Dipartimento di Statistica
"Giuseppe Parenti"

A Time-varying Mixing Multiplicative Error Model for Realized Volatility

Giovanni De Luca,
Giampiero M. Gallo

Università degli Studi
di Firenze

WORKING PAPER 2010/03

Dipartimento di Statistica "G. Parenti" – Viale Morgagni 59 – 50134 Firenze - www.ds.unifi.it

A Time-varying Mixing Multiplicative Error Model for Realized Volatility

Giovanni De Luca

Dipartimento di Statistica e Matematica per la Ricerca Economica

Università di Napoli *Parthenope.*

Via Medina, 40 – 80133 Napoli, Italy – (*giovanni.deluca@uniparthenope.it*)

Giampiero M. Gallo

Dipartimento di Statistica "G. Parenti"

Università di Firenze

Viale G. B. Morgagni, 59 - 50134 Firenze, Italy – (*gallog@ds.unifi.it*)

April 12, 2010

Abstract

In this paper we model the dynamics of realized volatility as a Multiplicative Error Model with a mixture of distributions for the innovation term with time-varying mixing weights forced by past behavior of volatility. The mixture considers innovations as a source of time-varying volatility of volatility and is able to capture the right tail behavior of the distribution of volatility. The empirical results show that there is no substantial difference in the one-step ahead conditional expectations obtained according to various mixing schemes but that fixity of mixing weights may be a binding constraint in deriving accurate quantiles of the predicted distribution.

Keywords: Multiplicative Error Models, Realized Volatility, Mixture Distributions

JEL: C22, C51, C53

1 Introduction

In financial econometrics, increasing effort has been devoted to the measurement of volatility of asset prices for the interest it has in risk management, derivative pricing, and asset allocation. While GARCH modeling with daily data has become common among practitioners, in recent times, a lot of work was poured into the potential for empirical applications presented by the availability of intra–daily data. The realized volatility literature has developed techniques for estimating the unobservable quadratic variation of an underlying continuous time process for the evolution of asset prices, thus suggesting a valid alternative to modeling the conditional variance of returns.

Although it is vulnerable to some phenomena affecting asset price formation such as jumps and microstructure noise, the main measure of reference is the daily realized variance as the sum of intradaily squared returns. In what follows we will not address the issues of sampling schemes for market microstructure noise reduction nor the possible presence of jumps, and limit ourselves to a realized variance based on five–minute returns. Rather, we want to show how a Multiplicative Error Model (Engle, 2002) coupled with a mixture of distributions approach provides a model for the dynamics of the time series of interest and a basis for forecasting, which can be applied to any other measure of volatility. In so doing, we pursue an alternative to Lanne (2006) by showing that a single conditional expectation may suffice and that a time-varying mixing coefficient introduces a rich extension of the specification.

The debate as to which strategy to follow to investigate the dynamics of the series is open to question: the realized volatility (or variance)[1] is a positive valued process, with a seeming long memory feature suggested by a slow, hyperbolic decay of the unconditional autocorrelation function. Some authors prefer to model realized volatility as a linear function of the past: a simple linear autoregressive model, an ARFIMA model (Andersen *et al.*, 2003), or models where data are sampled at different frequencies: daily, weekly and monthly in the heterogeneous Autoregressive (HAR) model by Corsi (2009); a weighted average of past daily realized volatility to predict longer period realized variance in the MIDAS model in Forsberg and Ghysels (2007). Other authors adopt a logarithmic transformation and apply linear

[1]From now on, we will refer to volatility although the considerations developed apply to its squares as well, unless otherwise noted.

models on the log of realized variance (Forsberg and Ghysels, 2007).

A different approach favored here is the one adopted by several authors, following the suggestion by Engle (2002; extending the logic of the Autoregressive Conditional Duration (ACD) models of Engle and Russell, 1998) who suggested the use of a Multiplicative Error Model (MEM) for positive valued processes, namely of the product of a conditional expectation that follows a GARCH–type dynamics and a positive valued innovation.

Engle and Gallo (2006) showed that the first order conditions of the likelihood function of an MEM with Gamma innovations do not depend on the shape of the Gamma distribution. This has the interesting consequence that, for example, all ACD (a special case of MEM) models estimated under an exponential distribution assumption can be reinterpreted as Gamma driven. A further point made by the same authors is that mixing information on different measures of volatility (as in their application on squared returns, daily range and realized variance) is beneficial for forecasting.

The properties of the innovation process are crucial, since they contribute to determine the capability of the model to capture the right tail behavior of the series distribution, especially in a purely univariate framework: to model durations, De Luca and Gallo (2004) show that a mixture of exponentials with fixed weights improve over a standard ACD model; De Luca and Gallo (2009) show further improvements and a greater flexibility in extending the mixture to have time-varying weights.

Parallelling this line of research, Lanne (2006) suggests an interesting flexible MEM specification in which the process for realized volatility can be seen as a mixture of two Gamma MEMs with different coefficients for the conditional expectation and different shape parameters for the Gamma (cf. also Ahoniemi and Lanne, 2009). We extend the approach even further, by adopting a MEM with mixing Gamma innovations with time–varying persistent weights driven by past realized variance multiplied by a common conditional expectation.

The basic message of the paper is that different mixing weights (a. constant with two Gammas; b. time varying with two Gammas; and c. constant with two Gammas and two conditional expectations) provide approximately the same results in a one-step ahead forecasting framework. The difference among the various approaches emerges in a stronger fashion when one shifts the attention to the predicted volatility of volatility and, more in general, to the shape of the density forecasts. The approach has a further element of interest in that the modeling strategy provides an alternative to modeling long–range dependence via fractionally integrated models.

The paper is organized as follows. In Section 2 we summarize some theoretical results on realized volatility in the context of volatility measurement and we summarize some of the features of the Multiplicative Error Model. In Section 3 we discuss at length the features of a mixture model, deriving some implications on the volatility of volatility and the coefficient of variation of the process from assuming mixtures of Gamma with fixed weights (as in Lanne's model) and with time-varying weights (as in the model suggested by us).

Section 4 illustrates the results obtainable with different models on the realized volatility of the return rates on the Deutsche Mark/US dollar, the Japanese Yen/USD and the Japanese Yen/Deutsche Mark. We show that, in spite of the better fit obtainable with a mixture of distributions for the innovations, the estimated parameters are fairly similar to one another under the different specifications. The real difference arises when extra information is added to the dynamics of the conditional expectations, either by adding two separate conditional expectations as in Lanne or extending the model to admit time varying mixing weights which depend on past realized volatility. Section 5 contains concluding remarks.

2 A Multiplicative Error Model for Realized Volatility

Various arguments can justify the interest in the high–frequency based measures of volatility[2]. Andersen and Bollerslev (1998) pointed out that squared daily returns are a noisy measure of variation: with simulation arguments they show that Mincer–Zarnowitz type regressions of squared returns on any conditional variance forecast would produce a very low R^2. Given that volatility or variance of returns is not observed, it has to be substituted with a proxy whose measurement error should vanish under certain conditions. One solution suggested is to refer to the availability of ultra–high frequency data on prices and to compute a variable called *realized variance*, constructed

[2] Cf. various survey papers by Andersen, Bollerslev and Diebold included in the references.

as

$$rv_t^2(\tau) = \sum_{i=1}^{1/\tau} r_{t-1+i\tau}^2 = \sum_{i=1}^{1/\tau} \left(p_{t-1+i\tau} - p_{t-1+(i-1)\tau}\right)^2 \tag{1}$$

where the generic term $r_{t-1+i\tau}$ is the return measured intra–daily as the log-price difference of an asset over a (very small) period τ so that its reciprocal is an integer value, representing the number of intradaily time intervals during the day. When $\tau = 1$ we get squared returns back; common choices are fractions of the trading day corresponding to five minutes or thirty minutes intervals. The theoretical support for such an approach stems from the fact that, under suitable conditions, as τ converges to zero, this measure converges to the integrated variance, that is the integral over a short period of the instantaneous (or spot) volatility of an underlying continuous time diffusion process. Other possible features of the phenomenon could be accommodated, such as the presence of jumps or of market microstructure noise (Barndorff–Nielsen and Shephard, 2002, 2004, 2006; Hansen and Lunde, 2006).

In what follows we will consider the square root of realized variance (referred to as realized volatility) and exploit its properties of being a non-negative valued series. We will model it as a multiplicative process of the form

$$rv_t = \mu_t \, \xi_t \tag{2}$$

where ξ_t is an iid stochastic process with unit conditional expected value and variance ϕ, and μ_t is the conditional expectation of realized volatility. By adopting a GARCH-type structure for μ_t, we get a Multiplicative Error Model to describe the dynamics of the conditional expectation of realized volatility,

$$\mu_t \;=\; \omega + \sum_{j=1}^{q} \alpha_j rv_{t-j} + \sum_{i=1}^{p} \beta_i \mu_{t-i} + \boldsymbol{\gamma}' \mathbf{z}_{t-1}, \tag{3}$$

in its general MEM(q,p) form. The vector $\mathbf{z}_{t-1}$ contains observable variables of possible interest in the information set.

Following Engle and Gallo (2006), it is now standard practice to consider a Gamma specification for ξ_t with one parameter (as a result of the unit mean constraint): in fact, such a specification turns out to entail the independence of parameter estimates for μ_t on the value of the shape parameter of the Gamma.

Since a MEM is a generalization of an ACD model (Engle and Russell, 1998), it is natural to extend the specification of the innovation term to accommodate more flexibility and allow for a better fit. As we will see, it is not necessarily an issue of obtaining a model which better predicts the conditional expectation (many different assumptions provide substantially equivalent forecasts), as much as one of having a flexible tool which can adapt to the varying market conditions. We are aiming at a better fit of the density of the distribution altogether, a task which proves useful when we need to derive confidence intervals for expected volatility or evaluate the probability of high values of volatility (say, in a scenario framework).

3 Mixture of distribution hypothesis

The strategy is one of generalizing expression (2) considering more flexible specifications for the innovation term ξ_t, with or without additional assumptions on the behavior of the conditional expectation term μ_t.

3.1 The Model with Mixing Gamma Innovations – MIX

We can assume for the innovation term ξ_t is a mixture of two Gammas $\xi_{1,t}$ and $\xi_{2,t}$ with a fixed mixing weight, π, as in De Luca and Gallo (2004, in an ACD framework). We would have

$$rv_t = \begin{cases} \mu_t \, \xi_{1,t} & \text{with prob } \pi \\ \mu_t \, \xi_{2,t} & \text{with prob } (1-\pi) \end{cases}, \tag{4}$$

with a corresponding density function for the innovation

$$f\left(\frac{rv_t}{\mu_t}\middle| I_{t-1}\right) = \pi \, g\left(\frac{rv_t}{\mu_t}\middle| I_{t-1}; \boldsymbol{\theta}_1\right) + (1-\pi) g\left(\frac{rv_t}{\mu_t}\middle| I_{t-1}; \boldsymbol{\theta}_2\right). \tag{5}$$

The function $f(\cdot)$ is constrained as to have conditional expectation equal to one. The parameters of two components are denoted, respectively, as $\boldsymbol{\theta}_1 = [\lambda_1 \quad \gamma_1]$ and $\boldsymbol{\theta}_2 = [\lambda_2 \quad \gamma_2]$. The unit expectation constraint implies that one of the five parameters is not free. We let

$$\gamma_2 = \frac{1 - \pi\lambda_1\gamma_1}{(1-\pi)\lambda_2}. \tag{6}$$

The variance is

$$\begin{aligned} Var(\xi_t) &= \pi\lambda_1^2\gamma_1 + (1-\pi)\lambda_2^2\gamma_2 + \pi\left(\lambda_1\gamma_1 - 1\right)^2 + (1-\pi)\left(\lambda_2\gamma_2 - 1\right)^2 \\ &= \frac{1}{1-\pi}(1 - \pi\lambda_1\gamma_1)^2 + \pi\lambda_1^2\gamma_1(1+\gamma_1) - \pi\lambda_1\lambda_2\gamma_1 + \lambda_2 - 1. \end{aligned}$$

From (2) the distribution $rv_t|I_{t-1}$, conditionally on the information set at time $t-1$, is a mixture of two distributions with conditional mean μ_t, that is

$$E(rv_t|I_{t-1}) = \mu_t.$$

Note that, given the conditional variance,

$$\text{Var}(rv_t|I_{t-1}) = \mu_t^2 \text{Var}(\xi_t),$$

the conditional coefficient of variation given by the ratio

$$\frac{\text{St.Dev.}(rv_t|I_{t-1})}{\mu_t} = \frac{\mu_t \text{St.Dev.}(\xi_t)}{\mu_t} = \text{St.Dev.}(\xi_t),$$

is constant, pointing out that the variability of the process does not have an independent source of variability other than its conditional expectation. In other words, with constant weights, the variance of realized volatility is just determined by its conditional expected value, μ_t, multiplied by a scale factor.

3.2 The Model with Time-Varying Mixing Gamma Innovations – TVM

We may want to go a step further and follow the mixture MEM approach adopting for the innovation term a mixture with time-varying weights, that is with π_t replacing π. We would have

$$rv_t = \begin{cases} \mu_t \, \xi_{1,t} & \text{with prob } \pi_t \\ \mu_t \, \xi_{2,t} & \text{with prob } (1-\pi_t) \end{cases}, \tag{7}$$

with a corresponding density function for the innovation

$$f\left(\frac{rv_t}{\mu_t}\middle| I_{t-1}\right) = \pi_t \, g\left(\frac{rv_t}{\mu_t}\middle| I_{t-1}; \boldsymbol{\theta}_1\right) + (1-\pi_t) g\left(\frac{rv_t}{\mu_t}\middle| I_{t-1}; \boldsymbol{\theta}_{2,t}\right). \tag{8}$$

Again, we denote the parameters of two components as $\boldsymbol{\theta}_1 = [\lambda_1 \quad \lambda_2]$, respectively, $\boldsymbol{\theta}_{2,t} = [\lambda_2 \quad \gamma_{2,t}]$. Note that $\boldsymbol{\theta}_{2,t}$ is time-varying, since we need unit expectation of the innovation at each period. We have allowed γ_2 to be time-varying so that

$$\gamma_{2,t} = \frac{1 - \pi_t \lambda_1 \gamma_1}{(1-\pi_t)\lambda_2}. \tag{9}$$

The variance of the mixture innovation ξ_t is now time-varying, that is $\xi_t \sim \text{iid}(1, \sigma_t^2)$. The conditional distribution of realized volatility is now characterized by time-varying mean, μ_t, as in (3), time-varying variance, $\mu_t^2\sigma_t^2$, and a time-varying coefficient of variation.

When $\text{Var}(\xi_t)$ is time-varying, the variance of realized volatility loses its dependency upon μ_t alone. In particular, under the assumption of two Gammas, the same conditional coefficient of variation is

$$\left[\pi_t\lambda_1\gamma_1(\lambda_1 + \lambda_1\gamma_1 - 2) + (1 - \pi_t)\lambda_2\gamma_{2t}(\lambda_2 + \lambda_2\gamma_{2t} - 2) + 1\right]^{1/2}.$$

In this formulation, equation (3) is used to capture the persistence in the conditional expectation (which corresponds to conditional heteroskedasticity in a GARCH framework) through μ_t, whereas the equation of π_t is used to capture the time-varying coefficient of variation (or varying volatility of volatility). In the equation of π_t, we enter those (lagged) elements that affect the conditional coefficient of variation. As a first suggestion, we can assume that the time-varying weights depend on some lagged volatility measure such as the realized volatility itself:

$$\pi_t = \frac{\exp\{\delta_0 + \delta_1 rv_{t-1} + \delta_2\pi_{t-1}\}}{1 + \exp\{\delta_0 + \delta_1 rv_{t-1} + \delta_2\pi_{t-1}\}}.$$

In the presence of a negative coefficient δ_1, a higher realized volatility involves a lower value of π_t, which implies a higher value of $\text{Var}(\xi_t)$. (the converse is true for a positive δ_1). As the empirical evidence shown later indicates, given a certain situation in $t-1$ summarized in a certain value of μ_t, a higher value of the realized volatility has an impact on the variance of the innovation term and hence is associated to a greater value of the volatility of the realized volatility.

3.3 The Model with Mixing Expectations and Gamma Innovations – LANNE

Finally, Lanne's (2006) suggestion is to take a mixture of two Gamma densities (with constant weights) and specify two equations with separate coefficients for the conditional expectations. In terms of notation this latter approach amounts to the following setup

$$rv_t = \begin{cases} \mu_{1,t}\,\xi_{1,t} & \text{with prob } \pi \\ \mu_{2,t}\,\xi_{2,t} & \text{with prob } (1-\pi) \end{cases}, \tag{10}$$

with a corresponding density function for the innovation

$$f\left(\frac{rv_t}{\mu_t}\middle| I_{t-1}\right) = \pi\, g\left(\frac{rv_t}{\mu_{1,t}}\middle| I_{t-1};\boldsymbol{\theta}_1\right) + (1-\pi)g\left(\frac{rv_t}{\mu_{2,t}}\middle| I_{t-1};\boldsymbol{\theta}_2\right). \quad (11)$$

$\mu_{1,t}$ and $\mu_{2,t}$ are the conditional expectations in each of the two regimes as in (3) with different parameters for each conditional expectation,

$$\begin{aligned}\mu_{1,t} &= \omega_1 + \sum_{j=1}^{q} \alpha_{1j} rv_{t-j} + \sum_{i=1}^{p} \beta_{1i}\mu_{1,t-i} \\ \mu_{2,t} &= \omega_2 + \sum_{j=1}^{q} \alpha_{2j} rv_{t-j} + \sum_{i=1}^{p} \beta_{2i}\mu_{2,t-i}\end{aligned}$$

Finally, $\xi_{1,t}$ and $\xi_{2,t}$ are Gamma random variables with unit expected values. As a result, the overall conditional expectation μ_t itself can be seen as a mixture of two conditional expectations.

We will perform a one–step ahead forecasting comparison among the following three models, denoted as:

1. a mixture of two Gammas with a fixed weight (MIX);
2. a mixture of two Gammas with a time–varying weight (TVM);
3. Lanne's model with fixed π but different specifications for $\mu_{1,t}$ and $\mu_{2,t}$.

4 Empirical Results

The data on intra-daily foreign exchange rates collected by Olsen and Associates have become a standard playing field to compare volatility models. Andersen et al. (2003, ABDL) and Lanne (2006), among others, have used them to illustrate the empirical properties of their contributions. In what follows we will use the daily realized volatility estimates from Dec. 2, 1986 to Nov. 27, 1996 constructed by ABDL, based on the return rates on the Deutsche Mark/US dollar (DEM/USD), the Japanese Yen/USD (JPY/USD), and the cross–rate Japanese Yen/Deutsche Mark (JPY/DEM) collected at thirty minutes intervals. Figure 1 displays the plots of annualized realized volatility for the three exchange rates. Lanne (2006) has discussed at length the issues surrounding long memory features and the fact that a MEM manages to capture the slow decaying feature of the correlogram.

Figure 1 approximately here

We comment the estimation results together across bilateral exchange rates (Table 1 for DEM/USD, Table 2 for YEN/USD and Table 3 for YEN/DEM). By and large we can say that there is no emerging dominance if one looks at the information criteria AIC and BIC, nor from the standardized residual (levels and squared) diagnostics (some marginal problems are present for MIX and TVM for JPY/DEM). The time-varying mixing weights add flexibility to the specification as we will see in forecasting. The past realized volatility does not get strong support in driving the weights (it is significant only for the JPY/USD case). There are some differences in parameter estimates for the second regime in Lanne's model where the dynamics of the conditional expectation in the second regime appears to be substantially stronger (in the case of JPY/DEM there is even a unit root). The overall dynamics must take into consideration the different parameters of the mixing Gamma (in Lanne's case the parameter λ_2 is always smaller than the corresponding parameter in the MIX or TVM).

Table 1 approximately here

Table 2 approximately here

Table 3 approximately here

Let us illustrate the behavior of the mixing weights across specifications by referring to Figure 2 where the estimated time-varying mixing weight, $\hat{\pi}_t$, is drawn together with the estimated π of the other models for the three exchange rates.

Figure 2 approximately here

In all the cases, the dynamics of $\hat{\pi}_t$ appears to be quite erratic around the fixed levels, showing some instances of substantial departure which affects mainly the tails of the one-step ahead predicted densities. We chose to illustrate this point by calculating such densities in correspondence to the lowest, respectively, the highest value of $\hat{\pi}_t$ for each exchange rate as detailed in Figure 3 (panels (a) and (b) for DEM/USD, (c) and (d) for JPY/USD, (e) and (f) for JPY/DEM). Unsurprisingly, the behavior of MIX and Lanne's models is quite similar, with the TVM showing some substantial departure in correspondence to the highest value of the time-varying weight.

Figure 3 approximately here

Finally, we calculated the mean square prediction error for horizons 1 to 10 for the three series (Figure 4). Once again, the evidence points to a substantial equivalence of the three models with a slight outperformance of Lanne's model by the MIX and TVM at longer horizons.

Figure 4 approximately here

5 Concluding Remarks

In this paper, we have introduced a specification for the Multiplicative Error Model which takes into account a mixture of Gamma distributions with time-varying parameters. The theoretical point we are stressing is that the specification allows us to have a time-varying volatility of volatility. In this specification we have chosen the forcing variable to be the lagged volatility. In spite of their exploratory flavor, the results are quite encouraging, since the specification improves above the simple Gamma specification and above a specification which has a mixture of two Gamma distributions each constrained to have unit mean, when measured by a metric of significance testing of nested hypotheses and characteristics of the residuals.

The unconstrained specification provides a better value of the likelihood function and very good diagnostic properties for the standardized residuals. The multiplicative specification seem to absorb completely the slow decay of the autocorrelation function. The analysis can be extended in a number of directions: the most immediate one is to explore whether other variables available in the information set have increased predictive power (lagged absolute returns, bipower variations or estimates of the jumps, overnight surprises).

References

[1] Ahoniemi, K. and Lanne, M. (2009), Joint Modelling of Call and Put Implied Volatility, *International Journal of Forecasting*, 25, 239–258.

[2] Andersen T.G. and Bollerslev T. (1997): Intraday Periodicity and Volatility Persistence in Financial Markets. *Journal of Empirical Finance*, 4, 115–158.

[3] Andersen T.G. and Bollerslev T. (1998), Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review*, 39, 885-905.

[4] Andersen T.G., Bollerslev, T., Diebold, F.X., and Labys, P., (2003) Modeling and Forecasting Realized Volatility, *Econometrica*, 71, 579-625.

[5] Andersen, T.G., Bollerslev, T., Christoffersen, P.F., and Diebold, F.X., (2006), Volatility and Correlation Forecasting, in G. Elliott, C.W.J. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting.* Amsterdam: North-Holland.

[6] Andersen T.G., Bollerslev, T., and Diebold, F.X., (2007) Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility, *Review of Economics and Statistics*, 89, 701-720.

[7] Bandorff-Nielsen, O., and Shephard, N., (2002), Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society*, Series B 64, 253280.

[8] Bandorff-Nielsen, O., and Shephard, N. (2004), Power and Bipower Variation with Stochastic Volatility and Jumps (with discussion) *Journal of Financial Econometrics*, 2, 1-48; ; doi:10.1093/jjfinec/nbh001.

[9] Bandorff-Nielsen, O., and Shephard, N. (2006), Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation, *Journal of Financial Econometrics*, 4, 1-30; doi:10.1093/jjfinec/nbi022.

[10] Bollerslev T. and Wooldridge J., (1992), Quasi-Maximum Likelihood Estimation of Dynamic Models with Time-Varying Covariances, *Econometric Reviews*, 11, 143-172.

[11] Corsi F. (2009), A Simple Approximate Long-Memory Model of Realized Volatility, *Journal of Financial Econometrics*, Vol. 7, No. 2, 174-196.

[12] De Luca G., Gallo G.M. (2004), Mixture Processes for Financial Durations, *SNDE*, 8.2. Available at `www.bepress.com/snde/vol8/iss2/art8`.

[13] De Luca, G. and Gallo, G.M. (2009). Time-varying Mixing Weights in Mixture Autoregressive Conditional Duration Models, *Econometric Reviews*, Vol. 28, 1-3, 102–120.

[14] Engle, R.F. (2002), New Frontiers for ARCH Models. *Journal of Applied Econometrics*, 17, 425-446.

[15] Engle, R.F. and Gallo G.M. (2006), A Multiple Indicators Model for Volatility using Intra-daily Data, *Journal of Econometrics*, 131, 3-27.

[16] Engle R.F. and Russell J.R. (1998): Autoregressive Conditional Duration: a New Model for Irregularly Spaced Transaction Data. *Econometrica*, **66**, 1127–1162.

[17] Forsberg L. and Ghysels, E. (2007), Why Do Absolute Returns Predict Volatility So Well? *Journal of Financial Econometrics*, 5, 31-67.

[18] Hansen, P.R. and Lunde, A. (2006), Realized Variance and Market Microstructure Noise (with discussion), *Journal of Business and Economic Statistics*, 24, 127–161.

[19] Lanne, M. (2006), A Mixture Multiplicative Error Model for Realized Volatility, *Journal of Financial Econometrics*, Vol. 4, No. 5, 594–616.

Table 1: Parameter estimates and diagnostics of the MIX, TVM and Lanne models for DEM/USD realized volatility – Dec. 2, 1986 to Nov. 27, 1996.

Parameters	**MIX**	**TVM**	**LANNE**	
Equation for the Expected Volatility				
			Regime 1	Regime 2
ω	0.1982	0.1914	0.1508	7.0756
	(0.0460)	(0.0452)	(0.0421)	(1.1613)
α_1	0.3412	0.3401	0.3250	0.2936
	(0.0197)	(0.0196)	(0.0198)	(0.1433)
α_2	-0.1848	-0.1872	-0.1790	0.4837
	(0.0255)	(0.0253)	(0.0250)	(0.1325)
β	0.8242	0.8283	0.8263	
	(0.0179)	(0.0177)	(0.0176)	
Shape Parameters of the Mixing Gamma				
λ_1	0.0528	0.0519	0.0577	
	(0.0029)	(0.0028)	(0.0028)	
γ_1	17.4650	17.7580	*17.331*	
	(0.8762)	(0.8725)		
λ_2	0.2325	0.2192	0.1500	
	(0.0258)	(0.0233)	(0.0184)	
γ_2	*6.4409*	*time-varying*	*6.6667*	
Parameters of the Mixing weights				
π	0.8647	*time-varying*	0.8698	
	(0.0255)		(0.0231)	
δ_0		-2.1239		
		(0.1546)		
δ_1		-0.0120		
		(0.0168)		
δ_2		4.7615		
		(0.1922)		
AIC	4.8419	4.8385	4.8373	
BIC	4.8608	4.8582	4.8570	
Std. Residuals	6.9050	6.5457	4.0734	
Squared Std. Residuals	1.2876	1.3587	1.6009	

Standard errors in parentheses. The three specifications are as follows: a mixture of two Gammas with a fixed weight (MIX); a mixture of two Gammas with a time–varying weight (TVM); Lanne's model with fixed π but different specifications for $\mu_{1,t}$ and $\mu_{2,t}$. Numbers in italics are constrained parameters. In the last two columns we reported the parameter estimates of the MEM(2,1) in the two regimes for Lanne's specification. Critical values for the Ljung-Box tests on standardized residuals and squared standardized residuals are 18.31 (5% significance), respectively 23.21 (1% significance).

Table 2: Parameter estimates and diagnostics of the MIX, TVM and Lanne models for JPY/USD realized volatility – Dec. 2, 1986 to Nov. 27, 1996.

Parameters	**MIX**	**TVM**	**LANNE**	
Equation for the Expected Volatility				
			Regime 1	Regime 2
ω	0.3041	0.3236	0.2107	0.2197
	(0.0688)	(0.0716)	(0.0648)	(0.2029)
α_1	0.4046	0.3979	0.3716	0.4963
	(0.0209)	(0.0208)	(0.0244)	(0.0809)
α_2	-0.2169	-0.2131	-0.1828	-0.4276
	(0.0304)	(0.0304)	(0.0336)	(0.1036)
β	0.7823	0.7828	0.7670	0.9285
	(0.0254)	(0.0257)	(0.0283)	(0.0478)
Shape Parameters of the Mixing Gamma				
λ_1	0.0502	0.0498	0.0543	
	(0.0043)	(0.0025)	(0.0041)	
γ_1	17.8854	18.1976	*18.1462*	
	(1.3779)	(0.8652)		
λ_2	0.2081	0.2124	0.1541	
	(0.0182)	(0.0148)	(0.0119)	
γ_2	*6.280*	*time-varying*	*6.489*	
Parameters of the Mixing weights				
π	0.7505	*time-varying*	0.7336	
	(0.0434)		(0.0412)	
δ_0		-2.4220		
		(0.0482)		
δ_1		0.0239		
		(0.0028)		
δ_2		4.4334		
		(0.0512)		
AIC	4.9225	4.9111	4.9201	
BIC	4.9415	4.9348	4.9461	
Std. Residuals	13.0839	10.9489	10.4672	
Squared Std. Residuals	9.0136	7.8873	9.9891	

Standard errors in parentheses. The three specifications are as follows: a mixture of two Gammas with a fixed weight (MIX); a mixture of two Gammas with a time–varying weight (TVM); Lanne's model with fixed π but different specifications for $\mu_{1,t}$ and $\mu_{2,t}$. Numbers in italics are constrained parameters. In the last two columns we reported the parameter estimates of the MEM(2,1) in the two regimes for Lanne's specification. Critical values for the Ljung-Box tests on standardized residuals and squared standardized residuals are 18.31 (5% significance), respectively 23.21 (1% significance).

Table 3: Parameter estimates and diagnostics of the MIX, TVM and Lanne models for JPY/DEM realized volatility – Dec. 2, 1986 to Nov. 27, 1996.

Parameters	**MIX**	**TVM**	**LANNE**	
Equation for the Expected Volatility				
			Regime 1	Regime 2
ω	0.1360	0.1337	0.0915	2.4356
	(0.0368)	(0.0361)	(0.0336)	(0.4025)
α_1	0.3575	0.3557	0.3370	0.4552
	(0.0207)	(0.0206)	(0.0235)	(0.0974)
α_2	-0.1649	-0.1653	-0.1667	0.5448
	(0.0300)	(0.0299)	(0.0316)	(0.0974)
β	0.7929	0.7952	0.8094	
	(0.0220)	(0.0218)	(0.0216)	
Shape Parameters of the Mixing Gamma				
λ_1	0.0417	0.0408	0.0453	
	(0.0030)	(0.0027)	(0.0024)	
γ_1	22.5270	23.0074	*22.0751*	
	(1.4856)	(1.3966)		
λ_2	0.1278	0.1247	0.1011	
	(0.0147)	(0.0128)	(0.0114)	
γ_2	*9.9383*	*time-varying*	*9.8912*	
Parameters of the Mixing weights				
π	0.8167	*time-varying*	0.8298	
	(0.0508)		(0.0343)	
δ_0		-2.1505		
		(0.0926)		
δ_1		-0.0080		
		(0.0124)		
δ_2		4.5783		
		(0.2250)		
AIC	4.3766	4.3732	4.3741	
BIC	4.3956	4.3969	4.3978	
Std. Residuals	18.9995	18.8704	10.3912	
Squared Std. Residuals	15.4696	15.3248	14.8376	

Standard errors in parentheses. The three specifications are as follows: a mixture of two Gammas with a fixed weight (MIX); a mixture of two Gammas with a time–varying weight (TVM); Lanne's model with fixed π but different specifications for $\mu_{1,t}$ and $\mu_{2,t}$. Numbers in italics are constrained parameters. In the last two columns we reported the parameter estimates of the MEM(2,1) in the two regimes for Lanne's specification. Critical values for the Ljung-Box tests on standardized residuals and squared standardized residuals are 18.31 (5% significance), respectively 23.21 (1% significance).

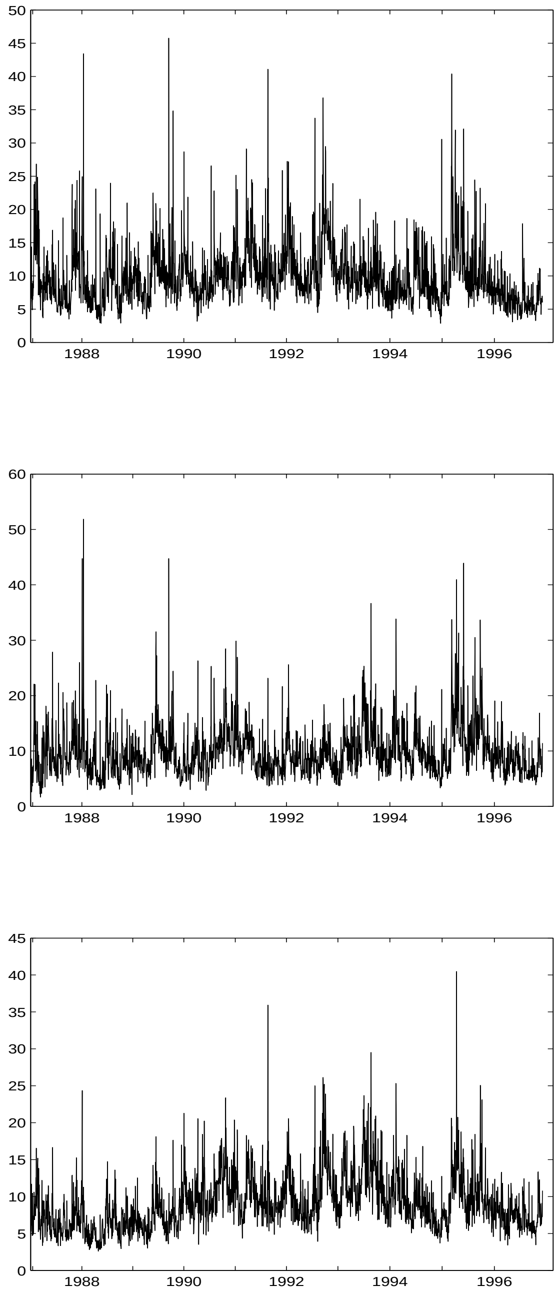

Figure 1: Annualized realized volatility for three exchange rates, DEM/USD (top), JPY/USD (middle), JPY/DEM (bottom) – Dec. 2, 1986 to Nov. 27, 1996.

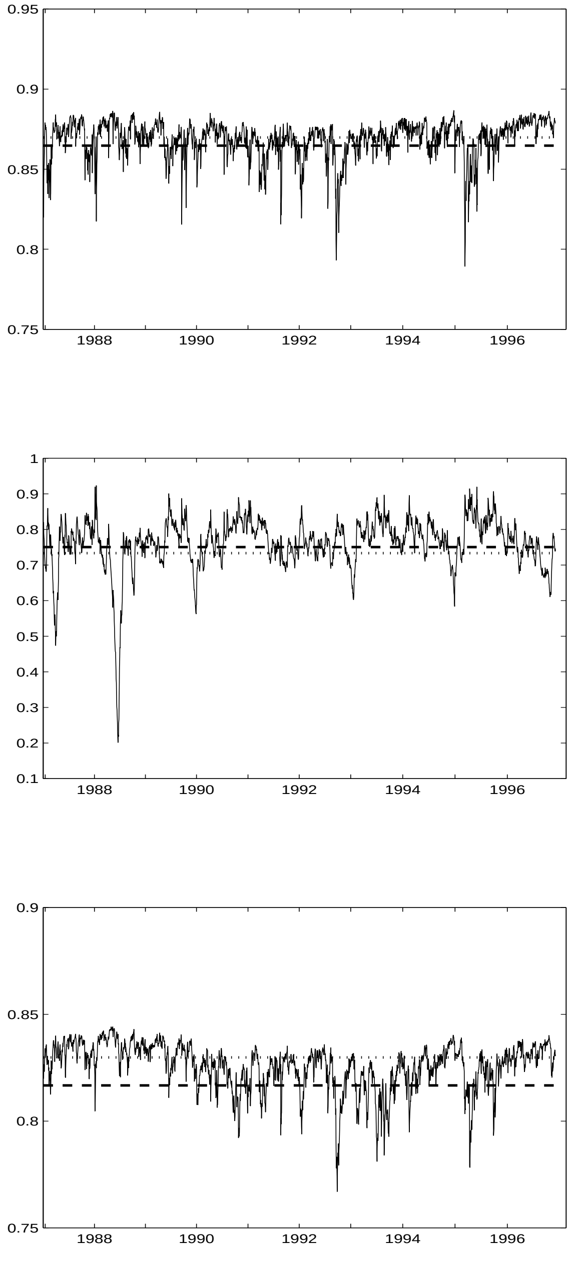

Figure 2: Estimated π for MIX (dashed line) and Lanne (dotted line) model and estimated π_t for TVM model (solid line), DEM/USD (top), JPY/USD (middle), JPY/DEM (bottom) – Dec. 2, 1986 to Nov. 27, 1996.

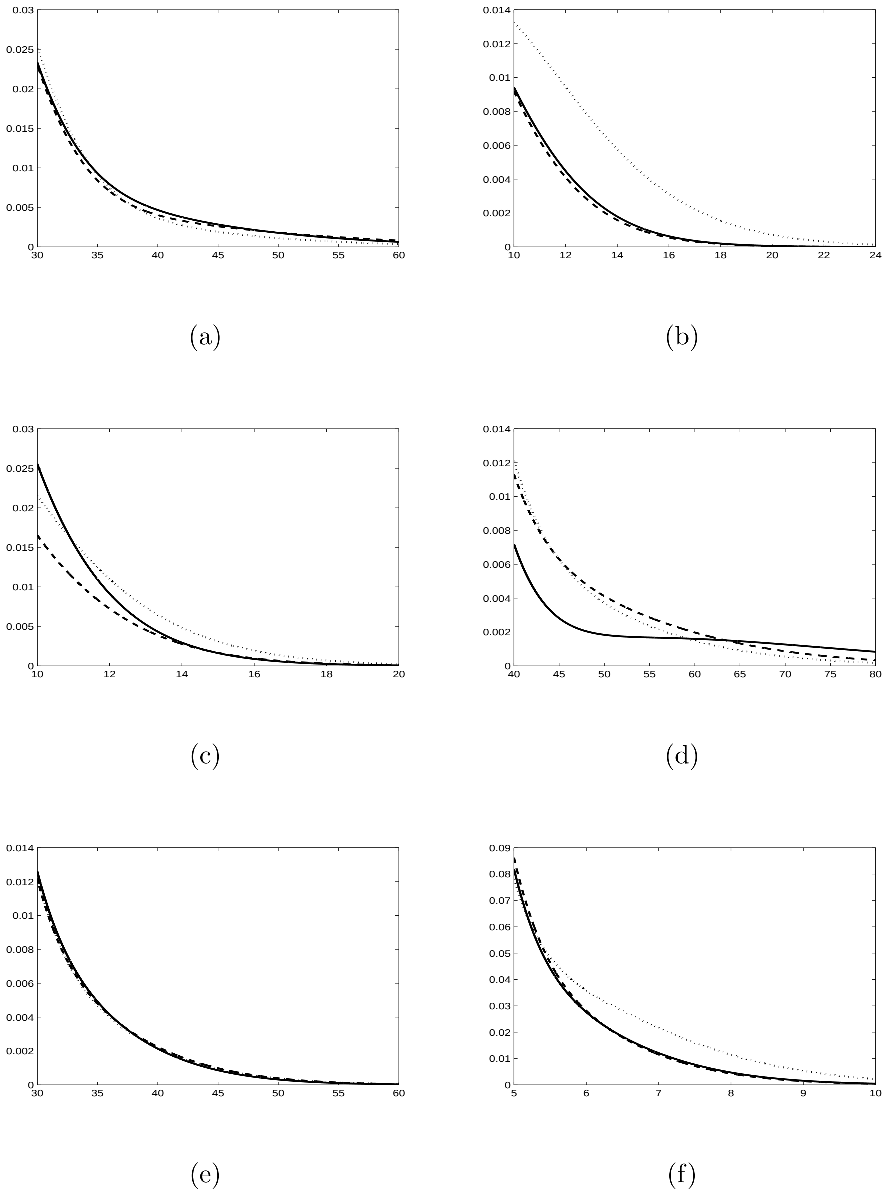

Figure 3: Right tail of the predicted one-step ahead probability density function for MIX (dashed line), TVM (solid line) and Lanne (dotted line) model in correspondence of the lowest and highest values of $\hat{\pi}_t$. Panel (a): DEM/USD, March 9, 1995, $\hat{\pi}_t = 0.7894$. Panel (b): DEM/USD, Dec. 21, 1994, $\hat{\pi}_t = 0.8867$. Panel (c): JPY/USD, June 22, 1988, $\hat{\pi}_t = 0.2013$. Panel (d): JPY/USD, Jan. 18, 1988, $\hat{\pi}_t = 0.9234$. Panel (e): JPY/DEM, Sept. 25, 1992, $\hat{\pi}_t = 0.7670$. Panel (f): JPY/DEM, May 17, 1988, $\hat{\pi}_t = 0.8442$.

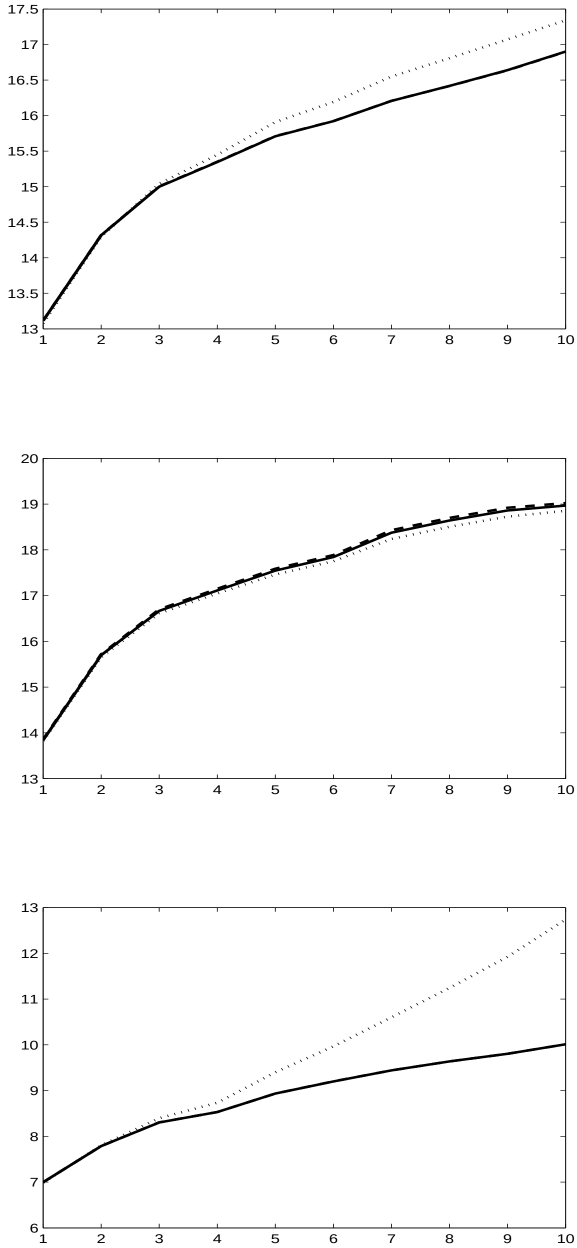

Figure 4: Mean Squared Error at forecasting horizons 1 to 10 for MIX (dashed line), TVM (solid line) and Lanne (dotted line) model, DEM/USD (top), JPY/USD (middle), JPY/DEM (bottom) – Dec. 2, 1986 to Nov. 27, 1996.