JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. (2012)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/jae.2292

SEMIPARAMETRIC VECTOR MEM

FABRIZIO CIPOLLINI,[a] ROBERT F. ENGLE[b] AND GIAMPIERO M. GALLO[a*]

[a] *Dipartimento di Statistica 'G. Parenti', Università di Firenze, Florence, Italy*
[b] *Department of Finance, Stern School of Business, New York University, New York, USA*

SUMMARY

Financial time series are often non-negative-valued (volumes, trades, durations, realized volatility, daily range) and exhibit clustering. When joint dynamics is of interest, the vector multiplicative error model (vMEM; the element-by-element product of a vector of conditionally autoregressive scale factors and a multivariate i.i.d. innovation process) is a suitable strategy. Its parameters can be estimated by generalized method of moments, bypassing the problem of specifying a multivariate distribution for the errors. Simulated results show the gains in efficiency relative to an equation-by-equation approach. A vMEM on several measures of volatility justifies a joint approach revealing full interdependence.

Received 14 January 2009; Revised 24 April 2012

Supporting information may be found in the online version of this article.

1. INTRODUCTION

In financial time series analysis we encounter several instances of non-negative-valued time series (volumes, trades, durations, realized volatility, daily range and so on). The multiplicative error model (MEM), as introduced by Engle (2002), exploits the stylized fact that these time series share similar persistence and clustering features as the absolute or squared returns. Mirroring the generalized autoregressive conditional heteroskedasticity (GARCH) approach (Bollerslev, 1986), a MEM is the product of the conditional expectation of the variable of interest (modeled as a time-varying function of the information set, including past values of the variable and past values of the conditional expectation) times an i.i.d. (unit mean) innovation process.

In this sense, the MEM is a generalization of the autoregressive conditional duration (ACD) model (Engle and Russell, 1998); other examples of MEMs are the conditional autoregressive range (CARR) model by Chou (2005) and the model by Manganelli (2005). Engle and Gallo (2006) suggest stacking three MEMs for three different measures of volatility (absolute return, daily range and realized volatility), allowing for the presence of lagged values of all variables in each conditional expectation equation. Consistent parameter estimation of such a model is very simple, since it proceeds equation-by-equation as if innovation terms were uncorrelated. The empirical results show a dynamic interdependence of the indicators, suggesting that a joint approach would be advisable to gain efficiency and to have a more reliable interpretation of the results.

Such a model is a vector MEM (vMEM), where the variables of interest are represented as the element-by-element product of a vector of conditional expectations times a vector of (unit mean) innovations with a general covariance matrix. A full parametric approach would complete the model with the specification of the joint probability density function (pdf) for the vector error term. A possible strategy is to adopt a copula approach as suggested by Cipollini *et al*. (2007).

* Correspondence to: Giampiero M. Gallo, Dipartimento di Statistica 'G. Parenti', Università di Firenze, Florence, Italy.
E-mail: gallog@ds.unifi.it

The alternative we pursue in this paper is a semiparametric one; that is, we bypass the issue of reconstructing a multivariate non-negative-valued pdf for the vector of innovations. We set the vMEM in a more general form, where the conditional expectation of one variable is a function not only of its own past conditional expectation but also of the past conditional expectations of all other variables; we adopt a novel generalized method of moments (GMM) estimation approach where consistency is ensured by the correct specification of the conditional expectation. One appealing feature of the properties of this estimator is that in the univariate case the moment conditions revert to the first-order conditions of the log-likelihood function maximization under a Gamma assumption for the error term. Moreover, simulation results show that the gain in efficiency is substantial when contemporaneous correlation of the error terms is medium to high, as one would expect. This entails more precise estimates and better statistical inference.

To provide some insights into the empirical relevance of this approach, we highlight the interdependence of several measures of volatility (absolute return, realized kernel volatility and daily range). In this respect, we add to the papers by Shephard and Sheppard (2010) and by Hansen *et al*. (2011) regarding the specification of a forecast model for volatility. In our application, the interdependence among the measures is very strong and refines the results of Engle and Gallo (2006), who suggested the need to make it explicit in forecasting return volatility. We show that statistical inference on cross-effects changes substantially in the joint GMM as opposed to the equation-by-equation framework.

Synthesizing the contribution:

1. we suggest a new model for non-negative processes without resorting to logs and we provide joint GMM-based statistical inference;
2. in modeling volatility, the full specification of the models receiving empirical support can be estimated just by our joint GMM procedure and not equation-by-equation;
3. gains in efficiency are supported by the degree of correlations among estimated residuals;
4. we show full dynamic interdependence and a relevant contribution of the daily range (a possible explanation is that such a variable acts as a proxy for the significant contribution of jumps (see Andersen *et al*., 2007)).

The structure of the paper is as follows: Section 2 presents the vMEM, summarizing its main features; Section 3 details GMM estimation and inference on vMEM parameters and illustrates the efficiency gain over the equation-by-equation approach in a simulation experiment; Section 4 discusses an application of the semiparametric vMEM for analyzing the dynamic interdependence among different measures of volatility; and Section 5 concludes.

2. MULTIPLICATIVE ERROR MODELS

The MEM extends the GARCH approach to processes x_t with non-negative support (Engle, 2002; Engle and Gallo, 2006; for a survey, see Brownlees *et al*., 2012). In the univariate case, the model is specified as

$$x_t = \mu_t \varepsilon_t \tag{1}$$

μ_t is a scale factor assumed to deterministically evolve conditional upon the information set $\mathcal{F}_{t-1}$: using a generic notation we denote this as $\mu_t = \mu(\mathcal{F}_{t-1}; \theta)$, where θ is a vector of unknown parameters ruling the dynamics of μ_t. ε_t is an innovation term assumed to have non-negative support, conditional mean 1 and unknown variance σ^2: $\varepsilon_t \big| \mathcal{F}_{t-1} \sim D(1, \sigma^2)$. A Gamma distribution assumption as in Engle and Gallo (2006) provides enough flexibility for practical purposes and includes the exponential and chi-square as special cases (other assumptions are of course possible).

DOI: 10.1002/jae

The previous assumptions on μ_t and ε_t give

$$E(x_t|\mathcal{F}_{t-1}) = \mu_t \tag{2}$$

$$V(x_t|\mathcal{F}_{t-1}) = \sigma^2\mu_t^2 \tag{3}$$

The ACD model by Engle and Russell (1998) is a special case of MEM, but also various versions of ultra-high-frequency-based measures of volatility, absolute return, high–low range, number of trades in a certain interval, and volume can all be modeled with MEMs. One of the advantages of such a model is to avoid the need to resort to logs (not possible when zeros are present in the data) and to provide conditional expectations of the variables of interest directly (rather than expectations of the logs). Empirical results show a good performance of these types of models in capturing the stylized facts of the observed series (e.g. for the daily range; Chou, 2005).

Jointly modeling the dynamics of two or more non-negative time series has the distinctive feature of allowing multi-step forecasting over a univariate model with predetermined variables. Examples in the literature are: volatility forecasting using different measures (Engle and Gallo, 2006); volatility spillovers for studying contagion among markets (Engle *et al*., 2012); order execution dynamics in order-driven markets (Noss, 2007); joint dynamics of duration, volume and volatility for the same asset (Manganelli, 2005); and volatility, volume and trading intensity (Hautsch, 2008).

In spite of all these examples providing convincing motivation, specification and estimation are limited by an equation-by-equation approach which can be made more general. In a full system representation for x_t, a K-dimensional process with non-negative components,[1] we define a vector MEM (or vMEM) as

$$x_t = \mu_t \odot \varepsilon_t = \mathrm{diag}(\mu_t)\varepsilon_t \tag{4}$$

where $\odot$ indicates the Hadamard (element-by-element) product and *diag*(.) denotes a diagonal matrix with the vector in the argument as its main diagonal. Conditional on $\mathcal{F}_{t-1}$, μ_t can be defined as above, that is, as

$$\mu_t = \mu(\mathcal{F}_{t-1};\theta) \tag{5}$$

except that now we are dealing with a K-dimensional vector depending on a (larger) vector of parameters θ. The innovation vector ε_t is a K-dimensional random variable defined over a $[0, +\infty)^K$ support, with unit vector 1 as expectation and a general variance–covariance matrix Σ:

$$\varepsilon_t|\mathcal{F}_{t-1} \sim D(1, \Sigma) \tag{6}$$

From the previous conditions:

$$E(x_t|\mathcal{F}_{t-1}) = \mu_t \tag{7}$$

$$V(x_t|\mathcal{F}_{t-1}) = \mu_t\mu_t' \odot \Sigma = \mathrm{diag}\,(\mu_t)\Sigma\,\mathrm{diag}\,(\mu_t) \tag{8}$$

[1] In what follows we will adopt the following convention: if x is a matrix, then $x_{i,.}$ and $x_{.,j}$ denote the ith row and the jth column, respectively, of x; if $x_1, \ldots, x_K$ are (m, n) matrices, then $(x_1; \ldots; x_K)$ indicates the (mK, n) matrix obtained by stacking the matrices x_is columnwise.

DOI: 10.1002/jae

follow, where the latter is a positive definite matrix by construction. Hence the second moments of the variables of interest in our model are time-varying; if we deal with volatilities, these would match conditional fourth moments of return (volatility of volatility), albeit with a constant conditional correlation.[2]

As per the scale term μ_t, the general specification adopted is

$$\mu_t = \omega + \sum_{l=1}^{L} \left[\alpha_l x_{t-l} + \gamma_l x_{t-l}^{(-)} + \beta_l \mu_{t-l} \right] \tag{9}$$

Among the parameters (whose non-zero elements are arranged in the vector θ), ω has dimension $(K, 1)$, whereas α_l, γ_l and β_l have dimension (K, K) and may have some elements equal to zero. In this sense, L denotes the largest lag at which at least one of these coefficients includes non-zero elements. The terms $\gamma_l x_{t-l}^{(-)}$ aim to capture asymmetric effects associated with the sign of an observed variable: the vector $x_t^{(-)}$ contains $x_{t,i}$s multiplied by a function related to a signed variable. For example, when different volatility indicators of the same asset are considered, such an indicator assumes value one when the daily return r_t is negative. In a market volatility spillover study, each market would have its own indicator function built from the sign of its own return $r_{t,i}$. Finally, in a microstructure context, we can think of assigning positive or negative values to volumes according to whether the trade was a buy or a sell.

With regard to the error term ε_t, a completely parametric formulation of the vMEM requires a full specification of its conditional distribution, which can be quite cumbersome for a number of reasons:

1. Multivariate distributions defined on the non-negative orthant are often not sufficiently flexible. Furthermore, sometimes they are defined via the characteristics function without an explicit pdf, thus complicating the parameter estimation considerably (see, for example, the discussion on the multivariate Gamma in Cipollini *et al.* (2007) and the work of Ahoniemi and Lanne (2009)).
2. Recurring to copulas bypasses these problems: some proposals are provided in Cipollini *et al.* (2007), where marginals for the components $\varepsilon_{t,i}$ (specified as before) are linked together using Gaussian copulas. However, although there is a one-to-one correspondence between the cumulative distribution function of an absolutely continuous multivariate random variable and its copula representation, there is no guidance for choices of a specific copula function except for convenience. For instance, although elliptical copulas are appealing because they can be employed also for moderately large K and can accommodate tail dependency, they have an elliptically symmetric behavior that may represent a limit in some contexts. Conversely, some Archimedean copulas can accommodate asymmetric dependency profiles but appear less useful when the dimension K tends to increase (McNeil *et al.*, 2005). As a consequence, even when combined with an appropriate specification of the marginals, copulas are not always able to model the association in the distribution.
3. It is not said that all components of the error term share the same marginal distribution. Even though, in principle, one may be able to select the appropriate pdf one by one, this would require a lengthy model tuning.
4. As a final point, if the dynamics of μ_t are the main focus of the analysis, a full specification of the distribution of ε_t may be of secondary importance.

[2] In order to have time-varying conditional correlations of volatilities across assets, extending the model to include a time-varying Σ seems to be quite a formidable task, given that we would need some assumption on the evolution of (own and mixed) fourth moments of return.

DOI: 10.1002/jae

3. MODEL INFERENCE

In what follows, we use a semiparametric specification of the vMEM where the distribution of the error component is left unspecified except for the conditional moments as in (6).

3.1. Efficient GMM

Let us start from the two conditional moments of the observed variable x_t, (7) and (8), to be seen as a function of μ_t (or, which is the same, of the p parameters θ) and of Σ (containing $K(K+1)/2$ nuisance parameters). Model estimation is based on the conditional moments of the martingale difference:

$$u_t = x_t \oslash \mu_t - 1 \tag{10}$$

namely, $E(u_t|\mathcal{F}_{t-1}) = 0$ and $V(u_t|\mathcal{F}_{t-1}) = \Sigma$, where $\oslash$ indicates the element-by-element division. Following a standard textbook GMM treatment (Newey and McFadden, 1994), we get p moment conditions (exact identification) by choosing any (p, K) matrix of instruments G_t as a deterministic function of the information set $\mathcal{F}_{t-1}$ such that

$$E(G_t u_t) = 0 \tag{11}$$

The corresponding sample moment conditions are

$$\frac{1}{T}\sum_{t=1}^{T} G_t u_t = 0 \tag{12}$$

Under a correct specification of μ_t and some regularity conditions, the GMM estimator $\hat{\theta}_T$, obtained by solving (12) for θ, is consistent, whatever the choice of G_t (Wooldridge, 1994, theorem 7.1).

An *efficient* choice G_t^* of the instrument matrix translates into the estimator having the smallest asymptotic covariance matrix. From

$$G_t^* = -E(\nabla_\theta u_t'|\mathcal{F}_{t-1}) V(u_t|\mathcal{F}_{t-1})^{-1} \tag{13}$$

(cf. Newey and McFadden, 1994, section 5.4), and given that in the vMEM context

$$\nabla_\theta u_t' = -\nabla_\theta \mu_t' \mathrm{diag}(\mu_t)^{-1} \mathrm{diag}(u_t + 1)$$

such a matrix becomes

$$G_t^* = \nabla_\theta \mu_t' \mathrm{diag}(\mu_t)^{-1} \Sigma^{-1}$$

Substituting it back into (12), this choice implies that the moment conditions are

$$\frac{1}{T}\sum_{t=1}^{T} \nabla_\theta \mu_t' [\mathrm{diag}(\mu_t) \Sigma \mathrm{diag}(\mu_t)]^{-1} (x_t - \mu_t) = 0 \tag{14}$$

and the asymptotic covariance matrix is

DOI: 10.1002/jae

$$\operatorname{avar}\left(\hat{\theta}_T\right) = \left[\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E\left[\nabla_\theta \mu_t'[\operatorname{diag}(\mu_t)\Sigma\operatorname{diag}(\mu_t)]^{-1}\nabla_{\theta'}\mu_t\right]\right]^{-1} \tag{15}$$

For $K = 1$, it is interesting to note that expression (14) simplifies to

$$\sigma^{-2}\frac{1}{T}\sum_{t=1}^{T}\nabla_\theta \mu_t \frac{x_t - \mu_t}{\mu_t^2} = 0 \tag{16}$$

and represents the first-order conditions of the univariate MEM under a Gamma distributed disturbance term (Engle and Gallo, 2006). However, this parallel between the parametric and the semiparametric framework is limited to the univariate case, since in the multivariate framework it is not possible to retrieve an explicit pdf for ε_t starting from (14) as first-order conditions. Moreover, while the solution to (16) in the univariate case does not depend on σ, in the vector version Σ is an integral part of the equation, i.e. a nuisance parameter to be estimated. This could have an impact on making inference about θ, but it is not a concern in the present context. In fact, following Newey and McFadden, 1994, section 6.2), a sufficient condition for θ being *nuisance parameter insensitive* (in the meaning of Jørgensen and Knudsen, 2004) is that

$$E\left[\nabla_{\sigma_{ij}}\left(\frac{1}{T}\sum_{t=1}^{T} G_t^* u_t\right)\right] = 0 \quad \forall i,\ j = 1,\ldots,K \tag{17}$$

This is indeed the case in a vMEM, since

$$E[\nabla_{\sigma_{ij}}\left(G_t^* u_t\right)|\mathcal{F}_{t-1}] = -\nabla_\theta \mu_t' \operatorname{diag}(\mu_t)^{-1}\Sigma^{-1}\nabla_{\sigma_{ij}}\Sigma\Sigma^{-1}\operatorname{diag}(\mu_t)^{-1}E(x_t - \mu_t|\mathcal{F}_{t-1})$$

is 0 for all t. Applying the law of iterated expectations, (17) follows.

Therefore, when Σ is estimated, (15) is still the asymptotic covariance matrix of $\hat{\theta}_T$. It can be consistently estimated by

$$\widehat{\operatorname{avar}}\left(\hat{\theta}_T\right) = \left[\frac{1}{T}\sum_{t=1}^{T}\widehat{\nabla_\theta \mu_t'}[\operatorname{diag}(\hat{\mu}_t)\hat{\Sigma}_T\operatorname{diag}(\hat{\mu}_t)]^{-1}\widehat{\nabla_{\theta'}\mu_t}\right]^{-1} \tag{18}$$

where a hat denotes quantities evaluated at $\hat{\theta}_T$, and $\hat{\Sigma}_T$ is a consistent estimator of Σ. A simple estimator, which is not compromised by zeros in the data, is

$$\hat{\Sigma}_T = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_t\hat{u}_t' \tag{19}$$

where $\hat{u}_t$ is the estimated counterpart to (10), computed using $\hat{\theta}_T$. The estimation algorithm will iterate between (14) and (19) until convergence.

3.2. Efficiency Gains: Simulation Results

In order to analyze the properties of the GMM estimator, it is instructive to run a simulation exercise. We adopt a trivariate model (9) with one lag ($K = 3, L = 1$). A fair comparison of the joint

DOI: 10.1002/jae

(i.e. considering a non-diagonal Σ) and the equation-by-equation approaches (i.e. with a Σ assumed to be diagonal) requires the same feasible specification for the dynamics of μ_t: the simulation design involves a full α_1 with zero restrictions, which could occur in practice (in our case $\alpha_{13}=\alpha_{21}=0$), a diagonal β_1 (with a non-diagonal matrix, equation-by-equation estimation is not admitted) and a diagonal γ_1. As per the disturbance term, we adopt a copula approach: marginals are simulated from Gamma distributions with unit mean and standard deviations equal to 0.5, 0.3 and 0.7, respectively. We choose a Student's t copula with 8 degrees of freedom and we envisage three scenarios, each corresponding to a different level of correlation in the copula function:

- the first, labeled *uncorrelated*, corresponds to all $\rho_{ij}=0, i\neq j$;
- the second, labeled *medium*, corresponds to $\rho_{12}=0.3$, $\rho_{13}=0.4$ and $\rho_{23}=0.5$;
- the third, labeled *high*, corresponds to $\rho_{12}=0.7$, $\rho_{13}=0.8$ and $\rho_{23}=0.9$.

Sample sizes are $T=1000$ and $T=3000$ over 1000 replications. In each scenario, we want to assess the efficiency gain of the joint approach over the equation-by-equation approach, measured as the relative improvement in the root mean square error (RMSE). The results are summarized in Table I: we report the true value of the coefficients in column (1). The results are arranged by correlation level and, within it, by sample size. The RMSE value is reported for the equation-by-equation approach; next to it we calculate the efficiency gain for each coefficient, evaluated (as a percentage) as one minus the ratio of the RMSE obtained by the joint approach to the RMSE obtained by the equation-by-equation approach.

As one would expect, the joint estimation method delivers efficiency gains which increase with the degree of correlatedness in the scenarios and remain approximately the same across sample size: from no gains under uncorrelatedness, there is a steady improvement as we move to the *strong* scenario. In order to synthesize results, we computed an average efficiency gain reported (as a percentage) in the last line of Table I as one minus the square root of the ratio between the traces of the coefficient MSE matrices obtained with the joint and with the equation-by-equation approaches, respectively. Such average gains are slightly below 10% in the *medium* scenario, and they reach 40% in the *high* scenario. Informal calculations show that the RMSE gains by sample size are in line with the root-T efficiency improvement in the coefficient covariance matrix augmented by a reduction of the bias due to consistency. Such gains are constant across estimation methods.

4. INTERDEPENDENCE ACROSS VOLATILITY MEASURES

Volatility measurement using intra-daily data (from the simplest daily range, Parkinson, 1980, to the plain vanilla realized volatility, Andersen *et al.*, 2006, to the more recent realized kernels, Barndorff-Nielsen *et al.*, 2008) has evolved in a natural complementary effort to build adequate models for volatility prediction (Corsi, 2009; Brownlees and Gallo, 2010; Shephard and Sheppard, 2010; Hansen *et al.*, 2011). The time series which are most commonly used in this respect are the squared close-to-close returns r_t^2, the realized variances rv_t^2 (in any of their flavors), the absolute returns $|r_t|$, the realized volatilities rv_t, and the daily ranges hl_t (computed as $hl_t = \sqrt{\pi/8}\, ln(\text{high}_t/\text{low}_t)$). They all show general features of persistence and an alternation between periods of high and low volatility. Such properties have long been discussed in the literature and will not be further commented upon here.

From an empirical point of view, we illustrate the characteristics of the vMEM in modeling the interaction among several volatility measures for the purpose of multiperiod forecasting. Among possible issues of interest, we concentrate on:

1. model variance versus volatility measures;
2. estimating the models equation-by-equation or jointly;

DOI: 10.1002/jae

Table I. Simulation results on GMM estimation: comparison between equation-by-equation and joint approaches. 1000 replications. True values are reported in the first column. Next to each correlation scenario and sample size, we report the coefficient RMSE of the equation-by-equation approach and the efficiency percentage gain obtained by the joint approach. The last row reports an average efficiency gain (AEG) measure (in percentage). See text for details

		Independent				Weak				Strong			
		$T = 1000$		$T = 3000$		$T = 1000$		$T = 3000$		$T = 1000$		$T = 3000$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	True	RMSE	R.Eff	RMSE	R.Eff	RMSE	R.Eff	RMSE	R.Eff	RMSE	R.Eff	RMSE	R.Eff
ω_1	2.2735	0.7147	0.1	0.3523	−0.2	0.7119	12.2	0.3407	7.9	0.7935	40.5	0.3437	33.4
α_{11}	0.08	0.0234	−0.0	0.0129	0.1	0.0245	6.9	0.0137	7.3	0.0258	24.9	0.0141	24.8
γ_{11}	0.07	0.0194	−0.4	0.0107	−0.0	0.0195	4.6	0.0108	2.4	0.0192	18.0	0.0106	16.0
α_{12}	−0.02	0.0096	0.1	0.0039	−0.1	0.0095	8.6	0.0040	10.1	0.0116	28.0	0.0046	35.4
β_{11}	0.8	0.0409	−0.0	0.0217	−0.1	0.0430	10.8	0.0222	6.8	0.0484	36.9	0.0228	31.1
ω_2	0.471	0.1412	0.1	0.0678	0.0	0.1395	1.4	0.0684	3.4	0.1421	18.1	0.0719	22.7
α_{22}	0.12	0.0234	−0.1	0.0125	−0.1	0.0246	4.1	0.0136	6.7	0.0428	27.1	0.0225	32.0
γ_{22}	0.02	0.0122	0.0	0.0066	0.1	0.0121	3.3	0.0068	6.4	0.0122	20.8	0.0069	23.4
α_{23}	0.06	0.0075	0.1	0.0041	0.0	0.0083	2.9	0.0047	5.4	0.0161	25.2	0.0086	29.9
β_{22}	0.78	0.0281	0.1	0.0144	−0.1	0.0249	4.5	0.0132	6.9	0.0305	26.9	0.0156	29.0
ω_3	0.7675	0.5112	0.0	0.2179	−0.0	0.4888	9.2	0.2237	11.7	0.5892	49.3	0.2676	46.8
α_{31}	−0.03	0.0152	−0.0	0.0067	−0.1	0.0167	14.4	0.0078	12.0	0.0242	60.8	0.0110	54.6
α_{32}	0.06	0.0411	0.2	0.0206	−0.1	0.0479	8.9	0.0242	7.4	0.0914	42.1	0.0426	40.3
α_{33}	0.1	0.0244	0.0	0.0133	−0.1	0.0260	7.1	0.0144	8.4	0.0420	34.7	0.0211	33.0
γ_{33}	0.05	0.0247	0.2	0.0134	−0.2	0.0242	4.1	0.0134	5.2	0.0248	25.2	0.0137	24.9
β_{33}	0.82	0.0432	0.1	0.0211	−0.2	0.0420	10.4	0.0211	9.5	0.0580	46.1	0.0266	41.9
AEG			0.0		−0.1		9.5		8.4		40.3		36.6

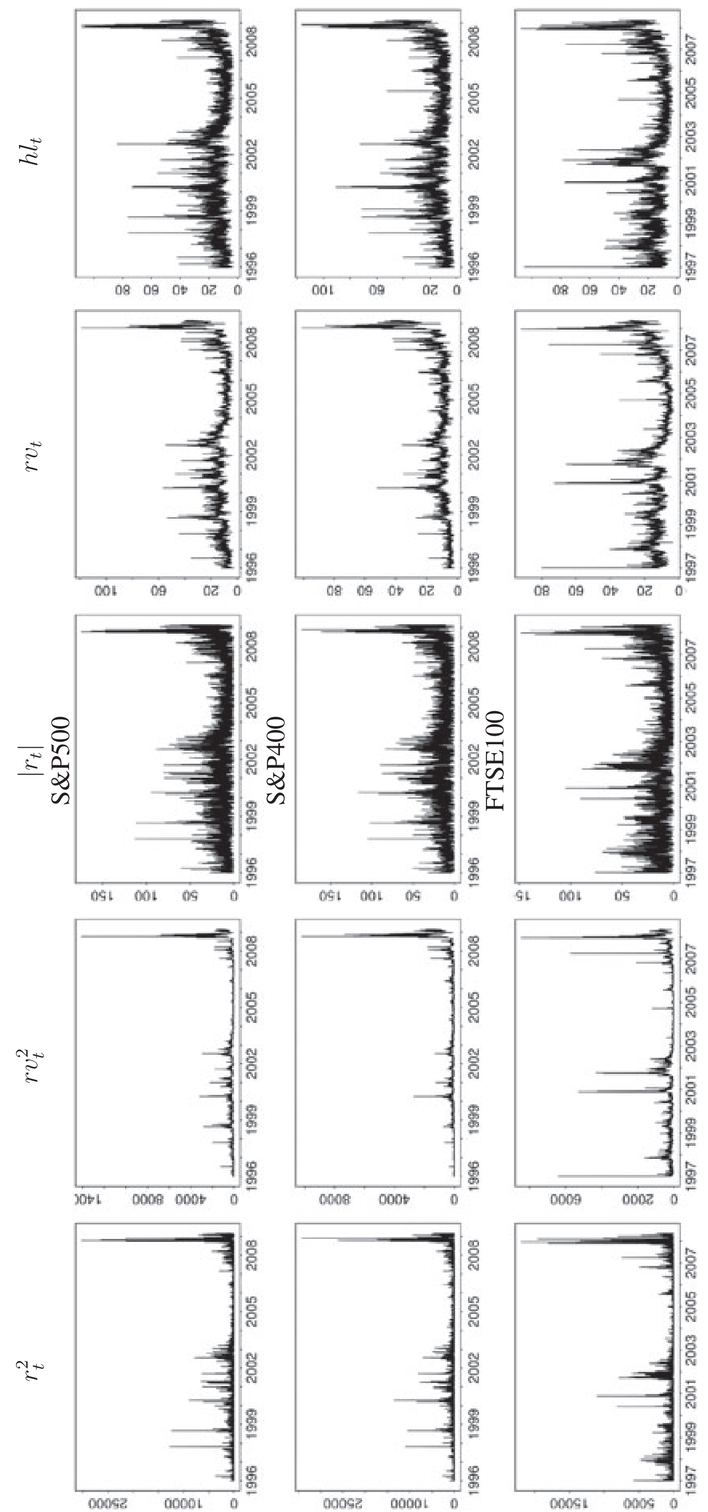

Figure 1. Time series plots of squared return (r_t^2), realized kernel variance (rv_t^2), absolute return ($|r_t|$), realized kernel volatility (rv_t), daily range (hl_t) for S&P500, S&P400, FTSE100

DOI: 10.1002/jae

3. test model restrictions and using residual diagnostics as a guideline to evaluate the specification;
4. evaluating the possible additional contribution of the daily range.

We cover the period between January 1996 and February 2009 for a total of $T = 3261$ observations for each series. We make use of the Oxford Man Institute (OMI) *Realized Library* (Shephard and Sheppard, 2010) for daily returns and realized kernels, together with the range computed using the daily highs and lows downloaded from Datastream (all measures are expressed in annualized percentage terms). We run our analysis on ten stock indices, but report the results for five of them, labeled DJ30 (Dow Jones Industrials), S&P500 (S&P 500), NASD100 (Nasdaq 100), S&P400 (S&P 400 Midcap), FTSE100 (UK FTSE 100), while moving to the Web Appendix (supporting information) the details for the remaining ones: TSE60 (S&P Toronto Stock Exchange), CAC40 (French CAC 40), DAX30 (German DAX), IBEX35 (Spanish IBEX), NIKKEI225 (Japanese Nikkei 225). As representative examples of the series appearance, see Figure 1 (just three tickers for the sake of space) where persistence and clustering are apparent.

4.1. Bivariate Modeling: The Basic vMEM

Paralleling the logic behind the HEAVY model (Shephard and Sheppard, 2010), we build a basic vMEM to jointly represent the conditional expectations of $x_t = \left(r_t^2; rv_t^2\right)$. From equation (9), taking one lag and $\gamma_1 = 0$, we specify

$$\alpha_1 = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \quad \beta_1 = \begin{pmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{pmatrix} \tag{20}$$

Such a specification allows for the simplest dynamic interaction between the two variables:

- concentrating on the first equation, we can recognize a GARCH-X (Engle, 2002), GARCH with realized variance as a predetermined variable);
- the imposition of $\alpha_{12} = 0$ in the same equation would deliver the basic GARCH;
- the HEAVY model (Shephard and Sheppard, 2010) is embedded in this vMEM and is obtained by imposing $\alpha_{11} = 0$ in the equation for the expected squared return and $\alpha_{21} = 0$ in the expression for the expected realized variance (HEAVY-RM).

The coefficients estimated equation-by-equation on the five tickers are reported in the leftmost section of Table II with *t*-statistics in parentheses underneath. The α_{11} parameters are statistically significant (with the exception of FTSE100), pointing to a generalized relevance of lagged squared returns in determining their own conditional expectation. By the same token, the coefficients α_{12}s are positive and highly significant for all five tickers. Together with this result, some α_{11} parameters take on negative values signaling that, at times, a high overall squared return may dampen the impact of a high intra-daily variance on the conditional expectation of r_t^2. For the second equation, we find a feedback of squared return on realized variance (all estimated α_{21}s are positive and significant at 5 %). All β_{11} estimates are larger than the corresponding β_{22} estimates, as discussed by Shephard and Sheppard (2010). As an additional remark, some numerical difficulties in convergence (especially for S&P500) were met for the variance series, possibly due to some occasional very large values. Similar results hold true for the other tickers (see Tables I and II in the Web Appendix (supporting information)).

A question that arises is whether the estimated dynamics change if one refers to the volatility, rather than to the variance. Assuming $x_t = (|r_t|; rv_t)$, a second round of estimates is obtained (still equation-by-equation) in the central section of Table II.

DOI: 10.1002/jae

There is a widespread tendency for the standard errors to increase, leading to a corresponding decrease in the t-statistics. As a consequence, the α_{11}s are now insignificant in two out of five cases (NASD100 joins FTSE100). The α_{12} estimates remain highly significant, while three α_{21}s lose significance and another one is borderline significant at 5%.

In the rightmost section of Table II we reproduce the parameter inference of the vMEM estimated in a GMM joint framework. As one would expect, the procedure achieves sharper standard errors. While the estimated coefficients do not change much relative to the corresponding equation-by-equation estimates (a possible exception being S&P500), the fact that volatility is a better driver of the dynamics of expected absolute return (insignificant α_{11}) is confirmed. On the other hand, the feedback effect from absolute return to volatility is present for all tickers (highly significant α_{21}s).

Further insights from the results can be gained by calculating the Wald test statistics for three joint hypotheses of interest which would strongly characterize the dynamics of the variables involved and the interactions between the two indicators. Relative to the unrestricted model with coefficients as in (20), we have

1. GARCH + HEAVY-RM, corresponding to a test of H_0: $\alpha_{12} = 0, \alpha_{21} = 0$;
2. HEAVY-R + HEAVY-RM, corresponding to a test of H_0: $\alpha_{11} = 0, \alpha_{21} = 0$;
3. GARCH-X + HEAVY-RM, corresponding to a test of H_0: $\alpha_{21} = 0$.

We show the p-values of such tests in Table III. The null hypothesis (GARCH + HEAVY-RM in column (1)) is strongly rejected across specifications, irrespective of measure or estimation method used: the realized measure significantly contributes to (squared or absolute) return dynamics. More mixed evidence is had on the other two hypotheses, in that when squared measures are involved both are always rejected (columns (2) and (3) in the top section), while the estimation method seems to be relevant for the absolute returns/volatilities: for example, GARCH-X + HEAVY-RM (column (3)) would not be rejected for any of the tickers at the 1% significance level when the equation-by-equation method is employed, while the joint estimation provides overwhelming evidence against all hypotheses, a result which can be connected to smaller standard errors in the second equation (cf. Table II). The same holds true for the remaining tickers (see Table III in the Web Appendix (supporting information)).

A complement to the analysis of the specification adopted is the diagnostics on the residuals. The results of a Ljung–Box joint residual autocorrelation test on the bivariate unrestricted vMEM (reported in Table IV of the Web Appendix (supporting information)) shows that further refinements are needed across measures and estimation methods: at lag 12, only the model for the variances of the DJ30 gives residuals with a Ljung–Box statistic that is not significant at 5%, with substantially similar results at lag 22 also. Taking equation (9) as a reference, we can increase the number of terms to be considered such as asymmetric responses to negative lagged return, more lags in the dynamics and a full β_1 matrix. We concentrate on absolute return and realized volatility (better numerical properties, less sensitivity to outliers, and direct measures of volatility) and adopt joint estimation throughout (better properties and a forced choice when β_1 is non-diagonal). Later, we extend the model to include the daily range as an additional measure of volatility.

4.2. Bivariate Modeling: Richer vMEM Formulations

Let us start by extending the dynamic structure of the bivariate model for $x_t = (|r_t|; rv_t)$. The possible computational problems from increasing the number of parameters must be taken into consideration when defining specifications. We explore four nested models:[3]

[3] When present, the matrix γ_1 is kept diagonal across specifications; a full version of it did not deliver any improvement (results not reported here for the sake of space).

DOI: 10.1002/jae

Table II. Unrestricted bivariate vMEM models: estimated coefficients (20) and t-statistics (in parentheses) across various choices of measures and estimation methods (variance measures equation-by-equation, volatility measures equation-by-equation, volatility measures joint estimation)

Ticker	$(r_t^2; rv_t^2)$, eq.-by-eq. estimation				$(\|r_t\|; rv_t)$, eq.-by-eq. estimation				$(\|r_t\|; rv_t)$, joint estimation			
	α_1		β_1		α_1		β_1		α_1		β_1	
DJ30	−0.0398	0.6243	0.6513	—	−0.0558	0.3817	0.6825	—	−0.0593	0.4083	0.6567	—
	(−4.53)	(12.43)	(23.68)		(−3.18)	(7.11)	(14.30)		(−4.01)	(8.87)	(15.88)	
	0.0263	0.3695	—	0.5604	0.0374	0.3568	—	0.5695	0.0379	0.3609	—	0.5642
	(3.47)	(10.66)		(16.01)	(1.86)	(5.66)		(8.31)	(7.20)	(21.93)		(31.64)
S&P500	−0.0281	0.4273	0.7621	—	−0.0851	0.5002	0.5771	—	−0.0672	0.3626	0.7146	—
	(−3.73)	(11.80)	(39.06)		(−4.54)	(8.52)	(10.54)		(−4.81)	(9.27)	(21.17)	
	0.0305	0.3717	—	0.5552	0.0391	0.3557	—	0.5684	0.0407	0.3558	—	0.5663
	(3.99)	(11.41)		(17.28)	(1.95)	(6.12)		(9.06)	(7.23)	(22.00)		(32.57)
NASD100	0.0244	0.4984	0.7255	—	0.0165	0.2650	0.7390	—	0.0140	0.3108	0.6960	—
	(2.59)	(9.94)	(29.72)		(1.03)	(5.26)	(15.92)		(1.04)	(6.94)	(16.74)	
	0.0212	0.4151	—	0.5236	0.0378	0.3845	—	0.5437	0.0396	0.3869	—	0.5387
	(3.99)	(12.71)		(16.14)	(2.30)	(6.69)		(8.70)	(8.56)	(23.96)		(30.66)
S&P400	0.0839	0.2546	0.7956	—	0.0639	0.1281	0.8146	—	0.0676	0.1232	0.8141	—
	(8.68)	(7.86)	(44.13)		(4.59)	(4.25)	(24.47)		(5.95)	(5.28)	(31.17)	
	0.0063	0.3906	—	0.5854	0.0119	0.3568	—	0.6096	0.0132	0.3416	—	0.6244
	(1.98)	(12.11)		(18.80)	(0.98)	(6.23)		(10.12)	(4.17)	(23.48)		(40.71)
FTSE100	−0.0131	0.6101	0.6511	—	−0.0236	0.3824	0.6585	—	−0.0254	0.3953	0.6471	—
	(−1.13)	(10.28)	(21.14)		(−1.13)	(5.62)	(11.04)		(−1.56)	(7.44)	(13.87)	
	0.0438	0.3306	—	0.5905	0.0406	0.3015	—	0.6380	0.0411	0.3127	—	0.6250
	(4.92)	(9.17)		(17.00)	(1.97)	(4.76)		(9.80)	(7.05)	(17.42)		(33.87)

DOI: 10.1002/jae

Table III. p-values of the Wald tests of three restricted specifications (headers GARCH + HEAVY-RM, HEAVY-R + HEAVY-RM and GARCH-X + HEAVY-RM, respectively) against the unrestricted bivariate vMEM (full α_1 and diagonal β_1). Various choices of measures and estimation methods (variance measures equation-by-equation, volatility measures equation-by-equation, volatility measures joint estimation)

Measures and estimation method	Ticker	(1) GARCH HEAVY-RM	(2) HEAVY-R HEAVY-RM	(3) GARCH-X HEAVY-RM
$(r_t^2; rv_t^2)$	DJ30	0.0000	0.0000	0.0005
	S&P500	0.0000	0.0000	0.0001
Eq.-by-eq.	NASD100	0.0000	0.0000	0.0001
estimation	S&P400	0.0000	0.0000	0.0480
	FTSE100	0.0000	0.0000	0.0000
$(\|r_t\|; rv_t)$	DJ30	0.0000	0.0011	0.0630
	S&P500	0.0000	0.0000	0.0511
Eq.-by-eq.	NASD100	0.0000	0.0412	0.0212
estimation	S&P400	0.0001	0.0000	0.3284
	FTSE100	0.0000	0.0763	0.0494
$(\|r_t\|; rv_t)$	DJ30	0.0000	0.0000	0.0000
	S&P500	0.0000	0.0000	0.0000
Joint	NASD100	0.0000	0.0000	0.0000
estimation	S&P400	0.0000	0.0000	0.0000
	FTSE100	0.0000	0.0000	0.0000

1. The first contains a full α_1 and a diagonal β_1 (Model 1).
2. Model 1 plus a diagonal α_2 and a diagonal γ_1 (Model 2).
3. Same as model 2, but with a full α_2 (Model 3).
4. Same as model 3, but with a full β_1 (Model 4).

Rather than reporting the values of the coefficients and the individual significance in the various cases, a natural comparison is a sequence of Wald-type tests, where richer models are tested against simpler models. By summarizing (details in Table V of the Web Appendix (supporting information)) the results point to an overall rejection of any restriction for all ten tickers, signaling the empirical support for richer dynamics.

Another good summary of the possible improvements in the dynamic properties of the model can be had by looking at Table IV, where we report the p-values of the Ljung–Box statistics for autocorrelation at lag 22 by each specification (each column corresponds to the models above). While the null of no autocorrelation is still rejected (at least at one horizon for each ticker with the exception of S&P500), we notice a marked improvement over the base specification (Model 1) and earlier reported results. In particular, the insertion of one more lag and asymmetric effects (column (2)) is already capable of increasing the p-values by several orders of magnitude (especially at higher lags) over those in column (1). A further improvement is had by specifying a full α_2 (column (3)). The addition of a full β_1 (column (4)) provides another contribution. Therefore, a partial conclusion is that richer dynamics are needed (more lags in the autoregressive part, asymmetric effects, possibly interaction among lagged conditional expectations). Full details at several lags for all tickers are reported in Table VI of the Web Appendix (supporting information) (marked improvement in residual autocorrelation moving from the simple to the full specification).

4.3. Trivariate Modeling

We want now to investigate the contribution of the daily range within a trivariate volatility vMEM (when appropriate, we denote *diagonal* matrices by superscript (d) and *full* matrices by superscript (f)). We need

DOI: 10.1002/jae

Table IV. Joint Ljung–Box statistics at lag 22 for different specifications of a bivariate vMEM on *absolute return* and *realized kernel volatility*. Matrices included in the specification appear with superscript (d) when they are *diagonal*, and with superscript (f) when *full*

	(1)	(2)	(3)	(4)
Ticker	$\alpha_1^{(f)}$ $\beta_1^{(d)}$	$\alpha_1^{(f)}$ $\alpha_2^{(d)}$ $\gamma_1^{(d)}$ $\beta_1^{(d)}$	$\alpha_1^{(f)}$ $\alpha_2^{(f)}$ $\gamma_1^{(d)}$ $\beta_1^{(d)}$	$\alpha_1^{(f)}$ $\alpha_2^{(f)}$ $\gamma_1^{(d)}$ $\beta_1^{(f)}$
DJ30	0.0048	0.1215	0.3479	0.1389
S&P500	0.0021	0.1133	0.3121	0.1076
NASD100	0.0000	0.0000	0.0019	0.0573
S&P400	0.0000	0.0000	0.0066	0.2187
FTSE100	0.0000	0.0093	0.0413	0.0431

to keep to a reasonable number the possible specifications to investigate (inclusion/exclusion; diagonal/full matrices); by the same token, we also need to deal with the fact that some individual coefficients may be zero. Thus we adopt a general-to-specific model selection procedure where (for identification purposes) we always keep the diagonal elements of α_1 and β_1, irrespective of their statistical significance (Cipollini and Gallo, 2010). The coefficients estimated by GMM with $\alpha_1^{(f)}$, $\alpha_2^{(f)}$, $\gamma_1^{(d)}$ and $\beta_1^{(f)}$ are reported in Table V. All full matrices retain the non-diagonality even after pruning – first and foremost α_1, but also α_2 and β_1, show many significant interactions, some with a negative sign. Asymmetric effects connected with the sign of the return are often absent (all details are in Tables VII and VIII in the Web Appendix (supporting information)).

In order to evaluate dynamic links, we resort to the calculation of Granger non-causality test statistics, the *p*-values of which are reported in Table VI. The weakest link seems to be the one from lagged absolute return (or the lagged conditional expectation thereof) to current daily range since in three cases such a link is non-existent (as represented by a dash in the table when the corresponding coefficients are set to zero); in other cases, the link is related to a single significant coefficient: for the DJ30 it is β_{31}; for the FTSE100 it is α_{31} at lag 2 (borderline significant for the latter). The other missing link is the vice versa for the FTSE100.[4]

With this proviso, for all tickers, there is general evidence of strong dynamic interdependence among the volatility measures, not just a unidirectional link from realized kernel volatility to absolute returns (see also Forsberg and Ghysels, 2007). This also seems to confirm the importance of separately including continuous and jump variation when modeling the dynamic dependencies in realized volatilities (e.g. Andersen *et al*., 2007). In this respect, the range could be serving as a proxy for jumps in the realized volatility equation.

In this richer model, it is also evident that moving from a simpler specification (full α_1 and diagonal β_1) toward one where full interactions are incrementally included brings about a strong improvement in the *p*-values of the Ljung–Box test statistics at various lags for all tickers: Table VII shows such statistics just at lag 22 (full details are in Table IX of the Web Appendix (supporting information)). This is particularly true for the last column in the table (full α_1, α_2 and β_1, diagonal γ_1), where four out of five tickers do not show significant autocorrelation at 5%, with some contribution to the improvement coming from the presence of a full β_1 matrix.

The fully interdependent vMEM accompanied by the joint GMM estimation method is well suited to capture the dynamic interdependence across volatility measures. From a substantive point of view, it emerges that the dynamics for absolute return involves many elements. Take, for example, the DJ30 equation; rewriting results from the first line in Table V yields

[4] These outcomes are strengthened by the inspection of the enlarged set of results in the Web Appendix (supporting information). The link from lagged absolute return and range shows up as missing in five out of ten cases (with another one, the CAC40, being borderline significant) and the link from range and absolute return is missing in two out of ten.

DOI: 10.1002/jae

Table V. Trivariate vMEM *absolute return* ($|r_t|$), *realized kernel volatility* (rv_t) and *high–low range* (hl_t): joint estimation. Coefficients and t-statistics (in parentheses)

Ticker		α_1			α_2			γ_1			β_1	
DJ30	−0.0723	0.5030	−0.0922	0.0447	—	—	0.0324	—	—	0.4270	−0.5225	0.6361
	(−6.77)	(9.84)	(−2.64)	(3.18)			(2.15)			(4.47)	(−4.83)	(7.14)
	0.0108	0.2706	0.0696	—	−0.1577	—	—	0.0532	—	0.0666	0.6760	—
	(2.05)	(43.36)	(5.18)		(−7.09)			(2.51)		(2.63)	(19.82)	
	—	0.3159	−0.0768	—	—	—	—	—	0.0760	−0.3075	—	1.0313
		(9.79)	(−9.61)						(3.61)	(−4.51)		(16.02)
S&P500	−0.1045	0.3956	−0.0559	0.0771	−0.1395	—	0.0464	—	—	0.8132	—	—
	(−12.99)	(6.63)	(−2.63)	(6.18)	(−2.28)		(3.03)			(48.47)		
	0.0166	0.2849	0.0506	0.0226	−0.1587	0.0618	—	0.0751	—	0.3274	0.2795	—
	(2.31)	(35.46)	(2.62)	(2.16)	(−3.97)	(3.19)		(2.94)		(5.12)	(3.14)	
	—	0.3548	−0.0931	—	−0.1934	0.0987	—	—	0.0719	—	—	0.7898
		(9.29)	(−11.86)		(−4.61)	(5.20)			(3.55)			(43.67)
NASD100	−0.0813	0.5035	—	0.0946	−0.2489	—	—	—	—	0.9737	−0.2514	—
	(−5.10)	(8.06)		(5.89)	(−2.96)					(127.80)	(−6.95)	
	0.0490	0.4384	—	−0.0467	−0.2858	—	—	—	—	−0.0092	0.8486	—
	(10.43)	(24.54)		(−10.01)	(−11.79)					(−2.39)	(55.63)	
	—	0.4019	0.0351	—	−0.2028	—	—	—	—	—	−0.2074	0.9619
		(6.99)	(6.62)		(−2.64)						(−5.60)	(166.69)
S&P400	0.0012	0.3422	—	—	—	0.0470	—	—	—	0.9245	−0.3336	—
	(0.23)	(10.40)				(3.77)				(69.29)	(−9.83)	
	—	0.3151	0.0652	—	−0.1320	−0.0552	—	—	—	−0.0224	0.8209	—
		(14.78)	(7.98)		(−5.00)	(−7.90)				(−4.33)	(49.49)	
	—	0.3517	0.0692	—	−0.1105	—	—	—	—	—	−0.2112	0.8901
		(9.39)	(6.42)		(−2.75)						(−6.52)	(57.66)
FTSE100	−0.0470	0.3594	—	0.0741	—	—	—	—	—	0.8258	−0.2154	—
	(−3.25)	(9.10)		(4.32)						(24.36)	(−4.89)	
	0.0163	0.2984	0.0838	—	−0.1757	−0.0592	—	—	—	—	0.9772	−0.1322
	(3.95)	(11.74)	(5.89)		(−5.92)	(−4.83)					(149.04)	(−8.14)
	—	0.3333	0.0487	0.0277	−0.1504	—	—	—	—	—	—	0.7573
		(9.64)	(3.71)	(3.87)	(−4.36)							(33.09)

DOI: 10.1002/jae

Table VI. p-values of the non-causality tests for a trivariate vMEM on *absolute return* ($|r_t|$), *realized kernel volatility* (rv_t) and *high–low range* (hl_t). The specification is $\alpha_1^{(f)}$, $\alpha_2^{(f)}$, $\gamma_1^{(d)}$ and $\beta_1^{(f)}$

Null hypothesis	DJ30	S&P500	NASD100	S&P400	FTSE100
$\|r\| \nleftarrow rv$	0.0000	0.0000	0.0000	0.0000	0.0000
$\|r\| \nleftarrow hl$	0.0000	0.0002	0.0157	0.0086	—
$rv \nleftarrow \|r\|$	0.0062	0.0000	0.0000	0.0000	0.0001
$rv \nleftarrow hl$	0.0000	0.0000	0.0000	0.0000	0.0000
$hl \nleftarrow \|r\|$	0.0000	—	—	—	0.0001
$hl \nleftarrow rv$	0.0000	0.0000	0.0000	0.0000	0.0000
$\|r\| \nleftarrow rv, hl$	0.0000	0.0000	0.0000	0.0000	0.0000
$rv \nleftarrow \|r\|, hl$	0.0000	0.0000	0.0000	0.0000	0.0000
$hl \nleftarrow \|r\|, rv$	0.0000	0.0000	0.0000	0.0000	0.0000

Table VII. p-values of the joint Ljung–Box statistics at lag 22 on four specifications of a trivariate vMEM on *absolute return* ($|r_t|$), *realized kernel volatility* (rv_t) and *high–low range* (hl_t). Matrices included in the specification appear with superscript (d) when they are *diagonal* and with superscript (f) when *full*

Ticker	$\alpha_1^{(f)}\ \beta_1^{(d)}$	$\alpha_1^{(f)}\ \alpha_2^{(d)}\ \gamma_1^{(d)}\ \beta_1^{(d)}$	$\alpha_1^{(f)}\ \alpha_2^{(f)}\ \gamma_1^{(d)}\ \beta_1^{(d)}$	$\alpha_1^{(f)}\ \alpha_2^{(f)}\ \gamma_1^{(d)}\ \beta_1^{(f)}$
DJ30	0.0005	0.0252	0.1972	0.4067
S&P500	0.0000	0.0003	0.1020	0.0877
NASD100	0.0000	0.0000	0.0000	0.0000
S&P400	0.0000	0.0000	0.0181	0.2376
FTSE100	0.0000	0.0004	0.1475	0.2672

$$\begin{aligned} E(|r_t||\mathcal{F}_{t-1}) = {} & \underset{(0.03)}{0.0158} - \underset{(-6.77)}{0.0723}\,|r_{t-1}| + \underset{(9.84)}{0.5030}\,rv_{t-1} - \underset{(-2.64)}{0.0922}\,hl_{t-1} + \underset{(3.18)}{0.0447}\,|r_{t-2}| \\ & + \underset{(2.15)}{0.0324}\,|r_{t-1}|I(r_{t-1}<0) + \underset{(4.47)}{0.4270}\,E(|r_{t-1}||\mathcal{F}_{t-2}) - \underset{(-4.83)}{0.5225}\,E(rv_{t-1}|\mathcal{F}_{t-2}) \\ & + \underset{(7.14)}{0.6361}\,E(hl_{t-1}|\mathcal{F}_{t-2}) \end{aligned}$$

to be compared with the simpler bivariate expression (cf. the rightmost block in Table II):

$$E(|r_t||\mathcal{F}_{t-1}) = \underset{(0.15)}{0.1007} - \underset{(-4.01)}{0.0593}\,|r_{t-1}| + \underset{(8.87)}{0.4083}\,rv_{t-1} + \underset{(15.88)}{0.6567}\,E(|r_{t-1}||\mathcal{F}_{t-2})$$

For the ticker at hand, volatility modeling (and forecasting) is fully dependent on the most recent data on all indicators at lag 1 and even on itself at lag 2, with asymmetric effects and full dependence on past conditional expectations. As discussed earlier, other tickers and other measures may present different patterns, but the overall conclusion is that this kind of relationship ($\hat{\beta}_1$ always non-diagonal) could not be estimated equation-by-equation. Moreover, further support for the joint estimation approach is given by the correlations between the residuals $\hat{\varepsilon}$. Referring to the simulation design of Section 3.2, the gains in efficiency by the joint approach can be seen as substantial since our estimates fall between the *medium* and the *high* scenarios: in fact, the residual components relative to $(|r|, rv)$, $(|r|, hl)$, (rv, hl) have correlations around 0.25, above 0.7 and above 0.6, respectively (for nine out of ten tickers, NASD100 being an exception, with substantially lower values).

DOI: 10.1002/jae

4.4. Responses to Shocks

We compare the different vMEM formulations by analyzing impulse response functions to shocks. For the purposes of this paper, a *shock* to the ith measure is a vector where the ith innovation $\varepsilon_{t,i}$ is set to one plus its (unconditional) standard deviation σ_i and the other components $\varepsilon_{t,j}$ ($j \neq i$), are adjusted in terms of the corresponding linear projections:

$$E\left(\varepsilon_{t,j}|\varepsilon_{t,i} = 1 + \sigma_i\right) = 1 + \frac{\sigma_{i,j}}{\sigma_i^2}\sigma_i$$

Accordingly, the sequence of shocked forecasts is

$$\mu_{t+\tau|t}^{(i)} = E\left(x_{t+\tau}|\mathcal{F}_t, \varepsilon_t = 1 + \frac{\Sigma_{.,i}}{\sigma_i}\right)$$

so that the corresponding impact can be evaluated by comparing them with a *baseline* prediction obtained by setting ε_t to the unit vector, namely, $\mu_{t+\tau|t} = E(x_{t+\tau}|\mathcal{F}_t)$.

We then plot the trajectory of

$$\mu_{t+\tau|t}^{(i)} \oslash \mu_{t+\tau|t} - 1$$

All such impulse response functions are time-dependent, in that they vary with the initial starting date. For the purpose of illustration, in Figure 2 we report the trajectories obtained for a single date (4 May 2005) for the DJ30 ticker for two models of reference: the first is a bivariate vMEM model for $|r_t|$ and rv_t with full α_1 and diagonal β_1 according to coefficients reported in Table II; the second is a trivariate vMEM for $|r_t|$, rv_t, and hl_t with full α_1, α_2, and β_1, and a diagonal γ_1, according to coefficients reported in Table V. The panels in Figure 2 summarize well the differences between the two specifications: together with the presence of the daily range, in the latter we have a richer structure of the coefficient matrices (two lags, presence of off-diagonal coefficients). This produces substantially different profiles of the responses, confirming the interdependence detected in Table VI and also introducing a momentum effect in the dynamics of volatility which is absent in the bivariate model for this day.

5. CONCLUDING REMARKS

Market activity is reflected by many observable indicators derived from ultra-high-frequency data (e.g. duration, volatility, volume, number of trades); as they all relate to information flow, their dynamics should be modeled taking interdependence into consideration. Since most of these variables are non-negative-valued, in this paper we chose to cast the problem in a vMEM framework. We model the vector of variables as the element-by-element product of a unit mean innovation term times the conditional expectation of each indicator (a linear function of past values of all observables and of all conditional expectations). This model is semiparametric in that the nature of the innovation process is left unspecified and its parameters can be estimated by GMM. Simulation results show that the gains in efficiency over an equation-by-equation approach can be relevant.

We use this approach to contribute to the debate about which model should be adopted when forecasting asset return volatility. We show that several volatility measures (absolute return, realized kernel volatility and daily range) capture different aspects of the dynamics generated by the

DOI: 10.1002/jae

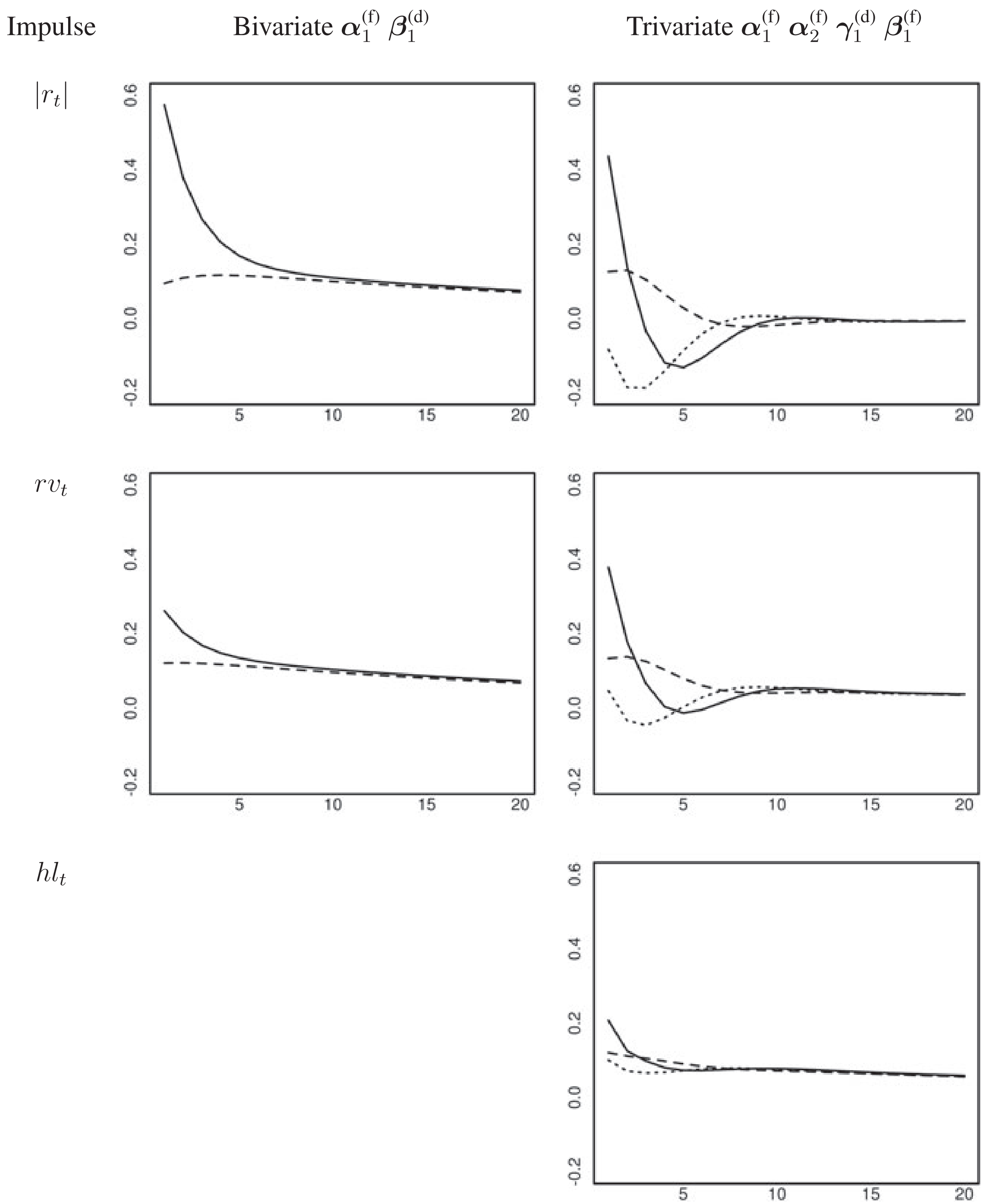

Figure 2. Impulse–response trajectories on the DJ30 ticker computed at 4 May 2005, where a row labels the variable giving the impulse according to the model by column (responses are: $|r_t|$ solid line; rv_t dashed line; hl_t dotted line)

information flow during market activity. The dynamic interdependence among measures is empirically supported by the significant non-diagonality of all estimated coefficient matrices, a result which can be achieved only by our joint GMM estimation. From an empirical point of view, the consequence is that

DOI: 10.1002/jae

realized measures do not dominate other measures in forecasting volatility: our results imply that all indicators give relevant contributions to the volatility predictions and suggest caution in adopting restricted formulations of this model.

The potential of the vMEM is far reaching since this model can be applied to the analysis of dynamic interdependence in volatilities among several assets, or among several markets or, yet, among indicators of market activity: volatility itself, volume, number of trades and average daily duration.

ACKNOWLEDGEMENTS

Without implicating, we acknowledge comments by Nour Meddahi and Kevin Sheppard which have led us to present in a more general GMM framework the Estimating Functions approach initially suggested by Cipollini *et al*. (2006). With the usual proviso, thanks are due to Monica Billio, Tim Bollerslev, Christian T. Brownlees, Nikolaus Hautsch, Peter R. Hansen, Neil Sheppard, David Veredas, the Associate Editor and three anonymous referees. Part of this research was funded by the Italian MIUR under grant PRIN 2006131140_004.

REFERENCES

Ahoniemi K, Lanne M. 2009. Joint modeling of call and put implied volatility. *International Journal of Forecasting* **25**: 239–258.

Andersen TG, Bollerslev T, Christoffersen PF, Diebold FX. 2006. Volatility and correlation forecasting. In *Handbook of Economic Forecasting*, Elliott G, Granger CWJ, Timmermann A (eds). North-Holland: Amsterdam; 777–878.

Andersen TG, Bollerslev T, Diebold FX. 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.

Barndorff-Nielsen OE, Hansen PR, Lunde A, Shephard N. 2008. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* **76**: 1481–1536.

Bollerslev T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**: 307–327.

Brownlees CT, Gallo GM. 2010. Comparison of volatility measures: a risk management perspective. *Journal of Financial Econometrics* **8**: 29–56.

Brownlees CT, Cipollini F, Gallo GM. 2012. Multiplicative error models. In *Volatility Models and their Applications*, Bauwens L, Hafner C, Laurent S (eds). Wiley: Hoboken, NJ; 223–247.

Chou RY. 2005. Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model. *Journal of Money, Credit, and Banking* **37**: 561–582.

Cipollini F, Gallo GM. 2010. Automated variable selection in vector multiplicative error models. *Computational Statistics and Data Analysis* **54**: 2470–2486.

Cipollini F, Engle RF, Gallo GM. 2006. Vector multiplicative error models: representation and inference. Technical Report 12690, National Bureau of Economic Research.

Cipollini F, Engle RF, Gallo GM. 2007. A model for multivariate non-negative valued processes in financial econometrics. Technical Report 2007/16, Dipartimento di Statistica, Università di Firenze.

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

Engle RF. 2002. New frontiers for ARCH models. *Journal of Applied Econometrics* **17**: 425–446.

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

Engle RF, Russell JR. 1998. Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica* **66**: 1127–1162.

Engle RF, Gallo GM, Velucchi M. 2012. Volatility spillovers in East Asian financial markets: a MEM based approach. *The Review of Economics and Statistics* **94**: 222–233.

Forsberg L, Ghysels E. 2007. Why do absolute returns predict volatility so well? *Journal of Financial Econometrics* **5**: 31–67.

Hansen PR, Huang Z, Shek HH. 2011. Realized GARCH: a complete model of returns and realized measures of volatility. *Journal of Applied Econometrics* (published online).

DOI: 10.1002/jae

Hautsch N. 2008. Capturing common components in high-frequency financial time series: a multivariate stochastic multiplicative error model. *Journal of Economic Dynamics and Control* **32**: 3978–4015.
Jørgensen B, Knudsen SJ. 2004. Parameter orthogonality and bias adjustment for estimating functions. *Scandinavian Journal of Statistics* **31**: 93–114.
Manganelli S. 2005. Duration, volume and volatility impact of trades. *Journal of Financial Markets* **8**: 377–399.
McNeil AJ, Frey R, Embrechts P. 2005. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press: Princeton, NJ.
Newey WK, McFadden D. 1994. Large sample estimation and hypothesis testing. In *Handbook of Econometrics*, Vol. **4**, Engle RF, McFadden D (eds). Elsevier: Amsterdam; 2111–2245.
Noss J. 2007. The econometrics of optimal execution in an order-driven market. Technical report, Masters thesis, Nuffield College, Oxford University.
Parkinson M. 1980. The extreme value method for estimating the variance of the rate of return. *Journal of Business* **53**: 61–65.
Shephard N, Sheppard K. 2010. Realising the future: forecasting with high frequency based volatility (HEAVY) models. *Journal of Applied Econometrics* **25**: 197–231.
Wooldridge JM. 1994. Estimation and inference for dependent processes. In *Handbook of Econometrics*, Vol. **4**, Engle RF, McFadden D (eds). Elsevier: Amsterdam; 2639–2738.

DOI: 10.1002/jae