DJF*	*SPC*	*SPF*	*DJC*	*DJF*	*SPC*	*SPF*
2001-2006												
SYMM	**121.83**	**364.82**	**216.95**	**501.11**	6.00	**1.95**	2.69	2.20	1.57	1.26	1.82	2.27
ASYMGJR	8.67	40.15	10.84	11.20	6.79	**2.55**	6.03	4.73	1.71	1.75	1.75	2.32
ASYMLS	**215.51**	**359.36**	**424.04**	**493.92**	3.39	**2.05**	2.61	2.10	1.60	1.73	1.60	**2.25**
2001-2008												
SYMM	4.87	**16.48**	2.53	**10.47**	6.10	3.50	5.48	**3.43**	2.84	2.12	2.50	1.58
ASYMGJR	4.77	3.55	4.72	4.02	**6.40**	5.80	5.34	**3.66**	2.08	2.32	2.52	1.30
ASYMLS	5.08	**68.92**	4.42	2.79	5.87	3.32	5.02	**3.42**	2.60	2.22	2.39	1.29

Figure 1: High frequency returns before and during the 2007-2008 crisis

The plots represent the 2005-2006 and 2007-2008 samples histograms - throughout the day in different time zones of the trading day - of the 5 minute returns.

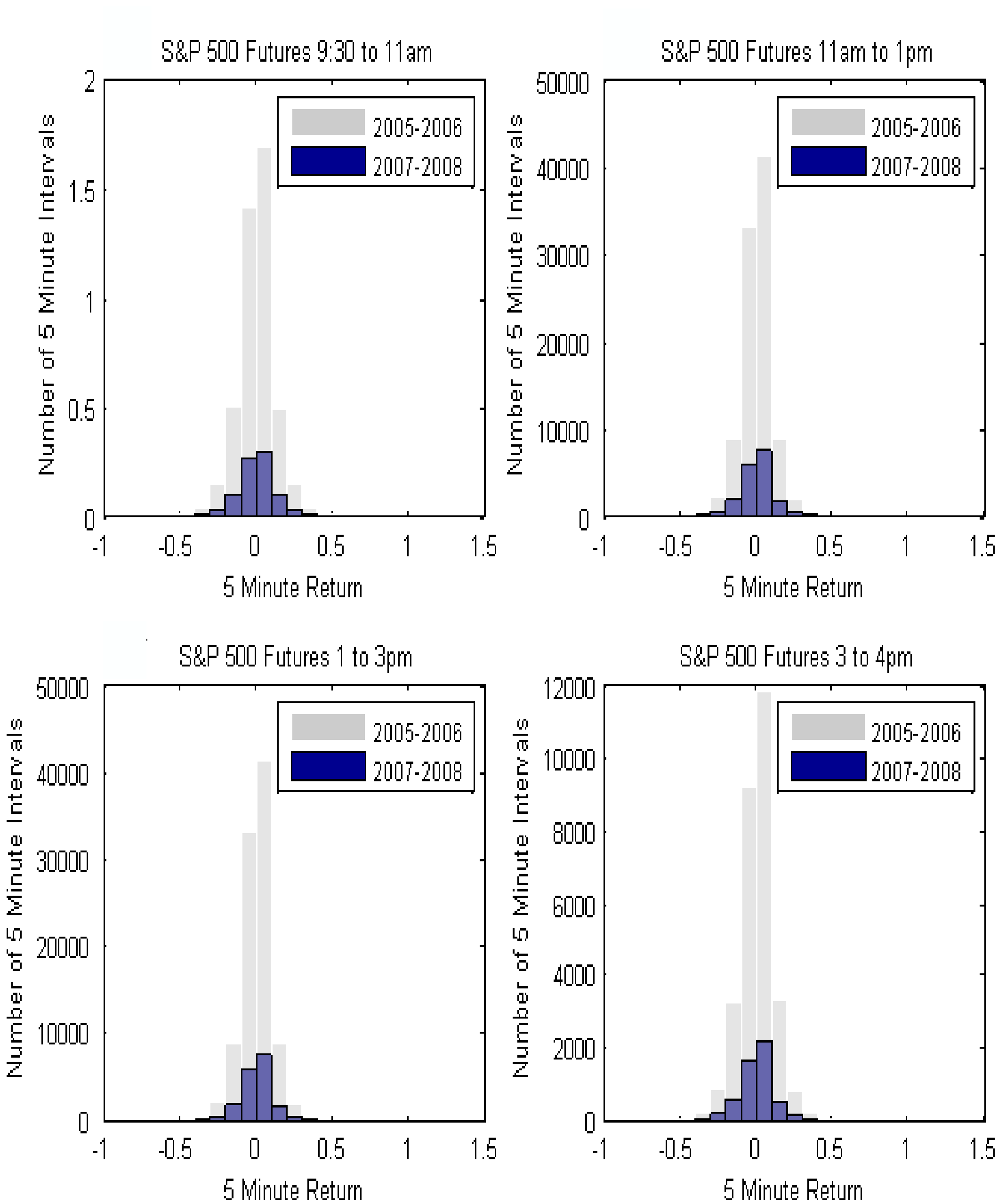


Figure 2: One-day ahead news impact curves for SP models

The plots represent estimates of semi-parametric news impact curves - i.e. the function NIC - as specified in equation (3.1). To facilitate the interpretation and display of results, the news impact is scaled by the mean of RV for each series/horizon - so that the news impact is in terms of fraction of average horizon volatility. Results for four series using five minute intra-daily returns are displayed. They are: (a) Dow Jones Cash Market; (b) Dow Jones Futures Market; (c) S&P 500 Cash Market; and (d) S&P 500 Futures Market. The confidence bands are computed according to formula (A.15) appearing in the Appendix.

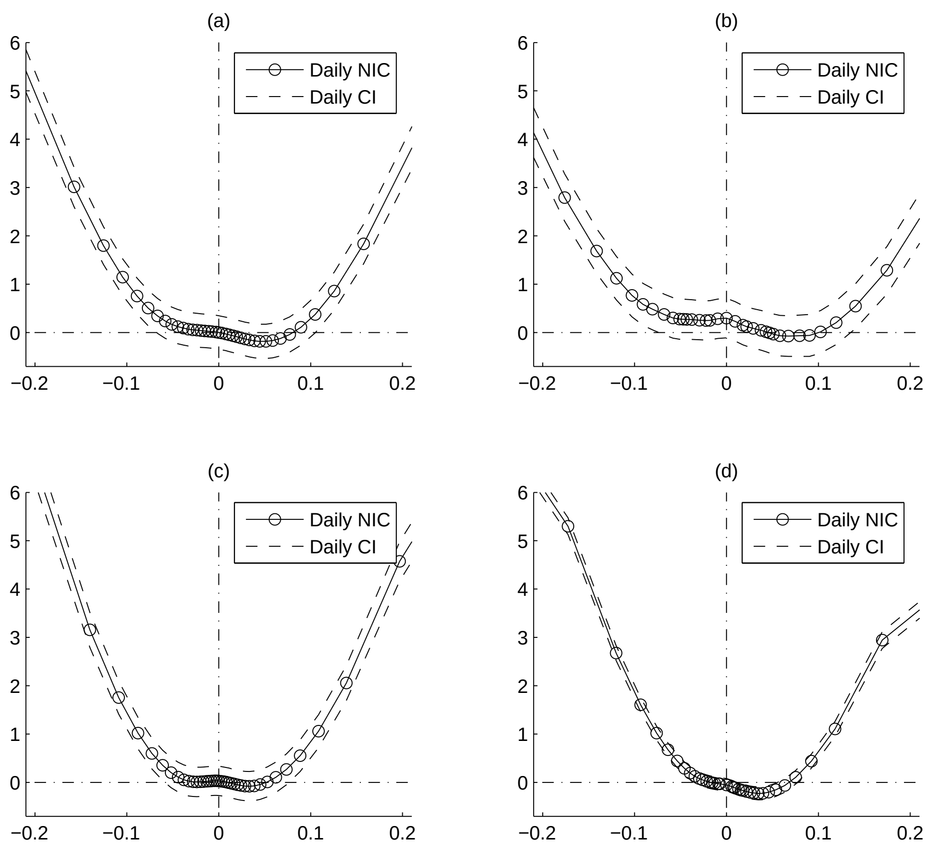


Figure 3: One-week ahead news impact curves for SP models

The plots represent estimates of semi-parametric news impact curves - i.e. the function NIC - as specified in equation (3.1). To facilitate the interpretation and display of results, the news impact is scaled by the mean of RV for each series/horizon - so that the news impact is in terms of fraction of average horizon volatility. Results for four series using five minute intra-daily returns are displayed. They are: (a) Dow Jones Cash Market; (b) Dow Jones Futures Market; (c) S&P 500 Cash Market; and (d) S&P 500 Futures Market. The confidence bands are computed according to formula (A.15) appearing in the Appendix.

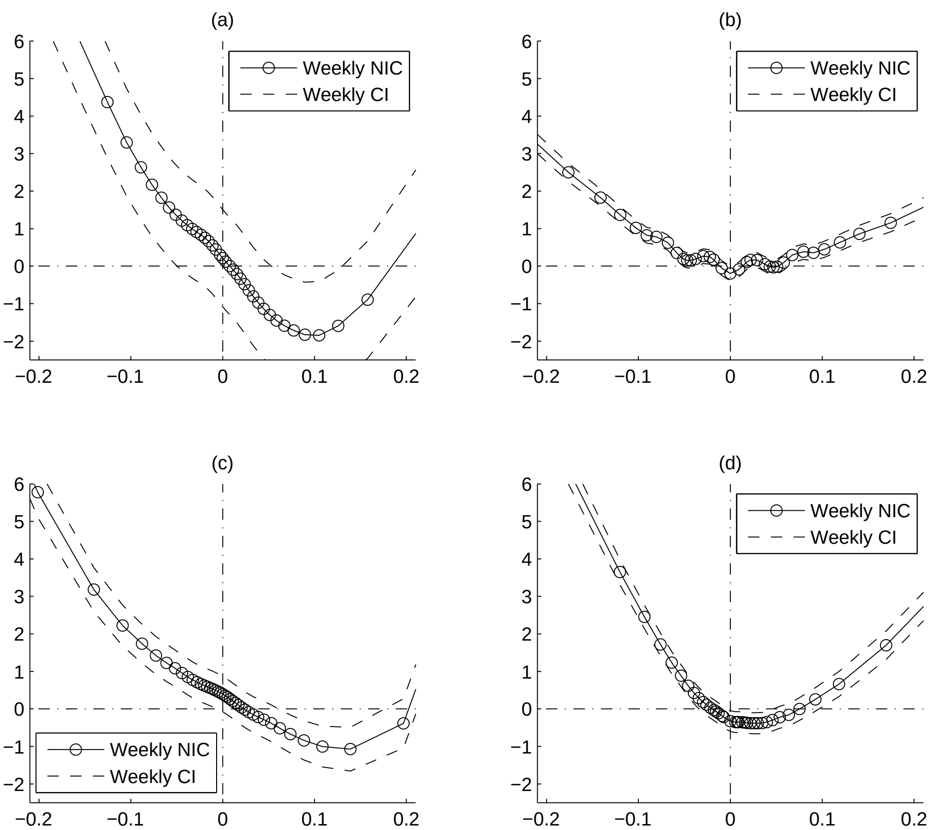


Figure 4: One-week and one-month ahead news impact curves for semi-parametric and parametric MIDAS models

The plots represent estimates of semi-parametric and parametric news impact curves for different horizons. Results for two series using five minute intra-daily returns are displayed. They are: (a) S&P 500 Futures Market and (b) Dow Jones Cash Market. The confidence bands are computed according to formula (A.15) appearing in the Appendix.

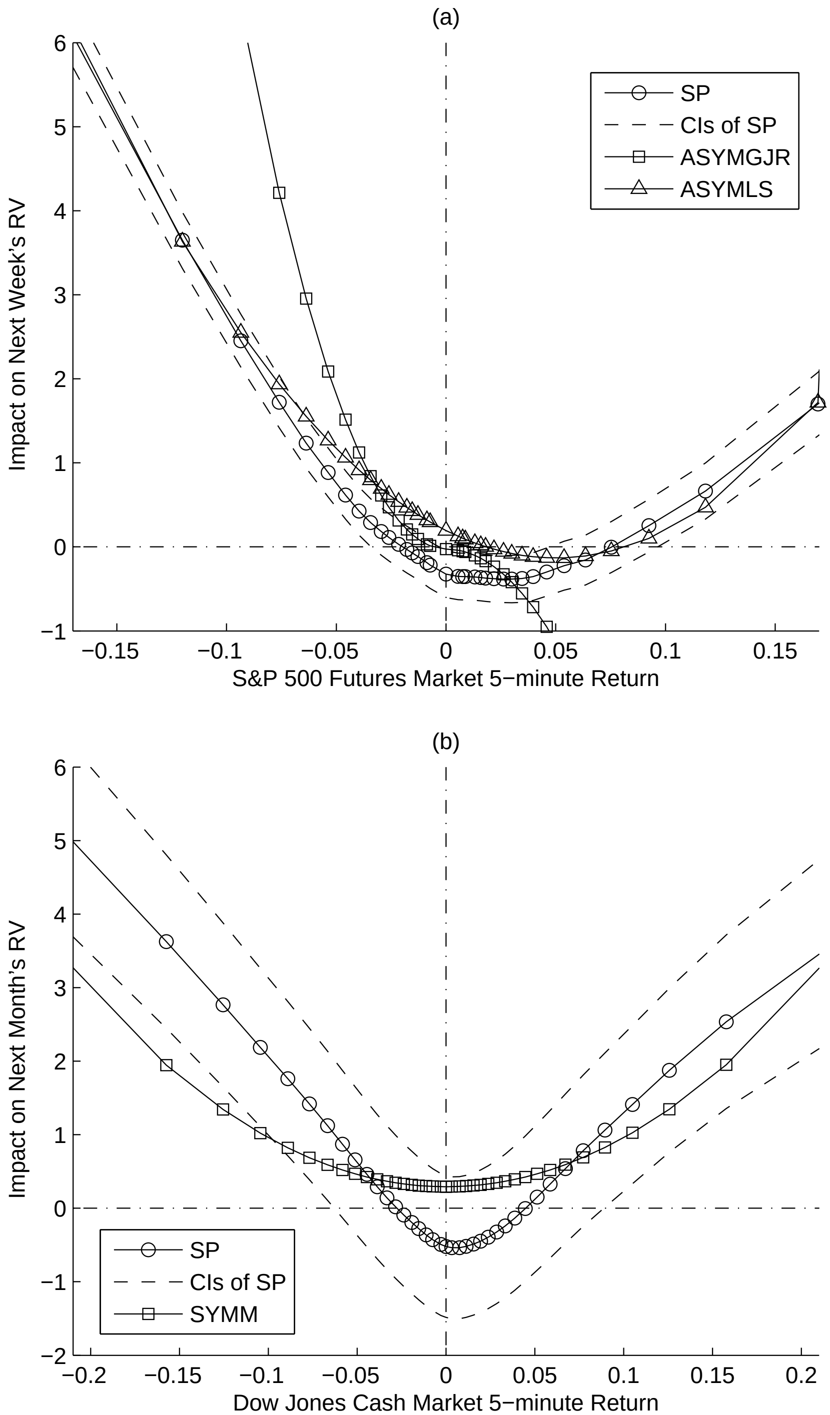


Figure 5: Parametric polynomial lag estimates of semi-parametric MIDAS

This figure shows the lag polynomials of the semi-parametric MIDAS regression using the Dow Jones futures data. The first plot provides the product of the daily and intra-daily Beta polynomials appearing in equation (3.2). The second contains only the daily polynomial whereas the third only the intra-daily.